Who is in the Family?: Grandparents, Stepkin and Reproduction of Social Inequality Across Generations

By
Jingying He

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Sociology)

at the
UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final oral examination: 9/21/2020

The dissertation is approved by the following members of the Final Oral Committee:

Felix Elwert, Professor, Sociology
Eric Grodsky, Professor, Sociology
Marcy Carlson, Professor, Sociology
Jenna Nobels, Professor, Sociology
Sarah Halpern-Meekin, Associate Professor, Human Development and Family Studies

Acknowledgements

I feel deeply thankful to all the people who have offered me inspiration, challenge, and help throughout my graduate school. The dissertation would not have been possible without my dear advisors, colleagues and friends, and it is truly a fruit of a collective effort.

On my visit day to Madison, a freaking cold and snowy day of March 2014, I had the first meeting with Felix Elwert, my advisor. His approach to research has enormously shaped my own research attitude and taste. Felix is a sociologist of extraordinary character and an advisor who inspires real growth in students. He pushed me to really think of causal inference as a way of sociological theory construction instead of merely a tool of testing hypotheses. This unique approach directly motivated the first two chapters of the dissertation. Felix showed me how to be earnest, honest and perfectionistic toward research. He was dedicated to me like a meticulous parent, writing with the child sentence-by-sentence. Felix continuously challenged me with critique of my ideas, arguments, models and writing, but at same time providing the warmest encouragement. Working with him, I learned how to cherish progress and growth while cultivating self-love.

I also owe my deep gratitude to Eric Grodsky, an outstanding scholar of social stratification and education. He provided me with my first job opportunity in my research career as his research assistant. To a newbie who just entered the department with little research experience, he not only demonstrated great patience but also excellent guidance. Throughout the five years, he constantly provided his insights and suggestions to improve my research, urged me to become a better storyteller, and encouraged me to connect and contribute to the stratification literature. I will never forget the heart-warming care that he offered and his help with navigating the US healthcare system when I was saddled with heart disease in my first year.

I especially thank my dissertation committee members: Marcy Carlson, Jenna Nobles and Sarah Halpern-Meekin who have provided many excellent suggestions, especially from the perspective of family sociology. I would also like to thank Michael Light, with whom I have learned so much about professional development: setting the goals of research, keeping a schedule, submitting to an academic journal, and engaging in the conversation with reviewers. It has been such a joy to have worked in a team with Michael and Jason Robey - we demonstrated the power of backing each other up and uniting as a team. I am grateful to Jason Fletcher, who has provided me with invaluable suggestions on my methods from the economist's perspective and with whom I worked in a research project on multigenerational processes which I enjoyed greatly. I am thankful Ted Gerber, with whom we chatted about our interest in social stratification over lunch.

I owe my deep thankfulness towards Xi Song, with whom I spent my best time in all the PAA meetings, discussing progress, sharing ideas and brainstorming new directions. Her research on multigenerational mobility gave me enormous inspiration. Moreover, she is an excellent role model and mentor for a female sociologist of international background. I am also grateful for Fabian Pfeffer, whose scholarship had a profound influence on me in the early stages of figuring out my field of interest. He also funded me for the presentation and visit to University of Michigan, through which I expanded my network of collaboration.

The dissertation would not be possible without the love and joy I had with numerous colleagues and friends. I am grateful for the contributions of James Raymo, Christine Schwartz, Pablo Mitnik, John Logan, Michael Bell, Michael Engelman, Erik Wright, Sida Liu, James Montgomery, Jia Wang, Yiyue Huangfu, Qian He, Jessica Polos, Ariane Ophir, Isabel Pike, Masoud Movahed, Julia Goodwin, Di Wang, Catherine Doren.

Last, I would like to thank my family members, my parents, Benzhong He and Ping Luo, my grandparents, Guangzhen Qin and Huafu Luo, and my other family members, Xiaomao Li, Vilius Kurauskas, Junlong Feng, Yuhua Li and Charlie the dog, who have given me unconditional love, trust and support.

Table of Contents

List of Tables

List of Figures

Introduction

Families play critical roles of rearing children, securing support for their members and transmitting culture across generations, through which social stratification is reproduced. Past sociological research on cross-generational inequality has been largely based on a nuclear family norm of Western societies consisting of two parents and their children (Becker and Tomes 1986; Blau and Duncan 1967; Conley and Glauber 2008; Erikson and Goldthorpe 2009; Hout and Hauser 1992; Jencks et al. 1983; Sewell et al. 1969). Recent demographic trends may challenge people's understanding of the family and shed new lights on social inequality and mobility, including for instance, the rising shared life-course overlap between generations and the expansion of step families.

This dissertation is concerned with the implications of the changing family on social inequality. The first two chapters address the theoretical and empirical questions regarding the rise of grandparent overlap and grandchildren's cognitive attainment. The third chapter deals with the expanding step families and intergenerational support and transfer to adult children and parents. Abstracts of each chapters are below:

Chapter One

Key problems in social stratification hinge on how the duration of exposure to social contacts affects individuals' outcomes. The overlap effect differs from traditional point estimates by its cumulative and endogenous nature. Conventional methods have failed to identify overlap effects because the effect of unobserved confounders interacts with overlap and constitutes the overlap effect. The first chapter of this dissertation addresses the methodological difficulties in three steps. First, I conceptualize the overlap effects as resulting from cumulative exposure to individuals' observed and unobserved characteristics over time. Second, I develop a flexible formal model in which an overlap effect is (a) confounded by, and (b) vary with individuals' observed and unobserved characteristics. Third, I show that the overlap effect is identifiable in this model and develop a new cumulative fixed effects (CFE) estimator to recover various overlap estimands, such as the average overlap effect, and length-specific overlap effects and the conditional overlap effects of one's baseline characteristics. I substantiate this method with the example of grandparent overlap effect and demonstrate the properties of the cumulative fixed effects estimator by simulation.

Chapter Two

The 20th century witnessed a rise in shared life-course exposure between the grandparent and the grandchildren as life expectancy increased (Song and Mare 2019). The increase of the number of generations alive at the same time and the improved health of the seniors grant more chances of interactions among generations. It is therefore important to investigate whether a longer grandparent overlap perpetuates the reproduction of social inequality across generations. In this chapter I conceptualize the grandparent overlap effect as the total effect of all grandparent time-varying and time-invariant characteristics which prolong with overlap, and formally model the grandparent overlap effect to reflect its cumulative and holistic nature. With the Danish register data, I find positive grandparent overlap effects on children's language ability at grade 6 for all grandparent lineages with a new method, the cumulative fixed effects models. I show that conventional fixed effects would underestimate the grandparent overlap effects tremendously while OLS regression would overestimate such effects. The effects of grandparent overlap effect are also heterogeneous across grandparent lineages, coresident status or health. Given the rising inequality in mortality and the effect heterogeneity of overlap of different social groups, grandparent overlap may amplify the influences of family background to enlarge the gap of children's status attainment of differential social groups.

Chapter Three

The growing prevalence of step families in the United States highlights the fundamental question of family relationship: who is in the family and who is not? The question of family boundary underscores our understanding of the nature of family relationship: why family members help each other intensively. Yet, we know little about how the ambiguity of norms and different motivations play out in family exchanges in complex families. This chapter

explores patterns of dynamics of intergenerational exchanges of interpersonal support, financial transfers and contact in step families; I estimate how parents respond to adult children's past signals of help depending on whether an adult child is step or biological, and how an adult child responds to the parents' past help differently depending on their step/biological status. In contrast to the theory of "biological premium" engaged by past studies which predicts consistently higher biological kin support regardless of the past support, I propose three hypotheses underlying the potential closure of "step gaps": the low-bar expectation, sensitive response and differential convergence. Using the HRS data from 1996 to 2014 and within-family fixed effect models, I find that in spite of step kin's lower interpersonal support, financial transfer and contact when neither shows signals of help, step-kin responds more sensitively to each other's signals of help with a larger increase of parents' monetary transfer, contact and children's senior care support, converging to the biological levels. If the step child has helped in the past, "Biological premium" is still observed for the likelihood of parents' provision of childcare and parents' expectation for future help.

Chapter 1

Identification and Estimation of Grandparent Overlap Effects

Introduction

Social mobility research has embraced the importance of grandparents for status attainment (see Anderson et al. 2018; Bengtson 2001; Mare 2011; Pfeffer 2014 for reviews). Numerous recent studies find that grandparental resources and characteristics are associated with grandchildren's education and occupational outcomes (Anderson et al. 2018). Some even report causal grandparent effects (e.g., Hällsten and Pfeffer 2017; Sharkey and Elwert 2011; Song 2016).

One open question in the literature on grandparent effects remains the role of multigenerational exposure. Most sociologists expect that grandparental influence should increase with the amount of contact between grandparents and grandchildren, where contact is often operationalized as coresidence, geographic distance, or shared life-course overlap (henceforth "grandparent overlap") (Bengtson 2001; Mare 2011; Knigge 2016; Song and Mare 2019; Zeng and Xie 2014). Grandparent overlap, in particular, has garnered attention. As life expectancy increases across cohorts, grandparent overlap increases. If multigenerational influence is proportional to grandparent overlap, then grandparent effects will increase over time. Furthermore, if grandparent effects are more positive in more privileged families, then the increase of grandparent effects with grandparent overlap will exacerbate inequality and diminish intergenerational mobility (Bengtson 2001; Knigge 2016; Lehti et al. 2018; Song and Mare 2019).

Surprisingly few studies, however, have assessed causal effects of grandparent overlap on grandchild outcomes empirically. Most research considers grandparent overlap only as a moderator of the effects of particular grandparent characteristics. For example, Sheppard and Monden (2018), Song and Mare (2019), and Daw et al. (2018) investigate how the association between grandparent education and grandchild education varies with grandparent overlap, and Knigge (2016) investigates the moderating role of grandparent overlap for the multigenerational transmission of occupational status. Results are weak. Neither Daw et al. (2018) nor Sheppard and Monden (2018) find statistically significant interactions between grandparent education and grandparent overlap. Song and Mare (2019) and Knigge (2016) find statistically significant, albeit miniscule interactions. For example, Song and Mare (2019) estimate that an additional year of grandparent overlap increases the effect of an additional year of grandparental education on grandchild education by 0.0013 years (11 hours). To Anderson et al. (2018), the paucity of evidence for larger grandparent effects when grandparents and grandchildren have more contact "casts doubt on a causal interpretation of the grandparent effect" (p.136).

This negative conclusion, however, may be premature. Grandparent contact, and grandparent overlap in particular, may not meaningfully amplify the effect of any single given grandparent characteristic. But since grandparents influence grandchildren via many processes, the overall effect of grandparent overlap on grandchild outcomes via all grandparental characteristics combined may in fact be large. In addition to searching for interactions between grandparent overlap and any one particular grandparental characteristic at a time, sociologists should therefore consider the total effect of grandparent overlap on grandchild outcomes.

Estimating the total effect of grandparental overlap, however, is methodologically challenging. To my knowledge, only one empirical study aims to estimate the overall causal effect of grandparent overlap on grandchild outcomes. Using a linear-probability sibling fixed effects model, Lehti et al. (2018) estimate that an additional year of grandparent overlap increases grandchildren's probability of high school graduation by 1 percentage point ($p < 0.001$), a non-trivial effect.

Their study, however, rests on two simplifying assumptions that cannot easily be relaxed in the conventional (sibling or panel) fixed-effects framework. First, like all conventional sibling fixed-effects models, the study must assume that there are no unmeasured sibling-varying covariates that affect both the exposure and the outcome (Wooldridge 2010). One candidate for such confounders might be grandparents' health. Since grandparents were younger at birth of the older sibling, they likely were also healthier. If healthier grandparents live longer and contribute more to their grandchild's cognitive development (because healthy grandparents can interact more with grandchildren), then failure to control for grandparent health would lead to an overestimation of the grandparent overlap effect. Lehti et al. (2018)'s analysis does not control for sibling-varying grandparent characteristics–nor is it obvious how it could. Since increasing grandparent overlap also increases grandparent's age, and since ageing leads to diminished health, grandparents' evolving health is not only a confounder but also a mediator of the overlap effect, such that controlling for grandparents' evolving health would control away part of the effect of interest.

Second, like all conventional fixed effects models, the model implicitly assumes that the effect of overlap does not vary across families with different unobserved characteristics. If, for example grandparent overlap effects are greater for high-wealth than for low-wealth grandparents (perhaps because of lower cultural and social capital of the latter, see Chan and Boliver 2013; Møllegaard and Jæger 2015), and if this heterogeneity is not explicitly modelled, then conventional fixed effects models at best recover variance-weighted average overlap effects (Wooldridge 2005, 2004). Unmodelled effect heterogeneity is of concern for two reasons. First, variance-weighted average effects are not typically of interest to sociologists (Morgan and Winship 2015). Second, sociologists are specifically interested in how overlap effects vary across grandparent characteristics, as this effect heterogeneity contributes to multigenerational reproduction of social inequality (Mare 2011; Pfeffer 2014; Song and Mare 2019).

This chapter takes up the threefold challenge of conceptualizing, modeling, and identifying the causal effects of grandparent overlap on grandchild outcomes in a novel way. Centrally, I newly conceptualize grandparent overlap effects as resulting from the cumulative exposure to a specific grandparent, and I operationalize the grandparent as the complete bundle of his or her observed and unobserved characteristics.

Conceptualizing the grandparent overlap effect in this way has major consequences for modeling and estimation. Instead of modeling grandparental overlap as a "main effect" term in a conventional regression or (panel or sibling) fixed effects model, I propose a cumulative fixed-effects model in which increases in grandparent overlap *accrue* the effects of all of the grandparent's observed and unobserved characteristics over time. My model permits two roles for fixed unobserved variables in the overlap effect. First, it allows fixed unobservables to confound the effect of grandparent overlap on grandchild outcomes. Second, it allows grandparent overlap effects to vary across fixed unobserved characteristics of the grandparents. Since every grandchild is exposed to a unique combination of grandparental characteristics, this naturally implies heterogeneous, grandchild-specific, overlap effects. This stands in stark contrast to existing work on grandparent effects in sociology, which typically constrains grandparent effects to be constant across most grandchildren, allowing variation only across certain broadly defined groups (e.g. race, family structure, or socio-economic status), and none theorize effect heterogeneity as a function of unobserved confounders.

Estimation of models in which the treatment effect varies with unobservables goes beyond the capabilities of conventional fixed effects methods. Therefore, I develop a two-step estimator for cumulative fixed effects (CFE) models that works both for individual panel data and for sibling panels.

My approach allows analysts to answer a set of new and sociologically interesting questions under fairly general conditions, while also engaging their inherent ambiguity. First, how does an increase in grandparent overlap affect grandchild outcomes on average across the population? Second, how does the effect of increasing grandparent overlap vary with the duration of overlap? Third, how does the grandparent overlap effect vary with grandparent characteristics, e.g., their race, education, or wealth? Fourth, how do these effects depend on the characteristics that the grandparents would possess during their hypothetically granted additional, final, year of life? For example, is the analyst interested in the effect of prolonging grandparents' life by one year in good health or on life support? My approach resolves this ambiguity by conceptualizing overlap effects such that the hypothetical intervention is not just on grandparents' survival, but also on their characteristics in the additional year of life. Thus, analysts will be given the opportunity to ask explicitly about grandparent overlap effects, for example, when grandparents would remain in good health, at death's door, or in whatever health state they are imagined to continue living had they not died.

Estimands

I propose two main estimands that answer two substantively interesting questions about the effects of increasing grandparental overlap on grandchild outcomes. To fix ideas, I focus on grandchild academic test scores as the outcome of interest. In keeping with prior literature, I study the effect of grandparental overlap with respect to only one grandparent, e.g., the paternal grandfather (called "grandparent" henceforth).[1]

Consider a same birth cohort of grandchildren that are observed longitudinally. Let Y_{it}

[1] Past research either focused on paternal grandfathers only (Song et al. 2015), picked any one grandparent among all grandparents (Chan and Boliver 2013; Song 2016; Zeng and Xie 2014), or aggregated over both paternal and maternal lineages (Hällsten and Pfeffer 2017). The extension section generalizes my model to capture the joint overlap effects of multiple grandparents.

be the test score (outcome) of grandchild i at age $t \in 0, ..., T$. The treatment of interest is grandparent overlap, A_{it}, defined as the length of shared life-course overlap between grandchild i and her grandparent up to grandchild's age t. If the grandparent survives to t, then $A_{it} = t$. If the grandparent dies before t then $A_{it} = d_i$, the grandchild's age at the grandparent's death. Hence, $A_{it} \in 0, ..., min(t, d_i)$.

I use potential outcomes notation to define my estimands (Rubin 1974). Let $Y_{it}(a)$ be grandchild i's outcome at age t that would be observed if the grandchild had been exposed to overlap $A_{it} = a$. The individual-level effect of hypothetically increasing grandchild i's actual overlap by one year on grandchild's test score at age t is the individual overlap effect, $IOE_{it} = \{Y_{it}(a_{it}+1) - Y_{it}(a_{it}) | A_{it} < t\}$. Clearly, all effects of increasing grandparent overlap are only defined among grandchildren whose grandparent has died before t, $A_{it} < t$, because, for example, a grandchild cannot have experienced $A_{it}+1 = 10$ years of overlap by age $t = 9$.

The first main estimand is the population average overlap effect ($PAOE_t$), defined as the average effect of increasing every grandchild's actual overlap by one year on grandchild's test scores at age t,

$$PAOE_t = E\big[Y_{it}(A_{it} + 1) - Y_{it}(A_{it}) | A_{it} < t\big]. \tag{1.1}$$

The $PAOE_t$ will often be the primary object of interest, if for no other reason than that it provides a one-number summary of the potentially heterogeneous grandparent overlap effects on grandchildren's outcomes at a specific age, t.

By the same token, however, the $PAOE_t$ likely averages over systematic and sociologically interesting effect heterogeneity. For example, the effect of an additional year of overlap may decrease with the length of overlap because grandparents are getting older, older grandparents are frailer, and frail grandparents may contribute less to their grandchildren's academic achievement. Or, perhaps, the $PAOE_t$ decreases with overlap because, as grandchildren are getting older, they are also getting relatively less receptive to family inputs (Heckman 2006).

To account for possible non-linearities in the grandparent overlap effect across the length of grandparent overlap, I define the length-specific average overlap effect, $LAOE_t(a)$, as the average effect of an additional year of overlap among grandchildren with a given length of overlap, a, on grandchildren's outcomes at age t,

$$LAOE_t(a) = E\big[Y_{it}(A_{it}+1) - Y_{it}(A_{it})\big|A_{it} = a < t\big] \tag{1.2}$$

The definitions of the IOE_{it}, $PAOE_t$, and $LAOE_t(a)$ are generic, in the sense that they do not depend on any particular theory about the data generation process (DGP) of these grandparent overlap effects. Both the $PAOE_t$ and the $LAOE_t(a)$ likely mask effect heterogeneity by the family's characteristics, such as grandparent's race, education, or age at grandchild's birth; or grandchildren's gender or place of birth. To account for possible heterogeneity across baseline characteristics, I define the conditional population-average overlap effect, $CPAOE_t^b$, and the conditional length-specific average overlap effect, $CLAOE_t^b(a)$, as the $PAOE_t$ and $LAOE_t(a)$ among the subgroup of grandchildren with baseline characteristics $B = b$.

When discussing identification and estimation of the grandparent overlap effects below, I will see that, regardless of the DGP, all estimands necessarily depend on the specific characteristics that the grandparent would possess during their additional year of life. Highlighting this fact, I believe, is a contribution of my conceptualization of grandparent overlap effects as the cumulative effects of exposing grandchildren to grandparents with specific fixed and time-varying characteristics. I explicate this notion next.

Table 1.1 summarises the notation and definitions used throughout the chapter, some of which will be introduced later.

Conceptualizing Grandparent Overlap as Cumulative Exposure

This section conceptualizes grandparent overlap effects as the cumulative effects of grandchildren's exposure to grandparent's observed and unobserved characteristics across the length of overlap. Equation 1.3 gives the assumed data generating process (DGP), i.e. my theory of how grandparents affect grandchildren's test scores.

$$Y_{it} = \beta_0 + \sum_{t'=0}^{A_{it}} C_{it'}\beta_{t'} + U_{i1} + U_{i2}A_{it} + \epsilon_{it} \tag{1.3}$$

In this model, grandchild test scores at age t, Y_{it}, are affected by grandparent's observed and unobserved characteristics, $C_{it'}$ and U_i, across the entire course of overlap, from 0 to A_{it}. As usual, β_0 and ϵ_{it}, represent a shared intercept and idiosyncratic, period-specific, mean-zero error terms. The heart of the model resides in the middle three terms. The term $\sum_{t'=0}^{A_{it}} C_{it'}\beta_{t'}$ captures the effect of the history of grandparent's observed time-varying characteristics, $\bar{C}_{it} = \{C_{i0}, ..., C_{iA_{it}}\}$, from grandchild's birth at $t' = 0$ across the length of overlap, A_{it}, where t' is the index of summation. At each period t', $C_{it'}$ is a vector of observed grandparent's characteristics that affect grandchild's test scores, including, for example, grandparent's health, income, labor force participation, marital status, and coresidence with the grandchild. Past $C_{it'-s}$, $s > 0$, are permitted to affect present $C_{it'}$. Time-varying effects, $\beta_{t'}$, acknowledge, that a given grandparental characteristic may have different effects depending on the age at which the grandchild experienced the characteristic. For example, if $C_{it'}$ is a time-varying indicator of grandparent's ill health at grandchild's age t', and $\beta_2 \neq \beta_{15}$, then grandparent's illness when the grandchild was age $t' = 2$, has a different effect on grandchild's test score at age $t \geq 15$, than does grandparent's illness at age $t' = 15$.

Grandchild's test scores are also affected by grandparent's fixed characteristics, which may be observed or unobserved, such as grandparent's education, values, genes, and class background. As a key innovation, I recognize that grandparent's fixed characteristics may exert both a constant ("fixed") effect on grandchild's test scores at a given child's age, captured by the term U_{i1}, and an effect that accumulates with overlap, captured by the term $U_{i2}A_{it}$. For example, the grandparent's class position may afford the grandchild a fixed bonus of privileged attention from teachers at every age, U_{i1}, and each additional year of grandchild's interaction with the exemplars of the family's privilege may confer additional advantage via building grandchild's sense of entitlement, $U_{i2}A_{it}$. The model remains agnostic whether any given fixed characteristic exerts both a fixed and a cumulative effect, only a fixed effect, or only a cumulative effect, i.e. whether it is an element of U_{i1}, U_{i2}, or both.

The total grandparent overlap is generated by the two terms containing overlap, A_{it}, namely $\sum_{t'=0}^{A_{it}} C_{it'}\beta_{t'}$ and $U_{i2}A_{it}$ which I call the observable and unobservable components of the grandparent overlap effect, respectively.[2] The individual-level causal effect of one additional year of grandparent overlap on grandchild i in my model is given by

$$IOE_{it} = \{Y_{it}(a_{it}+1) - Y_{it}(a_{it})|A_{it} < t\} = \{C_{ia_{it}+1}\beta_{a_{it}+1} + U_{i2}|A_{it} < t\}. \qquad (1.4)$$

The $LAOE_t(a)$ gives the average grandparent overlap effect among children who experience a given length of overlap, $A_{it} = a$,

$$\begin{aligned} LAOE_t(a) &= E\big[Y_{it}(A_{it}+1) - Y_{it}(A_{it})|A_{it} = a < t\big] \\ &= E\big[C_{iA_{it}+1}\beta_{A_{it}+1} + U_{i2}|A_{it} = a < t\big]. \qquad (1.5) \end{aligned}$$

[2]By contrast, U_{i1} is a grandparent effect but not part of the grandparent overlap effect, since it captures the grandparental influence that would exist even if the grandparent had died before grandchild's birth, $A_{it} = 0$. For example, having had a wealthy grandparent, U_{i1} may confer advantages to the grandchild even if the grandparent died before the grandchild's birth.

The $PAOE_t$ averages across all IOE_{it}'s in the entire population,[3]

$$\begin{aligned} PAOE_t &= E\big[Y_{it}(A_{it}+1) - Y_{it}(A_{it}) | A_{it} < t\big] \\ &= E\big[C_{iA_{it}+1}\beta_{A_{it}+1} + U_{i2} | A_{it} < t\big] \end{aligned} \tag{1.6}$$

I add three remarks. First, readers will notice that my expressions for grandparent overlap effects do not include a "main effect" for overlap, A_{it}. This visible departure from past sociological research on overlap effects is intentional. My model highlights that overlap generates cumulative exposures to grandparent's observed and unobserved characteristics. Overlap cannot exert an effect on grandchild test scores net of these characteristics, and hence does not merit a main-effects term.

Second, the grandparent characteristics during the hypothetically granted additional year of life, $C_{iA_{it}+1}$, are properly thought of as part of the hypothetical intervention of prolonging the grandparent's life. The values of $C_{iA_{it}+1}$ are by definition unobservable and must be explicated by the analyst. In other words, when defining the effect of hypothetically prolonging a grandparent's life, the analyst owes the reader a statement as to what characteristics the grandparent is meant to possess during their additional year of life. For some classes of grandparent deaths (e.g. freak traffic accidents), it is easy to imagine a counterfactual scenario in which the grandparent would naturally live another year in good health. In other cases (e.g. death from old age), however, the natural counterfactual would be living at death's door, perhaps in a coma. Instead of hiding the implicit counterfactual, my approach thus empowers–indeed, requires–analysts to specify the desired grandparental characteristics as additional hypothetical intervention variables.[4] The simplest way to specify these characteristics would be to let the analyst choose for themselves, as if by external intervention.

[3]The $PAOE_t$ is also the average of the $LAOE_t(a)$ s, $E_{A_{it}}\Big[E\big[Y_{it}(a+1) - Y_{it}(a) \,|\, A_{it} = a < t\big]\Big]$.

[4]The requirement of clearly stating the counterfactual is not unique to conceptualizing overlap effects. For example, in studies of the effect of parental divorce on child outcomes, analysts should similarly specify whether they imagine the counterfactual to divorce to be "happy marriage" or "marriage at the verge of divorce." Unfortunately, conventional estimation approaches do not typically force a detailed explication of the counterfactual.

That being said, analysts should pause to consider which values of $C_{iA_{it}+1}$ are both interesting and plausible. To this end, they might want to predict (or even structurally model) these characteristics from the history of grandparents' prior characteristics, as discussed below.

Third, honoring the growing sociological interest in effect heterogeneity (Brand and Thomas 2013; Elwert and Winship 2014; Xie 2013), my model allows individual-level grandparent overlap effects to vary for several reasons: The IOE_{it}s may vary across grandchildren, i, at a fixed age, t, because (1) grandparents may have different characteristics, $C_{ia_{it}+1}$, in the hypothetically granted additional year of life, (2) the effect of these characteristics may vary with the duration of overlap since $\beta_{a_{it}+1}$ varies with a_{it}, and (3) grandparents may vary with respect to their fixed characteristics, U_{i2}. The $LAOE_t(a)$ and $PAOE_t$ estimands simply average across these heterogenous individual-level effects in different ways but retain variation due to grandparents' observed and unobserved characteristics.

Like all models, my model contains simplifications. First, it asserts that the cumulative effects of grandparents' fixed unobserved characteristics are linear in overlap, $U_{i2}A_{it}$–although this assumption could be relaxed with respect to the observed characteristics in U_{i2}. Second, it rules out unobserved time-varying confounders–a property it shares with conventional panel fixed effects models. Third, it imposes a restriction on how past grandparental characteristics may affect test scores in the present via $\beta_{t'}$, in that the effect of grandparent inputs at a given age is not allowed to decay over time. I consider alternative specifications in Extensions section.

Identification

The parameters of the overlap model in Equation 1.3 are identified under the assumption of strict exogeneity. Remarkably, this is the same assumption that also underlies the identification of more conventional panel fixed effects models that do not contain cumulative fixed effects (Chamberlain 1982). Consider a balanced panel of $i \in 1, ..., N$ grandchildren observed for $t' \in 1, ..., T$ periods. The strict exogeneity assumption,

$$E(\epsilon_{it'}|\bar{C}_{iA_{iT}}, A_{iT}, U_i) = 0, \tag{1.7}$$

states that, for each grandchild i and each period t', the idiosyncratic period-specific error term, $\epsilon_{it'}$, is conditionally mean independent of (1) the past and future history of grandparent's observed time-varying characteristics across the entire period of overlap, $\bar{C}_{iA_{iT}}$, (2) overlap itself, A_{iT}, (i.e. grandparent mortality in each period), and (3) the fixed characteristics, $U_i = (U_{i1}, U_{i2})$.

Substantively, this assumption says that, 1) past outcomes may not directly affect the current exposures, and 2) there are no unobserved time-varying confounders after controlling for grandparent's observed covariate history, $\bar{C}_{iA_{it}}$ (Sobel 2012; Imai and Kim 2019). The first claim is immediately plausible: surely, grandchildren's past academic test scores do not cause grandparent's survival, A_{it}, or observed time-varying confounders, $C_{it'}$, such as income, or health. The second claim regarding the absence of time-varying unobservable confounders will have to be justified by arguing that the covariates, $\bar{C}_{it'}$, observed in a particular empirical application are sufficiently detailed to control for time-varying confounding.

I add three remarks on what variables should or could be controlled for. First, the analysis must control for time-varying grandparent characteristics, $C_{it'}$, that are (a) confounders of grandparent overlap (i.e. grandparent mortality) and grandchild test scores, or (b) confounders of future grandparent characteristics and grandchild test scores. Clearly, these factors should include, for example, detailed histories of grandparent's income, labor force participation, health, marital history, and coresidence.

Second, the analysis may (but need not) additionally control for grandparent characteristics, $C_{it'}$, that are mediators on the causal pathway from past observed or unobserved grandparent characteristics, even if these mediators are not themselves confounders.

Third, analysts should be careful about additionally controlling for time-varying characteristics of parents or grandchildren. Such characteristics must be included if they are confounders of (a) grandchild test scores and grandparent mortality, or (b) grandchild test scores and grandparent's observed time-varying characteristics. This inclusion is unproblematic if these parent or grandchild characteristics are not themselves mediators of the effects of past grandparent characteristics on grandchild test scores (otherwise, the analysis will incur overcontrol bias). Nonetheless, one should note that time-varying parent and grandchild characteristics that are confounders but not mediators do not contribute to the grandparent overlap effect and hence must be handled differently than grandparent characteristics (which do contribute to the overlap effect) in the estimation stage, as explained in the Extensions section.

The Cumulative Fixed Effects Estimator

Under the strict exogeneity assumption, the $PAOE_t$ and $LAOE_t(a)$ generated by the model in Equation 1.3 are identifiable over some range of overlap durations from individual-level panel data with at least three waves of outcomes and detailed histories of grandparent characteristics.[5] Estimation involves four steps. The first two steps exploit within-individual variation in the outcome to remove the main effect of the unobservables, U_{i1}, and their interaction with grandparent overlap, $U_{i2}A_{it}$, respectively. These first two steps identify the parameters, $\beta_{t'}$, on the observed time-varying grandparent characteristics, $C_{it'}$. The third step recovers the contribution of the fixed unobservables, U_{i2}, to the grandparent overlap effect. In the fourth step, the analyst chooses the time-varying characteristics, $C_{iA_{it}+1}$, that the grandparent should posses during their additional year of life. With these inputs, one can then compute the $LAOE_t(a)$ and $PAOE_t$ of an additional year of overlap with the chosen characteristics for certain durations of overlap.

[5]Alternatively, one could estimate the overlap effects from panel data on sibling pairs with two waves of outcomes. The extension is straight-forward. I focus on individual-level panels henceforth.

Data Structure

Consider individual-level panel data that follow one cohort of grandchildren and their grandparents from the time of grandchildren's birth, $t' = 0$, until the outcome, Y_{it}, has been measured three times. Let the grandchildren's age at the three outcome assessments ("test dates") be $q_1 < q_2 < q_3$. The test dates are shared across grandchildren in the cohort but do not need to be equally spaced. For reasons that will become apparent shortly, I restrict the analysis to grandchildren whose grandparent does not die until after the second test date, $d_i > q_2$. Therefore, $A_{iq_1} = q_1$, $A_{iq_2} = q_2$, and $q_2 < A_{iq_3} \leq q_3$. Examples of suitable data sets might include research panels or administrative data linking families across generations in areas that maintain regular educational testing programs, like those available in some states of the United States, Europe, or Asia (Song and Campbell 2017).

Steps 1 and 2: Estimation of $\beta_{t'}$

I start by transforming the data in order to eliminate the dependence of the outcome on the fixed unobserved confounders. Since my model permits that the unobservables exert not only a fixed effect, U_{i1} on the outcome, but also interact with the duration of overlap, $U_{i2}A_{it}$, the required transformations are somewhat more involved than in conventional panel fixed effects estimation.

The first step takes the conventional first difference of Equation 1.3 between consecutive outcome waves twice, in order to eliminate the main effect of the unobservables, U_{i1}, yielding

$$\triangle_{q_2q_1} Y_i = \sum_{t'=q_1+1}^{A_{iq_2}} C_{it'}\beta_{t'} + U_{i2} \triangle_{q_2q_1} A_i + \triangle_{q_2q_1} e_i, \tag{1.8}$$

and

$$\triangle_{q_3q_2} Y_i = \sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\beta_{t'} + U_{i2} \triangle_{q_3q_2} A_i + \triangle_{q_3q_2} e_i, \tag{1.9}$$

where $\Delta_{ab}V_i = V_{ia} - V_{ib}$. These first-differenced equations, cannot yet be consistently estimated by regression, because the terms involving U_{i2} are unobserved confounders.

The second step therefore removes the dependence on U_{i2} by first rescaling and then differencing Equation 1.8 and Equation 1.9. Specifically, I first divide Equation 1.8 by $\triangle_{q_2q_1}A_i$,

$$\frac{\triangle_{q_2q_1}Y_i}{\triangle_{q_2q_1}A_i} = \frac{\sum_{t'=q_1+1}^{A_{iq_2}} C_{it'}\beta_{t'}}{\triangle_{q_2q_1}A_i} + U_{i2} + \frac{\triangle_{q_2q_1}e_i}{\triangle_{q_2q_1}A_i}, \tag{1.10}$$

then divide Equation 1.9 by $\triangle_{q_3q_2}A_i$,

$$\frac{\triangle_{q_3q_2}Y_i}{\triangle_{q_3q_2}A_i} = \frac{\sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\beta_{t'}}{\triangle_{q_3q_2}A_i} + U_{i2} + \frac{\triangle_{q_3q_2}e_i}{\triangle_{q_3q_2}A_i}, \tag{1.11}$$

and finally subtract Equation 1.11 from Equation 1.10 ,

$$\frac{\triangle_{q_2q_1}Y_i}{\triangle_{q_2q_1}A_i} - \frac{\triangle_{q_3q_2}Y_i}{\triangle_{q_3q_2}A_i} = \frac{\sum_{t'=q_1+1}^{A_{iq_2}} C_{it'}\beta_{t'}}{\triangle_{q_2q_1}A_i} - \frac{\sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\beta_{t'}}{\triangle_{q_3q_2}A_i} + \frac{\triangle_{q_2q_1}e_i}{\triangle_{q_2q_1}A_i} - \frac{\triangle_{q_3q_2}e_i}{\triangle_{q_3q_2}A_i}. \tag{1.12}$$

Equation 1.12 no longer depends on U_{i1} or U_{i2}. Furthermore, the two terms involving the differenced error terms, e_{it}, are mean independent of the terms involving covariates, $C_{it'}$, under the assumption of strict exogeneity (Equation 1.7). Therefore, the $\beta_{t'}$ parameters in Equation 1.12 can be consistently estimated by OLS regression.

I add two remarks. First, it is obvious that this estimator requires that grandparents survive beyond the second test date. Otherwise, Equation 1.11 would involve dividing by $\triangle_{q_3q_2}A_i = 0$. Second, the estimator only uses information between the first and third test date, q_1 and q_3. Therefore, it can only recover the parameters $\beta_{t'}$ for observed grandparent characteristics between these dates, $t' \in [q_1 + 1, q_3)$. Clearly, this is not much of a limitation if the tests are widely spaced apart.

Estimation of the $PAOE_t$ and the $LAOE_t(a)$, however, requires not only the coefficients for the grandparent characteristics, $\beta_{t'}$, but also knowledge of the observable and unobservable components of the grandparent overlap effect, defined in the Estimands section. I turn to their estimation next.

