THE UNIVERSITY OF CHICAGO

RELATIONSHIP QUALITY, FAMILY STRUCTURE, AND CHILD OUTCOMES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY

DAVID J. MALISON

CHICAGO, ILLINOIS

AUGUST 2018

To Michelle, for always believing in my best self.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my advisers Stéphane Bonhomme, Alessandra Voena, Michael Dinerstein for their invaluable support and guidance. I am particularly indebted to Stéphane for his insightful advice, honest criticism, and persistent encouragement. I also thank my University of Georgia mentors David Mustard and Christopher Cornwell for inspiring me to pursue a career in economics. Jorge Luis García, Nina Nguyen, Melissa Tartari and seminar participants at the University of Chicago's Crossing Disciplinary Boundaries working group, the University of Chicago's Econometrics working group, and the University of Michigan's H2D2 Research Day provided helpful feedback for which I am grateful. I also gratefully acknowledge the support of the University of Chicago Department of Economics.

I thank Mohsen Mirtaher for his enduring friendship throughout graduate school. I also thank my parents for investing in my future and giving me so many of the tools I have needed to succeed. Without their support, none of this would have been possible. Lastly, I thank my wife Michelle for helping me find the courage to pursue my dreams.

ABSTRACT

This dissertation explores heterogeneity in how family structure affects children. Detailed measurements of parental interactions are exploited to estimate a dynamic economic model of parental relationship quality, parental decisions to continue their relationship, and child human capital development. The empirical specification takes into account measurement error, controls for observed demographics and initial conditions, and integrates out missing data. Parental separation is shown to have little effect on cognitive skill development during childhood, but is found to exert considerable influence on the development of non-cognitive skills such as the ability to control aggression. Parental separation's effect on non-cognitive skill development largely depends on parental relationship quality; in particular, children whose parents have a poor quality relationship on average benefit if their parents separate. If separated parents in the sample had instead chosen to stay together, their children would have on average been 21 percent more likely to receive special services because of Attention Deficit Disorder (ADD) and 20 percent more likely to have been suspended or expelled. Almost two-thirds of the difference in non-cognitive skills between children of separated and non-separated parents at age 9 can be explained by differences in parental relationship quality, while only 16 percent can be explained by differences in observable demographics.

CHAPTER 1

INTRODUCTION

The share of American children growing up in single-parent households has increased substantially over the past half-century. In 1960, less than 10 percent of births in the United States were to unmarried mothers and almost 90 percent of US children under the age of 18 lived with two parents (Ventura and Bachrach (2000), United States Census Bureau (2018)). By 2015, the non-marital birth rate had quadrupled to 40 percent while the fraction of children under the age of 18 living with two parents had declined to 70 percent (Martin et al. (2017), United States Census Bureau (2018)).

The increasing prevalence of single-parenthood, particularly in low-income communities, has attracted considerable attention from policymakers across the political spectrum (Lundquist et al. (2014), Solomon-Fears (2015)). This interest has largely been driven by concern for how family structure affects children, as several studies have shown that children who grow up with two married, biological parents have better cognitive, behavioral, and educational outcomes than children growing up in other living arrangements (Härkönen et al. (2017), McLanahan et al. (2013), Amato (2010), Brown (2010), McLanahan and Sandefur (1994)). Family institutions may therefore play an important role in the intergenerational transmission of inequality.

Although the extensive research on family structure and child development has helped inform policy discussions, papers in this literature have tended to overlook heterogeneity in how parental separations affect children. In particular, separating parents who frequently argue likely impacts children differently than the dissolution of a warm and loving parental relationship. Most previous studies either ignore parental relationship quality entirely or implicitly treat relationship quality as an unobserved fixed effect and use differencing methods to control for selection bias (Amato and Anthony (2014), McLanahan et al. (2013), Amato (2010)).

There are two problems with the fixed effect approach. First, it is difficult to rationalize why parents separate if the quality of their relationship is fixed over time. Indeed, virtually all economic models of divorce in part rely on match quality shocks to generate separations (Browning et al. (2014)). Second, first-differencing fails to eliminate selection bias when the effect of separation is heterogeneous (see Section A in the Appendix).

Accounting for heterogeneous separation effects would be straight-forward if parental relationship quality were observed. Although relationship quality is difficult to quantify, this disseration addresses the measurement challenge by exploiting a unique data set that contains detailed longitudinal information about parental interactions. The measurements are used to estimate a latent parental relationship quality process. Test scores and child behaviors are also treated as noisy signals of latent child skills.

The data are used to estimate a dynamic economic model of parental relationship quality, parental decisions to continue their relationship, and child outcomes. In the model, parents separate because of shocks to their relationship quality and to the values of their outside options. Parental relationship quality also influences child development. Heterogeneous effects of parental separation are identified by comparing children of parents with similar relationship qualities and other demographics who make different separation choices. Identification thus comes from matching individuals on observable characteristics and the dynamic latent variables for which noisy measurements are available.

The latent variables enter non-linear choice and measurement equations, complicating estimation. This dissertation adopts a Bayesian estimation approach by imposing weakly informative priors on the model parameters and drawing from the model's posterior distribution using a Hamiltonian Monte Carlo (HMC) algorithm. The HMC algorithm permits flexible (e.g. non-conjugate) prior specification and converges quickly to a stationary distribution. As long as the prior is not dogmatic, the posterior means of the model parameters are consistent and asymptotically normal. If the prior distribution of the latent variables is well-specified, functions of the model parameters and latent variable are also consistently estimated. Even under misspecification, these estimators converge to the true values as the number of measurements grows (Arellano and Bonhomme (2009), Arellano and Bonhomme (2011)). Missing data are easily integrated out from the likelihood under a missing-at-random assumption.

This dissertation has four main results. First, parental separation and parental relationship quality have little affect on cognitive skill development in children. This result is consistent with other recent work that shows cognitive abilities to be less malleable than non-cognitive skills (Heckman et al. (2013)).

Second, parental separation has an important effect on non-cognitive skills like child aggression. However, the effect is largely determined by the quality of the parental relationship. In particular, separation improves behavioral outcomes on average when parents have a poor relationship.

Third, if parents in the data who chose to separate had instead stayed together, their children would on average have had worse behavioral outcomes. These children would on average have been 34 percent more likely to receive a Behavior Intervention Plan (BIP) in school, 21 percent more likely to receive special services because of Attention Deficit Disorder (ADD), 15 percent more likely to purposely damage property, 10 percent more likely to have a fist fight with another person, and 20 percent more likely to have been suspended or expelled. Thus the observed parental separations on average significantly improved child behavior.

Finally, variation in parental relationship quality is a large reason why children of separated parents have worse non-cognitive outcomes than children whose parents remain together. Almost two-thirds of the difference in non-cognitive skills between these groups of children can be explained by differences in parental relationship quality, while only 16 percent can be explained by observable demographics. By contrast, 80 percent of the difference in cognitive skills can be explained by demographics and almost none can be attributed to parental relationship quality.

The rest of this dissertation proceeds as follows. Chapter 2 reviews previous research on this topic. Chapter 3 contains a description of the longitudinal data set used in the analysis. Chapter 4 develops a formal dynamic economic model to rationalize the data-generating process. Chapter 5 details the estimation procedure, while Chapter 6 discusses the results. Chapter 7 concludes.

CHAPTER 2

RELATED LITERATURE

Developmental psychologists and sociologists have long studied the impact of parental separation on children. Härkönen et al. (2017), McLanahan et al. (2013), Amato (2010), and Ribar (2004) review this literature. Studies in this area generally find that children who grow up with two biological parents have better average outcomes than children growing up in other family structures. A related body of work in family psychology analyzes the effect of parental conflict and intermarital discord on child outcomes (Barthassat (2014), Pendry and Adam (2013), Baxter et al. (2011), Musick and Meier (2010), Amato and Sobolewski (2001), Jekielek (1998), Davies and Cummings (1994), Cherlin et al. (1991), Grych and Fincham (1990), Emery (1982)). These papers typically find parental relationship quality to be positively associated with child well-being.

Many of these studies proceed by regressing a child outcome on indicators of family structure or inter-parental conflict while controlling for observed demographic variables. Some of the more recent papers use lagged outcomes or fixed effects to control for initial conditions and unobserved heterogeneity (e.g. Arkes (2015), Amato and Anthony (2014), Magnuson and Berger (2009)). The model in this dissertation extends this literature by accounting for the interaction between parental relationship quality and family structure during child development. The estimation procedure also uses a coherent statistical framework to correct for measurement error and missing data, problems largely ignored by the psychology and sociology literature.

Several economists have also made contributions in this area. Using French employment survey data, Piketty (2003) found that family structure has very little correlation with children's educational outcomes after controlling for pre-separation outcomes. His conclusions are consistent with the results reported here. Gruber (2004) found adults born after the introduction of unilateral divorce laws in the United States marry earlier, divorce more frequently, have slightly lower educational attainment, and are more likely to commit suicide. González and Viitanen (2018) found analogous results using changes in divorce legalization across Western Europe. None of these papers attempt to measure parental relationship quality.

Tartari (2015) is the most closely related work in the economics literature. In her model, marital conflict is a binary state that occurs stochastically. Parents can expend effort to decrease the probability of conflict occurring in a given period, and marital conflict affects child development by entering a child quality production function. She estimates her model via indirect inference on data from the National Longitudinal Survey of Youth (NLSY79). In counter-factual experiments, she finds that parents would invest more in their children and exert more effort to avoid conflict if they were not allowed to divorce.

The analysis in this dissertation differs from Tatari's work in three key respects. First, this dissertation analyzes the effect of parental separation on non-cognitive as well as cognitive development. The results indicate that parental interactions have a much stronger effect on behavioral outcomes than on test scores. Second, the child production function is estimated directly rather than via indirect inference. This empirical approach more transparently connects the model to the data. However, the estimation procedure does not recover preference parameters needed to conduct policy experiments. Third, the model in this dissertation more closely resembles other theoretical and empirical models of divorce (e.g. Voena (2015), Browning et al. (2014)) because it treats parental relationship quality as a latent continuous variable rather than a binary state.

This dissertation is also related to the literature on dynamic child skill production functions (e.g. Del Boca et al. (2014), Cunha et al. (2010), Cunha and Heckman (2008)). These papers tend to focus on how child development responds to time and good investments, while the work presented here assesses the impact of parental relationship quality and family structure. This dissertation thus broadens the set of inputs into the skill production process.

CHAPTER 3

DATA

This dissertation analyzes data collected from the Fragile Families and Child Wellbeing Study (FFCWS) (Reichman et al. (2001)), a survey that followed a cohort of 4,898 children born in 75 hospitals across the United States between 1998 and 2000. The parents of participating children were interviewed in person shortly after the participating child's birth and subsequently over the phone when the child reached ages 1, 3, and 5. Researchers also conducted home visits when the child was 3, 5, 9 and 15 years old.

Although studied extensively by psychologists, sociologists and demographers, this dataset has received relatively little attention from economists. The data were featured in a few economics articles about maternal and infant health (e.g. Currie et al. (2015), Carroll et al. (2007)). Aizer and McLanahan (2006) used these data to analyze how child support enforcement influences fertility and marriage patterns, while Liu and Heiland (2012) used propensity score matching to assess the impact of marriage on children born out-of-wedlock. Fletcher (2016) used some of the parental interaction measurements discussed below to study how maternal investment in children responds to changes in parental relationship quality.