Step 3: Estimation of the Unobservable Component

Estimating the contribution of the unobservable components–$E(U_{i2}|A_{it} < t)$ for the $PAOE_t$ and $E(U_{i2}|A_{it} = a < t)$ for the $LOAE_t(a)$–is difficult. Recognizing that these components are part of the residuals, however, I can recover them between the second and third test date, $E(U_{i2}|q_2 < A_{it} < q3)$ and $E(U_{i2}|A_{it} = a, q_2 < A_{it} < q_3)$. This implies that I can identify the $PAOE_t$ and $LAOE_t(a)$ only for durations of overlap corresponding to deaths between the second and third test date.

Subtracting the observables from Equation 1.9 yields

$$\triangle_{q_3q_2} Y_i - \sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\hat{\beta}_{t'} = U_{i2} \triangle_{q_3q_2} A_i + \triangle_{q_3q_2} e_i. \tag{1.13}$$

Taking the conditional expectation gives

$$\begin{aligned} &E\big(\triangle_{q_3q_2} Y_i - \sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\hat{\beta}_{t'} \big| A_{iq_3}, \bar{C}_{iq_3}\big) \\ &= E\big(U_{i2} \triangle_{q_3q_2} A_i \big| A_{iq_3}, \bar{C}_{iq_3}\big) + E\big(\triangle_{q_3q_2} e_i \big| A_{iq_3}, \bar{C}_{iq_3}\big). \end{aligned} \tag{1.14}$$

Notice that $E(\triangle_{q_3q_2} e_i|A_{iq_3}, \bar{C}_{iq_3})$=0 under the strict exogeneity assumption. Dividing equation Equation 1.14 by $\triangle_{q_3q_2} A_i$ and rearranging terms expresses the conditional residual in terms of the observables,

$$E(U_{i2}|A_{iq_3}, \bar{C}_{iq_3}) = E\frac{\big(\triangle_{q_3q_2} Y_i - \sum_{t'=q_2+1}^{A_{iq_3}} C_{it'}\hat{\beta}_{t'} \big| A_{iq_3}, \bar{C}_{iq_3}\big)}{A_{iq_3} - A_{iq_2}}. \tag{1.15}$$

Marginalizing over the conditioning arguments then gives the desired expressions.

For example, consider a model where $C_{it'}$ includes only two grandparent characteristics, income, $I_{it'}$, and health, $H_{it'}$. Recalling that the $PAOE_t$ is defined only for $A_{it} < t$, I would use Equation 1.15 to estimate the unobservable term $E\big[U_{i2}|q_2 < A_{iq_3} < q_3\big]$ in the $PAOE_t$ as

$$\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{Y_{i,q_3} - Y_{i,q_2} - \sum_{t'=q_2+1}^{t'=A_{iq_3}} I_{it'}\hat{\beta}_{t'} - \sum_{t'=q_2+1}^{t'=A_{iq_3}} H_{it'}\hat{\beta}_{t'}}{A_{i,q_3} - q_2}\Big), \tag{1.16}$$

where m is the number of grandchildren in the sample whose grandparent die after q_2 and before q_3. The analogous estimate for the $E\big[U_{i2}|A_{i,q_3} = a, q_2 < A_{iq_3} < q_3\big]$ term in the $LAOE_t(a)$ would be

$$\frac{1}{m_a}\sum_{i=1}^{m_a}\Big(\frac{Y_{i,q_3} - Y_{i,q_2} - \sum_{t'=q_2+1}^{t'=a} I_{it'}\hat{\beta}_{t'} - \sum_{t'=q_2+1}^{t'=a} H_{it'}\hat{\beta}_{t'}}{a - q_2}\Big), \tag{1.17}$$

where m_a is the number of observations of the subgroup with the length of overlap at the third test equal to a, i.e. $A_{i,q_3} = a$.

Step 4: Choosing Grandparent Characteristics During the Extended Overlap Period

As a last step, I need to choose values for the observable components: $E\big[C_{iA_{iq_3}+1}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3]$ for the $PAOE_t$, and $E\big[C_{iA_{iq_3}+1}|q_2 < A_{iq_3} < q_3]$ for the $LOAE_t(a)$, respectively. Because grandparents' characteristics are only measured until they die, the grandparent characteristics in the extended overlap period, $C_{iA_{iq_3}+1}$, are in nature hypothetical and unobserved. I regard the mean of the grandparent characteristics $E(C_{iA_{iq_3}+1}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3)$ as intervention variables, whose values have to be chosen by the analyst in order to explicate the precise counterfactual to having a dead grandparent.

Analysts have wide latitude in choosing theoretically interesting and substantively plausible values for $E\big[C_{iA_{iq_3}+1}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3]$ and $E\big[C_{iA_{iq_3}+1}|q_2 < A_{iq_3} < q_3]$. Here, I list three different approaches for illustration. First, sociologists could ask what the grandparent overlap effect would be if the deceased grandparent's characteristics were to remain unchanged from the last period prior to grandparent's actual time of death, $C_{iA_{iq_3}+1} = C_{iA_{iq_3}}$. Indeed, when the time series of the grandparents' time-varying characteristic are mean stationary, the last observation would be an unbiased prediction of the future realization (Hyndman and Athanasopoulos 2018), i.e. $\hat{E}\big[C_{iA_{iq_3}+1}|q_2 < A_{iq_3} < q_3\big] = E\big[C_{iA_{iq_3}}|q_2 < A_{iq_3} < q_3\big]$ and $\hat{E}\big[C_{iA_{iq_3}+1}\big|A_{iq_3} = a, q_2 < A_{iq_3} < q_3\big] = E\big[C_{iA_{iq_3}}\big|A_{iq_3} = a, q_2 < A_{iq_3} < q_3\big]$.

Second, when the time series of grandparent time-varying characteristics are not stationary, analysts might assume that the average characteristics in the extended overlap period of the grandparents who die at a would equal the average characteristics of the grandparents who die in the subsequent period, $a+1$. For example, $\hat{E}\big[C_{iA_{iq_3}+1}\big|A_{iq_3} = a, q_2 < A_{iq_3} < q_3\big] = \frac{1}{m_{a+1}} \sum_{i=1}^{m_{a+1}} C_{iA_{iq_3}+1}$, where m_{a+1} is the number of grandparents who die at $A_{it} = a + 1$.

Third, to better reflect each grandparent's individual life-course trajectory, sociologist could use the history of grandparent's time-varying characteristics, $\bar{C}_{iA_{iq_3}}$, to predict $C_{iA_{iq_3}+1}$, for example using auto-regressive models or machine learning approaches.

Clearly, methods of forecasting the grandparent's future characteristics (if they had stayed alive) are beyond enumeration. Different assumptions about the time-series of grandparent characteristics will generally lead to different estimates for the various grandparent overlap effects. It is an advantage that my method remains agnostic about the process that would generate future grandparent characteristics. Whatever the choice of future grandparent characteristics that the analyst wishes to defend as realistic of informative can be plugged into the estimator. What is more, analysts could even entirely eschew the prediction of future grandparent characteristics and freely choose desired reference values to answer hypothetical questions, e.g. about the effect of another year with a healthy grandparent or a rich grandparent. The choice of the research question-via the choice of the counterfactual characteristics of the grandparent-is not mandated by nature, but is the choice of the researcher.

Finally, in order to estimate the $PAOE_t$ and $LAOE_t(a)$, analysts would substitute the estimated parameters β_t and estimated unobservable components $E[U_{i2}|q_2 < A_{iq_3} < q_3]$ and

$E[U_{i2}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3]$ together with the chosen values for the observable components $E\big[C_{iA_{iq_3}+1}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3\big]$ and $E\big[C_{iA_{iq_3}+1}|q_2 < A_{iq_3} < q_3\big]$ into Equation 1.6 and Equation 1.5. Estimates of $PAOE_t$ and $LAOE_t(a)$ will be unbiased given the chosen values of $C_{A_{iq_3}+1}$ as long as the estimation of β_t and $E[U_{i2}|q_2 < A_{iq_3} < q_3]$ and $E[U_{i2}|A_{iq_3} = a, q_2 < A_{iq_3} < q_3]$ are unbiased. Standard errors could be computed using bootstrap methods.

In sum, one can estimate $LAOE_t(a)$ and $PAOE_t$ following the four steps introduced above. The above estimation strategy is easily extended to the estimation of conditional grandparent overlaps effects, $CPAOE_t^b$ and $CLAOE_t^b(a)$, that trace effect heterogeneity across groups defined by the family's fixed or baseline characteristics, B, including for instance, the grandparent's gender, race, education, income and health at the baseline wave, or grandchild's gender, or region of birth. The easiest way to estimate conditional overlap effects is to stratify the sample into groups, defined by B, and then apply the estimator separately to each group. I illustrate this subsample estimation approach in the simulations below.

Simulation

To evaluate the statistical properties of my estimator, I specify a model for the effect of grandparent overlap on grandchild test scores with known parameters and simulate the sampling distributions of the estimates. This section outlines the simulation procedure and presents key results. (Details of the simulation procedure are shown in the Appendix A.)

Data Generation

I specify a data-generating model that elaborates on the the endogenous process outlined above. The process evolves in annual increments from grandchildren's birth to age 20. For simplicity, each grandparent has exactly one grandchild. I first generate exogenous inputs for each grandparent-grandchild pair as i.i.d. draws from various distributions: The grandparent-level fixed unobservables that confound grandparent overlap and grandchild test scores are drawn from $U_i \sim N(10, 100)$, where each fixed unobservable is allowed to affect the outcome as a fixed effect irrespective of overlap and in interaction with overlap, so that $U_{i1} = U_{i2}$. The idiosyncratic age-specific errors in grandchildren's test scores are drawn from $\epsilon_{it} \sim N(0, 100)$. Grandparents' age at grandchild's birth is uniformly distributed between ages 50 and 60 as a function of U_i, since mothers' (and hence grandmothers') age of childbearing correlates, for example, with their socio-economic status (Fomby et al. 2014).

Next, I posit two time-varying grandparental characteristics, $C_{it'} = \{I_{it'}, H_{it'}\}$, income and health, that co-evolve as a function of (i) each other's baseline values, C_{i1}, (which are themselves function of age); (ii) each other's most recent values, $C_{it'-1}$; (iii) the fixed effects; and (iv) their own time-trends. The parameters of the process are tuned such that grandparents' health decreases over time and grandparents' income increases over time within individuals, with increasing variance across individuals, as is common in western countries (Deaton and Paxson 1994). The resulting time-series are non-stationary (i.e., with changing means and variances over time).

Grandparent overlap is endogenously determined by all of the above inputs. I generate grandparent's survival so that it negatively depends on grandparents' age and positively depends on their income and good health.

Finally, from these variables, I generate grandchild test scores, Y_{it} at three test dates, $q_1 = 6$, $q_2 = 10$ and $q_3 = 20$, according to the reduced-form DGP of Equation 1.3. I posit that grandparents' observed characteristics measured at times t' affect grandchildren's test scores with parameters $\beta_{t'I} = \beta_{t'H} = 5 + 0.5 * (t' - 1)$, so that grandparent's income and health become more important as the grandchild ages, and recent exposures matter more than earlier exposures. Approximately 89, 84, and 50 percent of grandparents survive beyond the first, second, and third test date, respectively.

Results

From this process, I generate 2,000 data sets, each containing $N = 100,000$ grandparent-grandchild pairs, i, to simulate the relevant sampling distributions. This sample size roughly corresponds to those available in typical population registers. Estimation follows the four-step procedure developed in the cumulative fixed effects estimator section.

Using the first two steps, I estimate the parameters of grandparents' time-varying characteristics, $\beta_{t'}$, between the first and third test date. Figure 1.1 shows the sampling distributions of the coefficients for the effects of grandparent's health and income at grandchildren's age $t' = 18$, $\beta_{18,I}$ and $\beta_{18,H}$, for illustration. (The sampling distributions for the effects of income and health experienced at the remaining ages are similar, available upon request.) The sampling distributions are approximately normal and centered around the true parameter values of the DGP. Thus, the parameter estimates are unbiased.

Using the third step, I estimate the contribution of the unobserved component, involving U_{i2}, to the grandparent overlap effect. Since the contribution of the unobserved component varies across different grandparent overlap effects, Figure 1.2 specifically illustrates the sampling distribution of the unobserved component $\widehat{E}(U_{i2}|A_{i20} = 17)$ of the $LAOE_{20}(17)$- the length-specific average overlap effect of experiencing 18 rather than 17 years of overlap on test scores at age 20 among grandchildren who actually experience 17 years of overlap. Again, the estimate is approximately normal and unbiased.

In the fourth step, I estimate various grandparent overlap effects. For illustration, I choose to estimate the effects that would obtain, if the grandparents who died had lived another year and, in their additional year of life, had possessed the observable characteristics of the grandparents who actually did live another year, $C_{i,a+1}$. Table 1.2 shows the true values and the average estimates of the three components constituting the length-specific average overlap effects on test scores at age 20, $LAOE_{20}(a) = E\big[C_{a+1}\beta_{a+1} + U_{i2}\big|A_{it} = a\big]$, for various specific lengths of overlap between the second and third test date, $10 < a < 20$: (i) the parameters on grandparent income and health, $\beta_{a+1,I}$ and $\beta_{a+1,H}$; (ii) the mean of the unobservable components, $E(U_{i2}|a)$; and (iii) the investigator-chosen value of $E(C_{i,a+1}|a)$, which contains the averages of the values to which I suppose grandparents' income and health would be set if grandparents' life would be extended by one year to age $a+1$. Figure 1.3 shows the sampling distribution for the average overlap effect on grandchild test scores at age 20 for grandchildren who in fact experienced 17 years of overlap, $LAOE_{20}(17)$. The mean of this sampling distribution can also be computed from the penultimate line of Table 1.2,

$$\overline{\widehat{LAOE}}_{20}(a=17) = \bar{\hat{E}}\big[C_{a+1}\beta_{a+1}+U_{i2}|A_{it}=17\big] = \bar{\hat{\beta}}_{18,I} * \bar{I}_{i,18} + \bar{\hat{\beta}}_{18,H} * \bar{H}_{i,18} + \bar{\hat{E}}\big[U_{i2}|A_{it}=17\big] \tag{1.18}$$
$$= 13.49 * 438.55 + 13.48 * -285.07 + 70.33 = 2143.62 \approx 2142.00^6$$

Figure 1.4 shows the true and estimated $LAOE_{20}(a)$s between the second and third test date graphically. All estimates are unbiased and approximately normal.

Figure 1.5 shows the sampling distribution of the population-average grandparent overlap effect, $PAOE_{20}$, which is the average effect of increasing grandparent overlap by one year on grandchild test scores at age 20 among grandchildren whose grandparents died in any year between the second and third test date (when grandchildren were aged 11 to 19), $PAOE_t = E\big[C_{A_{it}+1}\beta_{A_{it}+1} + U_{i2}|10 < A_{it} < 20\big]$. The $PAOE_{20}$ is simply the average of the $LAOE_{20}(a)$s between the ages of 11 and 19, shown in Table 1.2, weighted by the fraction of grandparents who died at each of these ages. The estimate for the $PAOE_{20}$ is unbiased and approximately normal.

The overlap effects estimated in Table 1.2 exhaust the average and length-specific grandparent overlap effects, $PAOE_t$ and $LAOE_t(a)$ that can be estimated when the second and third scores are measured at grandchild ages 10 and 20. However, these effects would change if the analyst chose to explore different counterfactual values than I did here for grandparents' additional year of life.

Analysts could further explore effect heterogeneity by conditioning the estimation on particular values of grandparent's observed covariate history, $\bar{C}_{it'}$, or by exploiting additional test scores measured at different ages.

To show an example of the effect heterogeneity by conditioning on different grandparent's observed covariates, I estimate the conditional grandparent overlaps effects, $CPAOE_t^b$ and $CLAOE_t^b(a)$ by stratifying the sample with baseline grandparent income and health, $\overline{\widehat{CPAOE}_t}^b(a) = \overline{\widehat{PAOE}}_{20}(a|C_{it'=1})$ and $\overline{\widehat{CLAOE}_t}^b(a) = \overline{\widehat{LAOE}}_{20}(a|C_{it'=1})$. Specifically, I follow the same four-step procedure and estimate with the same simulated sample which is stratified into four subgroups based on grandparent income and health at $t' = 1$ (0%-25%, 25%-50%, 50%-75%, and 75%-100% quantiles). All these subgroup estimates are unbiased to their true values. Besides, I find a greater total grandparent overlap effect $CPAOE_t^b$ for those with higher baseline income and fewer health problems, as is shown by Table 1.3. Figure 1.6 shows different trends of $CLAOE_t^b(a)$ for grandparents with different baseline health and income. Grandparents who start with higher income or fewer health problems show an increasing grandparent overlap effects as overlap increases, while grandparents with lower initial income or more health problems have declining grandparent overlap effects as they age.

In general, my cumulative fixed effects(CFE) approach eliminates unobserved variable biases to recover the unbiased estimation of the grandparent overlap effects and allows for exploration of effect heterogeneity which would provide sociologists with a detailed picture of different grandparent overlap effects across grandchildren's age, various durations of overlap, and different grandparental characteristics.

Extensions

The model for grandparent overlap effects of Equation 1.3, which I have discussed so far, inevitably makes simplifying assumptions compared to a more complex social reality. Because causal identification is only ever achieved relative to an assumed DGP, I next discuss several directions in which my model could fruitfully be generalized.

More Flexible Specifications of the Parameters

First, one could consider a more flexible parameterization for the effects of grandparental characteristics on grandchild outcomes. In Equation 1.3, the effect of each grandparental characteristic, $C_{it'}$, depends only on the age, t', at which the characteristics is experienced by the grandchild. This specification prevents the effects of exposures experienced at a particular age to change (increase or decrease) over time. For example, the effect of having had a healthy rather than sick grandparent at age 10, β_{10}, is stipulated to have the same effect on grandchild's test scores at ages 10 and at age 20, Y_{i10} and Y_{i20},

More generally, one could allow the effect of grandparental characteristics, $C_{it'}$, to depend both on the age at which the characteristic is experienced, t', and on the time that has elapsed between exposure and the measurement of the outcome Y_{it} (Todd and Wolpin 2003), so that each $\beta_{it'}$ is replaced by a series of β_t^{gap}, $gap = t - t'$. Hence, the effect of experiencing a healthy rather than a sick grandparent at age 10 on test scores at age 10 could differ from the effect on test scores at age 20, $\beta_{10}^{0} \neq \beta_{10}^{10}$.

Relaxing the DGP in this manner may increase sociological realism but also comes at a cost. Whereas the parameters, $\beta_{t'}$ can be identified between the first and third test date, the more flexible parameters $\beta_{t'}^{gap}$ can only be identified between the second and third test date (see Appendix C for details). This cost, however, is arguably minor, since even in the more restrictive DGP of Equation 1.3, the grandparent overlap effects, $LAOE_t(a)$ and $PAOE_t$, which are the primary focus of my analysis, are only identified between the second and third test date, and this is still the case in this more general DGP.

Death Effects

Second, one could elaborate on the role of grandparent's death in determining grandchildren's test scores. The DGP of Equation 1.3 describes a model in which grandparent-overlap effects originate exclusively from the grandchild's cumulative exposure to grandparental characteristics while the grandparent is alive. The death of the grandparent in this model simply marks the end of grandparental exposure, but does not exert an effect on grandchild test scores by itself. One could reasonably explore more general models, in which grandparent's death does affect grandchildren's test scores directly, for example, via inheritance (Hällsten and Pfeffer 2017) or grief (Silverman et al. 2000).

Theorizing such death effects would be interesting, because they might trade-off in complex fashion with the grandparent overlap effect understood, as before, as the effect of cumulative exposure. Specifically, dying one year later would increase the grandchild's exposure to the grandparent (likely a positive influence on the grandchild), but plausibly would also reduce the amount of the inheritance that the grandchild receives (likely a negative effect); then again, dying one year later would mean that more of whatever inheritance the grandchild may have received is left over at the time of the fixed subsequent test date (likely a positive effect), but grief is more acute, too (likely a negative effect). Hence, the total effect of cumulative grandparental exposure plus the postponement of death resulting from an additional year with the grandparent on the grandchild's test scores could be positive or negative.

Such considerations could be explored in elaborated models, such as that of Equation 1.19, which adds death effects via inheritance to Equation 1.3,

$$Y_{it} = \beta_0 + \sum_{t'=0}^{A_{it}} C_{it'}\beta_{t'} + D_{it}N_i\gamma + U_{i1} + U_{i2}A_{it} + \epsilon'_{it} \tag{1.19}$$

where $D_{it} = 1$ is an indicator for grandparent's death, $= 0$ if alive; N_i measures the amount of the inheritance and grief; and γ is the effect of the inheritance on grandchild test scores.[7]

When N_i is observed, my estimation approach can still identify grandparent overlap effects by adjusting for N_i explicitly. The interpretation of the estimates, however, would subtly change. Rather than estimating the total effect of postponing a grandparent's death, my estimation approach would now identify the controlled direct effect of grandparent overlap,

[7] To allow that the effect of the inheritance diminishes over time, one could define N_{it} as the present value of the inheritance (suitably defined) or permit γ_{t-d_i} to vary freely with time since grandparent's death.

net of fixing the amount of inheritance and grief. Hence, the $PAOE_t$ would no longer identify $E_{\bar{C}_t, A_{it}}\big[E\big[Y_{it}(a+1) - Y_{it}(a)|q_2 < A_{it} < t\big]\big]$ but $E_{\bar{C}_t, A_{it}, N_i}\Big[E\big[Y_{it}(a+1, N_i) - Y_{it}(a, N_i)|q_2 < A_{it} < t\big]\Big]$.

When N_{it} is unobserved, my approach cannot identify the model in Equation 1.19. This is because grandparent's death, D_{it}, is not random, but is a function of the history of past grandparent characteristics, $\bar{C}_{it'}$, and the fixed unobservables, U_i. When the term $D_{it}N_i\gamma$ is non-zero, it cannot be differenced out by the double differencing of estimation steps 1 and 2 above, so that a function of the unobserved confounders U_i would remain in the error term of the regression.

Since few grandparents leave significant bequests (Wolff and Gittleman 2014), and those who do often divide the bequest over multiple beneficiaries, it may be permissible to ignore the issue of inheritances. Once the Pandora's box of death effects is opened, however, sociologists should at least discuss the complications arising from typically unmeasured consequences of bereavement, such as grief or the possible lifting of parental caregiver obligations, on grandchild educational test scores. Further work should explore the likely magnitude of bias from ignoring death effects–work that would certainly benefit from elaborating sociological theories about the consequences of grandparental death and their temporal articulation with respect to specific grandchild outcomes.

Parental Time-varying Variables

Equation 1.3 made the simplification by assuming away any parental or grandchild time-varying confounders that simultaneous affect the grandparent's survival, characteristics and grandchild's outcomes. For instance, mother's marital status can affect grandparent's coresidence and employment status and affect the grandchild's cognitive outcomes. So mother's marital status is a time-varying confounder. I can and should control for these parental time-varying confounders, P_{it}:

$$Y_{it} = \beta_0 + \sum_{t'=0}^{A_{it}} C_{it'}\beta_{t'} + \sum_{t'=0}^{t} P_{it'}\gamma_{t'} + U_{i1} + U_{i2}A_{it} + e^*_{it} \tag{1.20}$$

where e^*_{it} is an exogeneous error.

By including P_{it}, the identification assumption may be relaxed to:

$$E(e^*_{it}|\{C_{it'}, A_{it'}, P_{it'}\}_{t'=0}^{t'=T}) = 0 \tag{1.21}$$

I should not control for parental characteristics that are mediators of the grandparent overlap effect if aiming at identifying the total effect of grandparent overlap, otherwise there may be an over-control bias (Elwert 2013; Elwert and Winship 2014). By adjusting for parent's marital status, I assume that grandparent's survival or characteristics will not cause a change in the parent's marriage to affect the grandchild's cognitive outcome, i.e. parent marital status is not a mediator.

Generally speaking, parent time-varying characteristics, P_{it}, are different from C_{it} in that P_{it} do not constitute the grandparent overlap effect and are not treatments. Controlling for P_{it} as a mediator would control away part of the total effect of overlap. Therefore, P_{it} can be either a mediator or a confounder, but not both. If P_{it} is a mediator (but not confounder), then analyst should not control for it. If P_{it} is a confounder (but not mediator), then analyst should control for it. Analysts should evaluate P_{it} specifically in relation to the other variables in the DGP and adjust for these parental characteristics with care. By contrast, C_{it} can be both confounders and mediators; since grandparent overlap effects are derived from the simultaneous interventions on A_{it} and C_{it}, involving C_{it} naturally serves the dual purposes of controlling for confounding and intervention.

Joint Grandparent Time-varying Confouders

The model of Equation 1.3 may be extended to include grandparent spouse's characteristics as potential time-varying confounders. Grandparent spouse's characteristics are potential time-varying confounders as the grandmother's death could affect her husband's survival through widowhood effects (Elwert and Christakis 2008) and the grandmother's health status may affect her husband's survival through her ability to provide care. When the grandparent spouse's characteristics are potential time-varying confounders, they should be adjusted for in the model.

It is feasible to model the grandmother's characteristics and the grandfather's characteristics jointly to capture their influences on each other and their joint effect on the grandchildren. One way is to regard the grandmother's characteristics as the grandfather's spousal characteristics, which constitute C_{it} with the effects accumulating across the grandfather's overlap, and vise versa for the grandmother's overlap. This approach models each grandparent separately while exploiting the information of his or her spouse. Another approach is to aggregate C_{it} and A_{it} over each lineage of grandparents, taking either their mean or max. This provides estimates of grandparent overlap effect of each grandparent family defined by a marriage union. The DGPs including joint grandparent time-varying confounders are straightforward and similar to the forms of either Equation 1.3 or Equation 1.20 with slightly different meanings of the notations. So I do not elaborate on this aspect in this chapter.

Conclusion

As the life-course overlap between the grandparent and grandchild's generation is continuously increasing, sociologists begin to investigate how this demographic trend would affect the achievement gaps and perpetuate the transmissions of social inequality across generations. My approach intervenes on the sociological literature by providing an new approach to identify and estimate the grandparent overlap effects. I posit realistic data generation models that reflect the nature of grandparent overlap effects as the cumulative and heterogeneous influences characterized by the unit's characteristics. I show that the overlap effects are identified with individual panel data and provide an feasible estimator (CFE) to recover estimands of sociological interests based on the assumptions that the cumulative effects of grandparents' fixed unobserved characteristics are linear in overlap and there are no unobserved time-varying confounders. My approach also allows for effect heterogeneity across various lengths of overlap, different grandparent lineage, characteristics and grandchildren's age.

Several authors have noted that conventional fixed effects models (which enter treatment as a main effect) at best recover variance-weighted average treatment effects if the treatment effect varies across individuals (Wooldridge 2005; Imai and Kim 2019). My model specifically departs from the conventional practice of entering treatment as a main effect and explicitly models the cumulative and heterogeneous treatment effect with interactions between treatment and the entire history of the observed time-varying covariates as well as the interaction between treatment and the fixed effect. This permits the effect of an additional year of overlap to vary flexibly across fine-grained classes defined by the grandparent's observed covariate history with individual variations characterized by the interacted fixed effect term. Thus, my method contributes to the growing literature on fixed effects models with heterogeneous treatment effects.

In various research contexts that are of interest to social scientists, an overlap effect uncovers the total influence of one unit on another. The analytical framework can be readily applied to other interesting sociological questions. For instance, with the increase in divorce and single motherhood, how does father's overlap affect the children's well-being (McLanahan 2004)? With increased teacher turnover in public schools in low-income neighborhoods, how does teacher overlap affect the students' achievement gap (Simon and Johnson 2015; Ingersoll 2001)?

Tables and Figures

Table 1.1: Table of Notations

i, t', t	The grandchild-grandparent pair, grandchild's age index, and the grandchild's age at the test. t is different from t' in that it indicates the grandchild's age whenever a test outcome is measured.
d_i	The grandchild's age at the grandparent's death
A_{it}	The duration of overlap of grandchild i with the grandparent at granchild's age t
Y_{it}	The test score outcome of individual grandchild i at time point t
$C_{it'}$	The time-varying grandparent characteristics at t'
P_{it}	The time-varying parent characteristics at t
$\bar{C}_{it'}$	The history (i.e. complete sequence) of grandparent characteristics, $C_{it'}$, from i's birth to age t'
$U = (U_{i1}, U_{i2})$	Grandparent's fixed unobservables U_{i1} comprises the fixed unobservables that affect Y_{it} marginally. U_{i2} comprises the fixed unobservables that affect Y_{it} in interaction with A_{it}. U_{i1} and U_{i2} may share elements in common.
e_{it}	i's exogeneous idiosyncratic time-varying unobservables (error term)
$D_{it'}$	Time-varying dummy indicator for the death of i's grandparent at t'
N_i	The amount of inheritance i would receive
$I_{it'}$	Grandparent income at t'
$H_{it'}$	Grandparent health at t'

Table 1.2: Elements of the four steps in estimating $LAOE_{20}(a)$

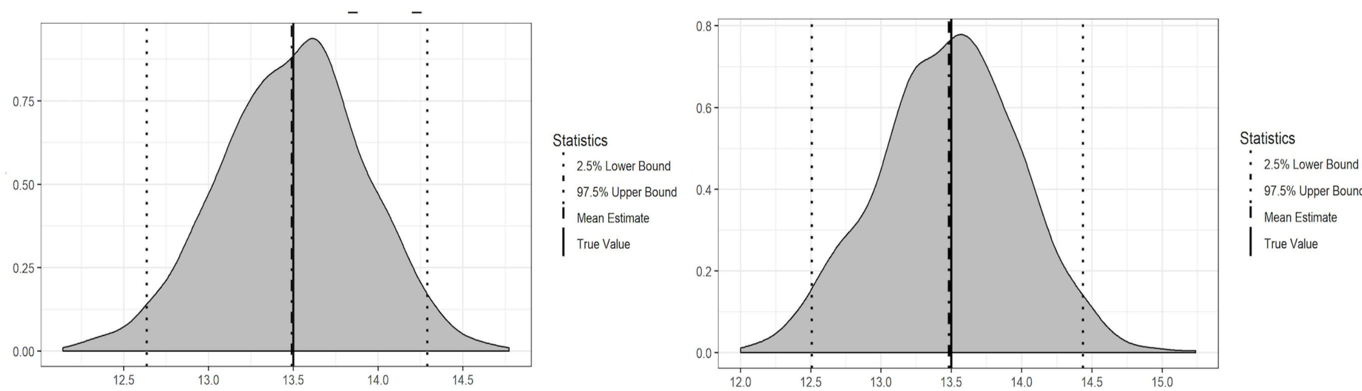
a	$\beta_{a+1,I}$	$\overline{\widehat{\beta}}_{a+1,I}$	$\beta_{a+1,H}$	$\overline{\widehat{\beta}}_{a+1,H}$	$\bar{I}_{ia+1}$	$\bar{H}_{ia+1}$	$E(U_{i2}\|a)$	$\overline{\widehat{E}}(U_{i2}\|a)$	$LAOE_{20}(a)$	$\overline{\widehat{LAOE}}_{20}(a)$
11	10.50	10.52	10.50	10.54	262.20	-93.82	-2.48	-3.69	1765.60	1765.58
12	11.00	10.99	11.00	10.97	286.92	-119.58	8.35	7.12	1847.22	1847.13
13	11.50	11.50	11.50	11.49	313.39	-148.00	19.37	18.11	1920.32	1920.54
14	12.00	12.02	12.00	12.03	341.55	-178.73	30.79	29.50	1984.21	1984.50
15	12.50	12.49	12.50	12.48	371.46	-212.03	42.98	41.63	2039.42	2039.14
16	13.00	13.02	13.00	13.02	403.47	-247.46	56.35	55.00	2089.12	2089.30
17	13.50	13.49	13.50	13.48	438.55	-285.07	71.72	70.33	2142.17	2142.00
18	14.00	14.02	14.00	14.02	477.07	-324.90	90.70	89.30	2219.76	2219.97
19	14.50	14.50	14.50	14.49	526.32	-367.96	117.1	115.65	2411.04	2411.00

Notes: a indicates the length of duration of overlap at $t = 20$, $A_{i20} = a$. $\beta_{a+1,I}$ and $\beta_{a+1,H}$ indicate the true values of the parameters of grandparent income and health at the extended overlap stage, $a+1$. $\overline{\widehat{\beta}}_{a+1,I}$ and $\overline{\widehat{\beta}}_{a+1,H}$ indicate the sample average of the estimates of the parameters of grandparent income and health at $a+1$. $\bar{I}_{ia+1}$ and $\bar{H}_{ia+1}$ are the average grandparent income and health at $a+1$ for those who die at this wave, which are predictions of the grandparent income and health at $a+1$ if those who die at a survive to $a+1$. $E(U_{i2}|a)$ is the true values of the unobservable at different a. $\overline{\widehat{E}}(U_{i2}|a)$ is its sample averages. $LAOE_{20}(a)$ are the true values of length-specific grandparent overlap effect at different a. $\overline{\widehat{LAOE}}_{20}(a)$ is the estimated sample average. Recall that $LAOE_{20}(a) = E\big[C_{a+1}\beta_{a+1}+U_{i2}|A_{it} = a\big]$, the values listed in each row demonstrate how $LAOE_{20}(a)$ correspond to the observed and unobserved components at a given a.

Table 1.3: Subgroup mean estimates and true values of conditional grandparent overlap effect, $\overline{\widehat{CPAOE}}^{b}_{t}(a)$

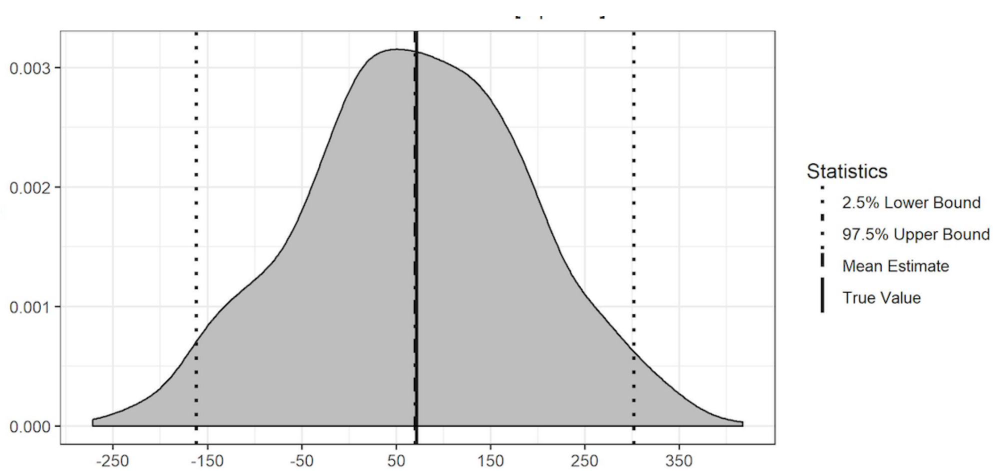
Subgroup	True Values	Mean Estimates
Baseline Health: 0-25%	2319.16	2319.22
Baseline Health: 25-50%	1978.14	1978.32
Baseline Health: 50-75%	1797.50	1797.66
Baseline Health: 75-100%	1671.46	1671.89
Baseline Income: 0-25%	1693.52	1694.17
Baseline Income: 25-50%	1820.03	1820.31
Baseline Income: 50-75%	1998.98	1999.04
Baseline Income: 75-100%	2299.17	2299.23

Figure 1.1: Sampling distributions for $\hat{\beta}_{18}$



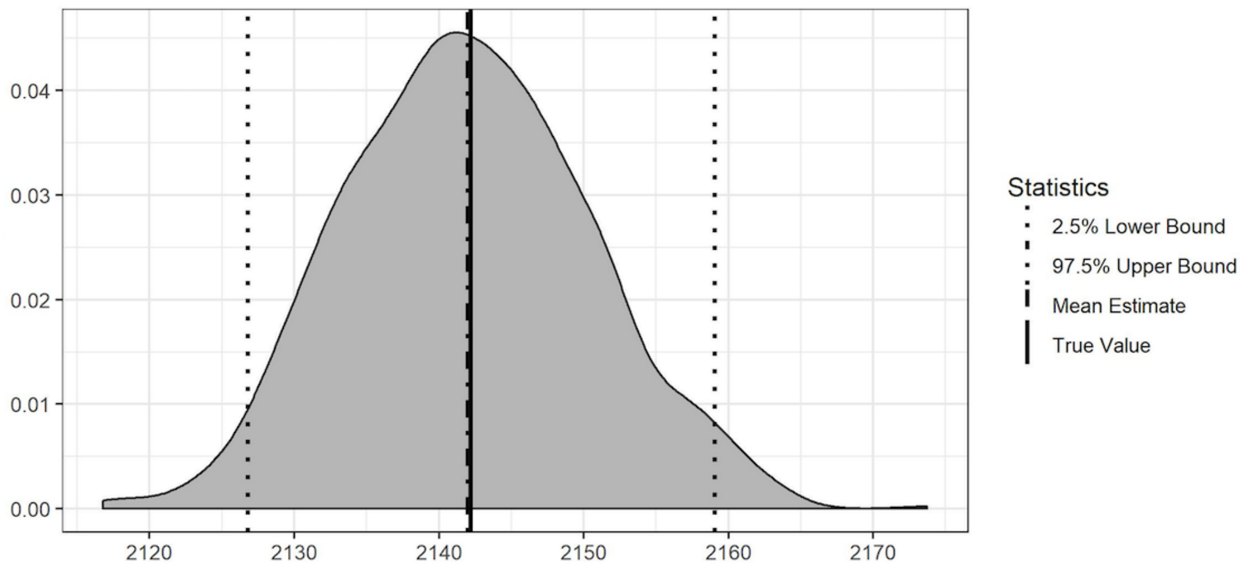
Notes: The left and right panels show the sampling distributions of the coefficients for grandparent income, $\hat{\beta}_{18,I}$, and grandparent health, $\hat{\beta}_{18,H}$, respectively, experienced at grandchildren's age 18 on test scores at age 20 from 2000 simulated samples of size $N = 100,000$. Both estimates are unbiased, as the means of the estimates equal the population parameters, $\bar{\hat{\beta}}_{18} = \beta_{18}$=13.5.

Figure 1.2: Sampling distribution of the estimated unobservable component $\widehat{E}(U_{i2}|a = 17)$



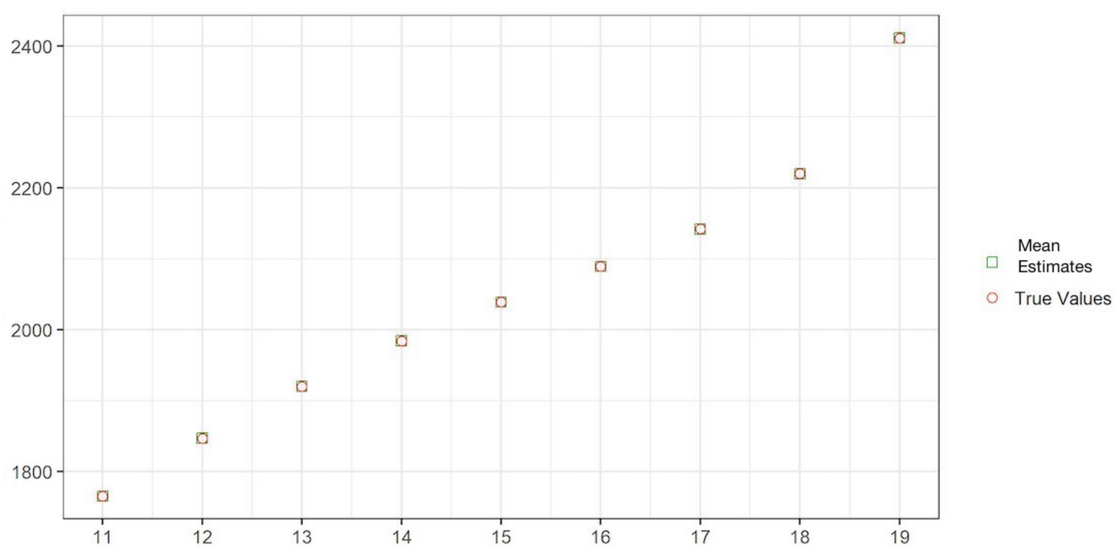
Notes: This graph shows the sampling distribution of the estimated unobservable component, $\widehat{E}(U_{i2}|a = 17)$, for the $LAOE_{20}(17)$ from 2000 simulated samples of size $N = 100,000$. The estimates are approximately unbiased, as the mean of the estimates, $\overline{\widehat{E}}(U_{i2}|a = 17) = 70.33$ is approximately equal to the true value $E(U_{i2}|a = 17)$=71.72.

Figure 1.3: Sampling distribution of $\widehat{LAOE}_{20}(a = 17)$



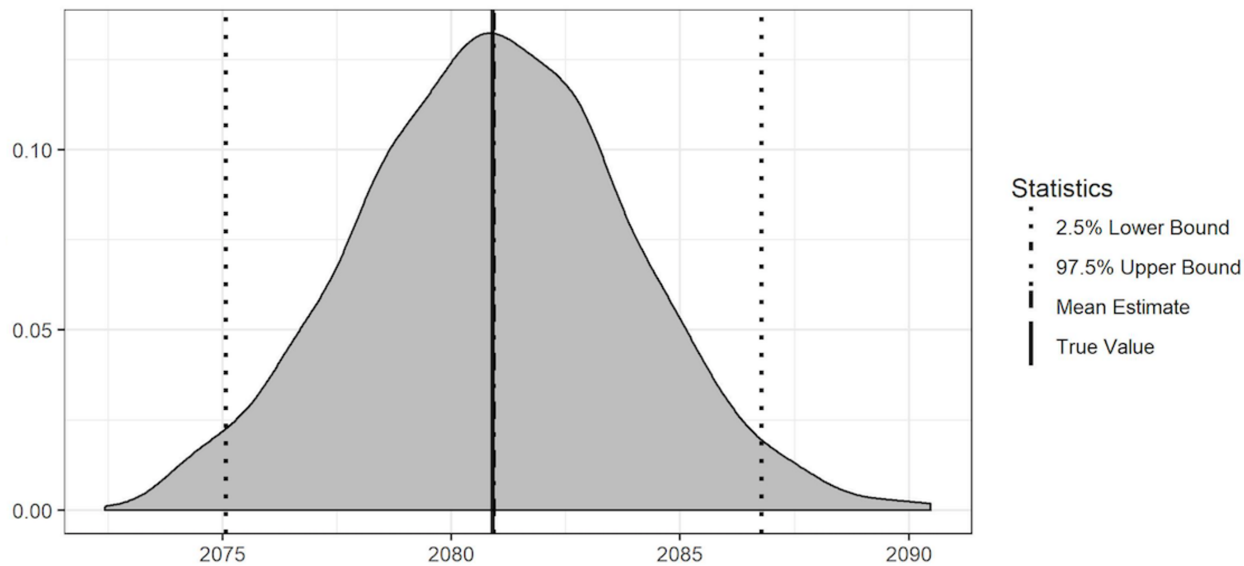
Notes: The sample density plot is drawn from 2000 samples. The sample average of the estimates, or $\overline{\widehat{LAOE}_{20}}(a = 17)$=2142.00 is approximately equal to the true value of $LAOE_{20}(a = 17)$, which is 2142.17.

Figure 1.4: Averages estimates, $\overline{\widehat{LAOE}_{20}}(a)$ compared to their true values



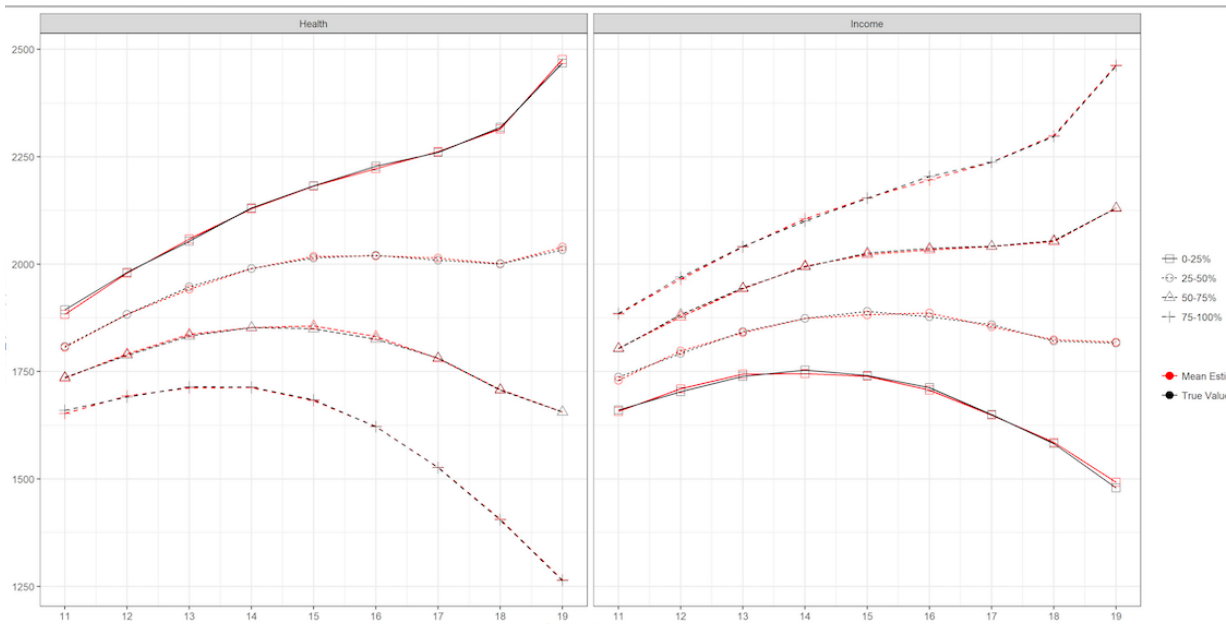
Notes: The squares indicate the average estimates, $\overline{\widehat{LAOE}_{20}}(a)$, and the circles indicate the true values, $LAOE_{20}(a)$, of the length-specific grandparent overlap effect on test scores at age 20 at different lengths of overlap $A_{i20} = a$, across 2000 samples of size $N = 100,000$.

Figure 1.5: Sampling distribution of $\widehat{PAOE}_{20}$



Notes: This is the sample density plot of estimated $\widehat{PAOE}_{20}$ from 2000 samples. The dash line indicate the sample average of the estimates, or $\overline{\widehat{PAOE}_{20}} = 2080.93$, which is approximately equal to the true value which is 2080.91.

Figure 1.6: Subgroup mean estimates and true values of conditional length-specific grandparent overlap effect, $\overline{\widehat{CLAOE}_t}^b(a)$



Notes: The left and right panels show the subgroup mean estimates and true values based on grandparent health $\overline{\widehat{LAOE}_{20}}(a|H_{it'=1})$ and income $\overline{\widehat{LAOE}_{20}}(a|I_{it'=1})$, separated by four quantile groups of grandparent income and health at $t' = 1$ (0%-25%, 25%-50%, 50%-75%, and 75%-100%), at different lengths of overlap $A_{i20} = a$ respectively, across 2000 samples of size N=100,000.

Appendix

A. Appendix on Simulation of Individual Panels

We first generate the endogenous grandparent's fixed effects U_{i1} and U_{i2} from $N(10, 100)$. Because the length of overlap, A_{it} can be affected by the grandparent time-varying observables. So we need the time series of $C_{it'}$ from $t' = 1$ to $t' = q_3 = 20$ before we generate A_{it}. The time series of $C_{it'}$ would eventually be truncated to a sequence from $t' = 1$ to $t' = A_{iq_3}$ for each individual as inputs in the date generation process.

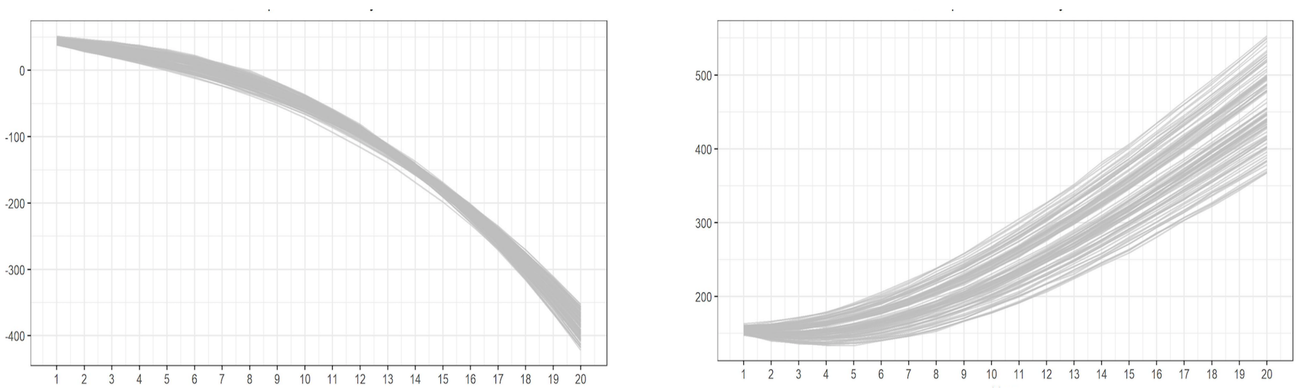
Then we generate the grandparent observables in the order of $Age_{it'=0}$, $I_{it'=1}$ and $H_{it'=1}$, and then the entire series of $I_{it'}$ and $H_{it'}$. This is because the initial grandparent health and income $I_{it'=1}$ and $H_{it'=1}$ can be affected by the grandparent's age at child's birth, $Age_{it'=0}$. And all of these characteristics $Age_{it'=0}$, $I_{it'=1}$ and $H_{it'=1}$ may be determined by their fixed characteristics U_i (Fomby et al. 2014). Specifically:

$$
\begin{aligned}
&Age_{it'=0} = percentile(U_i) * 10 + 50 \\
&I_{it'} = 0.95 I_{it'-1} + 0.035 H_{it'-1} + 0.015 U_{i1} + 0.015 U_{i2} + 0.02 I_{it'=1} t' + e_t \\
&H_{it'} = 0.85 H_{it'-1} + 0.005 I_{it'-1} - 0.01 U_{i1} - 0.01 U_{i2} + (-0.004 t' - 0.006 t'^2) * H_{it'=1} + e_t \\
&where \quad I_{it'=1} \sim N(100 + Age_{it'=0}, 1), \quad H_{it'=1} \sim N(100 - Age_{it'=0}, 1), \\
&and \quad e_t \sim N(0, 1)
\end{aligned}
$$

Over t' from 1 to 20, grandparent health decreases over time, while the grandparent income increases over time with diverging trajectories for grandparents with higher status and lower status. Their trajectories are shown as below:

In order to generate grandparent overlap, we first generate a time-varying survival indicator of grandparent at grandchild's age t', $D_{it'}$. $D_{it'}$ is generated as a function of U_{i1}, U_{i2}

Figure 1.7: Trend plots of grandparent income and health



Note: The left panel shows the trends of grandparent health and the right panel shows the trends of grandparent income from $t' = 1$ to $t' = 20$. They do not correspond to $\bar{I}_{ia+1}$ or $\bar{H}_{ia+1}$ in Table 2, $\bar{I}_{ia+1}$ or $\bar{H}_{ia+1}$ are taken for the grandparents who die on $a+1$.

and the grandparent income and health of the previous stage. Specifically,

$$D_{it'} = \begin{cases} 1, & if \quad I_{it'-1} + H_{it'-1} + 0.5U_{i1} + 0.5U_{i2} + e_{it} < 150 \\ 0, & if \quad I_{it'-1} + H_{it'-1} + 0.5U_{i1} + 0.5U_{i2} + e_{it} \geq 150 \end{cases} \qquad where\ e_{it} \sim N(0, 100)$$

The grandchild's age at the grandparent's death d_i is defined as $d_i = 20 - \{\min t' \big| D_{it'} = 1\}$. The duration of overlap, A_{it}, upon each test age t for i is generated as a function of d_i and t. Specifically, $A_{it} = \min(d_i, t)$.

Last, three observations of Y_{it} at three ages q_1, q_2 and q_3 for each individual i are generated following equation Equation 1.3.

B. Appendix on Simulation of Sibling Panels

We can also achieve the identification of the grandparent overlap effects by exploiting the sibling variations in addition to within-individual variations in the repeated measures of test scores.

Denote a younger sibling of i, as j, who was also born before the death of the grandparent. Besides the two test outcomes of i taken at q_1 and q_2, we also observe two test outcomes of

j, measured at i's ages q_2' and q_1', where $q_2' > q_1'$. The grandparent overlap for j at q_2' and q_1' are $A_{jq_2'}$ and $A_{jq_1'}$.

After generating the overlap of the first sibling i,A_{it}, we generate the birth gap between the two siblings as a function of U_i and the baseline $C_{it'}$. Since the birth year of i is $t' = 0$, the birth time of j is $birthgap_{i,j}$. Specifically:

$$
\begin{aligned}
&A_{iq} = \min(d_i, q) \\
&A_{jq'} = \min(d_i - birthgap_{i,j},\ q' - birthgap_{i,j}) \\
&where \quad birthgap_{i,j} = U_{i1} * 0.001 * X_i + 0.0005 I_{A_{it=1}} - 0.002 H_{A_{it=1}} \\
&where \quad X_i \sim poisson(80)
\end{aligned}
$$

Then generation of the model and estimation proceed in the same way as what we have shown in the main text with individual CFE. Recall that we rely on the sample of grandparents who survive at least the second test for individual CFE. Similarly, for sibling CFE, we require that the sample satisfies $q_1 < d$ and $q_1' < d$. In other words, we require the sample of grandparents who survive both the first tests of i and j, no matter whether j's first test q_1' occurs before or after i's second test, q_2. Notice that we do not require the gap between the two tests of j to be the same as that of i. We can identify β_t using sibling panel data. Estimation proceeds by an analogous two-step procedure, first taking first difference between waves within each individual, and second taking the sibling difference between the equations resulting from the first step.

C: Appendix Regarding More Flexible Parameters $\beta_{t'}^{gap}$

The Extensions section discussed the more flexible specifications of the parameters generalizes, permitting that the effect of grandparent characteristics $C_{it'}$ on Y_{it} vary both with the age t' at which the grandchild is exposed to the characteristic and the age t at which the outcome is measured, $\beta_{t'}^{gap}$, $gap = t - t'$. This specification permits, for example, that the effect of exposure to a given grandparental characteristic decays with time. In fact, it permits that the effect of a given grandparental characteristic evolves freely over time.

For outcomes measured at $q1 = 6$, $q2 = 10$ and $q3 = 20$, we have,

$$Y_{i6} = C_{i1}\beta_1^5 + C_{i2}\beta_2^4 + ... + C_{i6}\beta_6^0 + U_{i2}A_{i6} + U_{i1} + e_{i6}, \tag{1.22}$$

$$Y_{i10} = C_{i1}\beta_1^9 + C_{i2}\beta_2^8 + ... + C_{i10}\beta_{10}^0 + U_{i2}A_{i10} + U_{i1} + e_{i10}, \tag{1.23}$$

$$Y_{i20} = C_{i1}\beta_1^{19} + C_{i2}\beta_2^{18} + ... + C_{iA_{i20}}\beta_{A_{i20}}^{20-A_{i20}} + U_{i2}A_{i20} + U_{i1} + e_{i10} \tag{1.24}$$

We use the sample of grandparents that survive beyond the second test to identify the parameters.

Step 1 first takes the difference between equation 23 and 22 to remove U_{i2} and divides by $\triangle_{q_2 q_1} A_i = 4$,

$$\frac{Y_{i10} - Y_{i6}}{4} = \frac{C_{i1}\beta_1^9 + C_{i2}\beta_2^8 + ... + C_{i10}\beta_{10}^0 - (C_{i1}\beta_1^5 + C_{i2}\beta_2^4 + ... + C_{i6}\beta_6^0)}{4} + U_{i2} + \frac{e_{i10} - e_{i6}}{4}, \tag{1.25}$$

and then takes the difference between equation 23 and 24 and divides by $\triangle_{q_3 q_2} A_i = A_{i20} - 10$,

$$\frac{Y_{i20} - Y_{i10}}{A_{i20} - 10} = \frac{C_{i1}\beta_1^{19} + C_{i2}\beta_2^{18} + ... + C_{iA_{i20}}\beta_{A_{i20}}^{20-A_{i20}} - (C_{i1}\beta_1^9 + C_{i2}\beta_2^8 + ... + C_{i10}\beta_{10}^0)}{A_{i20} - 10} + U_{i2} \tag{1.26}$$

$$+ \frac{e_{i20} - e_{i10}}{A_{i20} - 10},$$

to remove U_{i1}.

Step 2 takes the difference between equation 26 and 25 to remove U_{i2}, and rearranges terms to combine the parameters containing the same $C_{it'}$,

$$\frac{Y_{i20}-Y_{i10}}{A_{i20}-10}-\frac{Y_{i10}-Y_{i6}}{4}=C_{i1}(\frac{\beta_1^{19}-\beta_1^9}{A_{i20}-10}-\frac{\beta_1^9-\beta_1^5}{4})+...+C_{i10}(\frac{(\beta_{10}^{10}-\beta_{10}^0}{A_{i20}-10}-\frac{\beta_{10}^0}{4}) \quad (1.27)$$
$$+\frac{C_{i11}\beta_{11}^9}{A_{i20}-10}+...+\frac{C_{iA_{i20}}\beta_{A_{i20}}^{20-A_{i20}}}{A_{i20}-10}+\frac{e_{i20}-e_{i10}}{A_{i20}-10}-\frac{e_{i10}-e_{i6}}{4}$$

From this, it is obvious that the parameters $\beta_{t'}^{gap}$ for grandparent characteristics $C_{it'}$ experienced between the first and second test date (i.e., for $t'=1,...10$) are not identified, because the number of parameters exceeds the number of variables. By contrast, the parameters $\beta_{t'}^{gap}$ for grandparent characteristics experienced between the second and third test date (i.e., for $t'=11,...,20$) are identified given that the sample contains grandparents who die at each time point of $t'=11,...,20$. The parameters can be estimated by running an OLS regression on Equation 27.

Chapter 2

The Grandparent Overlap Effect on Children's Cognitive Development

Introduction

Intergenerational overlap forms the demographic basis of reproduction of social inequality across generations. The long period of overlap inherent to humans allows parents to pass down their social skills, language, and culture to their children, thereby establishing and maintaining formal and informal institutions of social stratification across generations. Thus, overlap facilitates human capital development. With the sharp increase in human life expectancy occurring across the 20th century, overlapping life courses has changed from predominantly intergenerational to increasingly multigenerational. In 1900, the estimated average grandparent-grandchild overlap was five years. In 2010, it was 35 (Song and Mare 2019). As life expectancy continues to rise, the degree of multigenerational transmission of inequality and its implication for future opportunity structures may also increase in importance.

The recent decade saw a surge in studies of studies of multigenerational mobility. Researcher have documented the demographic trend of grandparenthood (Leopold and Skopek

2015; Margolis and Wright 2017; Song and Mare 2019), as well as adopted the perspective of grandparent-grandchild overlap to study grandparent effects (Daw et al. 2018; Lehti et al. 2018; Sheppard and Monden 2018). The research has conceptualized grandparental overlap in one of two ways. Either as a moderator of the relationship between specific set of grandparent-grandchild characteristics (e.g., educational level or occupation) (Daw et al. 2018; Sheppard and Monden 2018; Song and Mare 2019), or as an effect time in itself (Lehti et al. 2018). Studies from the two approaches produce mixed results, with some finding weak support for a significant relationship and others showing null findings. In a recent review, Anderson et al. (2018) concludes that evidence is scarce that a longer overlap can amplify the grandparent's influence, which casts doubts on the causal interpretation of the grandparent overlap effect (p.136). Yet, it may still be premature to put the grandparent overlap effect to rest.

In this study, I argue that the present conceptualization of grandparental overlap as either an effect moderator or an effect of more time in itself provides an incomplete depiction of the full grandparent overlap effect. A longer grandparent overlap enables more interactions between a grandparent and grandchildren, which in turn the amplifies effect that any and all grandparent characteristics may have on grandchildren. So, any effect of grandparent characteristics that get amplified with overlap should be seen as part of the overlap effect. For this reason, I propose a new conceptualization of the grandparent overlap effect. Instead of searching for the significant effect of grandparent overlap in moderating any particular effect of grandparent characteristics, I investigate the total effect of extending grandparent overlap – essentially asking whether a longer multi-generational life-course overlap amplifies the effect of all grandparent characteristics, be they time-varying or fixed, and observed or unobserved. I focus on children's cognitive development as an early indicator of human capital, and model a data generating process that closely corresponds to my conceptualization. The model has three main features. First, since a grandchild's current cognitive ability is the result of the entire history of family inputs, I model the grandparent overlap effect as a complete and holistic effect of the cumulative exposure to the entire grandparent across the shared life course. Second, the data generating process includes two types of effects of grandparents' fixed characteristics (such as the effect of grandparent social class or ability). A temporally fixed component whose impact is independent of grandparent's survival, and a component whose effect accumulates through overlap. For instance, a high social status of the grandparent may affect the grandchild via daily interactions when they are both alive through concerted cultivation, but may also affect the grandchild through social connection and elite college preferences and legacies even after the grandparent's death. Third, the grandparent's unobserved ability and social class are likely to determine their observable characteristics, such as income, health, and marital status. In my model, I allow the observable grandparent characteristics to be endogenous to these grandparent unobserved characteristics.

I argue that my proposed model provides a more complete conceptualization of the grandparent overlap effect. Yet, my preferred model is not identifiable using any standard estimations strategies. Instead, I turn to a new estimation strategy, cumulative fixed effect models (CFE), to be able to identify the effect of grandparent overlap on grandchildren's cognitive ability. To provide an estimate of the grandparent overlap effect, I rely on a large and rich dataset obtained from Danish administrative registries, which covers the entire Danish population. The data allows me to link information on all Danish grandparents' tax, hospital, health, residential, and socio-demographic characteristics to their grandchildren's standardized test scores across years. I pay special attention to, discuss, and address the identification challenges that mare conventional methods. I find positive grandparent overlap effects on the grandchild's cognitive development, equivalent to the effect size of educational interventions known to have a small to medium effect size. My CFE estimates larger than estimates from conventional panel fixed effects models, which does not take into account the contribution of grandparent observed and unobserved characteristics sufficiently. Further, my results are more robust than OLS estimates, which does not unobserved confounders into account. My findings suggest that past estimates using either OLS and fixed effects models likely are biased in opposite directions.

The study makes three contributions to the literature on social mobility and multigenerational relationships. First, departing from the past literature on grandparent overlap effects, I conceptualize the effect in a holistic way which captures the accumulation of exposure that occurs through life-course overlap. I formally model and empirically identify the grandparent overlap effects using novel methods and large-scale, high-quality data. Second, this holistic approach to overlap departs from previous work's focus on one or two dimensions of transmission of grandparent SES. Instead, I define a grandparent overlap effect that encompasses all major dimensions of the grandparent's social status and resources, and also allow for negative effect of, e.g., health shocks, thereby providing new lens to study the influences of family members through the shared life-course. Third, my framework is not only a more holistic reflection of the grandparent's social status, it is also highly flexible to incorporate the investigation of effect heterogeneity. I allow the grandparent overlap effects to be heterogeneous across socioeconomic status, family structure, and health. While a longer grandparent overlap may reinforce multigenerational immobility because of unequal mortality and life span, it may also perpetuate immobility because of the unequal benefits of a longer grandparent overlap across groups with different socioeconomic characteristics.

Theoretical Background

Social inequality is transmitted across multiple generations via demographic, social, economic and biological mechanisms (Hällsten and Pfeffer 2017; Mare 2011, 2014; Song et al. 2015; Song 2016; Pfeffer 2014). For the wealthy families, grandparents may affect the grandchild via the transmission of wealth (Hällsten and Pfeffer 2017; Mare 2011, college admission legacy (Mare 2011) and social networks (Clark 2015). For the majority of the population, grandparents affect grandchildren mostly as a kin resource (Bengtson 2001; Lehti et al. 2018). Although part of grandparent effect is associated with social institutions external to the family processes (such as the college legacy and wealth inheritance institutions), grandparents affect their grandchildren mostly via direct and indirect interactions in daily life. As a “mother-saver”, a safety net and role model, grandparents influence the grandchildren’s cognitive development by spending time with them, rearing them, complementing or substituting for the parent’s role, investing in the grandchild’s cognitive growth, and leveraging the extended family resources (Bengtson 2001; Lehti et al. 2018). But at the same time, they may also affect their grandchildren by demanding for material and personal support from the parent generation when they are poor, sick or widowed (Bengtson 2001; Preston 1984).

Past Work

Recent years have witnessed a surge of interest in multigenerational transmission of social economic status in sociology, demography and economics (e.g., Jæger 2012; Knigge 2016; Mare 2011, 2014; Song 2016; Solon 2018; Pfeffer 2014). Advancing from the two-generational framework of studying the parent’s influences on children’s status attainment (Becker and Tomes 1986; Blau and Duncan 1967; Sewell et al. 1969), the multigenerational mobility literature focuses on whether grandparents affect grandchildren’s status attainment with respect to education, occupation and income, and whether the grandparent effect is moderated by race, family structure and coresidence status (Song 2016). Because the increase in multigenerational overlap of life courses is one of the defining changes to multigenerational relationships across the 20th century (Song and Mare 2019), researchers have engaged specifically with the role of grandparent overlap in the production and reproduction of social inequality (Daw et al. 2018; Knigge 2016; Sheppard and Monden 2018). Past research has focused on two questions. First, how grandparent overlap moderates the effects of specific grandparent characteristics on grandchild’s outcomes (Daw et al. 2018; Sheppard and Monden 2018), and second, the direct grandparent overlap effect on grandchild outcomes net of the mediation of parental characteristics (Knigge 2016).

The first vein of literature considers grandparent overlap as a moderator of the effect of specific, singular grandparent characteristics. Analogous to geographical proximity, grandparent overlap is considered as a proxy of intensity of contact (Knigge 2016; Sheppard and Monden 2018). Studies reach mixed results looking at different grandparent characteristics and using various measures of overlap and contacts. Zeng and Xie (2014) find that multigenerational coresidence moderates the effect of grandparent education on grandchild's educational attainment. Examining grandparent's survival status to certain grandchild ages, Daw et al. (2018) and Sheppard and Monden (2018) find no statistically significant moderating effects of overlap on the association between grandparent and grandchild education. Knigge (2016) proxies the length of grandparent overlap with the age difference between the grandfather and grandson and finds a small positive moderating effect of overlap on grandfather's occupational status. Advancing from the past research, Song and Mare (2019) use precisely measured length of grandparent overlap and find the grandparent-grandchild association in education attainment is amplified with a longer multigenerational overlap.

The second line of literature considers grandparent overlap as a treatment of its own. Lehti et al. (2018) estimate the direct effect of grandparent overlap on grandchild's secondary education attainment net of family characteristics. They control for mediators of grandparent survival including family income, extended family networks (number of cousins, aunts and uncles), mother's age at child's birth and parent's occupational status. Using a linear probability sibling fixed effects model, they show that an additional year of grandparent overlap directly increases grandchildren's probability of high school graduation by 1 percentage point ($p < 0.001$), a small but non-trivial effect. Comparing separate models for each grandparent lineage, they find that the direct grandparent overlap effect is statistically significant for grandmothers but not grandfathers. They also test for the moderating effect of grandparent overlap and find some evidence that a longer shared overlap exposure may amplify the effects of number of cousins and family income on the grandchildren's probability of high school graduation for some grandparents but not all[1].

[1] Specifically, Lehti et al. (2018) find that an additional year of paternal grandmother's overlap increases the effect of the number of cousins on grandchildren's probability of high school graduation by 0.05 percent-

The contributions of prior work on grandparent overlap effect are profound. They advance findings of grandparent-grandchild associations and try to analyze the conditions under which grandparent effects can be considered "causal". At the heart of both approaches is an underlying notion that prolonged grandparent-grandchild interaction is the driving force behind the grandparent overlap effect. Yet, both approaches fall short of fully capturing the nature of prolonged grandparent-grandchild interaction—that is, falls short of fully defining and capturing the overlap effect.

Reconceptualizing Grandparent Overlap Effect

The total grandparent overlap effect is the totality of any effect of grandparent characteristics that is amplified with a longer grandparent overlap. Prolonged overlap means prolonged exposure to all aspect of a grandparent. The approach adopted by the past works, of picking one or two grandparent characteristics, such as education and occupation, and testing for the moderation effect of grandparent overlap on these characteristics fall short on the following three aspects.

The total grandparent overlap effect is by definition accumulative, holistic and endogenous. First, I stress that the children's cognitive development is affected by the entire history of grandparent characteristics cumulatively where the shared life-course overlap is the time axis on which both the grandchildren and the grandparent characteristics change as they age. Past literature mainly focused on the moderating effect of overlap on fixed grandparent education (Daw et al. 2018; Sheppard and Monden 2018; Song and Mare 2019), which has not taken into account the fact the grandparents themselves experience tremendously changes as they age. The intergenerational ambivalence theory and the role theory predicts a changing grandparent-grandchild relationship as they age. A younger grandparent may have higher income and are healthier but they may work longer and thus are less available for

age point ($p<0.05$), which is not consistent across the four grandparents. An additional year of maternal grandmother's overlap decreases the effect of family income on grandchildren's probability of high school graduation by 0.77 percentage point ($p<0.05$)

grandparenting. As the grandparents become older, some of them may become ailing, less wealthy, and widowed but they may have more time to help with childcare. These change may have important implications for children's development as their relationship changes.

Second, the total grandparent effect is holistic and the grandparent overlap amplified the effect of all grandparent characteristics, observed and unobserved. The multigenerational mobility literature emphasized that social institutions may pass down either in a intergenerational or skip-generational way (Mare 2011). Many social institutions, as embodied by individuals, are reflected as one's culture and social capital, which are mostly unobservable. Past literature on grandparent overlap has not accounted for the grandparent unobservables. For example, the grandparent's cultural capital may affect the grandchild's cognitive development when the family spend time together, via talking with each other in a specific way. Past research shows that the caregiver's vocabulary diversity increases the children's language ability significantly . The grandparent's child-care style (i.e.authoritarian, authoritative, negligent, permissive, see Baumrind) may also affect the children's non-cognitive and congitive ability.

Third, each factor constituting the total grandparent overlap effect may be endogenous to other observed and unobserved characteristics. The grandparent's survival and grandchild's cognitive outcomes may be determined by the grandparent's observed income, health and occupation, and the grandparent's unobserved characteristics, such as the grandparent ability, social and cultural capital. As a part of the grandparent overlap effect, the effect of grandparent's income is determined by the other grandparent observables, such as health and occupation as well as grandparent's unobserved characteristics, such as their ability and social class.

Not accounting for all the three aspects leads to failure to define and capture the entirety of the grandparent overlap effect. From this follows, the conclusion of Anderson et al. (2018) that a lack of findings in the grandparent overlap effect as a modifier of grandparent characteristics "casts doubt on a causal interpretation of the grandparent effect" (p.136) may be premature. Existing literature is limited in that none have estimated the total grandparent overlap effect on grandchild's outcomes.

Given that I define the total grandparent overlap effect as the accumulation of effect of all grandparent characteristics in their shared life course, the data generating model of how grandparent overlap affects the grandchild's cognitive ability is given as:

$$Y_{it} = \beta_0 + \sum_{t'=0}^{A_{it}} C_{it'}\beta + U_{i1} + U_{i2}A_{it} + \epsilon_{it} \tag{2.1}$$

In this model, the grandchild's cognitive outcome Y_{it} at age t is determined by 1) the accumulative effect of the grandparent's time-varying social-economic characteristics, family structure and health, C_{it} over the course of grandparent overlap A_{it}, 2) the time-invariant effect U_{i1} and 3) the overlap effect of time-invariant grandparent characteristics $U_{i2}A_{it}$. By convention, β_0 and ϵ_{it} are the intercept term and the idiosyncratic exogenous error term.

C_{it} is a vector of the grandparent's time-varying characteristics. It could potentially include the grandparent's salary, employment and occupational status, marital status, widowhood, coresidence, and health condition. As the grandparents gets older, they are more likely to get retired, become widowed, work shorter and earn less. The effects of C_{it} accumulate from the birth of the grandchild across the length of overlap A_{it}. C_{it} play three roles. First, C_{it} describes the nature of grandparent-grandchild interaction, and its effects constitute the grandparent overlap effect. Another way to put it is that C_{it} is a moderator of the grandparent overlap effect. Second, C_{it} contains confounders which may affect the grandparent's survival status and the grandchild's test outcomes at a future stage. Third, C_{it} contains mediators of the effect of previous survival status on the grandchild's test outcome. The grandparent's survival determines sample attribution and thus the missingness of all C_{it} in the future stage to affect the grandchild.