3.1 Demographics

As its title suggests, the Fragile Families and Child Wellbeing Study focuses on children born into unstable families. Non-marital births in urban areas were over-sampled, since previous work indicated such families were most at risk of future disruption (Reichman et al. (2001)). 648 observations (13 percent of the original sample) were not included in this analysis because they were missing baseline demographic information. The first column of Table 3.1 displays summary statis-

Table 3.1: Summary statistics

Variable	Sample	2000 Census
Maternal Demographics		
Highest level of education completed		
Less than high school grad	.332	.203
High school grad/GED	.306	.250
Some college/technical training	.249	.396
College degree	.113	.251
Race/Ethnicity		
Non-Hispanic Black	.478	.128
Hispanic	.264	.184
Non-Hispanic White or other	.258	.688
Age	25.2 (6.06)	28.5 (6.25)
Paternal Demographics		
Highest level of education completed		
Less than high school	.321	.184
High school/GED	.360	.253
Some college/technical training	.213	.280
College degree	.106	.283
Race/Ethnicity		
Non-Hispanic Black	.496	.087
Hispanic	.270	.184
Non-Hispanic White or other	.234	.729
Parental Relationship		
How long mother knew father before pregnancy (years)	4.85 (4.58)	
Do parents have other children together?	.359	
Do parents have other children with different partners?	.512	
Other		
Child gender (1 if female)	.476	.490
Number of Observations	4,250	190,378

First column reports means and standard deviations (in parentheses) of the sample used in analysis. Variables recorded at the time of the child's birth. Second column reports statistics of parents of newborn children (less than 1 year old) in the 2000 Census 5-Percent PUMS (calculated using IPUMS's person weights (Ruggles et al. (2010)).

tics for the remaining 4,250 cases used in the analysis. For comparison, the second column of Table 3.1 displays statistics of parents of newborn children in the 2000 Census 5-Percent Public Use Microdata Sample (PUMS) (Ruggles et al. (2010)).

At the time of the baseline interview, about one third of parents in the sample had less than a high school degree, one third had a high school degree or GED, and one third had attended at least some college. By contrast, more than half of parents of newborn children in the 2000 Census had attended some college. Approximately one-half of the sample were non-Hispanic Black and a quarter were Hispanic. By contrast, approximately two-thirds of mothers in the 2000 Census sample were non-Hispanic White. The average age of mothers in the sample at the time of the child's birth was 25.2, three and a half years younger than the average age of mothers in the 2000 Census.

The three variables under the "Parental Relationship" heading in Table 3.1 are unavailable in the Census data, but were included in the analysis as additional controls. Parents in the sample had on average known each other for a little less than 5 years prior to the pregnancy.[1] At the time of the baseline interview, about one third of parents had another child together and about half had children with different partners.

Although not nationally representative of the US population as a whole, the sample is particularly interesting for studying the causes and effects of parental separation. Children of minority and less educated mothers are particularly likely to grow up without the presence of a father in the home, and such children also fare worse in school and as adults (McLanahan and Sandefur (1994)). By providing access to a large and detailed sample of disadvantaged parents, the Fragile Families and Child Wellbeing Study presents a unique opportunity to assess how much growing up in a broken home is only correlated with poverty rather than a critical mechanism in its inter-generational transmission.

[1]This variable was top-coded at the age of the mother minus 5 to limit the influence of outliers.

Table 3.2: Parental relationship dynamics

	Baseline	Year 1	Year 3	Year 5	Year 9	Year 15
$R_t = 1$	3,782 (89.0%)	2,890 (68.0%)	2,268 (53.4%)	1,770 (41.7%)	1,220 (28.7%)	814 (19.1%)
$R_t = 0$	468 (11.0%)	1,360 (32.0%)	1,928 (45.4%)	2,350 (55.3%)	2,625 (62.2%)	2,961 (69.6%)
R_t missing			54 (1.3%)	150 (3.1%)	385 (9.1%)	475 (11.2%)
Annualized exit hazard		23.6%	10.6%	10.1%	5.3%	5.3%

Each column corresponds to a survey wave. R_t denotes an indicator for whether the parents were in a continuous relationship from the child's birth until wave t. See text for details on how R_t was constructed.

3.2 Family structure dynamics

Let R_t denote an indicator for whether parents have been in a relationship from the birth of their child until start of period $t = 0, 1, \cdots, 5$, where t indexes survey waves. These indicators do not distinguish between different types of parental relationships (e.g. cohabitation versus marriage) and separations (e.g. break-ups versus divorces), even though such distinctions are made in the data. Although these distinctions likely contain information about the quality of the parental relationship (Brien et al. (2006)), modelling them is beyond the scope of this dissertation. This dissertation also only considers the first observed parental separation, and does not analyze the parents' relationships with other partners. Thus $R_t = 0$ implies $R_{t+s} = 0$ for all $s > 0$. The estimated effects of separation should therefore be interpreted as averaging over all possible outcomes after an initial parental separation, including remaining single, returning to a relationship, or entering into relationships with other partners.

Table 3.2 tabulates the parental relationship indicators for each survey wave. When the children in the study were born, almost 90 percent of their parents were in a relationship. However, almost a quarter of these relationships ended before the child's first birthday. By the time the children reached age 15, only 20% of parents still in the sample were in a continuous relationship since the start of the study. The annualized exit hazards, displayed in the last row of Table 3.2, decreased significantly over the course of the survey, consistent with negative duration dependence and selection out of unhealthy relationships.

3.3 Parental relationship quality

Economic models of divorce typically rely on a stochastic relationship quality process (or stochastic signals about unobserved relationship quality) to explain separations (Browning et al. (2014)). While the relationship quality shocks in these models are usually presumed to be unobservable, the Fragile Families and Child Wellbeing Study contains detailed longitudinal measurements of parental interactions. These measurements permit an empirical examination of relationship quality dynamics, providing a deeper understanding of the forces governing family structure.