U_{i1} and $U_{i2}A_{it}$ are the fixed effects and interactive fixed effects that confound A_{it}, C_{it} as well as Y_{it}. Corresponding to U_{i1} and $U_{i2}A_{it}$, I allow time-invariant confounders, such as the cognitive ability and social and culture capital of the grandparents, to affect the grandchild's outcomes in two different ways; through constant effect and through the overlap effect. As in conventional fixed effects models (pp. 248, Wooldrige 2002), U_{i1} is included to capture the constant effect of the time-invariant unobservables which is not part of the grandparent overlap effect. U_{i1} also captures the mean effects of the time-varying unobservables for each individual. $U_{i2}A_{it}$ represents the effect of unobserved confounders which is accumulative with a longer duration of overlap. For instance, the effect of grandparent race and ability may be amplified with a longer overlap because its transmission relies partly on interpersonal interaction. [2] In general, $U_{i2}A_{it}$ serves as a unobserved analogue to the accumulation of the effect of grandparent observable characteristics, $\sum_{t'=0}^{A_{it}} C_{it'}\beta$.

The Estimand of Total Grandparent Overlap Effect

I can define the estimand of total grandparent overlap effect formally from the data generation model in equation 1.3. Let $Y_{it}(a)$ be the potential cognitive test score outcome of grandchild i if he/she is exposed to an overlap of $A_{it} = a$. The individual treatment effect of overlap is $IOE(i) = Y_{it}(a_{it}+1) - Y_{it}(a_{it})$, and the population average overlap effect (PAOE) can be defined by taking expectation of IOE:

For $A_{it} < t$,

$$PAOE = E\big[Y_{it}(A_{it}+1) - Y_{it}(A_{it})\big] = E[C_{A_{it}+1}\beta] + E[U_{i2}] \tag{2.2}$$

PAOE consists of two parts, $E[C_{A_{it}+1}\beta]$ and $E[U_{i2}]$. The former represents the average effect of all grandparent's time-varying observed characteristics (such as income, occupation and health) in the extended stage of overlap. Given that the grandparent time-varying

[2] For grandparent time-varying unobservables, $U_{i2}A_{it}$ captures their effect to the extent that their effects change linearly with A_{it}. For instance, $U_{i2}A_{it}$ may capture the effect of grandparent-grandchild closeness if that sense of closeness increases with the chance of interpersonal interaction and overlap linearly.

observed characteristics in their extended life are hypothetical for people who have died, I regard $C_{A_{it}}$ in $E[C_{A_{it}+1}\beta]$ as treatments to be externally decided. $E[U_{i2}]$ represents the effects of the unobservables whose mechanisms rely on shared-life course, where U_{i2} may encompass the effect of the grandparent's social and culture capital and cognitive ability.

Empirical Approach

Apparently, this notion of average grandparent overlap effect, $E[C_{A_{it}+1}\beta] + E[U_{i2}]$, cannot be recovered by regression or conventional fixed effects estimators. First, to identify β, I need to eliminate any unobserved confounders in the estimator, i.e. both U_{i1} and $U_{i2}A_{it}$ in equation 2.1. Regression approaches are apparently biased as they do not allow any forms of unobserved confounders. Conventional fixed effects may only eliminate the term of U_{i1} but not $U_{i2}A_{it}$ because the length of overlap A_{it} differs for a individuals at different time point t (and at the same t across siblings). Second, the grandparent overlap effect should include the effect from grandchildren's exposure to grandparents' unobserved characteristics ($U_{i2}A_{it}$ in equation 2.1), so I need to further recover this part of effect of $E[U_{i2}]$. Neither regressions or conventional fixed effects models provide a method to achieve that. If I assume away the component of $U_{i2}A_{it}$, then the model would likely be mis-specified, because it would implicitly assume the effect of grandparent overlap to be homogeneous across all the important grandparent fixed characteristics (Wooldridge 2004, 2005). See Appendix 1 for a full description of different data generation processes underlying the regression and conventional fixed effects models.

To compare how regression, conventional fixed effects model and my cumulative fixed effect models would produce different estimates which provides different implications on the concept of grandparent overlap effect, I present the estimates of OLS, conventional FE and CFE with respect to the same set of grandparent time-varying covariates using the same sample. Before discussing the sample, variables and results, I briefly discuss the idea of the

CFE approach and its identification assumption in this section.

The Cumulative Fixed Effects Approach

CFE strategy relies on a panel data covering at least three test outcomes of the grandchildren at time point q1, q2 and q3. The grandparent time-varying covariates measured across the first and the third tests (q1 to q3). The identification strategy exploits the variation contributed by grandparents who die after the second test and before the third test, thus $A_{iq1} = q1$, $A_{iq2} = q2$ and $q2 < A_{iq3} <= q3$. I will elaborate on my data in further details in the Data and Measure section.

To identify and estimate the overlap effect, the empirical approach of CFE includes three steps. First, I will eliminate all form of unobserved confounders, U_{i1} and $U_{i2}A_{it}$ through double differences (differencing the resulting equations achieved from the first difference estimator), so that I could identify the coefficients of grandparent observables β. Second, I will intervene grandparent time-varying characteristics in the extended stage of overlap $E[C_{A_{it}+1}]$ with values of substantive interest. Third, I will recover the effect of the unobserved confounders, $E(U_{i2})$, which constitutes the overlap effect by the residual method. To achieve statistical efficiency, I can use the GMM package in Stata 15 for estimation and inference. A more detailed description of these three steps will be included in the appendix.

Identification Assumptions

Strict exogeneity condition is assumed to identify the grandparent overlap effect. For $i \in (1, \ldots, N)$, and $t' \in (1, \ldots, t)$,

$$E(\epsilon_{it} | \bar{C}_{iA_{it}}, A_{it}, U_{i1}, U_{i2}) = 0. \tag{2.3}$$

It implies that 1) past test scores of grandchildren do not cause the death of grandparents or any current treatment, which appeals plausible immediately, 2) no time-varying unobserved confounders after eliminating U_{i1},U_{i2} and conditioning on the history of C_{it}, or $\bar{C}_{iA_{it}}$. Same as the assumption of conventional fixed effects models, the conditional means of the exogenous time-varying unobserved variables, ϵ_{it}, are assumed to be zero over time.

In the context of grandparent effect, the assumption suggests that the effect of time-varying unobserved confounders of both the grandparent overlap and grandparent characteristics should be entirely mediated by the observables. For instance, a car accident of the grandparent (or the grandparent's mental health status) may be a time-varying unobserved confounder which affects both their survival and the grandchild's test score. By assuming strict exogeneity, its effect on the grandchild's test score is assumed to be entirely mediated by the grandparent observables, such as their income, health, and marital status and the effects of their fixed characteristics, U_{i1} and U_{i2}. In another word, the length of overlap and each grandparent characteristic are "randomized" after conditioning on U_{i1} and U_{i2} and the history of all the grandparent characteristics.

Because I aim at identifying the total grandparent overlap effect originated from the grandparent time-varying and invariant characteristics, I should not control for parental characteristics which are mediators of the grandparent overlap effect. For instance, grandparent coresidence may affect the employment time and occupation of mothers to affect the grandchild's development. Because I aim at identifying the total effects of grandparents (which includes the effect of grandparent coresidence), I shall not control for the mother's employment. Otherwise I would risk inducing over-control biases or controlling away part of the effect of grandparent coresidence on the grandchild's development (Elwert and Winship 2014). As such, I assume that parental time-varying characteristics do not affect the grandparent's survival status and grandparent time-varying characteristics.[3] I note that this

[3]This is the inevitable trade-off for my inclusive conceptualization of grandparent overlap effect which encompasses all sorts of mechanisms intiated by grandparent characteristics or during the shared multigenerational life-course overlap. Alternatively, if I allow for parental time-varying confounders, I could adjust for them while assuming that they do not mediate the grandparent effects on the grandchildren, which I deem to be less likely.

may be a strong assumption imposed by the analytical framework. Given the social context of Denmark, this assumption may be more likely compared to the context of the US. Because the pension and health care system for the senior are more equalized across the family capacity.

Data and Measure

Data

I use data from the Danish administrative school and population registries from the year 2009 to 2016 to empirically identify and estimate the grandparent overlap effects. Since 2010, the Danish Ministry of Education has tested Danish children from grades 2 to 8 annually or biennially using standardized tests (Beuchert et al. 2014). Through the Danish fertility database and population registries, I can link all children to their parents and grandparents. From death records, I obtain precise dates of death for all grandparents. Through the hospital and income records, I obtain time-varying grandparent characteristics. The large sample size and high data quality allow us to robustly estimate the grandparent overlap effects.

The population of my analysis is the children who are born in Denmark on the year of 2001 and 2002 with available Danish language test score at Grade 2. I link the focal children to their grandparents who are alive and living in Denmark when their grandchildren attend the second grade. I are able to construct four samples of the lineages of maternal grandmother (MM), paternal grandmother (FM), maternal grandfather (MF) and paternal grandfather (FF), with 77,344, 70,449, 66,350, and 58,705 grandparent-grandchild pairs respectively.

For the analysis of the study, I rely on the sample of grandchildren who have taken three standard Danish language tests that are administrated at Grade 2, 4 and 6. During the years when grandparents are not present in Denmark, their residential addresses and family relationships are not recorded by the Danish Register, bringing in missingness into measures of coresidence and marital status (less than 0.34%). I remove the grandparents who ever resides outside Denmark. I retain the following number of grandparent-grandchild pairs: 68,614, 62,639, 58,793, 52,148 from the maternal grandmother, paternal grandmother, maternal grandfather and paternal grandfather lineages respectively (Sample 1). Besides, the CFE strategy requires variations in lengths of overlap at each of the three tests. So I use the sample with grandparents died after the year of the second test, so that the duration of overlap at the second test is different from the length of overlap at the third test. The final sample of grandparent-grandchild pairs are 66,333, 59,972, 55,148, 48,285 from the maternal grandmother, paternal grandmother, maternal grandfather and paternal grandfather lineages respectively (Sample 2). This is the sample I use for the analysis.

All the variables contributing to the measure of grandparent overlap are calculated on a yearly basis, including the grandparent death year, test year, the grandchild's birth year. Take for example a grandchild born in year 2001 whose grandmother died on 5/30/2012, which is assumably between when wave 3 register is recorded and when wave 4 information is recorded. Year 2012 is also the year when the second test takes place. His grandmother's overlap at the second test is calculated as the gap between the year of the second test (2012) and his birth year (2001), which is 11, while his grandmother's overlap at the third test is calculated as the gap between the year of the grandmother's death (2012) and his birth year (2001), which is also 11. So he and his maternal grandmother would not be included in the analysis. Therefore, as illustrated by the following Figure 2.1, only those grandparents who are dead between wave 4 and 5 (on the year of 2013) and after are included, amounting to missingness only in the observations of wave 5, but not wave 4. The number of grandparents who are died between wave 4 and 5 are 825, 908, 1327 and 1422 on the maternal grandmother, paternal grandmother, maternal grandfather, and paternal grandfathers lineages respectively, which is denoted as Sample 3.

The comparison of the descriptive statistics of the population and the three samples (Sample 1, Sample 2 and Sample 3) is shown in Appendix Table 2.5. Generally speaking, the samples represent the populations closely. But not surprisingly, grandparents of sample

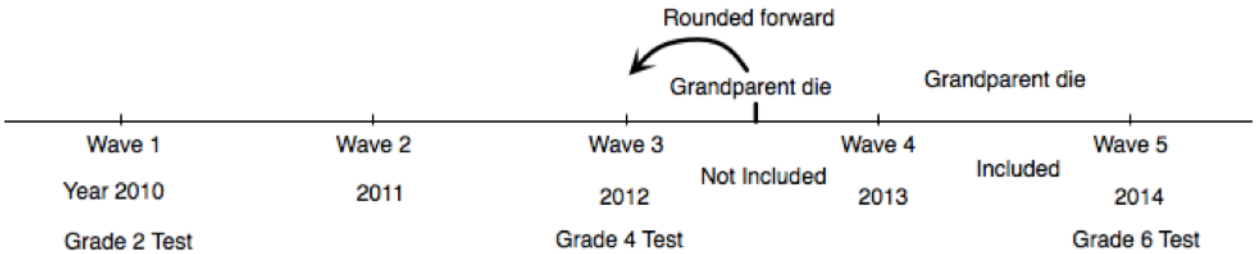


Figure 2.1: Illustration of Sample Inclusion and Exclusion Criterion

1 and sample 2 tend to be younger, employed, and are slightly more advantaged in terms of health, salary, marital status. The differences between Sample 1 and Sample 2 are rather small with respect to means and SDs of the grandparent covariates, with the grandparents represented in Sample 2 slightly more better-off compared to the grandparent in Sample 1. This is not surprising either given that Sample 2 only includes grandparents who die after the second test. In contract, grandparents of Sample 3 are older and more disadvantaged in terms of health, salary, and occupational status.

Measures

Grandchild test scores. I measure the grandchildren's cognitive test score by the standardized Danish language scores at the grandchild's grade 2, 4 and 6. The Danish national tests are compulsory and standardized for all public schools in Denmark. The tests are taken online and teachers are not involved in the test administration or evaluation. Thus, the test scores serve as a objective measure of the children's language ability. The reading tests involve questions such as word-to-picture matching, word splitting, or reading a text and answering questions. Earlier reading achievements have been shown to predict children's later educational outcomes with strong predictive validity (Beuchert and Nandrup 2017). For comparability and interpretation, I normalize the reading test scores to mean zero and unit standard deviation, consistent with the previous research (Beuchert and Nandrup 2017).

Grandparent overlap. Grandparent overlap is constructed by the grandchild's birth year, the test year, and the grandparent's death year. If the grandparent dies before the test year, grandparent overlap equals the grandchild's age at the grandparent's death. If the grandparent dies after the test year, grandparent overlap equals the grandchild' age at the tests. The three test outcomes correspond to three repeated measures of grandparent overlaps.

Grandparent characteristics. In my analysis framework, I believe that both grandparent's capacity and availability are important elements constituting the grandparent overlap effect because they could either affect the resources of the kin network, the nature of grandparent role or the intensity of grandparent-grandchild interactions. The multigenerational mobility literature emphasized the role of grandparent's social economic status and capacity in shaping the grandchildren's development (Mare 2011; Song 2016). In my analysis, grandparent's social and economic capacities are measured with their salary and occupation. The family literature, on the other, stresses more on family structure and grandparent's availability in affecting the patterns of multi-generational interactions which constrains their time and resources that can be directed to their children and grandchildren even when they are fully willing to (Cherlin and Furstenberg Jr 1992; Dunifon et al. 2014; Igel and Szydlik 2011). I consider working status, grandparent health, coresidence and marital status are important indicators of grandparent's availability. All these variables are measured from children's grade 2 to grade 6. As will be elaborated below, all these factors may affect the grandchild's development while these factors themselves may be endogenously affected by the fundamental causes of family's class and culture capital, which are usually unobserved.

Grandparent age. I include grandparent's age at the grandchild's tests as a time-varying measurement. The grandparents' age is determined by both their age of having their adult children (i.e. the parent's generation) and the parent's age at having the grandchildren. Older grandparents who give birth later (and/or whose children give birth later) may be more financially secure to invest in their children and grandchildren, and they may have less competing obligations at work or home and thus are more available to participate in child care if they are also healthy. On the other hand, older grandparents may be frailer with an increased chance of death. Although past literature finds older grandparents are more likely to give money and gifts to children (Silverstein and Marenco 2001), and their children show a higher language ability (Fomby et al. 2014), it is not clear if the association is due to the effect of grandparent age or the latent social-economic status associated with people and/or their children who give birth later.

Grandparent occupational status. Grandparent occupation is perhaps one of the most important characteristics in the multi-generational mobility literature to approximate their class. Grandparent's professional-management occupations have been found to positively associate the grandchild's occupation status after adjusting for the parental occupation (Chan and Boliver 2013; Hertel and Groh-Samberg 2014; Knigge 2016). The traditional two-generational framework of social mobility research assumes the social advantages of a high-standing grandparent is totally absorbed by the parents to pass on to the children (Becker and Tomes 1986; Blau and Duncan 1967). By contrast, the literature of multigenerational mobility investigates the direct effect of grandparent's class on their grandchildren net of the mediation of the parents (Song 2016). It is argued that grandparent class is important when they make various contacts with their grandchildren, during which the professional knowledge, grand-parenting styles and culture capital (such as vocabulary diversity) of a high-status grandparent can be transmitted to their grandchildren's developmental advantages (Cherlin and Furstenberg Jr 1992; Lareau 2011). Besides, the role of grandparent occupation manifests regardless of the specific social contexts; which has been shown to help their grandchildren return to their class origins (e.g. Jewish professional and bourgeois families in Hungary after the Communist regime, see Andorka 1997), transform their class standing during industrialization in Western Europe (Knigge 2016) and maintain their social privileges in modern democratic societies, such as in the UK and the US (Chan and Boliver 2013; Hertel and Groh-Samberg 2014). Because of the data limitations, most studies have only focused on the class status of the grandfathers but not grandmothers, and do not provide further causal evidence to separate the effect of occupations from other underlying family resources. In my study, I measure the occupation status of all grandparents with two binary variables, high-status occupation (low occupation=0, which is equivalent to management and professional class) and self-employmentnon self-employed=0).

Grandparent income. Today's grandparents are financially more secure than the elderly in the past. Research shows that grandparents are predominantly givers rather than receivers of money in modern European families (Albertini et al. 2007). This suggests that grandparent's income contributes to their adult children and grandchildren's well-being. Existing research in multigenerational mobility has emphasized the positive influence of grandparent income on grandchild's educational achievement and status attainment (Erola et al. 2018; Lindahl et al. 2015; Wightman and Danziger 2014). Although it is debatable whether part of the association actually represents the process producing it; such as the effect of grandparent cognitive and non-cognitive abilities (Carneiro and Heckman 2002) and the social environment on children's cognitive development (Sharkey and Elwert 2011), grandparent income itself may still contribute to their grandchildren's development via purchasing mechanisms (investing in stimulating activities and environment, for instance) (Cherlin and Furstenberg Jr 1992; Mayer 1997; Wightman and Danziger 2014). In my analysis, the grandparent incomes are measured as log transformed values of the total personal incomes.

Grandparent employment status. The multigenerational mobility literature has not investigated the relationship between grandparent's employment status and the grandchildren's cognitive ability extensively. But a few studies in the family literature show that working grandparents are less likely to provide child care on a regular basis (Hank and Buber 2009), and employed grandparents tend to participate in periodical grandparent tasks in a welfare state (Igel and Szydlik 2011). So I hypothesize that grandparent's employment status may affect the grandchildren's cognitive development via the forms of their interaction (remotely vs. inter-personally), the involved activities when they spend time together (Dunifon et al. 2018) as well as the interaction frequency and intensity (duration per time) (Igel and Szydlik

2011). In my study, the grandparent retirement status and employment status are coded as binary variables measured at each of the five waves. I also measure the days grandparents have worked in the past year (log transformed).

Grandparent family structure. There is little evidence of how grandparent's marital and widowhood status affects the grandchildren's cognitive development. Given that married people are more advantageous in various social-economic dimensions, I would expect better children's outcomes for whom with married grandparents than non-married grandparents. Grandparent's marital and widowhood status may affect the grandchild's well-being via their availability to help and possibly also the direction of intergenerational transfer. Past research shows that married grandparents have more frequent contact with their grandchildren compared to their nonmarried counterparts (Uhlenberg and Hammill 1998). On the other hand, widowed grandparents, especially grandmothers, may have more time and flexibility to devote to their children and grandchildren as they no longer have the competing need to care for their spouses. In spite of these, the association between grandparent's marital status and the grandchildren's well-being may be biased from their causal effect as they may reflect the latent social-economic status and cultural preference (McLanahan 2004; Elwert and Christakis 2008). In this chapter, grandparent's family structure is measured by three dummy indicators across the five waves; widowed, married and non-married (including divorced and never-married, which are combined because of the small number of cases).

Grandparent coresidence. A coresident grandparent has been shown to interact with the grandchild more intentively (Cherlin and Furstenberg Jr 1992; Dunifon et al. 2014), and some studies even find that multigenerational coresidence is necessary for grandparents to really affect the grandchildren's educational outcome (Zeng and Xie 2014). Although findings are mixed regarding whether grandparent coresidence affects the grandchildren's cognitive development positively or negatively given that multigenerational coresidence is highly selective in the western context (Dunifon et al. 2014), it is reasonable to believe that a coresident grandparent may interact with the grandchildren more frequently (Cherlin and

Furstenberg Jr 1992; Uhlenberg and Hammill 1998), engage in different activities (Dunifon et al. 2018; McKinney 2015), and change the family dynamics (such as parental stress, mother's employment Mueller et al. 2002) which may bring about different consequences on children's development. In my analysis, I measure grandparent coresidence as as dummy variable recovered from the de-identified residential addresses of each individual grandparent and the parents.

Grandparent health. Today's grandparents are not only able to liver longer, but also healthier, which enable the grandparents to participate in their grandchildren's lives more actively (Margolis and Wright 2017). In spite of the lack of direct evidence of how grandparent health affect children's cognitive ability, it is reasonable to believe that grandparent health may condition their ability to provide child care and affect their role as either a provider or recipient of cross-generational transfer of time and money (Hank and Buber 2009; Igel and Szydlik 2011). Even in Scandinivian countries like Denmark, health and life expectancies of the seniors are highly unequal across different social-economic groups (Aburto et al. 2018). I hypothesize that families with healthy grandparents may also have high-achieving children, both because grandparent health can causally affect the grandchildren via changing the allocation the family resource and time, and because grandparent's health may reflect the latent social class of the family (Link and Phelan 1995; McKinney 2015). In my study, grandparent's health status is measured with Charlson index (Charlson et al. 1987)and number of days in hospital. Charlson Indexes are coded into two binary variables; $Charlson\ Index = 1$ and $Charlson\ Index \geq 2$. Similar to *Days Works*, *Days in Hospital* is a log transformed value of the total number of days the grandparent has been admitted in hospital in the past year.

Empirical Result

Sample Description

Table 2.1 shows the descriptive statistics of the treatment, grandparent overlap and the outcome, the grandchildren's Standardized Danish score. First, grandparents from all of the four lineages share a similar length of life-course overlap with their grandchildren when they attend grade 2, 4 and 6. Relatively speaking, the grandmothers enjoy a longer overlap compared to the grandfathers, and maternal grandparents share a longer life-course exposure with their grandchildren compared to paternal grandparents ($Maternal\,Grandmothers > Paternal\,Grandmothers > Maternal\,grandfathers > Paternal\,Grandfathers$). Because most grandchildren take the first tests at age 9 (83% of all focal grandchildren take the first test at age 9, 16% take the first test at age 10, a total of less than 0.7% taking the first test at either age 8 or 11), the lengths of overlaps average 9,11 and 13 at the first, second and third tests. The standard deviations of lengths of overlaps are the same for the 1st and the 2nd test. The increased standard deviations at the 3rd test are associated with the grandparent's death between the 2nd and the 3rd test.

Table 2.2 presents the descriptive statistics of grandparent characteristics. First, the grandparents' age differs across the four lineages when their grandchildren are in the second grade. Grandmothers age an average of 63.70 and they are almost 4 years younger than the paternal grandfathers (average age 67.37) on the baseline wave. The life-course overlap shared between a grandchild and the maternal grandmother is longer than the rest and the difference enlarges over the five waves, which suggests both a age difference at the baseline and a survival advantage of the maternal grandmother compared to the other grandparents.

Second, there are important variation in grandparent's employment, occupational status and income as they age. The employment rate declines, the retirement rate increases, occupational status declines and the income decreases from wave 1 to wave 5. Among the four lineages of grandparents, the paternal grandmother has the highest retirement rate and lowest employment rate throughout the five waves, while the maternal grandfather has the smallest retirement rate and largest employment rate. This may be because paternal grandmothers tend to be older than maternal grandmothers and women are less likely to remain employed as they age. Occupational status (High/Low/Self-employed) is only measured for those who are employed, among which, the maternal grandmother is mostly like to have a high status occupation (28.04%), following by the paternal grandmother (27.73%), the maternal grandfather (24.38%) and the paternal grandfather (22.61%) measured at the baseline wave. The proportion of the employed grandparents who have high status occupation declines as they age. Grandfathers and grandmothers of same lineages tend to have similar numbers of working days, with the maternal side (log mean around 1.5) longer than that of the paternal side (log mean around 1.1). The maternal grandfathers have the highest income (log mean 6.23), followed by the maternal grandmothers (log mean 5.90), the paternal grandfathers (log mean 5.7) and the paternal grandmothers (log mean 5.19). Again, this may reflect the maternal grandfathers' age and gender advantages in the labor market.

Third, the grandparent's coresidence status and marriage and family relationship constitutes the decryption of their family structure. All grandparents become less likely to coreside with the parents and children as they age, although the variation is small. The coresidence rates of paternal grandparents (1.4%) are higher than that of maternal grandparents (0.7%), with little changes over time and little differences between grandmother and grandfather of the same lineage. Grandparents become less likely to remain married and more likely to be widowed as from wave 1 to wave 5. Greater variation in marital status of grandparents exists across grandparent gender than lineages; grandfathers of both lineages are more likely to be married, and grandmothers are more likely to be non-married (divorce and never-married) and widowed.

Finally, measured by the Charlson Index and Days in Hospital, grandparent health gets worse from wave 1 to wave 5 as they age. Both Chaslson Index and Days in Hospital reflect a great gender difference. Compared to grandfathers, both paternal and maternal grandmothers are more likely to have a Charlson Index which equals to one (one type of chronical disease). But interestingly, both grandfathers are more likely to have a Charlson Index that is larger than one. This pattern is perhaps because grandfathers are older to start with and tend to have worse health conditions. This is consistent with the past findings on gender differences in morbidity and mortality (Case and Paxson 2005). The total number of days in hospital increases from the first wave to the fifth wave. Paternal grandfathers spend the longest time in hospital, following by the maternal grandfather, the paternal grandmother and the maternal grandmother.

Grandparent Characteristics and Grandparent Overlap

Social scientists may want to know how grandparent social economic characteristics determines the length of grandparent overlap before asking about the grandparent overlap effect on grandchildren's cognitive development. The length of overlap is not random, but is likely a function of grandparent's social-economic characteristics, family structure and health (Margolis and Wright 2017; Song and Mare 2019). Table 2.3 shows the regression coefficients of duration of overlap at the child's third test (the 5th wave, Grade 6) on the baseline grandparent characteristics (the 1st wave, Grade 2). Not surprisingly, grandparent baseline age is negatively associated with the length of overlap as older grandparents are more likely to die during their grandchildren's primary school years. Consistent with the mortality compression theory (Myers and Manton 1984), the relationship between grandparent's baseline age and overlap is concave (with negative coefficient of square term of ages) showing that the increase in the grandparents chances of death decreases as they age. The indicator of grandparent's health, $Charlson\ Index \geq 2$, is most strongly correlated with the duration of overlap at the 5th wave, with a negative correlation coefficient ranging from -0.048 to -0.030 from paternal grandfathers to paternal grandmothers. Being widowed or non-married is weakly negatively associated with the length of grandparent overlap, with larger coefficients for grandfathers than for grandmothers. However, grandparent's employment, retirement, occupational status and salary, seem to be strongly correlated with the duration of grandparent overlap on the 5th wave.

Different Models, Assumptions, and Total Grandparent Overlap Effect

I examine how grandparent overlap affects the grandchild's test outcomes with least square regressions, fixed effects models and the cumulative fixed effects models and compare the estimates from each models. All the estimates on the grandparent overlap effects are summarized in Table 2.4. As is shown in row 1 of Table 2.4, grandparent overlaps are found to be positively associated with grandchild's test scores at the third test for paternal grandmothers when the baselines covariates are not adjusted for (except for the grandchild's age at the 2nd test). After I adjust for the baseline grandparent characteristics, the length of grandparent overlaps are found to be positively associated with the grandchild's test scores for maternal grandmothers (0.086 of a standard deviation, $p<0.01$), paternal grandmothers (0.102 of a standard deviation, $p<0.01$), and maternal grandfathers (0.070 of a standard deviation, $p<0.001$). This suggests that baseline grandparent characteristics are likely to confound the relationship between the duration of shared life-course overlap and the grandchild's test outcome simutaneouly, which is consistent with the past literature (Song and Mare 2019). The OLS regressions would still assume no time-varying confounders or unobserved confounders; that the grandparent time-varying income, health and family structure do not affect their future survival status and the grandchild's outcomes. Besides, unobserved confounders such as the grandparent ability, social and cultural capital are assumed away, which can be rather unrealistic. If the unobserved confounders affect the length of overlap and grandchildren's outcomes positively, then OLS regression estimates tend to be overestimated.

The third row of Table 2.4 shows the estimated coefficients of grandchild outcome on grandparent overlap from bivariate fixed effects models without adjusting for any covariates, and the fourth row shows the estimates from the same models adjusting for time-varying grandparent confounders. Compared to OLS regressions, the coefficients decrease extensively to around 0.005 for the four grandparents. Compared to OLS estimates, the coefficients are less volatile across grandparent's gender and lineage and remain largely stable even after furthur adjusting for grandparent time-varying characteristics. Eventually, the estimated grandparent overlap effect of maternal grandmother is 0.006 ($p<0.01$), that of paternal grandmothers is 0.005 ($p<0.01$), maternal grandfathers 0.004 ($p<0.01$) and paternal grandfathers 0.005 of a standard deviation ($p<0.01$). The difference between the OLS estimates and fixed effects estimates confirms the importance of considering unobserved fixed confounders. It is likely that these fixed unobserved grandparent characteristics are positively associated with the grandparent's survival and grandchildren's cognitive advantage and thus inflate the estimates of OLS regression. This echoes the premise of the multigenerational mobility literature, that grandparent's ability, cultural and social capital and their underlying social institution are largely unmeasured and outlive individuals (Mare 2011).

The last row of Table 2.4 shows the estimated coefficients of grandparent overlap effect using the cumulative fixed effects model. The estimates of grandparent overlap effects lie between the estimates from OLS and conventional fixed effects estimates, with that of maternal grandmothers 0.015 ($p<0.01$), that of paternal grandmothers 0.019 ($p<0.01$), maternal grandfathers 0.016($p<0.01$), and paternal grandfathers 0.022 of a standard deviation ($p<0.01$). This is below 0.05 SD which is equivalent to a small effect according to the benchmarks proposed by Kraft (2020), but with increase of exposure to all the four lineages of grandparents jointly, the jointly grandparent overlap effect is likely to be a medium effect (between 0.05 and 0.20SD). For ten years' increase in grandparent overlaps for all lineages, the grandchildren's cognitive test score is likely to increase by a large effect size (above 0.2 SD). [4]

The reason for the stronger effect of paternal grandfathers may be that they are most

[4]Kraft (2020) draw from the distribution of 1,942 effect sizes from 747 RCTs evaluating education interventions with standardized test outcomes during preK–12, and propose that a effect size less than 0.05 SD is small, 0.05 SD to less than 0.20 SD is medium, and 0.20 SD or greater is large.

positively selected on their social-economic status in order to survive their grandchild's grade 4 and be selected in my sample of analysis, because they tend to be the oldest among the four lineages. Besides, they show a larger $E[U_{i2}]$ which suggests they may have a larger effect on the grandchild due to their higher education and other cultural and social capital. The CFE estimates are based on a more realistic data generation model of grandparent overlap effect which takes into account effect heterogeneity of grandparent observed and unobserved characteristics. Details of the estimation results are shown in the Appendix Table 2.9 to 2.13.

Robustness Check

Sample selection and age effect

Child's age and grandparent age are perhaps the most important confounders of grandparent overlap effect. Older grandparents are likely less helpful to the grandchildren and are also less likely to survive. I have adjusted for grandparent age as a time-varying grandparent confounder in the models. Less obvious it may be, children's age may also confound the relationship between grandparent overlap and grandchild's cognitive achievement. This is because older children may have some cognitive advantages in school and may respond to the grandparent's inputs in different ways. Although the majority of the children in my sample aged 9 at grade 2 (83%), the range of children's age is from 8 to 11. This would not be a concern of the cumulative fixed effects models if the age effect is entirely captured by U_1 and $U_2 * A_{it}$. But if the children's age affect the outcome in a non-additive way, or as interaction with other grandparent time-varying variables, then it would not be entirely eliminated. Therefore, it is informative to test to what extent the result is robust to children of the same age group, or whether or not the result is driven by these outlier children who are either younger or older than the majority. I restrict the sample to grandchildren who take the 2rd grade test at age 9, I find that the estimates of grandparent overlap effects are consistent and stable. Due to the sample size restriction, I can not estimate the grandparent overlap effects for every age group.

Intervening the grandparent characteristics with different values

Recall from Equation 2.2 that estimation of PAOE involves intervention on grandparent's time-varying observed characteristics in the extended stage of overlap, $E[C_{A_{it}}+1\beta]$. I regard the grandparent characteristics C_{it} as treatment variables and fix on the values of $C_{A_{it}+1}$ in order to estimate $E\big[C_{A_{it}+1}\beta_{A_{it}+1}\big]$. For robustness check, I test if I fix the value of $C_{A_{it}+1}$ with either the value of $C_{A_{it}}$, i.e. the last observation in the sequence, or with the value predicted by auto-regression fitted by the past observations of the entire sequence. I find that the estimates of PAOE are robust and close to the estimates in the result section. Specifically, the estimates of PAOE of the maternal grandmother, maternal grandfather, paternal grandmother and paternal grandfather lineages are 0.015($p<0.01$), 0.018($p<0.01$), 0.015($p<0.01$) and 0.022($p<0.01$) with the last observation approach, while the estimates of PAOE with the autoregression approach are 0.014($p<0.01$), 0.019($p<0.01$), 0.013($p<0.01$) and 0.026($p<0.01$). Although the continuous variables, such as grandparent salary, days employed and days in the hospital may be better predicted using the autoregression approach, values of the dummy variables, such as grandparent marital and occupation status may fall out of the range of 0 and 1 if predicted by autoregressions. So in the main analysis, I combine these two approaches using the last observations to predict the dummy variables while using autoregressions to predict the continuous variables.

Effect heterogeneity of baseline grandparent characteristics

Sharing a same length of overlap with grandparents may suggest different consequences for children. High-SES grandparents may benefit the children more via economic, social and cultural capitals (Møllegaard and Jæger 2015). Coresident, retired and healthy grandparents may be more available and thus interact with grandchildren with higher frequency given the same length of shared life-course overlap. On the other hand, poor, elderly and ailing grandparents may affect the grandchildren's cognitive development less intensively. Although perhaps less prevalent in social democratic welfare states such as Denmark, the elderly could affect the grandchildren negatively if they have to compete for material or interpersonal support from their adult children(Tanskanen et al. 2016).

Notice that the estimated grandparent overlap effects in the main analysis are averaged over grandparents with heterogeneous characteristic of health, family structure and SES. To test for the effect heterogeneity, I conducted subgroup analysis for grandparents with different baseline health, income, occupation, retirement, coresident and widowhood statuses, where the sub-samples are defined based on the grandparent characteristics at wave 1 (grade 2). I found that although the grandmother's overlap effects are positive averaged across grandparents with different baseline characteristics, there are certain degrees of effect heterogeneity across grandparent lineages (See Appendix Table 2.14). For instance, I found that for most of the lineages (except for the paternal grandfathers), healthy grandparents have larger grandparent overlap effects on grandchildren compared to unhealthy grandparents. High or medium occupational paternal grandmothers and older maternal grandmothers may have larger overlap effects on the grandchildren. Similar to Lehti et al. (2018), I find that the patterns of effect heterogeneity are not always consistent across the four grandparents, and it would be interesting to further investigate the differences of interactions of the four grandparents and the grandchildren.

Conclusions and Discussion

As life expectancy has risen tremendously since the 20th century, the shared life-course exposure between the grandchildren and the grandparent generation has increased, which brings out the important question of how grandparent overlap affects the children's status attainment and shape the transmission of social inequality. I conceptualize the grandparent overlap effect in a holistic way which encompasses all the effects of grandparent time-varying and time-invariant characteristics, either observed and unobserved, and provide the first empirical analysis for such conceptualized grandparent overlap effect with danish register data. I also engage with the challenges of identifying the overlap effects as the conventional fixed effects models fail to adjust for grandparent time-varying confounders and capture the unobserved confounders that are interactive with grandparent overlap. Using the cumulative fixed effects model, I show that both OLS and fixed effects estimates are likely biased due to the unobserved heterogeneity and the failure to capture the cumulative nature of grandparent overlap. I find positive grandparent overlap effects on children's 6th grade language test score which are much larger than estimates from conventional fixed effects models. The estimates are robust to a series robustness checks.

This finding suggests that the rising life expectancy and grandparent overlap may have important implications for the reproduction of social inequality across multiple generations as an understudied mechanism. First, as the length of grandparent overlap continues to rise, social institutions underlying social immobility may get increasingly transmitted in a non-markovian way via direct multigenerational interactions. This suggests that the two-generational model of social mobility assuming markovian transmission may become increasingly unlikely. Second, the increase of grandparent overlap has been shown to be unequally distributed across different social groups. Groups with advantageous social status enjoy a longer life span and a larger increment in the life expectancy compared to the disadvantaged groups (Vierboom et al. 2019). While the increase in life expectancy for the low SES families is smaller, stagnated or even reversed (especially those who are born in the 50s, 60s, and live in the rural and southern US). Because the grandparent overlap effect is positive, a longer overlap may further reinforce the transmission of the social advantages of the privileged that they already have to perpetuate and enlarge the gaps of social inequality in the children's generation. Third, I find some evidence that the effect of the same increase in grandparent overlap may be heterogeneous given the different grandparent baseline social economic characteristics, such as health, salary and working status.

The study has the following limitations: first, I can not estimate the grandparent overlap effect on a wide range of grandchild's age. Given the data availability, the current analysis focused on the grandchild's cognitive achievement at the 6th grade and exploits the variation of the sample of grandparents who survive grade 4 but die before grade 6. Because grandparents may play a different role in the children's early childhood, and children may be more susceptible to family inputs, it is likely that the grandparent overlap effects may be even larger for the grandchildren's earlier life developmental outcomes. Also, the relative importance of different lineages of grandparents may change given different children's age so their grandparent overlap effects may also differ. Second, although I have controlled for a large range of grandparent time-varying characteristics underlying the reproduction of social inequality, it may not be exhaustive. To identify the effect of grandparent overlap, I assume no unobserved time-varying confounders after conditioning on the observables and taking the fixed effects. Besides, for grandparent time-varying unobservables which constitute the grandparent overlap effect, such as the sense of closeness between the grandparent and grandchild. I assume their effect changes linearly with the length of overlap thus is captured by $U_{i2}A_{it}$. Although this has relaxed the assumption of conventional fixed effects models greatly, it may still be rather restrictive. Third, I can not fully explore the patterns of effect heterogeneity of grandparent overlap effects given this sample of analysis. For instance, I find the estimates of grandparent overlap effects to be larger for the coresident grandparents than non-coresident grandparents, but the estimates are not statistically significant due to the limited amount of data. These can be interesting for future research.

The literature of multi-generational mobility uses either grandparent income, education or occupation to capture the latent SES and capacity of grandparents and find inconsistent "grandparent effect". However, the partial effects of grandparent characteristics are likely inconsistent because each category may only reflect one dimension of a person's real social capacity. Income and occupation are affected by market luck and people make trade-offs among categories; such as between high income and a graduate degree in humanities (Clark 2015). Multigenerational transmission of inequality may be underestimated by one or two specific dimensions because of the existence of such deviations. In addition, family lineages with higher crystallization of various categories (low income, poor health, low education, and low occupation) are likely to be even more immobile multi-generationally than families with less crystallized SES. In contrast to looking at partial effects of grandparent's SES, my perspective of grandparent overlap effect envelopes various dimensions of the grandparent SES to reflect the grandparent's real social capacity more closely. To the end, I provide a new landscape of how multigenerational transmission of inequality can be pictured and add to the understanding of grandparent effects on children's cognitive development and status attainment.

Tables and Figures

Table 2.1: Sample Descriptive Statistics of the Grandparent Overlap and Child's Test Scores

		Maternal Grandmother		Paternal Grandmother		Maternal Grandfather		Paternal Grandfather	
		mean	sd	mean	sd	mean	sd	mean	sd
Standardized Danish Test Score	3rd	0.103	0.936	0.107	0.932	0.106	0.933	0.112	0.931
	2nd	0.084	0.953	0.087	0.951	0.088	0.948	0.09	0.949
	1st	0.083	0.965	0.089	0.963	0.088	0.963	0.093	0.964
Grandparent Overlap	3rd	9.158	0.383	9.15	0.376	9.154	0.38	9.148	0.373
	2nd	11.158	0.383	11.15	0.376	11.154	0.38	11.148	0.374
	1st	13.145	0.398	13.135	0.394	13.13	0.409	13.119	0.41

Note: The 1st, 2nd and 3rd tests are language test taken at grade 2,4 and 6.

Table 2.2: Sample Descriptive Statistics of Grandparent Characteristics

		Maternal Grandmother		Paternal Grandmother		Maternal Grandfather		Paternal Grandfather	
		mean	sd	mean	sd	mean	sd	mean	sd
Grandparent Age	Wave 1	63.693	6.629	65.667	6.681	65.686	6.537	67.36	6.537
	Wave 2	64.693	6.629	66.667	6.681	66.686	6.537	68.36	6.537
	Wave 3	65.693	6.629	67.667	6.681	67.686	6.537	69.36	6.537
	Wave 4	66.693	6.629	68.667	6.681	68.686	6.537	70.36	6.537
	Wave 5	67.639	6.595	69.595	6.63	69.582	6.473	71.221	6.451
Retired	Wave 1	0.627	0.484	0.704	0.457	0.565	0.496	0.628	0.483
	Wave 2	0.672	0.469	0.744	0.436	0.608	0.488	0.666	0.472
	Wave 3	0.715	0.451	0.783	0.412	0.647	0.478	0.702	0.457
	Wave 4	0.752	0.432	0.813	0.39	0.681	0.466	0.734	0.442
	Wave 5	0.78	0.414	0.837	0.369	0.707	0.455	0.754	0.43
Unemployed	Wave 1	0.052	0.221	0.04	0.197	0.037	0.188	0.027	0.163
	Wave 2	0.045	0.208	0.034	0.181	0.03	0.172	0.022	0.146
	Wave 3	0.037	0.188	0.026	0.159	0.026	0.16	0.016	0.127
	Wave 4	0.031	0.173	0.021	0.142	0.021	0.144	0.012	0.11
	Wave 5	0.027	0.163	0.018	0.134	0.017	0.13	0.01	0.1
Employed	Wave 1	0.321	0.467	0.256	0.436	0.398	0.49	0.345	0.475
	Wave 2	0.283	0.45	0.222	0.416	0.361	0.48	0.312	0.463
	Wave 3	0.248	0.432	0.191	0.393	0.327	0.469	0.282	0.45
	Wave 4	0.216	0.412	0.166	0.372	0.297	0.457	0.254	0.435
	Wave 5	0.192	0.394	0.144	0.351	0.276	0.447	0.236	0.424
Coresident	Wave 1	0.007	0.083	0.014	0.119	0.006	0.074	0.014	0.117
	Wave 2	0.007	0.083	0.014	0.118	0.005	0.073	0.013	0.114
	Wave 3	0.007	0.082	0.014	0.116	0.005	0.07	0.013	0.112
	Wave 4	0.006	0.078	0.013	0.115	0.005	0.069	0.012	0.11
	Wave 5	0.006	0.078	0.013	0.113	0.005	0.068	0.012	0.107
Self-employed	Wave 1	0.021	0.144	0.02	0.14	0.091	0.288	0.09	0.286
	Wave 2	0.02	0.14	0.019	0.137	0.088	0.283	0.086	0.28
	Wave 3	0.019	0.135	0.018	0.135	0.084	0.278	0.082	0.274
	Wave 4	0.017	0.131	0.018	0.132	0.082	0.274	0.078	0.268
	Wave 5	0.017	0.128	0.017	0.128	0.08	0.271	0.076	0.265
Low-status Occupation	Wave 1	0.209	0.407	0.165	0.371	0.21	0.408	0.177	0.382
	Wave 2	0.183	0.387	0.142	0.349	0.189	0.392	0.16	0.366
	Wave 3	0.161	0.368	0.121	0.327	0.17	0.376	0.143	0.35
	Wave 4	0.141	0.348	0.105	0.307	0.155	0.362	0.127	0.333
	Wave 5	0.127	0.333	0.092	0.289	0.143	0.351	0.117	0.321
High-status Occupation	Wave 1	0.09	0.287	0.071	0.257	0.097	0.296	0.078	0.269
	Wave 2	0.079	0.27	0.061	0.239	0.084	0.278	0.067	0.25
	Wave 3	0.068	0.252	0.051	0.221	0.072	0.259	0.057	0.232
	Wave 4	0.058	0.233	0.043	0.202	0.061	0.239	0.048	0.214
	Wave 5	0.048	0.215	0.035	0.185	0.053	0.223	0.043	0.203

Charlson Index = 1	Wave 1	0.155	0.362	0.156	0.363	0.13	0.336	0.137	0.344
	Wave 2	0.159	0.366	0.162	0.369	0.139	0.346	0.144	0.351
	Wave 3	0.165	0.371	0.168	0.374	0.145	0.352	0.151	0.358
	Wave 4	0.175	0.38	0.174	0.379	0.155	0.362	0.159	0.365
	Wave 5	0.181	0.385	0.18	0.384	0.158	0.365	0.164	0.37
Charlson Index ≥ 2	Wave 1	0.065	0.247	0.07	0.255	0.085	0.279	0.094	0.292
	Wave 2	0.074	0.262	0.079	0.269	0.1	0.3	0.108	0.31
	Wave 3	0.085	0.278	0.09	0.286	0.115	0.319	0.128	0.334
	Wave 4	0.096	0.294	0.103	0.304	0.133	0.34	0.148	0.356
	Wave 5	0.103	0.305	0.111	0.315	0.142	0.349	0.16	0.366
Non-married	Wave 1	0.134	0.341	0.124	0.329	0.097	0.296	0.084	0.277
	Wave 2	0.136	0.343	0.124	0.33	0.099	0.299	0.086	0.28
	Wave 3	0.138	0.345	0.126	0.332	0.101	0.301	0.087	0.282
	Wave 4	0.139	0.346	0.127	0.333	0.101	0.302	0.089	0.285
	Wave 5	0.141	0.348	0.128	0.334	0.103	0.304	0.089	0.284
Widowed	Wave 1	0.117	0.322	0.148	0.355	0.045	0.208	0.052	0.222
	Wave 2	0.127	0.333	0.159	0.366	0.049	0.217	0.058	0.233
	Wave 3	0.137	0.344	0.171	0.377	0.054	0.227	0.063	0.244
	Wave 4	0.148	0.355	0.183	0.387	0.06	0.237	0.07	0.256
	Wave 5	0.158	0.364	0.195	0.396	0.063	0.243	0.075	0.263
Married	Wave 1	0.748	0.434	0.729	0.445	0.858	0.349	0.864	0.343
	Wave 2	0.737	0.44	0.716	0.451	0.852	0.356	0.857	0.351
	Wave 3	0.725	0.447	0.703	0.457	0.845	0.362	0.849	0.358
	Wave 4	0.713	0.452	0.689	0.463	0.839	0.368	0.841	0.366
	Wave 5	0.701	0.458	0.677	0.468	0.834	0.372	0.837	0.37
Log (Salary)	Wave 1	5.899	4.72	5.187	4.424	6.272	4.812	5.705	4.61
	Wave 2	5.498	4.576	4.821	4.231	5.847	4.686	5.294	4.442
	Wave 3	5.114	4.404	4.466	4.001	5.483	4.545	4.963	4.267
	Wave 4	4.763	4.209	4.168	3.774	5.158	4.384	4.672	4.086
	Wave 5	4.471	4.016	3.917	3.558	4.914	4.242	4.478	3.943
Log (Days worked)	Wave 1	1.576	2.567	1.173	2.315	1.466	2.509	1.114	2.273
	Wave 2	1.352	2.44	0.976	2.156	1.239	2.367	0.92	2.108
	Wave 3	1.148	2.299	0.796	1.983	1.037	2.21	0.741	1.926
	Wave 4	0.952	2.137	0.643	1.81	0.848	2.04	0.582	1.734
	Wave 5	0.802	1.99	0.519	1.645	0.702	1.882	0.47	1.574
Log (Days in Hospital)	Wave 1	0.677	3.762	0.723	4.122	0.886	4.813	0.946	4.783
	Wave 2	0.694	4.05	0.767	4.137	0.918	4.472	1.037	5.594
	Wave 3	0.751	3.987	0.822	4.402	1.024	4.852	1.132	5.352
	Wave 4	0.851	4.481	0.92	4.696	1.24	5.77	1.366	6.439
	Wave 5	0.882	4.563	0.955	4.933	1.231	5.715	1.371	6.307
N		66333	66333	59972	59972	55148	55148	48285	48285

Table 2.3: Estimated coefficients from OLS regression of the duration of overlap at the third test on grandparent covariates at the first wave

Covariates	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Unemployed	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)
Retired	0.00	(0.00)	-0.008**	(0.01)
	(0.00)	(0.00)	(0.00)	(0.00)
Corresidence	0.00	0.00	(0.00)	0.00
	(0.01)	(0.00)	(0.01)	(0.01)
Self-employed	0.00	0.01	0.00	0.01
	(0.00)	(0.01)	(0.00)	(0.00)
High-status Occupation	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Charlson Index=1	-0.005***	-0.006***	-0.004**	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)
Charlson Index≥2	-0.031***	-0.030***	-0.042***	-0.048***
	(0.00)	(0.00)	(0.00)	(0.00)
Non-married	-0.005***	-0.006***	-0.017***	-0.028***
	(0.00)	(0.00)	(0.00)	(0.00)
Widowed	(0.00)	-0.005***	-0.014***	-0.018***
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Salary)	0.000*	0.00	0.001**	0.001**
	0.00	0.00	0.00	0.00
Log (Days Worked)	0.00	0.00	-0.001**	-0.002**
	0.00	(0.00)	(0.00)	(0.00)
Log (Days in Hospital)	-0.001***	-0.001***	-0.002***	-0.001***
	0.00	0.00	0.00	0.00
Standardized Age	-0.006***	-0.007***	-0.012***	-0.016***
	(0.00)	(0.00)	(0.00)	(0.00)
Age Square	-0.004***	-0.005***	-0.006***	-0.009***
	0.00	0.00	0.00	(0.00)
Child Age at 1st Test	1.000***	0.998***	0.999***	1.002***
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	3.996***	4.016***	3.998***	3.969***
	(0.01)	(0.01)	(0.02)	(0.02)
N	66333.00	59970.00	55148.00	48284.00
R2	0.92	0.91	0.86	0.84

Note: Standard Errors in parenthesis
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.4: Comparison of estimated grandparent overlap effects across models

Model	Maternal Grandmother	Paternal Grandmother	Maternal Grandfather	Paternal Grandfather
OLS[a]	0.051 (0.033)	0.076** (0.031)	0.038 (0.026)	0.008(0.025)
OLS[b]	0.086*** (0.032)	0.102*** (0.031)	0.070*** (0.026)	0.037***(0.025)
Fixed Effects [c]	0.005***(0.001)	0.005***(0.001)	0.005***(0.001)	0.005***(0.001)
Fixed Effects [d]	0.006***(0.001)	0.005***(0.001)	0.004***(0.001)	0.005***(0.001)
CFE [e]	0.015***(0.003)	0.019***(0.003)	0.016***(0.003)	0.022***(0.003)

Note: Standard Errors in parenthesis
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$
[a] : Controlling for grandchild's age at the 2nd Grade test.
[b] : Controlling for grandchild's age and grandparent characteristics at the 2nd Grade including retired, employed, coresidence, self-employed, high-status occupation, Charlson Index=1, Charlson Index larger than 1, Non-married, widowhood status, salary, days of working, days in hospital, grandparent age (and its squared term).
[c] : No covariates.
[d] : Controlling for grandparent time-varying characteristics including retired, employed, coresidence, self-employed, high-status occupation, Charlson Index=1, Charlson Index larger than 1, Non-married, widowhood status, salary, days of working, days in hospital.
[e] : Controlling for grandparent time-varying characteristics including retired, employed, coresidence, self-employed, high-status occupation, Charlson Index=1, Charlson Index larger than 1, Non-married, widowhood status, salary, days of working, days in hospital and grandparent age.

Appendix Tables

Table 2.5: Comparison of Population and Sample Characteristics (Maternal Grandmother

Maternal Grandmother	Population mean	sd	Sample 1 mean	sd	Sample 2 mean	sd	Sample 3 mean	sd
z_score	0.054	0.988	0.083	0.965	0.083	0.965	0.043	0.94
Age	63.797	6.756	63.846	6.728	63.693	6.629	67.947	7.809
Retired	0.633	0.482	0.635	0.481	0.627	0.484	0.828	0.378
Unemployed	0.053	0.223	0.051	0.22	0.052	0.221	0.04	0.196
Employed	0.314	0.464	0.314	0.464	0.321	0.467	0.132	0.339
Coresidence	0.008	0.086	0.007	0.083	0.007	0.083	0.006	0.078
Self-employed	0.021	0.144	0.021	0.143	0.021	0.144	0.012	0.109
Low-status Occupation	0.205	0.404	0.205	0.404	0.209	0.407	0.095	0.293
High-status Occupation	0.088	0.284	0.088	0.284	0.09	0.287	0.025	0.158
Charlson Index=1	0.158	0.365	0.157	0.364	0.155	0.362	0.2	0.4
Charlson Index≥2	0.075	0.264	0.075	0.263	0.065	0.247	0.255	0.436
Non-married	0.139	0.346	0.136	0.342	0.134	0.341	0.177	0.382
Widowed	0.121	0.326	0.122	0.327	0.117	0.322	0.207	0.406
Married	0.739	0.439	0.743	0.437	0.748	0.434	0.616	0.487
Log(Salary)	5.825	4.698	5.821	4.697	5.899	4.72	3.787	3.454
Log(Days_worked)	1.541	2.548	1.539	2.547	1.576	2.567	0.555	1.694
Log(Days in Hospital)	0.877	4.861	0.87	4.887	0.677	3.762	3.095	9.913
N	77344		68614		66333		825	

Table 2.6: Comparison of Population and Sample Characteristics (Paternal Grandmother)

Paternal Grandmother	Population mean	sd	Sample 1 mean	sd	Sample 2 mean	sd	Sample 3 mean	sd
z_score	0.059	0.984	0.088	0.964	0.089	0.963	0.013	1.008
Age	65.854	6.847	65.898	6.836	65.667	6.681	70.326	8.193
Retired	0.711	0.453	0.713	0.452	0.704	0.457	0.872	0.334
Unemployed	0.04	0.197	0.039	0.194	0.04	0.197	0.021	0.143
Employed	0.249	0.432	0.248	0.432	0.256	0.436	0.107	0.309
Coresidence	0.015	0.122	0.014	0.119	0.014	0.119	0.014	0.119
Self-employed	0.02	0.139	0.02	0.139	0.02	0.14	0.006	0.074
Low-status Occupation	0.16	0.366	0.16	0.366	0.165	0.371	0.079	0.27
High-status Occupation	0.069	0.253	0.068	0.253	0.071	0.257	0.022	0.147
Charlson Index=1	0.16	0.367	0.159	0.366	0.156	0.363	0.202	0.401
Charlson Index≥2	0.083	0.275	0.081	0.273	0.07	0.255	0.231	0.422
Non-married	0.127	0.332	0.125	0.331	0.124	0.329	0.161	0.368
Widowed	0.155	0.362	0.155	0.362	0.148	0.355	0.283	0.451
Married	0.719	0.45	0.72	0.449	0.729	0.445	0.556	0.497
Log(Salary)	5.103	4.386	5.098	4.382	5.187	4.424	3.537	3.213
Days_worked	1.137	2.288	1.134	2.286	1.173	2.315	0.437	1.528
Days in Hospital	0.943	5.019	0.931	4.953	0.723	4.122	3.012	9.744
N	70449		62639		59972		908	

Table 2.7: Comparison of Population and Sample Characteristics (Maternal Grandfather)

Maternal Grandfather	Population		Sample 1		Sample 2		Sample 3	
	mean	sd	mean	sd	mean	sd	mean	sd
z_score	0.057	0.985	0.086	0.963	0.088	0.963	0.071	0.945
Age	65.914	6.733	65.971	6.719	65.686	6.537	69.936	7.606
Retired	0.58	0.494	0.581	0.493	0.565	0.496	0.792	0.406
Unemployed	0.038	0.191	0.036	0.186	0.037	0.188	0.023	0.149
Employed	0.382	0.486	0.383	0.486	0.398	0.49	0.185	0.389
Coresidence	0.006	0.076	0.006	0.074	0.006	0.074	0.007	0.082
Self-employed	0.088	0.284	0.088	0.284	0.091	0.288	0.057	0.232
Low-status Occupation	0.202	0.401	0.203	0.402	0.21	0.408	0.093	0.291
High-status Occupation	0.092	0.289	0.092	0.29	0.097	0.296	0.035	0.183
Charlson Index=1	0.134	0.34	0.134	0.34	0.13	0.336	0.16	0.367
Charlson Index≥2	0.103	0.304	0.102	0.303	0.085	0.279	0.267	0.442
Non-married	0.105	0.306	0.102	0.302	0.097	0.296	0.159	0.366
Widowed	0.048	0.214	0.048	0.215	0.045	0.208	0.097	0.296
Married	0.847	0.36	0.85	0.357	0.858	0.349	0.744	0.437
Log(Salary)	6.114	4.775	6.121	4.776	6.272	4.812	4.093	3.656
Days_worked	1.398	2.47	1.401	2.472	1.466	2.509	0.526	1.649
Days in Hospital	1.192	6.061	1.191	6.022	0.886	4.813	3.331	11.433
N	66350		58793		55148		1327	

Table 2.8: Comparison of Population and Sample Characteristics (Paternal Grandfather)

Paternal Grandfather	Population		Sample 1		Sample 2		Sample 3	
	mean	sd	mean	sd	mean	sd	mean	sd
z_score	0.064	0.984	0.092	0.964	0.093	0.964	0.076	0.968
Age	67.692	6.782	67.744	6.764	67.36	6.537	71.964	7.601
Retired	0.643	0.479	0.645	0.479	0.628	0.483	0.829	0.377
Unemployed	0.028	0.165	0.027	0.161	0.027	0.163	0.013	0.115
Employed	0.329	0.47	0.328	0.47	0.345	0.475	0.158	0.364
Coresidence	0.014	0.119	0.014	0.116	0.014	0.117	0.013	0.112
Self-employed	0.086	0.28	0.086	0.281	0.09	0.286	0.048	0.213
Low-status Occupation	0.168	0.374	0.168	0.374	0.177	0.382	0.088	0.283
High-status Occupation	0.075	0.263	0.074	0.262	0.078	0.269	0.022	0.146
Charlson Index=1	0.141	0.348	0.142	0.349	0.137	0.344	0.181	0.386
Charlson Index≥2	0.115	0.319	0.114	0.318	0.094	0.292	0.271	0.445
Non-married	0.09	0.286	0.089	0.284	0.084	0.277	0.155	0.362
Widowed	0.057	0.232	0.057	0.231	0.052	0.222	0.113	0.317
Married	0.853	0.354	0.855	0.353	0.864	0.343	0.732	0.443
Log (Salary)	5.548	4.558	5.542	4.551	5.705	4.61	3.819	3.449
Days_worked	1.058	2.229	1.051	2.223	1.114	2.273	0.408	1.469
Days in Hospital	1.28	5.988	1.285	5.98	0.946	4.783	2.697	9.128
N	58705		52148		48285		1422	

Table 2.9: Estimated coefficients from OLS regression of the third test on grandparent overlap at the third test

	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Grandparent Overlap	0.05	0.076**	0.04	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Child Age at 1st Test	-0.356***	-0.374***	-0.344***	-0.305***
	(0.03)	(0.03)	(0.03)	(0.03)
Constant	2.696***	2.529***	2.757***	2,801***
	(0.16)	(0.16)	(0.14)	(0.14)
N	66333.00	59972.00	55148.00	48285.00
R^2	0.02	0.02	0.02	0.01

Note: Standard Errors in parenthesis. We adjust for the child's age in the model

* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.10: Estimated coefficients from OLS regression of the third test on the duration of overlap at the third test and grandparent covariates at the first wave

	Maternal grandmother	Paternal grandmother	Maternal grandfather	Paternal grandfather
Overlap	0.086***	0.102***	0.070***	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
Retired	0.03	0.081***	0.067***	0.110***
	(0.03)	(0.03)	(0.03)	(0.03)
Unemployed	0.03	0.049**	0.077***	0.081***
	(0.02)	(0.02)	(0.02)	(0.02)
Coresidence	-0.137***	-0.304***	-0.091*	-0.275***
	(0.04)	(0.03)	(0.05)	(0.04)
Self-employed	0.180***	0.139***	0.167***	0.158***
	(0.03)	(0.03)	(0.02)	(0.02)
High-status Occupation	0.222***	0.234***	0.255***	0.214***
	(0.01)	(0.02)	(0.02)	(0.02)
Charlson Index Equals 1	(0.02)	-0.019*	-0.020*	(0.01)
	(0.01)	(0.01)	(0.01)	(0.01)
Charlson Index Larger Than 2	-0.073***	-0.097***	-0.043***	-0.044***
	(0.02)	(0.02)	(0.02)	(0.02)
Non-married	-0.043***	-0.052***	-0.089***	-0.088***
	(0.01)	(0.01)	(0.01)	(0.02)
Widowed	-0.064***	-0.080***	-0.067***	-0.036*
	(0.01)	(0.01)	(0.02)	(0.02)
Log (Salary)	0.014***	0.012***	0.012***	0.013***
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days Worked)	-0.010***	(0.00)	(0.00)	(0.00)
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days in Hospital)	0.00	(0.00)	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Standardized Age	0.183***	0.163***	0.162***	0.135***
	(0.01)	(0.01)	(0.01)	(0.01)
Age Square	-0.038***	-0.035***	-0.035***	-0.031***
	(0.00)	(0.00)	(0.00)	(0.00)
Child Age at 1st Test	-0.371***	-0.392***	-0.358***	-0.325***
	(0.03)	(0.03)	(0.03)	(0.03)
Constant	2.318***	2.309***	2.365***	2.501***
	(0.16)	(0.16)	(0.14)	(0.15)
N	66333.00	59970.00	55148.00	48284.00
R^2	0.05	0.04	0.05	0.04

Note: Standard Errors in parenthesis
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.11: Summary of estimated coefficients from conventional fixed effects regression of grandchild outcome on grandparent overlap

	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Grandparent Overlap	0.005***	0.005***	0.005***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.034***	0.044***	0.044***	0.048***
	(0.09)	(0.09)	(0.10)	(0.10)
N	198174.00	179004.00	164113.00	143431.00
R^2	0.00	0.00	0.00	0.00

Note: Standard Errors in parenthesis. We adjust for the child's age in the model

* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.12: Estimated coefficients from conventional fixed-effect regression of grandchild outcome on grandparent overlap and characteristics

	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Overlap	0.006***	0.005***	0.004***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)
Retired	0.01	0.00	0.01	(0.01)
	(0.01)	(0.01)	(0.01)	(0.01)
Unemployed	0.01	(0.01)	(0.02)	(0.01)
	(0.01)	(0.02)	(0.02)	(0.02)
Coresidence	0.02	(0.03)	(0.02)	(0.02)
	(0.03)	(0.02)	(0.04)	(0.03)
Self-employed	0.01	(0.03)	(0.02)	-0.033*
	(0.02)	(0.03)	(0.02)	(0.02)
High-status Occupation	(0.01)	(0.01)	(0.01)	-0.023*
	(0.01)	(0.01)	(0.01)	(0.01)
Charlson Index Equals 1	0.01	0.01	0.01	(0.00)
	(0.01)	(0.01)	(0.01)	(0.01)
Charlson Index Larger Than 2	0.01	0.016*	(0.01)	(0.01)
	(0.01)	(0.01)	(0.01)	(0.01)
Non-married	(0.00)	(0.02)	(0.01)	-0.036*
	(0.02)	(0.02)	(0.02)	(0.02)
Widowed	(0.02)	0.01	(0.01)	0.01
	(0.01)	(0.01)	(0.02)	(0.02)
Log (Salary)	0.003***	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days worked)	0.00	0.004*	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days in Hospital)	0.00	0.00	0.001***	0.00
	0.00	0.00	0.00	0.00
Constant	0.00	0.036**	0.045***	0.060***
	(0.02)	(0.02)	(0.02)	(0.02)
N	198174.00	179004.00	164113.00	143431.00
R^2	0.00	0.00	0.00	0.00

Note: Standard Errors in parenthesis. We adjust for the child's age in the model

* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.13: CFE estimates of sample parameters and PAEO

	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Coresidence	0.00	(0.00)	(0.01)	(0.00)
	(0.04)	(0.03)	(0.05)	(0.04)
Self-employed	0.04	0.02	0.01	-0.062**
	(0.03)	(0.04)	(0.02)	(0.03)
High-status Occupation	0.02	0.01	0.050***	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Charlson Index=1	0.02	(0.01)	0.01	(0.00)
	(0.01)	(0.01)	(0.01)	(0.01)
Charlson Index≥2	0.01	0.02	(0.00)	(0.02)
	(0.01)	(0.01)	(0.01)	(0.01)
Widowed	0.062***	0.02	0.03	-0.054**
	(0.02)	(0.02)	(0.02)	(0.02)
Non-married	(0.03)	0.04	(0.03)	0.02
	(0.02)	(0.03)	(0.03)	(0.03)
Log (Salary)	(0.00)	(0.00)	(0.00)	(0.00)
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days Worked)	0.00	0.00	(0.00)	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Log (Days in Hospital)	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	0.00	0.00
Retired	(0.01)	(0.00)	0.02	-0.026*
	(0.01)	(0.02)	(0.01)	(0.02)
Grandparent Age	0.003***	0.005***	0.004***	0.007***
	(0.00)	(0.00)	(0.00)	(0.00)
PAEO	0.015***	0.019***	0.016***	0.022***
	(0.00)	(0.00)	(0.00)	(0.00)
N	66331	59967	55143	48279

Note: Standard Errors in parenthesis. We adjust for the child's age in the model

* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.14: Effect Heterogeneity of PAEO Across Grandparent's Baseline Health or Coresidence

	Maternal grand-mother	Paternal grand-mother	Maternal grandfather	Paternal grandfather
Ever Sick	0.010*	0.01	0.018***	0.026***
	(0.01)	(0.01)	(0.01)	(0.01)
Not sick	0.017***	0.019***	0.019***	0.021***
	(0.00)	(0.00)	(0.00)	(0.00)
Coreside	0.04	0.06	0.04	0.02
	(0.04)	(0.05)	(0.03)	(0.03)
Not coreside	0.015***	0.015***	0.019***	0.023***
	(0.00)	(0.00)	(0.00)	(0.00)

Note: Standard Errors in parenthesis. We adjust for the child's age in the model. The same set of covariates are adjusted as in the Table 2.13.
* $p < 0.1$, ** $p < 0.05$, ***$p < 0.01$

Table 2.15: Comparison of Data Generating Models

Data Generating Models and the Variables or Effects Allowed in Each Model						
Data Generating Model	1	2	3	4	5	6
Length of grandparent overlap, A_{it}	Yes	Yes	Yes	Yes	No	No
Grandparent's baseline characteristics, C_{i0}	No	Yes	Yes	Yes	Yes	Yes
Time-variant confounders, C_{it}	No	No	No	Yes	Yes	Yes
Constant effect of time-invariant unobserved confounder, U_{i1}	No	No	Yes	No	Yes	Yes
Time-variant effect of time-invariant unobserved confounder, $U_{i2}A_{it}$	No	No	No	No	Yes	Yes
Estimator	LS[a]	LS[b]	FE[c]	MSM[d]	CFE[e]	CFE

Note:
[a] : LS refers to least square regressions. The identification assumption is perfect randomization of overlap.
[b] : The identification assumption is strict exogeneity given baseline covariates.
[c] : FE refers to conventional panel or sibling fixed effects models. The identification assumption is strict exogeneity given fixed effects.
[d] : MSM stands for marginal structural models. The identification requires Strict exogeneity given time-varying observed confounders.
[e] : CFE refers to the cumulative fixed effects estimator. The identification assumption is strict exogeneity given fixed effects (U_{i1} and $U_{i2}A_{it}$) and observables.

Appendix: Comparison of Data Generation Models

Estimating the total grandparent overlap effect is difficult. Engaging with these methodological difficulties promotes a deeper sociological understanding of the dynamics of multigenerational influence. Indeed, it leads to a precise definition of grandparent overlap effect as the cumulative effect of grandchild's exposure to the entire history of grandparent's observed and unobserved, fixed and time-varying characteristics across the full length of multigenerational overlap. I justify my definition by leading up to it through a discussion of various competing models of grandparent overlap effects, listed in stylized fashion in Table 2.15.

Table 2.15 serves two purposes. First its columns describe different substantive sociological theories of grandparent overlap. Second, the last row shows what statistical models should be used to estimate the total grandparent overlap effect if the corresponding substantive model of grandparent overlap were true. In short, Table 2.15 summarizes both competing data generating models and the identification strategies that they imply.

To fix ideas, I focus on the total effect of grandparent overlap, A_{it} on grandchild's academic test scores Y_{it} measured at age t. Starting from the most simplistic model of grandparent overlap, if the remaining duration of grandparent's survival were randomized at grandchild's birth in a proper randomized controlled trial (RCT), identifying the grandparent overlap effect would be trivial. I would simply regress grandchild's outcome, Y_{it}, on grandparent overlap, A_{it}[5]

$$Y_{it} = \alpha + \beta_{RCT} A_{it} + \epsilon_{it} \tag{2.4}$$

The coefficient β_{RCT} could be interpreted as the total effect of an additional year of overlap.

Of course, grandparent overlap is not random. One might next allow, that grandparent overlap is random conditional on observed baseline characteristics at grandchild's birth. For example, grandparent education, occupation, income and health are found to affect their

[5] Grandparent overlap A_{it} is equal to the grandchild's age at the test t if the grandparent dies after the test. If the grandparent dies before the test, overlap is the grandchild's age at the grandparent's death, d_i.

own survival and the grandchild's educational outcomes (Modin et al. 2012; Møllegaard and Jæger 2015; Song 2016; Wightman and Danziger 2014), and thus should be included in the model. This would license estimating a conventional OLS regression:

$$Y_{it} = \alpha + \beta_{OLS} A_{it} + X_{i0} + \epsilon_{it} \tag{2.5}$$

The coefficient β_{OLS} could be interpreted as the total effect of an additional year of overlap.

Lehti et al. (2018) have convincingly argued that grandparent overlap is not only a function of observed baseline characteristics, X_{i0}, but also of fixed unobserved characteristics, U_i, for which they named cultural capital, genetic factors and physical proximity. If so, β_{OLS} would be biased, but a conventional panel (or sibling) fixed effects model would work if the effect of these time-invariant unobserved confounders are also fixed,

$$Y_{it} = \alpha + \beta_{FE} A_{it} + U_i + \epsilon_{it} \tag{2.6}$$

where U_i is the individual-level fixed effect that would subsume X_{i0}.

Although model 2.6 is more realistic than model 2.5, it is limited in two subtle respects. First, it is important but not feasible to include grandparent time-varying (or sibling-varying) observed confounders in the fixed effects models. The grandparent's survival and the grandchild's developmental outcomes are affected by many time-varying confounders, such as the grandparent's health, income, marital status and coresidence. For instance, when the grandparents are younger, they tend to be healthier, and they get frailer as they age. At a given child's age, the grandparents tend to be healthier for the older sibling compared to the younger sibling because they are younger when the older sibling is born[6]. Failing to control for these time-varying confounders will lead to biases in the estimation of the grandparent overlap effect. However, controlling for these grandparent time-varying confounders would

[6]The sibling fixed effects strategy compares the sibling grandchildren's outcomes at the same age (Lehti et al. 2018).

lead to over-control biases. This is because of selective sample attribution on grandparent survival status (see Heckman (1976) selection model and discussion from Elwert and Winship (2014)). I can only measure grandparents' health and income at time $t+1$ if the grandparents survive until $t+1$. As part of the effect of past grandparent survival status is mediated by grandparent health, so controlling for health controls away part of the overlap effect. Hence, these grandparent time-varying confounders should be controlled for to identify the total grandparent overlap effect. Yet at the same time, the characteristics are also mediators, so they should not be controlled for. In total, this makes model 2.6 likely to suffer from either omitted variable bias or over-control bias.