Each survey wave, parents in the study answered a series of questions about the quality of their relationship. The analysis in this dissertation uses data from maternal interviews.[2] Table B.1 in the Appendix lists the full set of relationship quality survey items used in the analysis.

Figure 3.1 illustrates dynamic variation in relationship quality by displaying average responses to four survey items over time. Parents in Figure 3.1 are grouped by when they first separated during the study. Mothers who did not separate were consistently more likely to report that the child's father compromised during disagreements, was affectionate and supportive, and was not critical or insulting. Positive responses declined for all groups after the first period, suggesting a shared time effect for all parents. Positive responses also declined prior to separation, suggesting that a measurable deterioration in relationship quality predicts separation in the future.

Table 3.3 further illustrates variation in relationship quality by cross-tabulating responses to the survey item "In general, would you say that your relationship with the child's father is excellent, very good, good, fair, or poor?" with an indicator for whether the parents separate before the

[2]Paternal interviews had significantly higher non-response rates.

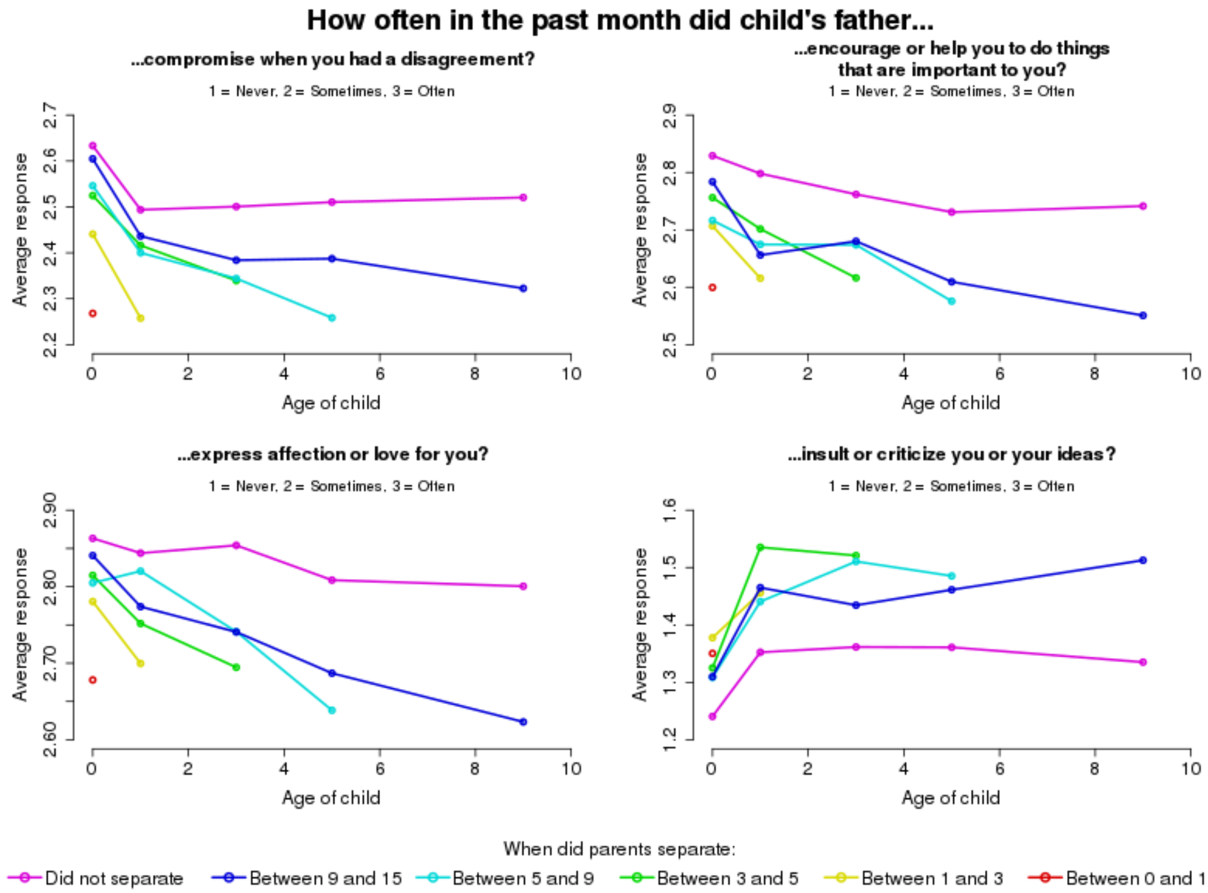


Figure 3.1: Longitudinal measurements of relationship quality

next survey wave. Although parental relationship quality is negatively correlated with separation, the correlation is far from perfect: fifteen percent of parents who reported very good or excellent relationships still separated before the next survey wave and two-thirds of parents who reported poor, fair, or good relationships did not separate. It is therefore possible to compare outcomes for parents who had similar quality relationships but who made different separation choices. The identification strategy essentially exploits these types of comparisons to infer how parental separation and relationship quality interact during child development.

3.4 Child skills

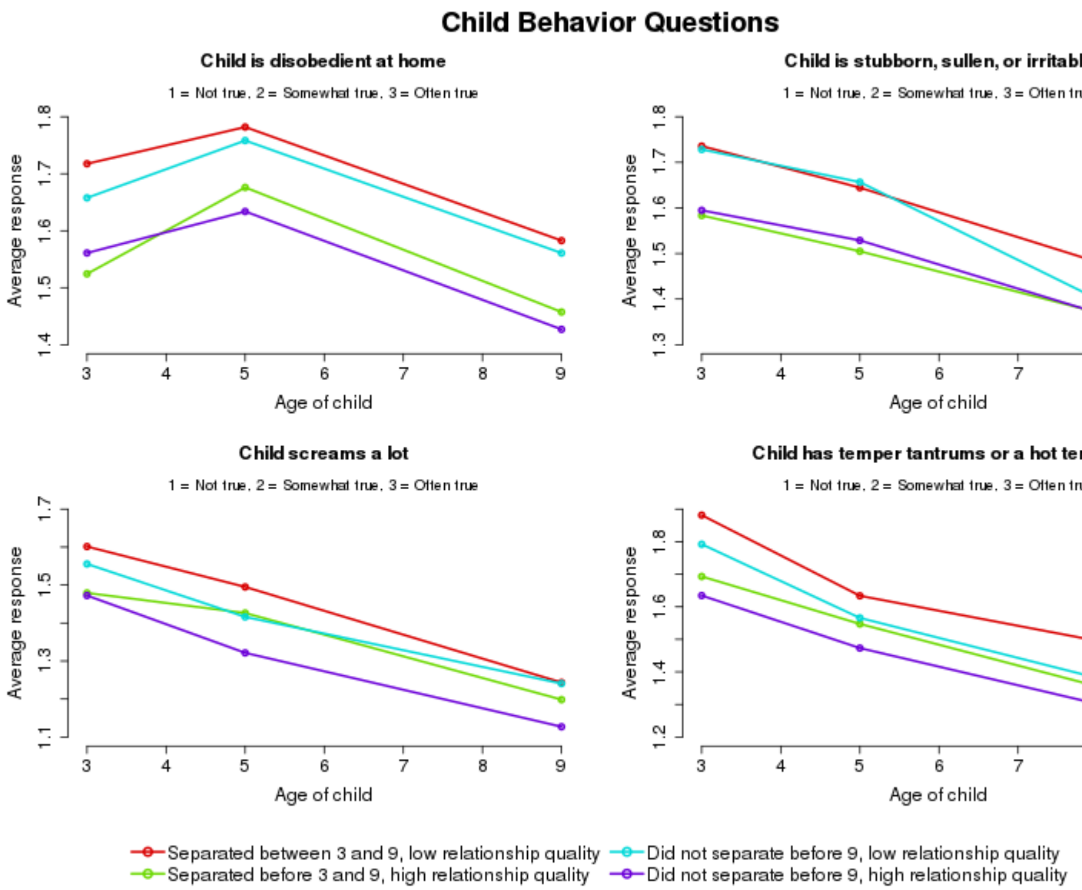
Table 3.3: Cross-tabulation between self-reported relationship quality and separation

In general, would you say that your relationship with the child's father is excellent, very good, good, fair, or poor? (Row frequencies reported in parenthesis.)

	Stayed together until next survey wave	Separated before next survey wave	Total
Poor	.004 *(0.408)*	.007 *(0.592)*	.011
Fair	.037 *(0.590)*	.026 *(0.410)*	.063
Good	.136 *(0.705)*	.057 *(0.295)*	.193
Very Good	.312 *(0.815)*	.071 *(0.185)*	.383
Excellent	.308 *(0.879)*	.042 *(0.121)*	.350
Total	.798	.202	

Economists have recently begun to appreciate the multidimensional nature of human capital development during childhood (Heckman and Mosso (2014), Almlund et al. (2011)). Cognitive skills, such as the ability to think abstractly, reason analytically, or recall information efficiently, develop early in a child's life. Non-cognitive skills like perseverance and self-control develop later and are more responsive to interventions (Cunha et al. (2006), Heckman et al. (2013)). Several studies have shown that the development of one type of skill fosters subsequent development of the other and that both cognitive and non-cognitive skills are important determinants of adult outcomes (Cunha et al. (2010), Cunha and Heckman (2008), Cunha and Heckman (2007)).

In contrast to previous research that has focused on the impact of parental time and goods investment, this dissertation emphasizes how parental relationship quality affects child development. The results reported here suggest that finding a high quality match may be an important avenue

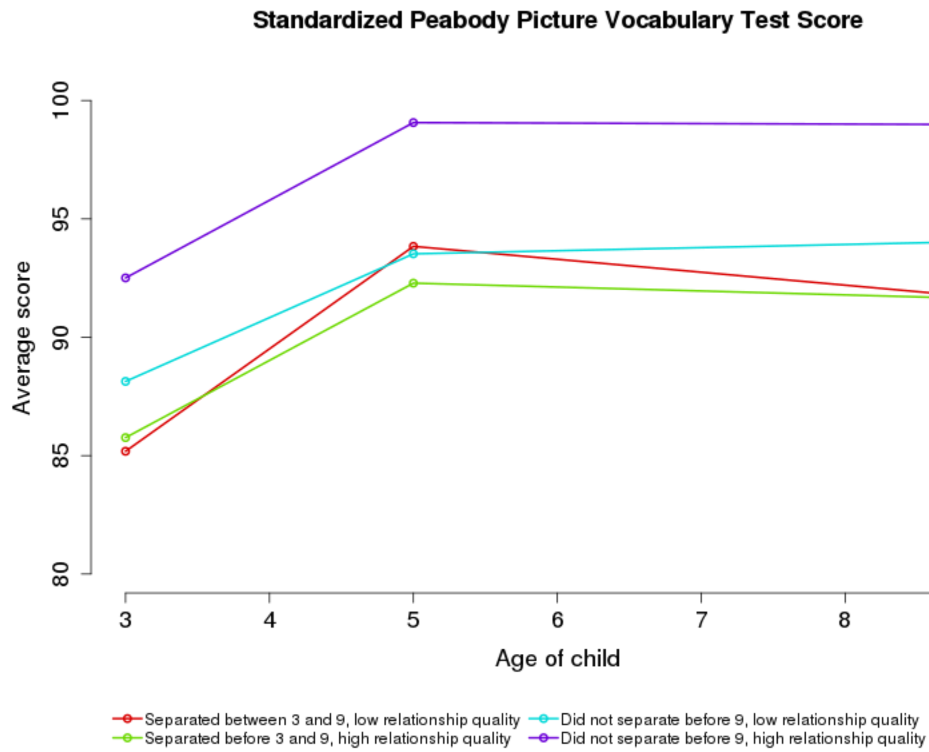


Parents who averaged < 4 (≥ 4) on a 5-point self-reported relationship quality scale (see Table 3.3) were categorized as being in a low (high) quality relationship.