Second, model 2.6 unrealistically assumes that the grandparent overlap effect is homogeneous to grandparents with different fixed unobserved confounders. Past literature has insightfully proposed that important social institutions and grandparent's unobserved traits may determined both the length of grandparent overlap and grandchild's cognitive outcomes (Lehti et al. 2018; Mare 2014). They outlive individual lives and transmit across multiple generations (Mare 2011). Lehti et al. (2018) explicitly mentioned grandparent's social and cultural capital and genes. I can also think about unobserved confounders such as grandparenting motivation and grandparent's ability. The effect of these grandparent's time-invariant unobserved confounders can be time-varying, especially in a way that is amplified with a longer grandparent overlap. For instance, stronger motivation of involving in grandparenting would affect the grandparent's survival via their physical and mention health. Stronger motivation of grandparenting may affect the grandchildren's development via the number of visits and ways of interpersonal interactions (Dunifon et al. 2014). These effects of grandparent time-invariant traits may be amplified with a longer duration of grandparent overlap because they rely on mechanisms involving grandparent-grandchildren interactions.

These two problems root in the fact that grandparent overlap effect is not a conventional point estimate but is an accumulative effect by nature. Overlap is a measure of time. However, the overlap effect is not drawn from the length of time itself but is derived from the effect of all the grandparent characteristics which characterize the grandparent-grandchild interactions. This discussion naturally leads to my new definition of grandparent overlap. I define *any* effects of grandparent characteristics that are amplified by a longer overlap to *constitute* the grandparent overlap effect. In other words, grandparent overlap is a delivery mechanism of effects of grandparent characteristics that require contacts.

PAEO corresponds to different quantities given different models of data generation. If the length of grandparent overlap is modelled to affect the grandchild's cognitive test as a main effect term as is shown by model 1-3 in Table 2.15, then PAOE would equal β_{RCT}, β_{ols} and β_{fe}. The identification of these quantities of PAOE assumes either grandparent overlap is randomized, is randomized conditional on baseline confounders and is randomized given the fixed effect, as is suggested by the last two rows of Table 2.15.

Chapter 3

On the Fence of a Family: Dynamics of Inter-generational Transfers, Contacts and Support Between Parents and Adult Children in Step Families

Introduction

The expansion of step families in the US challenges the very notion of family. Such demographic trends may have important implications on how families operate and pose important questions regarding who is in the family and who is not (Seltzer 2019). A family relationship can be delineated by its strong solidarity among family members (Bengtson and Roberts 1991; Bengtson and Oyama 2007) and the strong intergenerational support network in traditional conjugal families(Swartz 2009; Seltzer and Bianchi 2013). In contrast, the support network of step family is weaker. It was demonstrated that step-kin have fewer transfers, less interpersonal support and lower frequency of contact compared to biological kin in step families (Wiemers et al. 2019; Kalmijn 2013; Kalmijn et al. 2019; Becker et al. 2013).

In explaining for such "step gaps" of intergenerational cohesion, social scientists have proposed or adopted the "biological premium hypothesis" (Becker et al. 2013; Kalmijn et al. 2019). It argues that kin's preferential investment is associated with the symbolic meaning of blood ties and the strong norms of support of the biological relations in comparison to the weaker norm of support of step-kin(Buss 2016; Daly and Wilson 2000), which implicates that step-kin is not in the family. However, step-kinship is better described as a metastable state of "sitting on the fence" due to the ambiguous norms regarding the roles and obligations of the fathers, mothers and adult children in step families (Cherlin 1978). Considering the metastability of step-kin relationship, it would be interesting to examine the different dynamics of intergenerational interactions, i.e. how step-kin respond to each other's past support with providing future transfers, contacts, interpersonal supports, similarly or differently to biological kin.

The dynamics of intergenerational interactions involve sending signals of helpfulness, evaluation of the other's signals based on one's perceptions of the norm, and choosing among ways to respond. Studying the dynamics of intergenerational interactions in step families is meaningful, theoretically and empirically. Theoretically, it relates to the fundamental debates in the sociology and family economics literature about the motivations of intergenerational interactions. Economists and sociologists have argued that intergenerational transfers and supports may be altruistic, exchange-based or normative (Becker 1974; Becker and Becker 2009; Bengtson and Roberts 1991; Bianchi et al. 2006; Logan and Spitze 1995; Lye 1996; Rossi and Rossi 1990; Swartz 2009; Van Gaalen and Dykstra 2006). These different motivations are often behaviorally indistinguishable, especially in fully institutionalized traditional families. Step families provide a unique opportunity to better understand these fundamentals, especially by allowing the comparison of interactions of step-kin and biological kin. Because how people respond to others' actions reveals their motivations and perceptions of the social norm underneath their relationship.

Departing from the "biological premium hypothesis", which predicts a lower level of response of step-kin regardless of the other's past signals, I argue that the step-kin's ways of interactions and provision of support may display greater sensitivity to the other's signals of helpfulness. Specifically: 1) step-kin may have a lower transfer, contact and support when no past signals of helpfulness are present. 2) Conditional on each other's past signals of helpfulness, step-kin may respond more sensitively by providing more supports, transfers or making contacts in the future. So the "step gaps" can be reduced with positive signals in the past. 3) Future transfer, contacts and interpersonal support have different implications on the family boundary. If step-kin may have less tolerance of risk due to the more ambiguous norm, they may provide less senior care or child care which involves more under-defined responsibilities and potential conflicts.

These hypotheses may shed light on the theoretical debates of motivations of intergenerational transfer, contact and support. If they are supported, then biological kinship in adulthood is more "altruistic" and step-kin is more "exchange-based" in the sense that the latter is more dependent on each other's past signals of helpfulness and allows fewer risks. Also, step-kin may be not only "sitting on the fence" but also sensitively watching for the chance to jump in. From the evolutionary perspective of the norm, stable norms are built because they help reduce costs in repeated human interactions (Axelrod 1986). Family environment and the conjugal relationship of the parents predispose step-kin to intensive interactions, which may grant a motivation for step-kin to converge to a relationship that is more cohesive and family-like. Therefore, step-kin may respond to each other's positive signals of helpfulness with future support to a level that is closer to the levels of biological kin. If this is true, it may suggest for an eagerness of step-kin to converge their norm to the "biological kin norm" at least in some of its functionality of providing intergenerational instrumental and emotional supports.

Whether the "step gaps" in intergenerational support and cohesion can be closed also has critical empirical implications. Its social and policy relevance becomes evident when considering the roles of family in intergenerational solidarity and social inequality (Carlson and Meyer 2014; Seltzer and Bianchi 2013; Sweeney 2010). Step-kin network is expanding with the size of biological kin contracting with fertility decline. Traditionally, adult biological children provide a support network to seniors irreplaceable by the public safety net; the emotional bonds from adult children are important for parents' well-being, and they help with the parent's daily lives as the latter's physical ability decays. At the same time, adult children need to rely on the parents more, and for longer as labor market insecurity increases, tertiary education becomes more prevalent, and the unmet need of childcare support grows (Seltzer and Bianchi 2013). So to strengthen the support network of the nontraditional family network, including step-kinship, could become increasingly important for both the seniors' well-being and the adult children's successful transition to adulthood and parenthood.

In this essay, I investigate the different dynamics of interactions of step-kin and biological kin in step families. Rather than comparing their different levels of intergenerational transfers (Anderson et al. 1999; Berry 2008; Davey et al. 2007; Henretta et al. 2014; Kalmijn et al. 2019; Wiemers et al. 2019), I focus on how parents and adult children respond to the other party's signals of helpfulness differently with future transfers, support and contacts. This question translates to the moderation effect of step/biological status on the effect of the past signal of support of one on the future response of the other. It is important to adjust for parents and adult children's needs of support, ability to help and their baseline relationship before the past signal of helpfulness. It is debated whether it is the type of kin relationship that matters or the individual traits that select parents into remarriage and affect their bonds with close kin (Evenhouse and Reilly 2004; McLanahan et al. 2013). Within-family fixed effects model facilitates the elimination of such unobserved family confounders that select parents into the specific family structure, which helps recover the causal moderation effects of interest. Most data of American step families are either cross-sectional, not nationally representative or have insufficient statistical power to allow for within-family estimates. I analyze the RAND

Health and Retirement Survey (HRS) from wave 3 (1996) to wave 12 (2014) to obtain a sample 1,687 step families with 6,728 adult children who are aged 20 to 60 in 1996. With the rich HRS data and fixed effect strategy, I estimate the effect of both the adult children's and parents' past support on the other's future response, specifically, the future financial transfer and interpersonal help, future contacts and parent's future expectation of help.

Altruism, Exchange or Norms

Why do adult children and parents help each other? In spite of the numerous theoretical perspectives proposed in social science literature, arguments largely fall into three groups. The first claims that family members behave altruistically in a family, which distinguishes family behaviors from market behaviors (Becker 1974; Becker and Becker 2009). They provide financial transfer and interpersonal support when the other is in need without expectation of return and make frequent contacts out of affection and emotional bonds. This theory predicts that parents and adult children lend support to each other unconditionally, even when there is no anticipated bequest or need of interpersonal help. This theory was echoed by the fact that the downward instrumental transfers from parents to children are predominant in the US and other western countries.

The second group argues that intergenerational transfer, support and contact are self-interested and exchange-based. Actors explicitly or implicitly expect some forms of reciprocity. From this point, adult children provide care and support to their parents to either "pay back" parents' inputs in their childhood, or to exchange for parents' transfer in the future, such as bequest and grandchild care service. But the "currency" of exchange in intergenerational relationships can be diversified and perhaps asymmetric. Parents may invest in their children materially and interpersonally, paying for the schooling and providing care, in exchange for more subtle paybacks in the future, such as a sense of fulfillment, emotional support, and transmissions of their personal attributes. The type of currency in an exchange may be chosen based on the actors' preferences and currency's relative value and cost. Assuming the value of care is the same, to provide care for a frail parent, an adult child who has a high salary may choose to purchase professional care because of the high opportunity cost of his or her interpersonal support, while one with lower salary may offer personal help. Empirical research shows that neither the exchange theory nor the altruistic theory can explain for the entirety of intergenerational exchange, which has both rational and emotional elements.

Bridging the exchange and altruism theory, the third group relates to the theory of norm. Like other interactions, intergenerational transfer, support and contacts are governed by the norm. Norms are stabilized and internalized rules that help a social group to survive by reducing the costs of exchanges. A family norm is a shared belief of a family member's obligations and anticipated behavior in specific situations towards another as a parent, spouse or a child (Rossi and Rossi 1990). Family norms explain why family members may habitually act altruistically towards each other and support each other more than strangers or even friends (Andreoni 1989).

Different from the first two perspectives, the concept of norm emphasizes the co-evolving process of social environment, interactions and the rules of behaviors (Voss 2001). While family behaviors may appear altruistic and unconditional, the formation of family norms may be rooted in numerous repeated exchanges in history (Sethi 1996). With a computational simulation, Axelrod (1986) shows that a stabilized norm can emerge and persist among self-interested actors given a high chance of future interaction. Family members are usually bonded to interact repeatedly, which explains why family norms are strong and stable.

The motivations and rationales of intergenerational transfers and the formation of intergenerational norms are difficult to study empirically. The step family provides a unique opportunity for sociologists to study these fundamental concepts of family theory. In their review of the family literature, Bianchi et al. (2006) calls for the study of stepfamilies- "Examining how individuals adapt to these situations and how these quasi-kin relationships develop may help us understand how norms are formed and affect the behavior of individuals". It is interesting to investigate the distinctive dynamics of intergenerational interactions of step-kin; how step parents and adult children adjust to ambiguous family context and react to the actions of the other party.

Stepfamily Relationship

In his seminal work on step families, Cherlin argues that remarriage is an incomplete institution. He wrote that "where no adequate terms exist for an important social role, the institutional support for this role is deficient, and general acceptance of the role as a legitimate pattern of activity is questionable" (p. 643) (Cherlin 1978). Besides the view of incomplete institution, other sociologists have described step families as a deviant family norm (Ganong and Coleman 1997), or reconstituted nuclear family (Levin and Sussman 1997). Recent works of psychologists rejects such characterization of step families as deviant. They argue that step families are normative adaptive; albeit being different to a nuclear family, they are quite resilient in daily lives. With good communication and proper interventions, step families are not necessarily detrimental to children's or adults' well-being (Visher and Visher 2013).

Nowadays, nearly 30% American families have a step kin (either parent's or child's generations) (Wiemers et al. 2019). Such prevalence of step families may provide more opportunities for American people to either experience step family for themselves or interact with one (Troilo and Coleman 2008). Increased contacts with a stereotyped group may diminish stereotypes (Leyens et al. 1994). However, recent studies suggest that a clear norm about the expectations and obligations of family members in step families in the U.S is still lacking (Raley and Sweeney 2020). Formal institutions are largely absent; step parents have few legal responsibilities and rights toward their step children according to either the federal or state laws in the US, and the school and health care system still make little allowance for the presence or involvement of step kin (Ganong and Coleman 2017). Informal institutions are ambiguous; adolescents in step families discuss step-relationships using inconsistent terms, such as referring to a step father as "my mother's husband" (Chapman et al. 2016; Thorsen and King 2016). The confusion of labels reflect the lack of a complete institutionalization (Cherlin 1978). Besides, studies of public opinions (Pew research center) found that only one third of Americans agree that step families are a good environment for children and parents, the other two thirds either disagree or "tolerate but had concerns about them" (Morin 2011).

Sociologists find that the stepkin relationship is more distant and they share less resources. In their childhood this is reflected by children of step families experiencing more challenges in their behavioral and cognitive development (Amato et al. 1995; Hadfield et al. 2018; Lee and McLanahan 2015; Sweeney 2010) and adolescents being less close to the parents compared to their counterparts from traditional families (Jensen and Howard 2015). Recently, increased attention has been paid to relationship between grown-up children and their older parents (Kalmijn et al. 2019; Van Der Pas et al. 2013). The relationship between step parents and adult children is found to be more distant in the US (Aquilino 2006; Killian 2004; Henretta et al. 2014), Germany (Becker et al. 2013), Netherlands (Kalmijn 2013). Using PSID, Wiemers et al. (2019) describes that although step kin has significantly increased American household member size, step parents are less likely to both give and receive supports, especially the transfer of time. Interestingly, the "step gap" of intergenerational support is shown to be highly gendered and dependent on the duration of childhood coresidence. Using Dutch data, Kalmijn et al. (2019) finds that stepmothers are less likely to provide support for their step children. Besides, past interactions and relationship matter for the future transfers in a step-kin relationship. Step parents (especially stepfathers) may give more support to their step children when they have a longer shared duration of coresidence. Although the parents also receive less interpersonal support from their step children, the "step gap" of upward flow is smaller than that of the downward flow from parents to adult children.

Generally speaking, recent research provides insightful snapshots of the "step gap" of intergenerational support using cross-sectional data. They generally conclude that the expansion of step network is insufficient to compensate for the lower probability of transfers and supports in step families due to the weak norms between step kin(Wiemers et al. 2019). However, a weaker norm of support is not definitive to step-kin relationships. There is a gap between the theory and quantitative empirical research regarding step families. Step-kin norm has been theorized as "incomplete institution", "sitting on the fence", resilient and adaptive. But it is impossible to recover such dynamics by examining static status of relationships empirically using cross-section data. Existent studies have not studied the dynamics of intergenerational interaction and exchange in step families or the ambiguity of step norms empirically.

Four Hypotheses of Inter-generational Interactions

In this chapter, I ask how step-kin responds to each other's signals of support differently from biological kin in terms of providing future transfer, contacts and supports. I explore the following competing theories. The first hypothesis is implicitly assumed by past research. I propose three other competing theories to explicate the nuances of the different patterns of family interactions between them.

The Biological Premium Hypothesis

The evolutionary psychology theory suggests that parental transfer increases the evolutionary advantage of their genes (Trivers 1972; Buss 2016). Sociologists argue that there is a symbolic meaning of a blood tie and a strong normative obligations to support biological family members (Daly and Wilson 2000; Fine et al. 1998). However, the social norms governing step-kin relationship are less established (Cherlin 1978). Given the parents' preference, parents are more likely to provide support to their biological children compared to their step children in a step family. Although perhaps less potent, the social norm of adult children to support their biological parents are also stronger than that of supporting their step parents. Although they have not explored effects of exchange and family dynamics, this is the argument implicitly posited in both Kalmijn et al. (2019) and Wiemers et al. (2019). When this argument applies to family interactions, it would imply a "biological premium" of the return of the biological kin given the same level of past support, compared to the return of step-kin. In other words, when step children and biological children give the same amount of support in the past, the parents' transfer to the biological ones and contacts would always be higher. Given the same parents' past support, the future support of biological children would also be higher.

The Low-bar Expectation Hypothesis

Although the social norms of a step relationship are more uncertain, step parents and children may have a default perception that the other are more distant and usually offer less support compared to biological kin. However, such subjective prior may be changed if the step-kin shows signals of support and closeness to prove otherwise. Whereas without any signals of support, step parents would generally set a low-bar expectation of the step children based on such "statistical discrimination". Step children would also expect less from the parents.

In HRS data, the parent's expectation is measured, which is helpful to understand how past supports of the parents and children shape their belief. If this hypothesis is correct, I would expect a lower reported expectation on the step child to help in the future by a step parent when neither their step child nor the biological child shows support in the past. Due to the lower expectations, the amount of downward transfers, grandchild care and contacts from a step parent to a step child can be lower. Different from the first hypothesis, which assumes lower step parent transfer regardless of the children's past behavior, the low-bar hypothesis only expects a lower step-parent transfer when there is no signal from the step child, while remaining open to situations when the step child does demonstrate helpfulness, or the parent help the adult children in the past.

The Sensitive Response Hypothesis

Without a fully institutionalized social norms, the step-kin relationship may resemble an individualized contract which is conditioned on the past interactions between the parent and child. The step-kin may respond more sensitively to received signals of closeness and support associated with the other's past support compared to the signals sent between biological kin.

This idea draws on the key theoretical debate central to the family literature of exchange or altruism (Bianchi et al. 2006; Swartz 2009). Although both notions can be broad and multifaceted, a distinctive feature of altruistic support is its unconditionality. While neither of the two might entirely explain the motivations of family transfer alone, it is possible that a step-kin relationship is more "conditioned" than a biological relationship. In other words, step-kin interaction may be more strongly dependent on the past actions of the other. On the other hand, biological kin may be less sensitive to each other's signals because of the inertia under the strong institution of kinship norms, which may grant a sense of security of their relationship.

If this is true, I expect a larger effect of a step-kin's signal of support compared to the effect of a biological kin's signal, i.e. a positive moderating effect of step status on the effect of adult children or the parents' past help. In other words, this suggests a larger premium in a step-kin's transfer in response to the other's help, which is contrary to the first hypothesis. With the larger premium of "changes", the level of step-kin's response may converge to the level of biological kin as past support increases.

The Deferential Convergence Hypothesis

Although the sensitive response hypothesis predicts that intergenerational transfers, contact and support between step-kin may converge to that of biological kin given each other's positive signals, it leaves open whether the convergence is uniform in these categories. From instrumental to emotional solidarity, do past signals of help bring step-kin closer and similar to biological kin in all these ways? Kalmijn et al. (2019) and Wiemers et al. (2019) examine different dimensions of instrumental and emotional solidarity, such as by the frequency of contact and interpersonal support. But past research has not explicated their substantive implications on the sense of family boundary for step-kin and biological kin. Step-kin may respond to the same signal of support differently due to the more ambiguous norm underneath given each other's needs and ability to help. Different trust, risk tolerance and attachment are involved in adopting each form of transfer, contact and support.

If step-kin share less emotional attachment and associational solidarity (Bengtson and Roberts 1991), then they are likely to respond to each other's signals with more monetary transfer but not more contacts or support, because sending a gift card takes less time, requires less persistence, and involves low risk. In that case, step-kin may just want to "pay back", and so would not be regarded as really "converging" to a more family-like norm. As such, frequency of contact and interpersonal support are considered as important indicators of a more familism norm after controlling for their needs and ability to help and relationship. Nevertheless, given the obscure norm, it is possible that step-kin relationship is less risk-tolerant, so that step parents may be less likely to help with grandchild care for step children, and step children are less likely to provide senior care because personal care involve more under-defined responsibilities and potential conflicts.

The four hypothesises are summarized in Table 3.1, the latter three are competing hypothesises opposed to the first one.

Data and Measures

Data

I use the data of HRS Longitudinal File and HRS Family Data Files that RAND have derived from all available waves (1992 to 2014) and cleaned in a manner to ensure consistency across waves. HRS is a nationally representative longitudinal study of the United States starting from the year of 1992 and followed every two years. The majority of the respondents are Americans born between 1910 to 1950 who are older than 50 years when participating.

HRS data are collected at both respondent and the household levels. A family respondent is designated for each household to answer questions regarding himself/herself, the spouse and the children consistently over all the waves, regardless of the respondent's marital and partnership status. As such, all the information pertaining to the children's characteristics and intergenerational transfers of money and interpersonal support is reported by the family respondent. The HRS Longitudinal File contains information about the respondent's generation, and the HRS Family Data File includes the social-economic characteristics and transfers of each child and stepchild of the family respondents. Respondent's HRS Longitudinal File record is merged with a record in HRS Family Data File if he/she has at least one child. [1]

Because adult children and their parents are the focus of this chapter, I draw on the sample of step families with the parent respondents married/partnered, aged between 50 and 86 and children aged between 20 and 60 in 1996. The treatment variables of past transfers from either the parents or the adult children are measured from wave 3 (1996) to wave 6 (2002), and I select the respondents who have consistently participated in the surveys from wave 3 to wave 6 for the analytical sample.[2] The outcome variables of the parent's expectation, contact and transfers are measured from wave 6 (2002) to wave 12 (2014), so both the parent respondents and the children should survive wave 6 to be included in the study.[3]

[1] Among all the parents who have at least one child, 90% parent respondents are born between 1910 and 1946, with a mean of 1927. 55% are females and 45% are males. 84% respondents are white, 8% are black and 5% Hispanic, and they have obtained 11.8 years of education on average. Notably, the marriage status and marital history of these HRS respondents represent the US population closely (Kreider and Ellis 2011). For instance, 63% of the HRS parents are currently married (including 2% partnered), 24% are widowed, 9% are divorced and 3% are never married. 71% have married once, 20% twice, 4% three times, and 1% four times and more. They have 0.44 step children and 3.23 biological children on average.

[2] This is identified by variables indicate whether the records are available in that respective wave. To avoid the over-representation of coupled households, I select one parent respondent per household.

[3] The transfers and contacts are coded missing after the parent or the child's death.

Among all the families who meet these survival and participation criteria, there are 1,333 step families with 5,565 children, which is defined by the presence of at least a step child. Among these children, there are 4,680 step children who have either a step mother (N=2,139) or a step father (N=2,541), and 885 biological children who have two biological parents in a step family. I use this analytical sample because of the chapter's focus on step families.

Data with sufficient statistical power to study the intergenerational relationships of step-kins are scarce. Kalmijn et al. (2019) relies on the over-sampling strategy of the OKiN Survey based on the Dutch population register to study step families in the Netherlands. Wiemers et al. (2019) uses the 2013 Rosters and Transfers Module of PSID to study step families in the US. However, data of both aforementioned studies are cross-sectional and thus can not be used to answer questions regarding family dynamics and interactions, which are the focus of this chapter. HRS is unique in its coverage of all adult children of each family over a long period of time, which provides the rare opportunity to study family interaction and exchange between parents and their adult biological and step children in the US, while allowing for adjustment of unobserved family heterogeneity underlying the selection into this family structure. The major limitation of this data is its lack of documentation of the intergenerational relationship in the children's childhood, such as co-residential history, which are shown to be an important factor shaping step-kin relationship in children's adulthood (Kalmijn et al. 2019).

Measurement

Unit of Analysis

Consistent with some recent studies (Wiemers et al. 2019), this chapterfocuses on step families. 79% of the parent respondents who meet the participation criteria are married in 1996. When parents are in a marital union, it is often difficult or unrealistic to separate the financial transfer, contacts and sometimes even interpersonal support between the parents if the parents share joint financial account or live together. Indeed, most questions on intergenerational transfers in HRS pertain to the child and the parents (the respondent and the spouse jointly), except for questions regarding parent's expectation, which is related to the parent respondent himself/herself.[4] So this study focus on the financial transfers, contacts and interpersonal support between adult children and "parents" rather than parent-child dyads.

A step child is identified as either the respondent parent's own step child, or the respondent's biological child who is born before the start of the respondent's current marital/partnership relationship; this child would be a step child to the respondent's spouse. Correspondingly, a biological child should have two biological parents in the family, which is indicated as the biological child of the respondent and was born since the parent respondent's current marriage.

Given the longitudinal scope of this study, the family structure may change considerably during the course of the survey from year 1998 to 2014. In fact, although the divorce rate remains largely unchanged at around 8%, the percentage of parent respondents who are widowed increased from 26% in 1998 to 44% in 2014. If the respondent gets divorced or widowed at some point from 1998 to 2014, I code all the measures of family interpersonal and financial transfers and contacts (all outcome variables except for expectation) as missing subsequently. There are two reasons. 1) A step family is inherently defined by the parents' remarriage. Step-kin relationship usually breaks down as the parents' marriage endsSeltzer 2019). So whether or not a family would qualify the definition of a step family can be dubious if the parent respondent is not married or partnered. 2) In cases where the parent respondents are divorce or widowed, the definition of a household unit in HRS would change to only include the respondent parent himself/herself thus the meaning, targets and scope of the intergenerational transfers can change accordingly.

Although the co-residential history before 1992 is not known, the closeness of the step-kin relationship can be proxied with the age of the step child when entering the respondent's

[4]For instance, the question regarding parent-to-child financial transfer is framed as "the amount of financial transfer to the child from you, your spouse/partner, or jointly."

current or most recent marital/partnership relationship. I also include control variables of the coresidence status and residential proximity to account for the closeness between the parents and children.

How Adult Children's Past Support Affect Parent's Future Transfer and Contact

Independent Variables

The first set of analyses focuses on the effect of past interpersonal supports from the adult children. To measure the *adult children's past interpersonal support*, respondents were asked whether each of their children has helped with chores, errands and transportation in the past year from wave 3 to wave 6. Both dichotomous and continuous measures (the number of waves such help is reported) are included.

Dependent Variables

The dependent variables include the two dimensions of practical and emotional support from the parents to the children, as distinguished in the literature on intergenerational solidarity (Silverstein and Bengtson 1997; Swartz 2008). In addition, parent's expectations may play an intermediate role bridging the relationship between children's past support and the parent's future support and contacts in response. All dependent variables are measured temporarily posterior to the treatments.

Parent's expectation is a dichotomous measure of whether the parent expects help from a particular child in the future "if they need help with basic personal care activities like eating or dressing over a long period of time". This is reported at wave 6 , which I interpret as an indicator for the parent's belief regarding the child's willingness and ability to provide interpersonal support.

Parents' future monetary transfers and interpersonal support are measured from 2006 to 2014 (wave 7 to wave 12). The *average yearly parent monetary transfer* is calculated as the total transfer and value of gifts from the parents to a child throughout wave 7 to 12 divided by the number of non-missing waves. *Parent interpersonal help* is measured by the provision of care to the grandchildren of a specific adult child of more than 100 hours in the past two years. 86% adult children in the sample has at least one grandchild by 2014. Due the low prevalence of intensive grandparent childcare support (84% of all adult children who have grandchildren do not receive grandparent childcare support in any wave, 7-12), I code the variable as 1 (yes) as long as such grandparent childcare is provided in one survey wave from 7 to 12.

Contact between the parent (or the spouse) and the adult children is a commonly used indicator of emotional solidarity. Unlike financial transfer, contact is always reciprocal. It includes any intergenerational interactions which may be in person, by phone or online. Similarly to the measure of average monetary transfer, the average yearly contact is calculated as the total number of contacts from wave 7 to wave 12 divided by the number of nonmissing observations for each adult child.

How Parent's Past Support Affect Adult Children's Future Transfer and Contact

Independent Variables

The second set of analyses focuses on the effect of parent's past interpersonal support. To measure the *parent's past interpersonal support*, parent's provision of grandparent childcare of more than 100 hours in the past two years from wave 3 to wave 6 is used. Similar to the measures of child's past support, both dichotomous and continuous measures are included. The dichotomous measure indicates whether such grandparent childcare has been provided in any of these waves, and the continuous measure evaluates the number of waves such childcare has been provided.

Dependent Variables

To examine how children respond, in terms of transfer and contacts, to the parent's past support, I include outcomes of children's interpersonal help, monetary transfer and contacts which are measured from wave 7 to 12, in addition to the parent's expectation. The parent's past support may affect their own expectation of receiving help from the children in the future when they are in need as well as the frequency of contact, which are the same outcome variables of *Parent's expectation* and *Contact* as discussed above.

Children's future monetary transfers and interpersonal support are measured from 2006 to 2014 (wave 7 to wave 12), symmetrically to the measurement of parent's future transfer. The *average yearly children's monetary transfer* is calculated by the sum of child-to-parent transfer throughout wave 7 to 12 divided by the number of non-missing waves. *Children's interpersonal support* is measured by the number of days an adult child has helped the respondent in a month. Similarly it is averaged across wave 7 to 12. Both of the two measures of the children's monetary transfer and interpersonal help can be directed to either the respondent or the spouse.

Controls

All the control variables are measured at wave 3 (year 1996) to be temporarily prior to the treatments and outcomes.

The adult children's social economic status and characteristics may also affect their demand of support and ability to give assistance. Important confounders include the child's age, gender, employment status (working full time, part time and not working), marital status (married, not married and partnered), residential distance and income. If the children are employed full-time, live further away, have higher salary and are married, then they may have more competing obligations. They may be less likely to give interpersonal support when their opportunity cost is higher (Seltzer and Bianchi 2013; Swartz 2009). So they might choose to support their parents financially when their parents have a need. Their higher economic and social sufficiency may lower their need of the parents' interpersonal and financial support at the same time.

Besides, the quality of the relationship between parents and adult children is an important confounder. Parents may feel closer to a child than another, and so as the child. The relationship quality may affect both the chance of one's past support and the other's future response. Therefore, I adjust for measures of relationship quality, including the parent's expectation of future help as of the year 1996, the amount of parent-to-child monetary transfers measured on 1994, and the amount of child-to-parent monetary transfers measured on 1994 in order to control for the closeness between the parent and the adult child at the baseline level. The duration of childhood coresidence is found to be the most important factor shaping the step-kin closeness (Becker et al. 2013; Kalmijn et al. 2019). To account for it, I rely on the measure of the step child's age at the start of parents' current marriage. The variable of *Child Older Than 10 at Parents' Marriage* is a dichotomous variable coded from the continuous measure of age which indicates whether the children are older than 10 at the parents' current marriage, with 1 indicates yes, and 0 indicates the otherwise.

Some control variables have a larger number of missing cases, such as Income and whether the child lives within 10 miles. To test how the missing pattern may affect the results, I use multiple imputation with chained equation (MICE) approach as robustness check, and all the results remain unchanged (See Appendix section I).

Analytical Framework

Model

I denote an adult child as i who is nested in family j. When analyzing the effect of upward support, I let $X_{i,j}$ denote children's past support, while let Y_{ij} denote the parents' future transfers, expectations or contacts. Conversely, when analyzing the effect of downward support, I let $X_{i,j}$ represent the parent's past support, while Y_{ij} represent children's future transfer, contacts and the parents' own expectation. To explore the effect heterogeneity of past support on future responses given the kin relationship, I further include the interaction term $X_{ij} * Step_{ij}$.

$$Y_{ij} = \beta X_{ij} + \alpha Step_{ij} + \theta X_{ij} * Step_{ij} + \phi P_j + \kappa C_{ij} + U_j + \epsilon_{ij} \quad (3.1)$$

Family members' availability and ability to offer help, their needs of support and their baseline relationship qualities are important confounders which may affect both the past support they would have received and the future transfer and contact they can provide. The adult children's social economic characteristics, C_{ij}, may determine both their ability to lend a hand when their parents are in need and their need of parents' continued support. It may include age, gender, education, income, residential distance, marital and employment status. Besides, C_{ij} include important measures of their baseline relationship measured by the age of the step child at the start of the parents' current marriage, the baseline parents' expectation and the amount of downward and upward financial transfers. Similarly, the parents' characteristics, P_j, including the parent's age, gender, race, education, marital status, family income, wealth, and family size may determine their ability to support their adult children and their potential need of assistance, although these are shared across siblings and absorbed by the fixed effect term in fixed effect models.

A central concern of this study is confounding by unmeasured family-level characteristics, U_j. These may include both the family culture, which affects patterns of intergenerational exchange, and the unobserved heterogeneity underlying the selection into step families. U_j, is largely unobserved, and thus may be an unobserved confounder of both the effect of $X_{i,j}$ and $X_{ij} * Step_{ij}$. First, the closeness of a family network may affect the intergenerational transfers of instrumental and emotional support in both directions. Additionally, the quality of a family network can be determined by the personal traits of parents, their familism value, beliefs about intergenerational solidarity, and the way they raise the children (such as different parenting style schemes: authoritarian or disciplinarian, permissive or indulgent, involved, and authoritative) (Darling and Steinberg 1993). Second, U_j may also include unobserved factors selecting parents into step families. Sociologists argue that some traits of the parents may lead to both their marital instability and challenge in developing close and secure relationships with their children (Kalmijn et al. 2019). These factors may also affect the strength of the intergenerational ties of a family.

Considering the important influence of the unobserved family confounder, U_j, I adopt within-family fixed effects estimation strategy to estimate the effect of past intergenerational transfer, or β and θ, in the equation. When the parent and child's capacity and needs are controlled for, the effect of X_{ij} may reflect the parent and child's response to the other's willingness to help and signals of closeness and support. There may be an interaction effect between X_{ij} and $Step_{ij}$, because both the step adult child and the step parent may respond to the signals of support by the other differently.

Assumptions

There are two assumptions to identify the effects of past transfer on the future returns of either the downward or the upward intergenerational transfer: the absence of sibling-specific unobserved confounders, and the absence of spillover effects between siblings (Sjölander and Zetterqvist 2017). Without loss of generality, assume that there are two sibling, i and i' in the same family j.

First, $Y_{ij} \perp\!\!\!\perp Y_{i'j} | X_{ij}, X_{i'j}, U_j, P_j, Step_{ij}, Step_{i'j}, C_{ij}, C_{i'j}$. That is, the outcomes of the siblings are conditionally independent given the past treatments of the siblings, the sibling-specific confounders and the family fixed effect. In other words, there are no unobserved sibling-specific confounders which may explain the difference in their outcomes of either downward or upward transfers, contacts and parent's expectations. I control for the child-specific confounders including relationship quality and their capacity and their need of family support which affect both their past and future transfer.

Second, $Y_{ij} \perp\!\!\!\perp X_{i'j} | X_{ij}, U_j, P_j, Step_{ij}, Step_{i'j}, C_{ij}, C_{i'j}$. This suggests that the outcome of sibling i is independent of the past treatment of the other sibling, i', conditional on the treatment status of i, the siblings' observed characteristics and the family fixed effects. This is commonly known as "no spill-over effect of the treatment" condition. In this context of intergenerational exchange, parents react to each adult child's past behavior independently, without evaluating and comparing the adult children's past support among each other. Besides, an adult child makes decisions regarding future interpersonal support to the parents based on what the parents transferred to him/her in the past, regardless of how much the other siblings have supported the parents in the past. Nevertheless, this condition does not exclude the situation that one sibling's past help with the parents ($X_{i'j}$) can encourage or dis-encourage another one (i) to provide help to the parents (X_{ij}) as well, which turns out to affect the parents' future financial transfer to i (Y_{ij}).

Results

Descriptive Results

Table 3.2 describes the characteristics of parents and the step families with both biological and step children. The majority of parent respondents in the sample are female respondents (87%), the biological or step mothers of the adult children. The majority of them are white (77%), with 15% black, 7% Hispanic and 1% of other races. I select the birth cohorts from 1910 to 1946, with the average birth year 1934 and a standard deviation of 8 years. All of them are married. The mean total household income for the last calendar year (including the respondent and spouse, including earnings, pensions and annuities, and other social security and government transfers) is 54.59 thousand dollars. The mean household wealth is 204.14 thousands dollars (total non-housing assets). Most families have one (24%), or two (17%) adult step children, and most step families do not have biological children from the remarriage (less than 30%) in the defined age group between 20 and 60. The number of step children and biological children in the same family is not strongly correlated (corr=0.04).