Figure 3.2: Longitudinal measurements of non-cognitive skill

through which parents invest in their children.

In this dissertation, non-cognitive skill can be interpreted as the ability to control aggression and follow instruction. This skill was measured by the reported frequency of several disruptive behaviors listed in Table B.2 in the Appendix. The survey items at ages 3, 5, and 9 corresponded to the aggressive subscale of the Child Behavior Checklist (CBCL), a popular child psychology assessment (Achenbach (1991); Achenbach and Rescorla (2001)). Previous research has found these behaviors to be associated with several outcomes later in life, including educational attainment, labor force attachment, substance use, and the likelihood of committing a criminal offense



Parents who averaged < 4 (≥ 4) on a 5-point self-reported relationship quality scale (see Table 3.3) were categorized as being in a low (high) quality relationship.

Figure 3.3: Longitudinal measurement of cognitive skill

(Washbrook et al. (2013)).

The average responses to four non-cognitive skill measurements are displayed in Figure 3.2. Children in Figure 3.2 are grouped by when their parents separated and by average self-reported parental relationship quality. Parents in both low- and high-quality relationships reported somewhat similar levels of problematic behavior, regardless of their separation decisions. The figure therefore suggests that parental relationship quality explains some of the behavioral differences between children of separated and non-separated parents.

Figures 3.1 and 3.2 also illustrate how multiple survey items capture common trends in the data. The latent variable model discussed in Section 5 summarizes these trends by placing the panoply of available measurements into an organized, low-dimensional framework.

The child's cognitive skills were measured by several standardized tests listed in Table B.3 in the Appendix. Figure 3.3 displays average standardized scores from the Peabody Picture Vocabulary Test (PPVT), a frequently-used assessment of verbal ability and scholastic aptitude. Parents in Figure 3.3 are divided into the same categories displayed in Figure 3.2. The test scores only appear correlated with parental relationship quality if the parents did not separate. This pattern could indicate an interaction between relationship quality and parental separation.

Figures 3.2 and 3.3 also demonstrate that differences in child behavior across groups of parents persist over time. These figures therefore indicate the importance of accounting for initial conditions when assessing the causal impact of separation and relationship quality on child development.

3.5 Outcomes

In addition to the skill measurements discussed above, the data also contain information about several outcomes of interest. A major goal of the subsequent analysis is to predict how these outcomes would have changed had parents made different separation decisions.

Outcomes measured at age 9 include teacher ratings of the child's academic abilities, whether or not the child received different special education services, and the child's responses to a series of questions about early delinquency. Table C.1 in the Appendix compares age 9 outcomes between children whose parents stayed together through age 9 and children whose parents separated prior to age 9. The children of separated parents had worse academic ratings and were twice as likely to receive various special education services. These children were also more than twice as likely to have been suspended or expelled from school and seventy-five percent more likely to have had a fist fight with another person.

Outcomes measured at age 15 covered topics like grades, involvement with the criminal justice system, and substance use. Table C.2 in the Appendix performs a comparison analogous to the comparison in Table C.1. On average, children of separated parents scored almost half a letter grade lower across different subjects than children whose parents stayed together. These children were also significantly more likely to report participating in vandalism, more likely to have been arrested, and more likely to have used alcohol, tobacco and marijuana.

CHAPTER 4

MODEL

This section develops a dynamic economic model of parental relationship quality, family structure, and child human capital development. The model serves two purposes. First, it is used to derive equations for estimation. The model thus ties the empirical results in Section 6 to an explicit economic framework. Second, the model provides insight into the tradeoffs parents face when deciding whether or not to continue their relationship. In so doing, the model helps organize thinking about family structure dynamics.

4.1 Setup and definitions

The model takes place in discrete time, with periods indexed by t from 0 to $T+1$. A child is born at the end of period 0 and leaves the house at the end of period $T+1$. The child's parents begin each period either separated or in a relationship. As in Section 3.3, R_t denotes an indicator for whether parents have been in a relationship from child's birth until start of period t. The only endogenous variable in the model is the parents' binary decision to stay together or separate each period. The sequence of parental choices can be considered the solution to an optimal stopping time problem.

Children are characterized each period by a scalar human capital level θ_t^H (vector-valued human capital is discussed in Section 4.6). The parents' relationship is characterized by a scalar relationship quality θ_t^R. If $R_t = 0$, θ_t^R is set equal to zero.

4.2 Laws of motion

The parents start the model in a relationship with an initial relationship quality draw θ_0^R. Relationship quality then evolves according to a first-order autoregressive process (AR(1)) as long as the parents remain in a relationship:

$$\theta_{t+1}^R = \begin{cases} \alpha_{t+1}^R + \gamma_{t+1}^R \cdot \theta_t^R + \epsilon_{t+1}^R & \text{if } R_{t+1} = 1 \\ 0 & \text{if } R_{t+1} = 0 \end{cases} \tag{4.1}$$

α_{t+1}^R and γ_{t+1}^R are fixed parameters and ϵ_{t+1}^R is a stochastic shock.

The child's human capital sequence starts with an initial draw θ_0^H, which may be correlated with θ_0^R. θ_t^H then evolves according to

$$\theta_{t+1}^H = \begin{cases} \alpha_{t+1}^H + \gamma_{t+1}^{H,H} \cdot \theta_t^H + \epsilon_{t+1}^H & \text{if } R_t = 0 \\ \alpha_{t+1}^H + \beta_{t+1}^H + (\delta_{t+1}^H + \xi_{t+1}^H \cdot \theta_t^H) \cdot R_{t+1} + \gamma_{t+1}^{H,H} \cdot \theta_t^H + \gamma_{t+1}^{H,R} \cdot \theta_t^R + \epsilon_{t+1}^H & \text{if } R_t = 1 \end{cases} \tag{4.2}$$

where the intercept and coefficients are fixed parameters and ϵ_{t+1}^H is a stochastic term. According to (4.2), child human capital evolves as an AR(1) when $R_t = 0$. If $R_t = 1$, the evolution of the child's human capital depends on the parents' relationship quality and on the parents' decision to separate or stay together. The β_{t+1} and δ_{t+1} terms capture effects from parents starting and ending period t in a relationship, respectively, while $\xi_{1,t+1}^H$ is an interaction coefficient that allows parental relationship quality to influence the effect of separation (and vice versa).

4.3 Preferences

If the parents separate prior to the end of period t, the mother and father receive "singles" payoffs of S_t^M and S_t^F in period t, respectively. These payoffs are determined by the expressions

$$\begin{aligned} S_t^M &= \rho \cdot g_t^S(\theta_t^H) + \alpha_t^M + \eta_t^M \\ S_t^F &= (1-\rho) \cdot g_t^S(\theta_t^H) + \alpha_t^F + \eta_t^F \end{aligned} \tag{4.3}$$

where $g_t^S(\theta_t^H)$ denotes total transferable utility when parents are separated, $\rho \in [0,1]$ denotes the fraction of $g_t^S(\theta_t^H)$ that the mother receives, α_t^M and α_t^F are parameters, and η_t^M and η_t^F are stochastic. Parental relationship quality does not enter the singles payoffs. Parents determine ρ once and for all before the beginning of period 0.

If parents remain together through the end of period t, then mother and father receive "union" payoffs of U_t^M and U_t^F in period t, respectively. The total union payoff $U_t = U_t^M + U_t^F$ is given by

$$U_t = g_t^U(\theta_t^R, \theta_t^H)$$

The mother receives $(\rho \times 100)\%$ of the period t surplus when the parents are together. The period t union payoffs are therefore given by

$$\begin{aligned} U_t^M &= S_t^M + \rho \cdot (U_t - S_t) \\ U_t^F &= S_t^F + (1-\rho) \cdot (U_t - S_t) \end{aligned} \tag{4.4}$$

where $S_t = S_t^M + S_t^F$ represents the parents' total payoffs as singles. Parents discount future payoffs at rate $r \in (0,1)$.

4.4 Information and timing

Define $\boldsymbol{\theta}_t = (\theta_t^R, \theta_t^H)'$, $\boldsymbol{\epsilon}_t = (\epsilon_t^R, \epsilon_t^H)'$, and $\boldsymbol{\eta}_t = (\eta_t^M, \eta_t^F)'$. The vector of shocks $(\boldsymbol{\epsilon}_t, \boldsymbol{\eta}_t)$ is mean zero and drawn independently over time and independent of the vector of initial values $\boldsymbol{\theta}_0$. The law of motion shocks $\boldsymbol{\epsilon}_t$ are also drawn independently of the outside option shocks $\boldsymbol{\eta}_t$:

$$\boldsymbol{\epsilon}_t \perp\!\!\!\perp \boldsymbol{\eta}_t$$

Parents know the joint distributions of the shocks but do not know future realizations. Parents also know all non-stochastic functions and parameters of the model. The decision to continue the relationship in period $t+1$ is made after period t shocks have been revealed but before period $t+1$ shocks are known.

4.5 Solution

The period t state vector is given by $\varsigma_t = (\boldsymbol{\theta}_t, \boldsymbol{\eta}_t, R_t)$. Let $R_{t+1}(\varsigma_t)$ denote the period t policy function, which specifies the optimal choice of R_{t+1} given any possible realization of ς_t. Solving the model consists of constructing the policy function sequence $\{R_{t+1}(\varsigma_t)\}_{t=0}^{T}$. Following standard practice in finite-time dynamic programming problems, the sequence is derived recursively by first computing $R_{T+1}(\varsigma_T)$ and then working backwards.

4.5.1 Period *T* problem

It is convenient to define $\alpha_t^S = \alpha_t^M + \alpha_t^F$ and $\eta_t^S = \eta_t^M + \eta_t^F$ for $t = 0, 1, \cdots, T+1$. If $R_T = 1$, both parents are better off staying together through period T if their union utilities exceed their payoffs as singles:

$$
\begin{aligned}
U_T^M &> S_T^M \\
U_T^F &> S_T^F
\end{aligned}
\tag{4.5}
$$

From (4.4), we see that (4.5) holds if and only if $U_T > S_T$. The period T policy function is therefore given by

$$
\begin{aligned}
R_{T+1}(\varsigma_T) &= R_T \cdot \mathbb{1}\left\{U_T - S_T > 0\right\} \\
&= R_T \cdot \mathbb{1}\left\{p_T(\boldsymbol{\theta}_T) > \eta_T^S\right\}
\end{aligned}
\tag{4.6}
$$

where $\mathbb{1}\{\cdot\}$ is an indicator function and

$$
p_T(\boldsymbol{\theta}_T) = g_T^U(\boldsymbol{\theta}_T) - \alpha_T^S - g_T^S\left(\theta_T^H\right) \tag{4.7}
$$

Thus parents choose to stay together in period T if and only if their relationship has positive surplus in the last period. The $p_T(\cdot)$ function is a choice index that can be estimated from data on parental separations.