Table 3.3 compares the outcome variables by step status. It shows that the intergenerational relationship between step children and their parents features lower parent expectation and contacts, and less downward financial transfer and interpersonal support. This is consistent with findings of past research (Kalmijn et al. 2019; Wiemers et al. 2019). As is shown in this table, parents are less likely to expect that a step child will help them with daily lives in a long term as their biological child (25% vs 10%). Step children have less contacts (an average of 87 vs 154 in one year), suggesting a lower emotional solidarity. Besides, downward intergenerational transfer is less intensive and frequent. Parents provide more financial transfer to their joint biological children compared to their step children (an average of 836 vs 642 dollars in 2 years). Parents are not very likely to provide grandchild care for more than 100 hours in any wave from 7 to 12. But they are more than twice as likely to care for the grandchild of a biological child than of a step child (22% vs 12%). However, upward transfer from the adult step children to the parents is increased. Biological children transfer an average of 44 dollars to their parents, as compared to 64 dollars from step children in the past 2 years. Biological children spent an average of 0.36 days (in the past month) in supporting their parents inter-personally, as compared to 0.33 days from step children.

Table 3.3 also shows the large discrepancy in the past support measured from 1996 to 2002 (the treatment variables) between biological and step-kin, both upwardly and downwardly. Most step children are not reported to have provided interpersonal support (19% have helped). Biological children are more likely to provide interpersonal support (45%). As for the parents, 33% have helped with child care for more than 100 hours for the biological children, as compared to 19% for the step children.

Table 3.4 describes the demographic and social-economic characteristics of adult children in the sample of analysis who are aged between 20 and 60 as of 1996. The gender composition of the adult children is equally divided into females and males. Adult children have around 13 years of education on average, which is comparable to the average of the American population. Around 60% are married and 40% are not married which includes divorced, widowed and never married and the majority of them have a grandchild by 2014. Around

70% of the adult children work full time, 10% have part-time jobs and around 20% are not employed. Adult children receive considerably more monetary transfer from the parents than the transfer they give to the parents, which is consistent with past literature which has documented the dominance of downward flows in intergenerational transfers among American adult children and parents (Swartz (2009)). The most notable difference between the step children and the biological children is perhaps that step children are 3.5 years older than the biological children on average, which is expected because step children are from the fertility of the parents' previous marriage(s). Besides, biological children are more likely to live within 10 miles to the parents, compared to the step children (45% vs. 28%). Step children also give and receive less financial transfer as of the year 1996. Otherwise, they are very similar in gender, education, marital status, income and working status. There is no difference in social-economic achievements observed for the step children.

Past studies suggest that the duration of coresidence in the step children's childhood is important in explaining their closeness in adulthood. The step-gap in (step)father's transfers can even diminish after adjusting for the duration of coresidence (Kalmijn et al. (2019)). In order to account for the parents-children closeness, I calculated the children's age at the start of the parents' current marriage (measured at 1996), and generated an binary indicator of whether the step children are older than 10 years. Around 75% step children are older than 10 years when the parents get married.

Results From Fixed Effects Models

How Adult Children's Past Support Affect Parent's Future Transfer and Contact

Table 3.5 shows the estimates of the effects of children's signals of interpersonal support on parents' future monetary transfer, contacts, childcare and parents' expectation. For each outcome, two models with the dichotomous treatment (Child Ever Helped) and the continuous measure (Child Waves Helped) are presented next to each other. Results with dichotomous treatment are summarized and visualized by the margins plots of Figure 1.

Predicted Future Parents' Money Transfer Models 1 and 2 in Table 3.5 shows the effects of children's signals of interpersonal support on future parents' money transfer. First, the predicted amount of future money transfer is substantively lower for step children compared to for biological children when neither helped with the parents in the past, which can be seen from the blue line. This supports the lower-bar expectation hypothesis. Second, for the biological children, there is no evidence for a higher parent-to-child financial transfer in response to the children's past provision of support. In contrast, parents would give more transfer to the step children if they have provided interpersonal help. The larger premium in the parent's return for step child's past support is also obvious from the positive and statistically significant interaction term in model 1. This supports the sensitive response hypothesis. Due to such premium, for adult children who do support the parents in the past, biological and step children are expected to receive an equivalent amount of parents' financial transfer, which provides evidence against the biological premium hypothesis.

Predicted Future Parent Expectation Models 3 and 4 in Table 3.5 show the estimates of the effects of children's signals of support on parent's expectations. [5] It shows that the step parent expects less from the step child especially when the step child has not shown support in the past, which is consistent with the low-bar expectation hypothesis. But even when the step child has shown signs of support, and as the waves increase, the parent still expects less from the step children compared to the biological children. This suggests that the parents may have a stronger attachment to their biological children and do not expect as much from the step children. This lends support to the biological premium hypothesis.

Predicted Future Contact

Models 5 and 6 in Tables 3.5 show the effects of children's support on the yearly contact they have with the parents. There is a strong evidence supporting the lower-bar hypothesis since the average number of contacts between a step child and the parents is much lower than

[5] Different to the other three outcomes, the future expectation regards the belief of the family respondent himself/herself. Correspondingly, the step status refers to the relationship of the adult child to the respondent oneself. The majority of the "family step children" who are either step children of the respondents or the spouse are biological child of the respondents (4294, or 63.82%).

that of a biological child, when neither of them has helped. Whether the child has provided support and the number of waves show a positive effect on the frequency of contacts for both the biological child and the step child. But the gap for the step child is much larger as is indicated by the positive and statistically significant interaction term between the dummy treatment of Ever helped. So the result supports the sensitive response hypothesis.

Predicted Future Parent Childcare

Models 7 and 8 in Tables 3.5 show the effects of children's signals of support on the chance that parents provide child care support for more than 100 hours in a year. Although the predicted probabilities are generally low for both step and biological children when they do not provide help in the past, as the biological children provide support and the number of waves increases, parents become more likely to provide child care support to them. Although the chance of child care support to the step children also increases with past children's support, the moderation effect is small and not statistically significant from zero. This provide evidence for the low-bar expectation hypothesis and the biological premium hypothesis.

How Parent's Past Support Affect Adult Children's Future Transfer and Contact

Predicted Future Contact

Model (1) and (2) in Table 3.6 show the estimates of the effects of parents' signals of support on the yearly contact they have with the children. The lower-bar hypothesis is confirmed since the average number of contacts between a step child and the parents is much lower than that of a biological child, when no help was demonstrated. Parents' help tends to lead to more contacts with both biological child and step child. In addition, the gap for the step child is much larger which is indicated by the positive and statistically significant interaction term between the dummy treatment of Parent Ever helped. This is a strong evidence in support of the sensitive response hypothesis. The result resembles the predicts of future contact with child's past support. This suggests that the demonstrated willingness to help and closeness from both the parents' side and the children's side may play important role in shaping their own family norm and mutual emotional bonds.

Predicted Future Child's Interpersonal Help

Model (3) and (4) in Table 3.6 show the effects of parents' support on future chance that children provide interpersonal help in days. When parents provided no help, biological children seem to be provide more interpersonal help compared to step children, but the difference is not significant. This is a weak evidence for the low-bar hypothesis. When parents did provide help, the step children are likely to give more interpersonal help than the biological children, which lends support to sensitive response hypothesis.

Predicted Future Children's Money Transfer Model (5) and (6) in Table 3.6 summarizes the estimates of the effects of parents' interpersonal support. Figure 3.2 demonstrates the interaction effects of being step children and the presence of parent's help in at least one wave. First, the predicted amount of children' future money transfer does not differ much for the step children compared to for biological children no matter if parents helped them in the past, as the blue line is not significantly lower than the red line for the two types of children. Second, there is no evidence for both biological child and step children to give significantly more money in response of parents' help they previously received. While this result is apparently contradictory to both the biological premium hypothesis and the sensitive response hypothesis, this is consistent with previous findings that intergenerational financial transfer is mainly downward flow than upward, perhaps suggesting a weak norm for adult children to pay back to the parents' investment in the form of monetary return, regardless of the step status.

Predicted Future Parent Expectation Model (7) and (8) in Table 3.6 show the estimates of the effects of parents' support on their own expectations of children's future help. From the estimates, the parent respondents are more likely to expect future help from a biological child than from a step child, which is a strong evidence of lower-bar hypothesis. The parents tend to expect slightly more from their biological adult children given that the children provided support, which is not true for step children. This result supports the biological premium hypothesis and contradicts the sensitive response hypothesis. It may also suggest that the parents are indeed quite altruistic when they offer help, in the sense that they would not adjust their level of the expectation for reciprocal response simply because of their own give.

Summary

Comparison across transfer, support and contact supports the differential convergence hypothesis, as illustrated in Table 3.7. step-kin responds to each other's past support differently from biological kin. First, since parents' future monetary transfers do not seem to increase with biological children's past interpersonal help, monetary transfer manifests as the unique way of parents to convey positive feedback to step children's signals of help. Parents' response by financial means only suggests that the step norm is more practical than the real biological kinship. However, beyond instrumental solidarity, significantly strengthened emotional bonding is also found, as is demonstrated by the increase of contacts, which reflects emotional attachment and cohesion. In terms of interpersonal support, step children do react to parents' signals by increasing interpersonal help, which suggests that they not only feel closer, but are also willing to take more responsibilities. From the parents' side, although they provide more childcare support to step children in response to their past help as well, childcare support toward biological children is still more likely. Perhaps the ambiguity and responsibility involved in grandparent childcare entails some hesitation by parents or step children in provision or reception of such support. This is echoed by the finding that respondent parents consistently hold higher expectation towards their biological children given the same level of child's help. In spite of the convergence of the actual stepkin's network of support in terms of the functions of transfer, support and contacts, the step parent's perceptions appears to be more rigid.

Step Parent's Gender and Dynamics of Exchange

Past research finds that the "step gaps" in intergenerational support and contacts are highly gendered (Kalmijn et al. 2019; Wiemers et al. 2019). They show that the contact and support between step mother and step children are much weaker than between step father and step children. This points to the essential role of kin keeping by biological mothers regardless of the father type. In this section, I examine how the patterns of step-kin dynamics of interactions differ by step parent's gender.

In this analysis of step families, three groups of adult children can be identified in the data: children who are born in the parents current marriage who have both biological mothers and fathers (885, or 16%), children with step mothers and biological fathers (2,139, or 38%), and children with biological mothers and step fathers (2,541, or 46%). The effects of children's past signals of support on the future parents response by step/biological status is shown in Figure 3.3, and the effects of parents' past signals of support on the future children's response by step status is shown in Figure 3.4. Parent expectations in Figure 3.1 and Figure 3.2 are omitted in the three-group framework because they are reported based on the step/biological status to the parent respondent himself/herself, rather than the step status in the family.

As I can see from the margin plots, the low-bar expectation for step-kin is highly gendered. When the step-kin have not shown signals of helpfulness, children with step fathers are more similar to biological children compared to children with step mothers; the parents and children's future transfer, support and contacts are higher for children of step fathers than children of step mothers. This suggests the essential role of biological mothers in securing the children's resources.

The pattern of sensitive response and differential convergence is also highly gendered. Specially, there are three interesting findings. First, the typical gender roles of father and mother of nuclear families are likely carried over to step parents given past children's support, which is consistent with the adaptive normative theory. From Figure 3.3, parents' monetary transfers are more sensitive to the past signal of children with step fathers and the frequency of contact increases more sensitively to the signals of children with step mothers than the other two groups. A possible explanation may be step fathers respond sensitively by providing monetary support and step mothers respond sensitively by increasing contacts. This mirrors gender role of parents in intergenerational transfers (Manning and Smock 2000; Rossi and Rossi 1990; Swartz 2009). Nevertheless, this analysis shows that fathers and mothers motivations to transfer money and make contacts to a biological child and a step child are likely different; while the monetary transfer to and contact with the biological children is entirely independent of the children's past signals, the monetary transfer to children of step fathers and the contacts with children of step mothers are both more "exchange-based".

Second, parents' provision of childcare is highly selective on the mothers' blood tie, but its occurrence do require more signals of support from the adult children's than parents' monetary transfer. Parents are the same likely to provide child care support to their biological children as the children with step fathers (and biological mothers). But parents' childcare responds to mothers' biological children's past support sensitivity. This may be because that grandchild care is mainly provided by mothers and may involve most under-defined responsibilities and ambiguity, which requires both a blood tie and a closer mother-child relationship to overcome the barriers.

Third, children with step fathers are most sensitive to the parent's past help that they would provide the most personal care to the parents in response to the parents' past support. Previous research shows that biological children of divorced parents are as likely to provide support to their mothers as children of widowed parents, but they are less likely to provide support to divorced fathers(Lin 2008). Different from that, this result suggests that children of either the step fathers and mothers are equally less likely to provide interpersonal support than biological children if there are no parents' past signals of help. In spite of the closer relationship of children with divorced mothers(Lin 2008), children who are born in the mother's past marriage may develop a different relationship to the parents compared to the children born in the mother's current marriage. The former's provision of senior care is more "exchanged based" than the latter, which reflects a more insecure relationship and a larger variance in mother-child closeness.

Conclusion and Discussion

The high rates of divorce and remarriage in the US highlights a fundamental question regarding family relationship: Who is in the family and who is not (Seltzer 2019)? In this chapter, I study the dynamics of intergenerational interactions comparing step-kin and biological kin in step families in order to better understand the dynamics of family boundary and the motivations of intergenerational exchanges. Using the HRS data assembled from 1996 to 2014 and within-family estimation, I estimate 1) how adult step children's past signals of support affects the parents' future expectations, monetary transfer, childcare support and their contacts as compared to biological children, and 2) how parents' past signals of support to the adult step children affects children's future monetary transfer, personal care support, contacts, and the parents' own expectation as compared to biological adult children. Past research has investigated the levels of intergenerational supports and contacts and found a weaker network of support by step-kin, but has not examined the dynamics of family exchanges.

In contrast to the theory of biological favoritism which predicts a consistently lower step-kin support and contacts regardless of each other's past signals of help, I argue that a weaker norm of support is not definitive to step-kin intergenerational relationship, and propose three competing hypotheses of how step-kin family exchange may unfold differently: low-bar expectation, sensitive response and deferential convergence. From the analysis, I show that when neither of the step children nor the parents have shown any signs of support in the past, their future responses are lower and support network is weaker compared to biological kin (*low-bar expectation*). But as the adult children provide signals of helpfulness, parents would increase monetary transfers to step children substantially more as compared to their biological children who also helped (*sensitive response*). As past signals of help from either the parents or the children are present, the frequency of contacts between step-kin also rises significantly. Remarkably, when past support is present, the level of parents' monetary transfer and contact frequency becomes equivalent to biological kin. This suggests that past signals of supports between step-kin can increase both instrumental and emotional solidarity across generations. In addition, the ways that step-kin norm converge to biological kin norm is not uniform and the *deferential convergence hypothesis* is supported. Step-kin relationship converges to biological kin in terms of parent's monetary transfer, contact and children's senior care support, but not in likelihood of parents' childcare provision.

The study has several limitations. First, although I control for the step children's age when entering the step family, I am not able to measure their coresidence history as in Kalmijn et al. (2019). Even though I control for the parent-to-child and child-to-parent transfers and parent expectations to reflect their adulthood closeness prior to the signals of past help, these parameters might not completely capture the relationship quality. More detailed measures of the parent-children relationship in the childhood could be helpful. Second, the future response of parents and adult children are surveyed every other year, which provides a coarse measure of the intergenerational transfer, contact and support. Besides, HRS does not measure the adult children's expectation of future help from the parents since parents are the respondents, which does not allow to evaluate the children's perceptions of the norms. Third, due to the inconsistent participation of respondent parents in surveys from 1996 to 2014, I have to aggregate and average across multiple years to measure the signal sending and future responses to study the dynamics of intergenerational exchange. The time elapsed between the parents' or adult children's signal and the future response is relatively long. Therefore, the effects of the past signals may involve the process of many rounds of intergenerational exchanges which are triggered by the initial signal. Data with more complete history of supports, contacts and transfers in each wave would allow to investigate multiple stages of intergenerational exchange and interactions in the future. That would be helpful to recover the full process of the dynamics of step-kin intergenerational exchange, change of expectations and norms over time.

Investigating family dynamics under different norms sheds some light on the motivations of intergenerational exchange and interaction. If I interpret that altruistic help is independent of others' past support or potential future return, and exchange-based support are in expectation of reciprocity, then the higher sensitivity of step-kin exchange suggests that biological kin relationship is indeed more altruistic, while step-kinship more exchange-based.

Besides, different dynamics reflect different norms. Among the biological child, biological parent, step child and the step parent, the biological child is the least sensitive to past support, followed by biological parents, step parent and the step child. This suggests that the biological norms grant more sense of security so that reciprocity is not necessary to maintain their cohesion. Besides, due to the predominance of downward transfers in the Western countries, including the US (Eggebeen and Hogan 1990), the biological child may feel normal in a taker's role. Step norms are perhaps rapidly forming given the expansion of step families in recent decades. From an evolutionary perspective (Axelrod 1986), a stable norm benefits the survival of a group by reducing the cost in repeated social interactions. Given the importance of conjugal relationship in nuclear family systems in the US, step-kin may have the motivation to build a more solidary norm. In this process, the step child seems even more proactive. If the existing norm of downward transfer is carried over to step-kinship, perhaps step children can benefit more.

Table and Figures

Table 3.1: Summary of the hypothesis

Biological Premium	Symbolic meaning of a blood tie and strong biological norm. Biological kin have more transfer, support and contact, regardless of past signals of help.
Low-bar Expectation	Default perception (a distant step norm) subject to change. Biological kin have more transfer, support and contact, but only when there is no signals of help.
Sensitive Response	Ambiguous norm of step-kin, so future response more "conditioned". A larger effect of a step-kin's signal of help. The "step gaps" of transfers, support and contact may be reduced and even cancelled.
Deferential convergence	Transfer, contact and support, different implications on family boundary. Contact and support indicate how much step norm is family-like than transfer. But step-kin may response with less interpersonal support due to a lower risk tolerance.

Table 3.2: Descriptive Statistics of Family and Parent Characteristics

Variables	Mean	SD	Min.	Max.	Count
Years of Schooling	11.95	3.02	0	17	1333
Male	0.13	0.34	0	1	1333
White	0.77	0.42	0	1	1333
Black	0.15	0.36	0	1	1333
Hispanics	0.07	0.25	0	1	1333
Other Races	0.01	0.11	0	1	1333
Birthyear	1934	8.26	1910	1946	1333
Married [1]	1	0	1	1	1333
Household Wealth	204.14	610.43	-388.20	16008.93	1333
Household Income	54.59	58.15	0	932	1333
Sibling Size	Frequency	Percentage			
Number biological children: 0	951	71.34			
Number biological children: 1	150	11.25			
Number biological children: 2	103	7.73			
Number biological children: 3	57	4.28			
Number biological children: 4	72	5.40			
Number step children: 1	315	23.63			
Number step children: 2	230	17.25			
Number step children: 3	194	14.55			
Number step children: 4	594	44.56			

Notes: 1. For the within-family fixed effects analysis, I identify step families by parents' remarriage status. So I miss-coded the measures of intergenerational transfer, contacts and support which constitute the outcome variables if parents are not married including in wave 3 and sequential waves.

Table 3.3: Descriptive statistics of the Outcomes, by step status

Variables	Status	Mean	SD	Min.	Max.	Count
Outcomes						
Parents' Future Transfer-Outcomes of Child's Past Support:						
Parent Money Transfer[1]	Biological	836.14	2722.3	0	32521.91	609
	Step	642.25	3206.47	0	79949.69	4115
Non-zero Parent Money Transfer[2]	Biological	1229.17	4337.14	16.67	32521.91	191
	Step	975	6285.67	16.67	79949.69	923
Parent Grandchild Care[5]	Biological	0.22	0.42	0	1	609
	Step	0.12	0.33	0	1	4201
Children's Future Transfer-Outcomes of Parents' Past Support:						
Child Money Transfer[1]	Biological	44.4	388.29	0	7500	608
	Step	64.33	711.72	0	31250	4103
Non-zero Child Money Transfer[2]	Biological	322.92	1469.74	50	7500	34
	Step	436	3055.81	41.67	31250	185
Child Days Helped[5]	Biological	0.36	2.13	0	30.5	818
	Step	0.33	2.22	0	30.33	4388
Shared Outcomes:						
Parent Expectation for Future Help[4]	Biological	0.25	0.43	0	1	3401
	Step	0.1	0.29	0	1	1997
Contact[3]	Biological	154.21	127.33	0	399.67	570
	Step	86.99	107.37	0	399	3989
Treatments:						
Adult Children's Past Support:						
Child Ever Helped	Biological	0.45	0.5	0	1	885
	Step	0.19	0.39	0	1	4680
Child Waves Helped	Biological	0.76	1.02	0	4	885
	Step	0.3	0.72	0	4	4680
Parents' Past Support:						
Parent Ever Helped	Biological	0.33	0.47	0	1	885
	Step	0.19	0.39	0	1	4680
Parent Waves Helped	Biological	0.63	1.07	0	4	885
	Step	0.35	0.85	0	4	4680

Notes: 1. Monetary measures are adjusted to 2014 dollars. This is a yearly measure averaged from wave 7 to wave 12. In the models, they are normalized by log-transformation. 2. Results are reported for those whose money transfers values are larger than zero. Medians are reported instead of means. 3. Yearly contact frequency is averaged from wave 7 to wave 12. In the models, it is normalized by log-transformation. 4. For parent expectation for future help, the step status indicates if the adult children are step children to the respondent himself/herself. Because this question regards the respondent's individual expectation. 5. Child Days Helped is a monthly measure, which is averaged from wave 7 to wave 12.

Table 3.4: Descriptive statistics of the control variables, by step status

Variables	Status	Mean	SD	Min.	Max.	Count
Age(1996)	Biological	33.95	7.95	20	59	885
	Step	37.37	8.13	20	60	4680
Male	Biological	0.52	0.50	0	1	885
	Step	0.51	0.50	0	1	4680
Years of Schooling	Biological	13.34	2.32	0	17	884
	Step	13.07	2.36	0	17	4657
Married	Biological	0.56	0.50	0	1	885
	Step	0.60	0.49	0	1	4680
Non-married	Biological	0.42	0.49	0	1	885
	Step	0.38	0.49	0	1	4680
Partnered	Biological	0.01	0.12	0	1	885
	Step	0.01	0.10	0	1	4680
Work Full-time	Biological	0.70	0.46	0	1	885
	Step	0.69	0.46	0	1	4680
Work Part-time	Biological	0.09	0.29	0	1	885
	Step	0.08	0.27	0	1	4680
Not Working	Biological	0.18	0.39	0	1	885
	Step	0.16	0.37	0	1	4680
Live Within 10 Miles	Biological	0.45	0.50	0	1	751
	Step	0.28	0.45	0	1	4425
Money Transfer Children Received (1994)	Biological	725.38	2528.64	0	37000	853
	Step	618.55	6616.64	0	160000	4360
Non-zero Money Transfer Children Received (1994)[1]	Biological	1000	4264.78	5	37000	236
	Step	800	16076.20	5	160000	706
Expect the Child's Help (1996)	Biological	0.28	0.45	0	1	860
	Step	0.18	0.39	0	1	4608
Money Transfer Parents Received (1994)	Biological	8.57	63.27	0	800	840
	Step	3.98	66.12	0	2400	4348
Non-zero Money Transfer Parents Received (1994)[1]	Biological	200	234.06	100	800	24
	Step	400	559.26	25	2400	30
Income	Biological	36310.94	27874.53	0	200000	705
	Step	44089.16	32259.50	0	400000	3583
Have a Grandchild	Biological	0.86	0.35	0	1	885
	Step	0.86	0.35	0	1	4678
Child Older Than 10 at Parents' Marriage	Biological	0	0	0	0	885
	Step	0.75	0.43	0	1	4680

Notes: 1. Results are reported for those whose money transfers values are larger than zero. Medians are reported instead of means. 2. All monetary measures were adjusted to 2014 US dollars. In the models, they are normalized by log-transformation.

Table 3.5: Sibling fixed effects estimation of parent feedbacks on adult children's help by biological status

	Parent Money Transfer[2]		Contact[3]		Parent Childcare[4]		Parent Expectation[5]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Step Child	-0.59*	-0.50*	-0.54***	-0.53***	-0.04	-0.05	-0.10***	-0.10***
	(0.26)	(0.25)	(0.15)	(0.14)	(0.04)	(0.04)	(0.01)	(0.01)
Child Ever Helped	-0.15		0.56**		0.16**		0.09***	
	(0.35)		(0.20)		(0.05)		(0.02)	
Step Child *Child Ever Helped	0.77*		0.60**		-0.02		0.03	
	(0.37)		(0.21)		(0.05)		(0.04)	
Child Waves Helped		0.04		0.32**		0.06*		0.05***
		(0.18)		(0.10)		(0.02)		(0.01)
Step Child *Child Waves Helped		0.31		0.39***		0.01		0.03
		(0.19)		(0.11)		(0.03)		(0.02)
Constant	11.20***	11.24***	3.50*	3.72*	2.20***	2.23***	0.11	0.12
	(3.04)	(3.04)	(1.70)	(1.69)	(0.40)	(0.40)	(0.35)	(0.35)
N	3256	3256	3165	3165	2860	2860	3644	3644
N_g	926	926	916	916	896	896	1039	1039

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marrige, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Parents' money transfer adjusted for inflation and is log transformed 3. Contact amount is log transformed. 4. Sample is restricted to the adult children who have grandchildren by wave 12 (2014).5. Parent expectation regards the step child of himself/herself.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3.6: Sibling fixed effects estimation of adult children's feedbacks on parental interpersonal support

	Contact[2]		Child's Interpersonal Help		Child's Money Transfer[3]		Parent Expectation[4]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Step Child	-0.63***	-0.63***	-0.22	-0.17	-0.02	0	-0.10***	-0.10***
	(0.18)	(0.17)	(0.22)	(0.21)	(0.15)	(0.14)	(0.01)	(0.01)
Parent Ever Help[5]	0.32		-0.10		0.07		0.03^{+}	
	(0.21)		(0.25)		(0.17)		(0.02)	
Step Child *Parent Ever Help	0.48*		0.50^{+}		0.02		-0.03	
	(0.22)		(0.28)		(0.19)		(0.04)	
Parent Waves Help[5]		0.08		-0.02		0.05		0.01
		(0.07)		(0.10)		(0.06)		(0.01)
Step Child *Parent Waves Help		0.21*		0.18		0		-0.02
		(0.08)		(0.12)		(0.07)		(0.02)
Constant	2.83	2.99^{+}	3.04	3.02	4.23**	4.17**	0.09	0.11
	(1.75)	(1.76)	(2.54)	(2.54)	(1.52)	(1.51)	(0.37)	(0.37)
N	2615	2615	2917	2917	2678	2678	3020	3020
N_g	868	868	945	945	874	874	980	980

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marrige, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Contact amount is log transformed. 3. Kids' money transfer adjusted for inflation and is log transformed 4. Parent expectation regards the step child of himself/herself. 5. Parent's help and waves of help measures grandchild care. Sample restricted to adult children who have grandchildren by wave 3 (1996).

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3.7: Summary of the results

	Biological premium	Low-bar expectation	Sensitive response	Deferential convergence
How children's past support affect parent's future transfer and contact				
Parent Money Transfer	×	✓	✓	✓
Grandchild Care	✓	✓	×	
Parent Expectation	✓	✓	×	
Contact	×	✓	✓	
How parents' past support affect children's future transfer and contact				
Child Money Transfer	×	×	×	✓
Child Interpersonal Help	×	×	✓	
Contact	×	✓	✓	
Parent expectation	✓	✓	×	

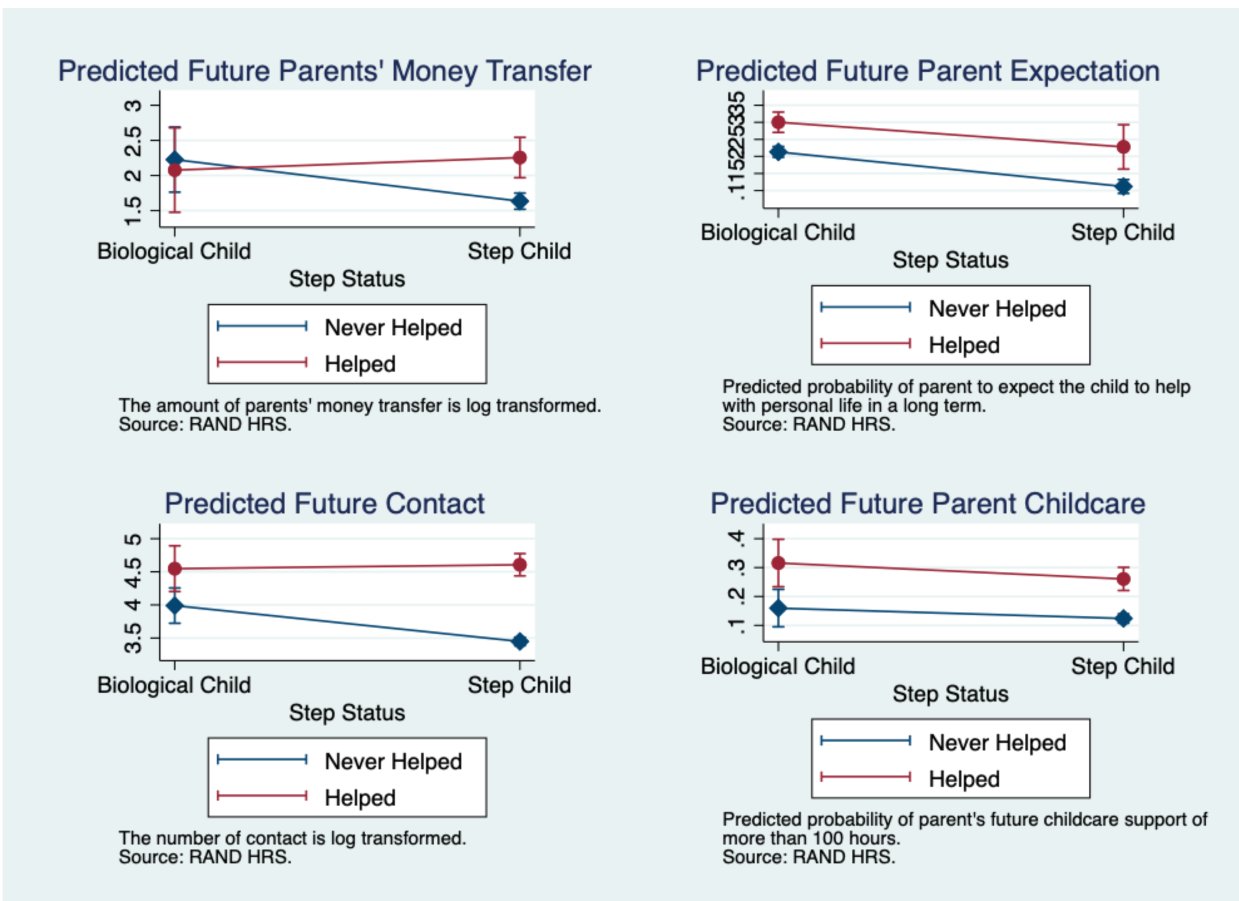


Figure 3.1: The margin plots of predicted parent's transfer, expectation and contacts by past child's past support and step/biological status

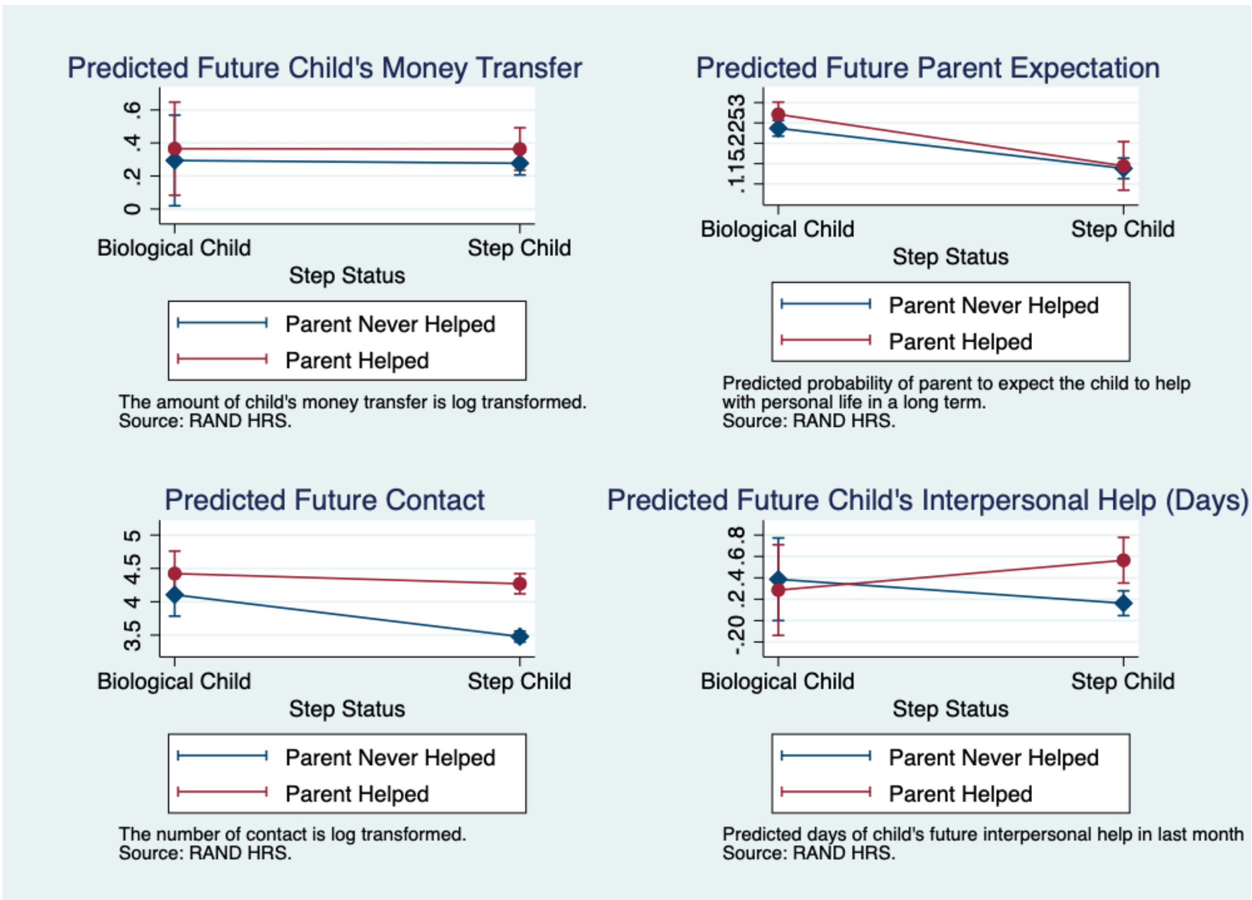


Figure 3.2: The margin plots of predicted child's transfer, interpersonal help and contacts and parent's expectation by parents' past support and step/biological status

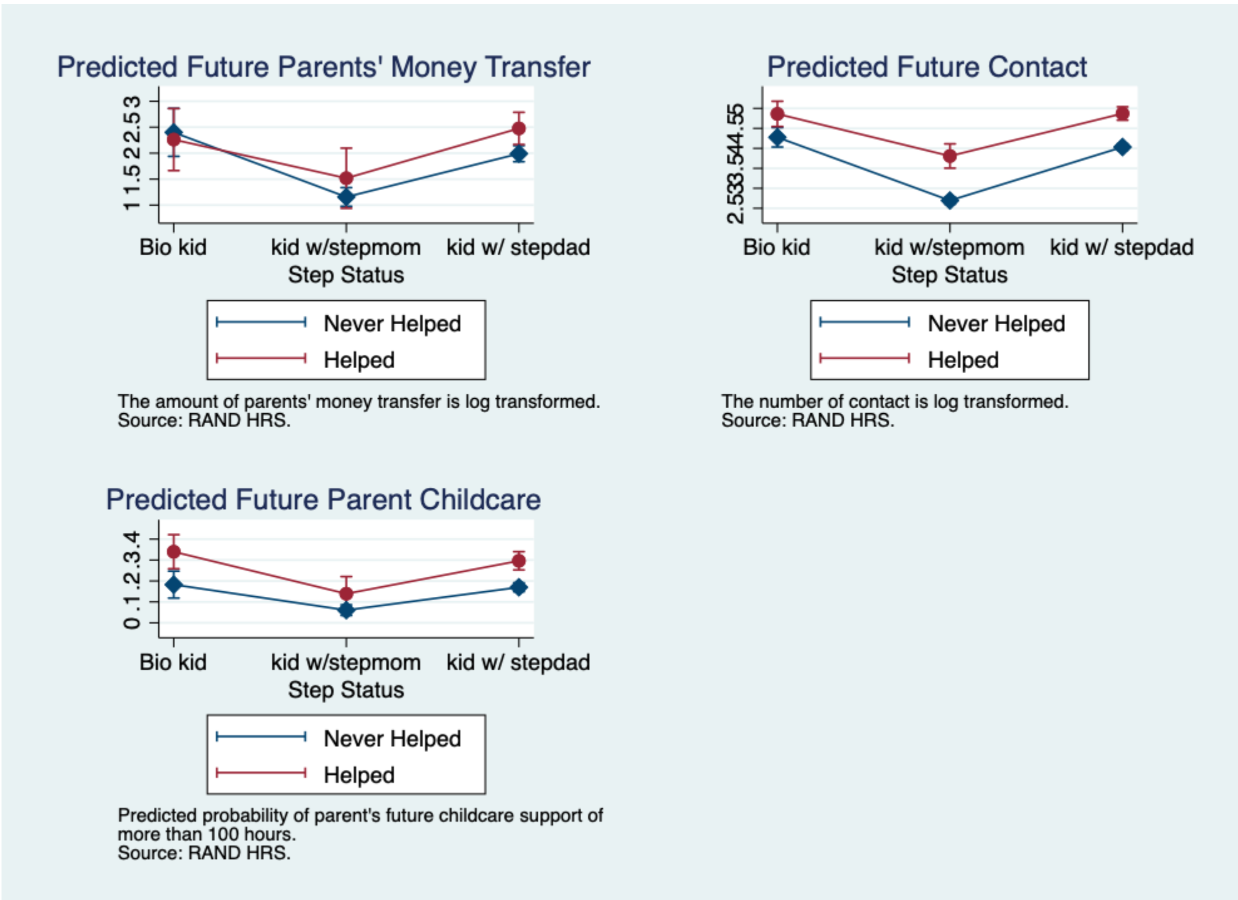


Figure 3.3: The margin plots of predicted parent's transfer and contacts by past child's past support and step/biological status

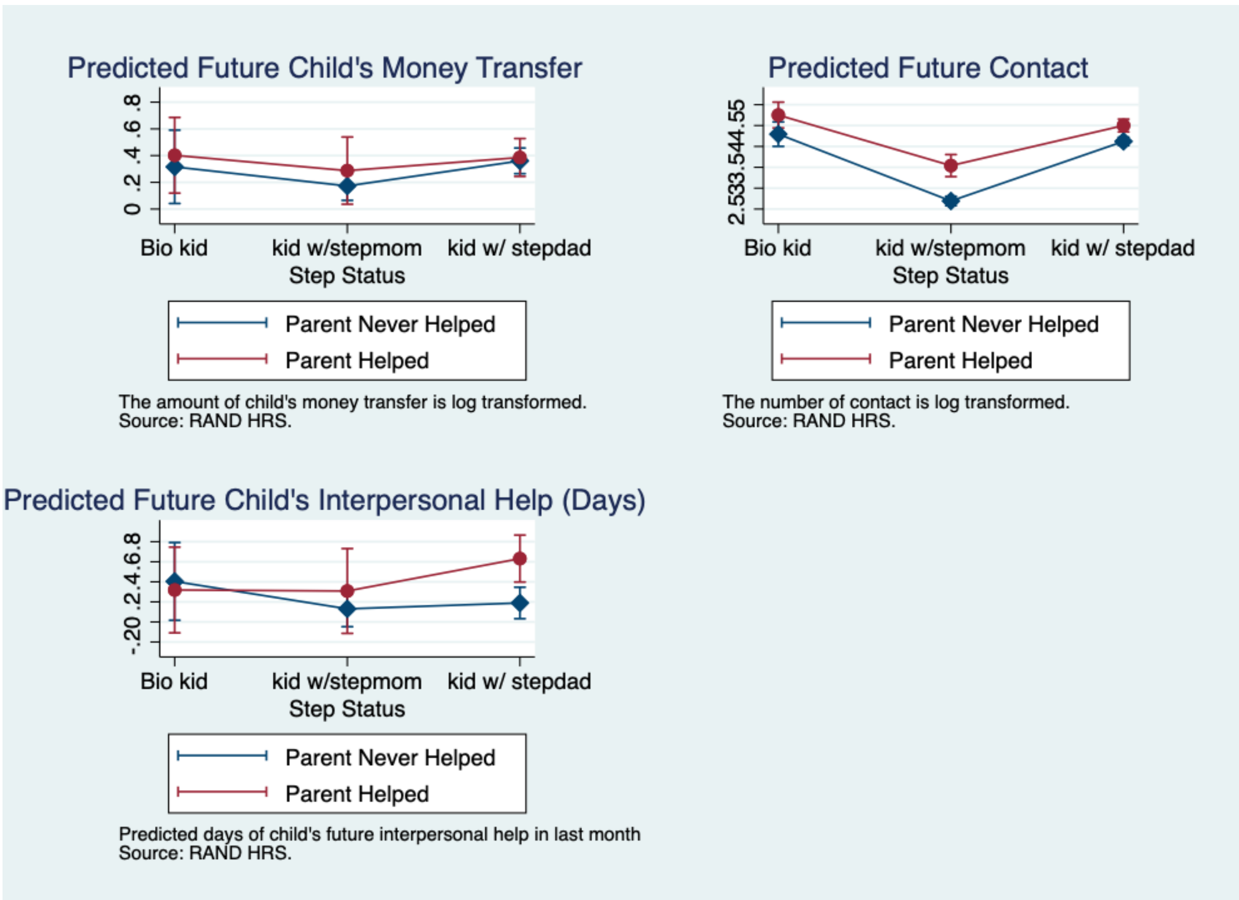


Figure 3.4: The margin plots of predicted child's transfer and contacts by parents' past support and step/biological status

Table 3.8: Sibling fixed effects estimation of parent feedbacks on adult children's help by biological status

	Parent Money Transfer[2]		Contact[3]		Parent Childcare[4]	
	(1)	(2)	(3)	(4)	(5)	(6)
Child of Step Mother	-1.24***	-1.18***	-1.58***	-1.58***	-0.12**	-0.14***
	(0.27)	(0.27)	(0.14)	(0.14)	(0.04)	(0.04)
Child of Step Father	-0.41	-0.31	-0.24^{+}	-0.24^{+}	-0.01	-0.02
	(0.26)	(0.25)	(0.14)	(0.13)	(0.04)	(0.03)
Child Ever Helped	-0.14		0.59**		0.16***	
	(0.34)		(0.18)		(0.05)	
Child of Step Mother *Child Ever Helped	0.50		0.53*		-0.08	
	(0.45)		(0.24)		(0.06)	
Child of Step Father *Child Ever Helped	0.62^{+}		0.25		-0.03	
	(0.37)		(0.20)		(0.05)	
Child Waves Helped		0.05		0.34***		0.06*
		(0.18)		(0.09)		(0.02)
Child of Step Mother *Child Waves Helped		0.33		0.53***		0.00
		(0.28)		(0.15)		(0.04)
Child of Step Father *Child Waves Helped		0.20		0.19^{+}		0.00
		(0.19)		(0.10)		(0.03)
Constant	11.91***	12.10***	5.09**	5.29***	2.29***	2.35***
	(3.02)	(3.01)	(1.55)	(1.54)	(0.40)	(0.40)
N	3256	3256	3165	3165	2860	2860
N_g	926	926	916	916	896	896

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marrige, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Parents' money transfer adjusted for inflation and is log transformed 3. Contact amount is log transformed. 4. Sample is restricted to the adult children who have grandchildren by wave 12 (2014).

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3.9: Sibling fixed effects estimation of adult children's feedbacks on parental interpersonal support

	Contact[2]		Child's Interpersonal Help		Child's Money Transfer [3]	
	(1)	(2)	(3)	(4)	(5)	(6)
Child of Step Mother	-1.601***	-1.669***	-0.274	-0.245	-0.143	-0.126
	(0.171)	(0.164)	(0.236)	(0.228)	(0.161)	(0.153)
Child of Step Father	-0.173	-0.223	-0.215	-0.146	0.045	0.046
	(0.165)	(0.156)	(0.224)	(0.213)	(0.155)	(0.146)
Parent Ever Help[4]	0.457*		-0.086		0.086	
	(0.187)		(0.257)		(0.174)	
Child of Step Mother *Parent Ever Help	0.390		0.264		0.028	
	(0.238)		(0.346)		(0.223)	
Child of Step Father *Parent Ever Help	-0.077		0.528^{+}		-0.061	
	(0.207)		(0.287)		(0.192)	
Parent Waves Help[4]		0.118		-0.020		0.056
		(0.072)		(0.107)		(0.066)
Child of Step Mother *Parent Waves Help		0.323**		0.100		-0.009
		(0.108)		(0.167)		(0.102)
Child of Step Father *Parent Waves Help		0		0.175		-0.027
		(0.081)		(0.121)		(0.075)
Constant	4.770**	4.953**	3.120	3.163	4.473**	4.401**
	(1.601)	(1.603)	(2.550)	(2.550)	(1.522)	(1.520)
N	2615	2615	2917	2917	2678	2678
N_g	868	868	945	945	874	874

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marrige, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Contact amount is log transformed. 3. Kids' money transfer adjusted for inflation and is log transformed 4. Parent's help and waves of help measures grandchild care. Sample restricted to adult children who have grandchildren by wave 3 (1996).

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix

Section I: Robustness Check With Imputed Samples

In this section, I present the within-family analysis with the missing cases of the covariates completed by 100 times of multiple imputations using chained equations(MICE). As is demonstrated by Table 3.10 and Table 3.11, the sample sizes increase by around 30% to 40%, and all the results regarding the low bar expectation, sensitivity response and differential convergence remain unchanged.

The parent's future transfer, contact, chance of providing childcare and expectation for future help are lower for step children when neither the biological or step adult children show signals of support in the past. But when step-kin do show support, parents would increase the monetary transfer to the step children to a level that is comparable to the biological children. Intergenerational contact also increases sensitively to both the adult children's or the parents' past support. Parents remain less likely to provide childcare to step children in spite of their past help, but step children respond to the parents' past help with more interpersonal care. Again, parents consistently expect less from the step children than their biological children given the same past supports offered by either the children or themselves.

Table 3.10: Sibling fixed effects estimation of parent feedbacks on adult children's help by biological status, with multiple imputation of covariates

	Parent Money Transfer[2]		Contact[3]		Parent Childcare[4]		Parent Expectation[5]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Step Child	-0.58**	-0.57**	-0.59***	-0.58***	-0.08**	-0.08**	-0.10***	-0.10***
	(0.20)	(0.19)	(0.12)	(0.11)	(0.03)	(0.03)	(0.01)	(0.01)
Child Ever Helped	-0.06		0.63***		0.11**		0.09***	
	(0.25)		(0.15)		(0.04)		(0.01)	
Step Child *Child Ever Helped	0.58*		0.60***		0.03		0.02	
	(0.27)		(0.16)		(0.04)		(0.03)	
Child Waves Helped		-0.02		0.35***		0.04*		0.04***
		(0.13)		(0.08)		(0.02)		(0.01)
Step Child *Child Waves Helped		0.34*		0.37***		0.03		0.04*
		(0.14)		(0.08)		(0.02)		(0.02)
Constant	3.88**	3.94**	2.92***	3.03***	1.19***	1.20***	0.24^{+}	0.25^{+}
	(1.31)	(1.31)	(0.87)	(0.86)	(0.18)	(0.18)	(0.13)	(0.13)
N	4612	4612	4450	4450	4038	4038	5396	5396
N_g	1117	1117	1108	1108	1097	1097	1295	1295

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marrige, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Parents' money transfer adjusted for inflation and is log transformed 3. Contact amount is log transformed. 4. Sample is restricted to the adult children who have grandchildren by wave 12 (2014).5. Parent expectation regards the step child of himself/herself.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3.11: Sibling fixed effects estimation of adult children's feedbacks on parental interpersonal support, with multiple imputation of covariates

	Contact		Child's Interpersonal Help		Child's Money Transfer		Parent Expectation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Step Child	-0.581***	-0.562***	-0.166	-0.064	0.152	0.136	-0.087***	-0.091***
	(0.15)	(0.14)	(0.19)	(0.18)	(0.12)	(0.12)	(0.01)	(0.01)
Parent Ever Help	0.318^{+}		-0.206		0.218		0.061***	
	(0.17)		(0.22)		(0.14)		(0.02)	
Step Child *Parent Ever Help	0.470**		0.670**		-0.05		-0.050^{+}	
	(0.18)		(0.25)		(0.15)		(0.03)	
Parent Waves Help		0.097		-0.071		0.087		0.021**
		(0.07)		(0.09)		(0.06)		(0.01)
Step Child *Parent Waves Help		0.200**		0.191^{+}		-0.005		-0.024
		(0.07)		(0.11)		(0.06)		(0.02)
Constant	2.862**	2.852**	1.586	1.521	0.307	0.309	0.189	0.202
	(1.01)	(1.01)	(1.29)	(1.29)	(0.76)	(0.76)	(0.15)	(0.15)
N	3616	3616	4136	4136	3721	3721	4391	4391
N_g	1063	1063	1163	1163	1072	1072	1238	1238

Note: 1. All models are sibling fixed effects regressions controlling for the relationship quality measured by baseline downward and upward financial transfer, parent's expectation and the step child's age at the parents' marriage, and children's covariates, including gender, education, income, age, marital status and employment status. 2. Kids' money transfer adjusted for inflation and is log transformed 3. Parent expectation regards the step child of himself/herself. 4. Contact amount is log transformed. 5. Parent's help and waves of help measures grandchild care. Sample restricted to adult children who have grandchildren by wave 3 (1996).

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Bibliography

Aburto, José Manuel, Maarten Wensink, Alyson van Raalte, and Rune Lindahl-Jacobsen. 2018. "Potential gains in life expectancy by reducing inequality of lifespans in Denmark: an international comparison and cause-of-death analysis." *BMC public health* 18:831.

Albertini, Marco, Martin Kohli, and Claudia Vogel. 2007. "Intergenerational transfers of time and money in European families: common patterns—different regimes?" *Journal of European social policy* 17:319–334.

Amato, Paul R, Sandra J Rezac, and Alan Booth. 1995. "Helping between parents and young adult offspring: The role of parental marital quality, divorce, and remarriage." *Journal of Marriage and the Family* pp. 363–374.

Anderson, Kermyt G, Hillard Kaplan, and Jane Lancaster. 1999. "Paternal care by genetic fathers and stepfathers I: Reports from Albuquerque men." *Evolution and Human Behavior* 20:405–431.

Anderson, Lewis R, Paula Sheppard, and Christiaan WS Monden. 2018. "Grandparent Effects on Educational Outcomes: a Systematic Review." *Sociological Science* 5:114–142.

Andorka, Rudolf. 1997. "Social Mobility in Hungary since the Second World War: Interpretations through surveys and through family histories." *Pathways to social class: A qualitative approach to social mobility* pp. 259–298.

Andreoni, James. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of political Economy* 97:1447–1458.

Aquilino, William S. 2006. "The noncustodial father–child relationship from adolescence into young adulthood." *Journal of Marriage and Family* 68:929–946.

Axelrod, Robert. 1986. "An evolutionary approach to norms." *The American political science review* pp. 1095–1111.

Becker, Gary S. 1974. "A theory of social interactions." *Journal of political economy* 82:1063–1093.

Becker, Gary Stanley and Gary S Becker. 2009. *A Treatise on the Family*. Harvard university press.

Becker, Gary S and Nigel Tomes. 1986. "Human capital and the rise and fall of families." *Journal of labor economics* 4:S1–S39.

Becker, Oliver Arránz, Veronika Salzburger, Nadia Lois, and Bernhard Nauck. 2013. "What narrows the stepgap? Closeness between parents and adult (step) children in Germany." *Journal of Marriage and Family* 75:1130–1148.

Bengtson, Vern L. 2001. "Beyond the Nuclear Family: The Increasing Importance of Multigenerational Bonds: The Burgess Award Lecture." *Journal of Marriage and Family* 63:1–16.

Bengtson, Vern L and Petrice S Oyama. 2007. "Intergenerational solidarity: Strengthening economic and social ties." *New York: United Nations Headquarters* .

Bengtson, Vern L and Robert EL Roberts. 1991. "Intergenerational solidarity in aging families: An example of formal theory construction." *Journal of Marriage and the Family* pp. 856–870.

Berry, Brent. 2008. "Financial transfers from living parents to adult children: Who is helped and why?" *American Journal of Economics and Sociology* 67:207–239.

Beuchert, Louise and Anne Nandrup. 2017. "The Danish national tests at a glance." *Available at SSRN 2996431* .

Beuchert, Louise Voldby, Anne Brink Nandrup, et al. 2014. "The Danish national tests–A practical guide." Technical report.

Bianchi, Suzanne, Judith Seltzer, V Joseph Hotz, and Kathleen McGarry. 2006. "Intergenerational ties: Alternative theories, empirical findings and trends, and remaining challenges." *UCLA CCPR Population Working Papers* .

Blau, Peter M and Otis Dudley Duncan. 1967. "The American occupational structure." .

Brand, Jennie E and Juli Simon Thomas. 2013. "Causal effect heterogeneity." In *Handbook of causal analysis for social research*, pp. 189–213. Springer.

Buss, David M. 2016. *The evolution of desire: Strategies of human mating*. Basic books.

Carlson, Marcia J and Daniel R Meyer. 2014. "Family complexity: Setting the context." *The ANNALS of the American Academy of Political and Social Science* 654:6–11.

Carneiro, Pedro and James J Heckman. 2002. "The evidence on credit constraints in post-secondary schooling." *The Economic Journal* 112:705–734.

Case, Anne and Christina Paxson. 2005. "Sex differences in morbidity and mortality." *Demography* 42:189–214.

Chamberlain, Gary. 1982. "Multivariate regression models for panel data." *Journal of econometrics* 18:5–46.

Chan, Tak Wing and Vikki Boliver. 2013. "The Grandparents Effect in Social Mobility: Evidence From British Birth Cohort Studies." *American Sociological Review* 78:662–678.

Chapman, Ashton, Marilyn Coleman, and Lawrence Ganong. 2016. ""Like my grandparent, but not": A qualitative investigation of skip-generation stepgrandchild–stepgrandparent relationships." *Journal of Marriage and Family* 78:634–643.

Charlson, Mary E, Peter Pompei, Kathy L Ales, and C Ronald MacKenzie. 1987. "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." *Journal of chronic diseases* 40:373–383.

Cherlin, Andrew. 1978. "Remarriage as an incomplete institution." *American journal of Sociology* 84:634–650.

Cherlin, Andrew and Frank F Furstenberg Jr. 1992. *The new American grandparent*. Harvard University Press.

Clark, Gregory. 2015. *The son also rises: Surnames and the history of social mobility*, volume 49. Princeton University Press.

Conley, Dalton and Rebecca Glauber. 2008. "All in the family?: Family composition, resources, and sibling similarity in socioeconomic status." *Research in Social Stratification and Mobility* 26:297–306.

Daly, Martin and Margo I Wilson. 2000. "The evolutionary psychology of marriage and divorce." *The ties that bind: Perspectives on marriage and cohabitation* pp. 91–110.

Darling, Nancy and Laurence Steinberg. 1993. "Parenting style as context: An integrative model." *Psychological bulletin* 113:487.

Davey, Adam, David J Eggebeen, and Jyoti Savla. 2007. "Parental marital transitions and instrumental assistance between generations: A within-family longitudinal analysis." *Advances in Life Course Research* 12:221–242.

Daw, Jonathan, S Michael Gaddis, and Anne Roback Morse. 2018. "3Ms of 3G: Testing

Three Mechanisms of Three-Generational Educational Mobility in the US." *Available at SSRN 2710883* .

Deaton, Angus and Christina Paxson. 1994. "Intertemporal choice and inequality." *Journal of political economy* 102:437–467.

Dunifon, Rachel E, Christopher E Near, and Kathleen M Ziol-Guest. 2018. "Backup parents, playmates, friends: Grandparents' time with grandchildren." *Journal of marriage and family* 80:752–767.

Dunifon, Rachel E, Kathleen M Ziol-Guest, and Kimberly Kopko. 2014. "Grandparent Coresidence and Family Well-being: Implications for Research and Policy." *The ANNALS of the American Academy of Political and Social Science* 654:110–126.

Eggebeen, David J and Dennis P Hogan. 1990. "Giving between generations in American families." *Human Nature* 1:211–232.

Elwert, Felix. 2013. "Graphical Causal Models." In *Handbook of Causal Analysis for Social Research*, pp. 245–273. Springer.

Elwert, Felix and Nicholas A Christakis. 2008. "Wives and Ex-wives: A new Test for Homogamy Bias in the Widowhood Effect." *Demography* 45:851–873.

Elwert, Felix and Christopher Winship. 2014. "Endogenous selection bias: The problem of conditioning on a collider variable." *Annual review of sociology* 40:31–53.

Erikson, Robert and John H Goldthorpe. 2009. "Social class, family background, and intergenerational mobility: A comment on Mcintosh and Munk." *European Economic Review* 53:118–120.

Erola, Jani, Elina Kilpi-Jakonen, Irene Prix, and Hannu Lehti. 2018. "Resource compensation from the extended family: grandparents, Aunts, and Uncles in Finland and the United States." *European Sociological Review* 34:348–364.

Evenhouse, Eirik and Siobhan Reilly. 2004. "A sibling study of stepchild well-being." *Journal of Human Resources* 39:248–276.

Fine, Mark A, Marilyn Coleman, and Lawrence H Ganong. 1998. "Consistency in perceptions of the step-parent role among step-parents, parents and stepchildren." *Journal of social and personal relationships* 15:810–828.

Fomby, Paula, Patrick M Krueger, and Nicole M Wagner. 2014. "Age at childbearing over two generations and grandchildren's cognitive achievement." *Research in social stratification and mobility* 35:71–88.

Ganong, LH and M Coleman. 2017. "Stepfamily relationships: Development, dynamics, and interventions. Boston, MA, US."

Ganong, Lawrence H and Marilyn Coleman. 1997. "How society views stepfamilies." *Marriage & Family Review* 26:85–106.

Hadfield, Kristin, Margaret Amos, Michael Ungar, Julie Gosselin, and Lawrence Ganong. 2018. "Do changes to family structure affect child and family outcomes? A systematic review of the instability hypothesis." *Journal of Family Theory & Review* 10:87–110.

Hällsten, Martin and Fabian T Pfeffer. 2017. "Grand Advantage: Family Wealth and Grandchildren's Educational Achievement in Sweden." *American Sociological Review* 82:328–360.

Hank, Karsten and Isabella Buber. 2009. "Grandparents caring for their grandchildren: Findings from the 2004 Survey of Health, Ageing, and Retirement in Europe." *Journal of family Issues* 30:53–73.

Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." In *Annals of Economic and Social Measurement, Volume 5, number 4*, pp. 475–492. NBER.

Heckman, James J. 2006. "Skill formation and the economics of investing in disadvantaged children." *Science* 312:1900–1902.

Henretta, John C, Matthew F Van Voorhis, and Beth J Soldo. 2014. "Parental money help to children and stepchildren." *Journal of Family Issues* 35:1131–1153.

Hertel, Florian R and Olaf Groh-Samberg. 2014. "Class mobility across three generations in the US and Germany." *Research in Social Stratification and Mobility* 35:35–52.

Hout, Michael and Robert M Hauser. 1992. "Symmetry and hierarchy in social mobility: a methodological analysis of the CASMIN model of class mobility." *European sociological review* 8:239–266.

Hyndman, Rob J and George Athanasopoulos. 2018. *Forecasting: principles and practice*. OTexts.

Igel, Corinne and Marc Szydlik. 2011. "Grandchild care and welfare state arrangements in Europe." *Journal of European Social Policy* 21:210–224.

Imai, Kosuke and In Song Kim. 2019. "When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?" *American Journal of Political Science* 23:467–490.

Ingersoll, Richard M. 2001. "Teacher Turnover and Teacher Shortages: An Organizational Analysis." *American Educational Research Journal* 38:499–534.

Jæger, Mads Meier. 2012. "The Extended Family and Children's Educational Success." *American Sociological Review* 77:903–922.

Jencks, Christopher, James Crouse, and Peter Mueser. 1983. "The Wisconsin model of status attainment: A national replication with improved measures of ability and aspiration." *Sociology of Education* pp. 3–19.

Jensen, Todd M and Matthew O Howard. 2015. "Perceived stepparent–child relationship quality: A systematic review of stepchildren's perspectives." *Marriage & Family Review* 51:99–153.

Kalmijn, Matthijs. 2013. "Adult children's relationships with married parents, divorced parents, and stepparents: Biology, marriage, or residence?" *Journal of Marriage and Family* 75:1181–1193.

Kalmijn, Matthijs, Suzanne G de Leeuw, Maaike Hornstra, Katya Ivanova, Ruben van Gaalen, and Kirsten van Houdt. 2019. "Family complexity into adulthood: The central role of mothers in shaping intergenerational ties." *American Sociological Review* 84:876–904.

Killian, Timothy S. 2004. "Intergenerational monetary transfers to adult children and stepchildren: A household level analysis." *Journal of Divorce & Remarriage* 42:105–130.

Knigge, Antonie. 2016. "Beyond the Parental Generation: The Influence of Grandfathers and Great-grandfathers on Status Attainment." *Demography* 53:1219–1244.

Kraft, Matthew A. 2020. "Interpreting effect sizes of education interventions." *Educational Researcher* 49:241–253.

Kreider, Rose Marie and Renee Ellis. 2011. *Number, timing, and duration of marriages and divorces, 2009*. US Department of Commerce, Economics and Statistics Administration, US . . . .

Lareau, Annette. 2011. *Unequal childhoods: Class, race, and family life*. Univ of California Press.

Lee, Dohoon and Sara McLanahan. 2015. "Family structure transitions and child development: Instability, selection, and population heterogeneity." *American sociological review* 80:738–763.

Lehti, Hannu, Jani Erola, and Antti O Tanskanen. 2018. "Tying the Extended Family Knot-Grandparents Influence on Educational Achievement." *European Sociological Review* 35:29–48.

Leopold, Thomas and Jan Skopek. 2015. "The demography of grandparenthood: An international profile." *Social Forces* 94:801–832.

Levin, Irene and Marvin B Sussman. 1997. *Stepfamilies: History, research, and policy*, volume 2. Psychology Press.

Leyens, Jacques-Philippe, Vincent Yzerbyt, and Georges Schadron. 1994. *Stereotypes and social cognition.* Sage Publications, Inc.

Lin, I-Fen. 2008. "Consequences of parental divorce for adult children's support of their frail parents." *Journal of Marriage and Family* 70:113–128.

Lindahl, Mikael, Mårten Palme, Sofia Sandgren Massih, and Anna Sjögren. 2015. "Long-term Intergenerational Persistence of Human Capital an Empirical Analysis of Four Generations." *Journal of Human Resources* 50:1–33.

Link, Bruce G and Jo Phelan. 1995. "Social conditions as fundamental causes of disease." *Journal of health and social behavior* pp. 80–94.

Logan, John R and Glenna D Spitze. 1995. "Self-interest and altruism in intergenerational relations." *Demography* 32:353–364.

Lye, Diane N. 1996. "Adult child–parent relationships." *Annual review of sociology* 22:79–102.

Manning, Wendy D and Pamela J Smock. 2000. "“Swapping" families: Serial parenting and economic support for children." *Journal of Marriage and Family* 62:111–122.

Mare, Robert D. 2011. "A Multigenerational View of Inequality." *Demography* 48:1–23.

Mare, Robert D. 2014. "Multigenerational Aspects of Social Stratification: Issues for Further Research." *Research in Social Stratification and Mobility* 35:121–128.

Margolis, Rachel and Laura Wright. 2017. "Healthy grandparenthood: How long is it, and how has it changed?" *Demography* 54:2073–2099.

Mayer, Susan E. 1997. *What money can't buy: Family income and children's life chances*. Harvard University Press.

McKinney, Chelsea O. 2015. "Early childhood overweight and obesity in multigenerational households." *GrandFamilies: The Contemporary Journal of Research, Practice and Policy* 2:2.

McLanahan, Sara. 2004. "Diverging Destinies: How Children are Faring Under the Second Demographic Transition." *Demography* 41:607–627.

McLanahan, Sara, Laura Tach, and Daniel Schneider. 2013. "The causal effects of father absence." *Annual review of sociology* 39:399–427.

Modin, Bitte, Robert Erikson, and Denny Vågerö. 2012. "Intergenerational Continuity in School Performance: do Grandparents Matter?" *European Sociological Review* 29:858–870.

Møllegaard, Stine and Mads Meier Jæger. 2015. "The effect of grandparents' economic, cultural, and social capital on grandchildren's educational success." *Research in Social Stratification and Mobility* 42:11–19.

Morgan, Stephen L and Christopher Winship. 2015. *Counterfactuals and causal inference*. Cambridge University Press.

Morin, Rich. 2011. "The Public Renders a Split Verdict On Changes in Family Structure." Technical report, Washington, DC: Pew Research Center.

Mueller, Margaret M, Brenda Wilhelm, and Glen H Elder Jr. 2002. "Variations in grandparenting." *Research on aging* 24:360–388.

Myers, George C and Kenneth G Manton. 1984. "Compression of mortality: myth or reality?" *The gerontologist* 24:346–353.

Pfeffer, Fabian T. 2014. "Multigenerational Approaches to Social Mobility. A Multifaceted Research Agenda." *Research in Social Stratification and Mobility* 35:1.

Preston, Samuel H. 1984. "Children and the elderly: Divergent paths for America's dependents." *Demography* 21:435–457.

Raley, R Kelly and Megan M Sweeney. 2020. "Divorce, repartnejavascript:void(0)ring, and stepfamilies: A decade in review." *Journal of Marriage and Family* 82:81–99.

Rossi, Alice S and Peter Peter Henry Rossi. 1990. *Of human bonding: Parent child relations across the life course*. Transaction Publishers.

Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology* 66:688.

Seltzer, Judith A. 2019. "Family change and changing family demography." *Demography* 56:405–426.

Seltzer, Judith A and Suzanne M Bianchi. 2013. "Demographic change and parent-child relationships in adulthood." *Annual review of sociology* 39:275–290.

Sethi, Rajiv. 1996. "Evolutionary stability and social norms." *Journal of economic behavior & organization* 29:113–140.

Sewell, William H, Archibald O Haller, and Alejandro Portes. 1969. "The educational and early occupational attainment process." *American sociological review* pp. 82–92.

Sharkey, Patrick and Felix Elwert. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." *American journal of sociology* 116:1934–81.

Sheppard, Paula and Christiaan Monden. 2018. "The Additive Advantage of Having Educated Grandfathers for Children's Education: Evidence From a Cross-national Sample in Europe." *European Sociological Review* 34:365–380.

Silverman, Phyllis R et al. 2000. *Never too young to know: Death in children's lives*. Oxford University Press, USA.

Silverstein, Merril and Vern L Bengtson. 1997. "Intergenerational solidarity and the structure of adult child-parent relationships in American families." *American journal of Sociology* 103:429–60.

Silverstein, Merril and Anne Marenco. 2001. "How Americans enact the grandparent role across the family life course." *Journal of Family Issues* 22:493–522.

Simon, Nicole S and Susan Moore Johnson. 2015. "Teacher Turnover in High-poverty Schools: What we Know and Can do." *Teachers College Record* 117:1–36.

Sjölander, Arvid and Johan Zetterqvist. 2017. "Confounders, Mediators, or Colliders." *Epidemiology* 28:540–547.

Sobel, Michael E. 2012. "Does Marriage Boost Men's Wages?: Identification of Treatment Effects in Fixed Effects Regression Models for Panel Data." *Journal of the American Statistical Association* 107:521–529.

Solon, Gary. 2018. "What do we know so far about multigenerational mobility?" *The Economic Journal* 128:F340–F352.

Song, Xi. 2016. "Diverging Mobility Trajectories: Grandparent Effects on Educational Attainment in one-and two-parent Families in the United States." *Demography* 53:1905–1932.

Song, Xi and Cameron D Campbell. 2017. "Genealogical Microdata and Their Significance for Social Science." *Annual Review of Sociology* 43:75–99.

Song, Xi, Cameron D Campbell, and James Z Lee. 2015. "Ancestry Matters: Patrilineage Growth and Extinction." *American Sociological Review* 80:574–602.

Song, Xi and Robert D Mare. 2019. "Shared Lifetimes, Multigenerational Exposure, and Educational Mobility." *Demography* pp. 1–26.

Swartz, Teresa Toguchi. 2008. "Family capital and the invisible transfer of privilege: Intergenerational support and social class in early adulthood." *New directions for child and adolescent development* 2008:11–24.

Swartz, Teresa Toguchi. 2009. "Intergenerational family relations in adulthood: Patterns, variations, and implications in the contemporary United States." *Annual review of sociology* 35:191–212.

Sweeney, Megan M. 2010. "Remarriage and stepfamilies: Strategic sites for family scholarship in the 21st century." *Journal of Marriage and Family* 72:667–684.

Tanskanen, Antti Olavi, Mirkka Danielsbacka, and Jani Erola. 2016. "Educational test scores among adolescents in three-generational households in 20 countries." *Finnish Yearbook of Population Research* 51:3–22.

Thorsen, Maggie L and Valarie King. 2016. "My mother's husband: Factors associated with how adolescents label their stepfathers." *Journal of social and personal relationships* 33:835–851.

Todd, Petra E and Kenneth I Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal* 113:F3–F33.

Trivers, Robert. 1972. "Parental investment and sexual selection." *Sexual Selection & the Descent of Man, Aldine de Gruyter, New York* pp. 136–179.

Troilo, Jessica and Marilyn Coleman. 2008. "College student perceptions of the content of father stereotypes." *Journal of Marriage and Family* 70:218–227.

Uhlenberg, Peter and Bradley G Hammill. 1998. "Frequency of grandparent contact with grandchild sets: Six factors that make a difference." *The Gerontologist* 38:276–285.

Van Der Pas, Suzan, Theo G Van Tilburg, and Merril Silverstein. 2013. "Stepfamilies in later life." *Journal of Marriage and Family* 75:1065–1069.

Van Gaalen, Ruben I and Pearl A Dykstra. 2006. "Solidarity and conflict between adult children and parents: A latent class analysis." *Journal of Marriage and Family* 68:947–960.

Vierboom, Yana C, Samuel H Preston, and Arun S Hendi. 2019. "Rising geographic inequality in mortality in the United States." *SSM-population health* 9:100478.

Visher, Emily B and John S Visher. 2013. *How to win as a stepfamily*. Routledge.

Voss, Thomas. 2001. *Game-theoretical perspectives on the emergence of social norms*. na.

Wiemers, Emily E, Judith A Seltzer, Robert F Schoeni, V Joseph Hotz, and Suzanne M Bianchi. 2019. "Stepfamily structure and transfers between generations in US families." *Demography* 56:229–260.

Wightman, Patrick and Sheldon Danziger. 2014. "Multi-generational income disadvantage and the educational attainment of young adults." *Research in Social Stratification and Mobility* 35:53–69.

Wolff, Edward N and Maury Gittleman. 2014. "Inheritances and the Distribution of Wealth or Whatever Happened to the Great Inheritance Boom?" *The Journal of Economic Inequality* 12:439–468.

Wooldridge, JM. 2004. "Fixed effects estimation of the population-averaged slopes in a panel data random coefficient model." *Econometric Theory* 20:428–429.

Wooldridge, Jeffrey M. 2005. "Fixed-effects and Related Estimators for Correlated Random-coefficient and Treatment-effect Panel Data Models." *Review of Economics and Statistics* 87:385–390.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT press.

Xie, Yu. 2013. "Population heterogeneity and causal inference." *Proceedings of the National Academy of Sciences* 110:6262–6268.

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