The parents' period T value functions can be written as

$$
\begin{aligned}
V_T^M(\varsigma_T) &= S_T^M + R_T \cdot \rho \cdot \max\{U_T - S_T, 0\} \\
V_T^F(\varsigma_T) &= S_T^F + R_T \cdot (1-\rho) \cdot \max\{U_T - S_T, 0\}
\end{aligned}
$$

Each value function is the sum of an outside option and a share of an option value for continuing the relationship. Define the net value function as

$$
\begin{aligned}
V_T(\varsigma_T) &= V_T^M(\varsigma_T) + V_T^F(\varsigma_T) - \alpha_T^S - \eta_T^S \\
&= g_T^S(\theta_t^H) + R_T \cdot \max\{U_T - S_T, 0\}
\end{aligned}
\tag{4.8}
$$

$V_T(\varsigma_T)$ equals the sum of the parents' period T value functions net of period T opportunity costs.

Each parent's value function can be written as

$$\begin{aligned} V_T^M(\varsigma_T) &= \alpha_T^M + \eta_T^M + \rho \cdot V_T(\varsigma_T) \\ V_T^F(\varsigma_T) &= \alpha_T^F + \eta_T^F + (1-\rho) \cdot V_T(\varsigma_T) \end{aligned} \tag{4.9}$$

This formulation facilitates solving the model by backwards induction.

4.5.2 Period *t* problem, *t* = 0, 1, ..., *T* - 1

Before proceeding, it is useful to write the laws of motion as

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, R_t, R_{t+1}, \boldsymbol{\epsilon}_{t+1})$$

where $\boldsymbol{\theta}_{t+1}(\cdot)$ is a vector-valued function representing the rules in (4.1) and (4.2). As a simplification, assume the utility functions $g_t^S(\theta_t^H)$ and $g_t^U(\boldsymbol{\theta}_t)$ can be written as

$$\begin{aligned} g_t^S(\theta_t^H) &= c_t\theta_t^H \\ g_t^U(\boldsymbol{\theta}_t) &= \phi_t(\theta_t^R) + c_t\theta_t^H \end{aligned}$$

where $c_t > 0$ and $\phi_t(\cdot)$ is continuously differentiable and strictly increasing. In this case, the surplus of the parents' relationship only depends on the parents' relationship quality and not on the child's human capital. An analysis of the more general case is considered in Section D in the Appendix.

Plugging into (4.8), we obtain

$$V_T(\varsigma_T) = c_T\theta_T^H + R_T \cdot \max\left\{\phi_T(\theta_T^R) - \alpha_T^S - \eta_T^S, 0\right\} \tag{4.10}$$

Now suppose $V_{t+1}(\varsigma_{t+1})$ has the form

$$V_{t+1}(\varsigma_{t+1}) = a_{t+1} + b_{t+1} \cdot \theta^H_{t+1} + R_{t+1} \cdot h_{t+1}(\theta^R_{t+1}, \eta^S_{t+1}) \tag{4.11}$$

where a_{t+1} and b_{t+1} are known constants and $h_{t+1}(\cdot)$ is a known function. It is apparent from (4.10) that (4.11) is satisfied for $t = T-1$. By using the law of motions in (4.1) and (4.2), it is possible to show

$$V_t(\varsigma_t) = a_t + b_t\theta^H_t + R_t \cdot h_t(\theta^R_t, \eta^S_t) \tag{4.12}$$

where

$$\begin{aligned}
a_t &= r \cdot (a_{t+1} + b_{t+1} \cdot \alpha^H_{0,t+1}) \\
b_t &= c_t + r \cdot b_{t+1} \cdot \gamma^{H,H}_{t+1} \\
h_t(\theta^R_t, \eta^S_t) &= r \cdot b_{t+1} \cdot \left(\beta^H_{t+1} + \gamma^{H,R}_{t+1} \cdot \theta^R_t\right) + \max\Big\{\phi_t(\theta^R_t) - \alpha^S_t - \eta_t + \\
&\quad r \cdot \left(\mathbb{E}\left(h_{t+1}\left(\alpha^R_{t+1} + \gamma^{R,R}_{t+1} \cdot \theta^R_t + \epsilon^R_{t+1}, \eta^S_{t+1}\right) \mid \theta^R_t\right) + b_{t+1} \cdot \left(\delta^H_{t+1} + \xi^H_{t+1} \cdot \theta^R_t\right)\right), 0\Big\}
\end{aligned} \tag{4.13}$$

We can conclude by induction that (4.12) holds for all t. The a_t term in (4.12) captures the deterministic evolution of the total net value function, while b_t represents the expected present value of an additional unit of child human capital in period t. $h_t(\cdot)$ is the value of starting period t in a relationship, which equals the sum of a discounted expected gain from changes in the evolution of θ^H_t and an option value for continuing the relationship.

The choice index takes the form

$$\begin{aligned}
p_t(\boldsymbol{\theta}_t) - \eta^S_t &= \phi_t(\theta^R_t) - \alpha^S_t - \eta^S_t + \\
&\quad r \cdot \left(b_{t+1} \cdot \left(\delta^H_{t+1} + \xi^H_{t+1}\theta^R_t\right) + \mathbb{E}\left(h_{t+1}\left(\alpha^R_{t+1} + \gamma^{R,R}_{t+1} \cdot \theta^R_t + \epsilon^R_{t+1}, \eta_{t+1}\right) \mid \theta^R_t\right)\right)
\end{aligned} \tag{4.14}$$

This expression can be broken down into three parts:

- $\phi_t(\theta_t^R) - \alpha_t^S - \eta_t^S$: current period surplus of the relationship
- $r \cdot b_{t+1} \cdot \left(\delta_{t+1}^H + \xi_{t+1}^H \cdot \theta_t^R\right)$: expected discounted gain from changing the evolution of child human capital
- $r \cdot \mathbb{E}\left(h_{t+1}\left(\alpha_{t+1}^R + \gamma_{t+1}^{R,R} \cdot \theta_t^R + \epsilon_{t+1}^R, \eta_{t+1}\right) \mid \theta_t^R\right)$: discounted expected value of starting the next period in a relationship

This decomposition summarizes the trade-offs parents face when deciding whether or not separate. Forward-looking parents not only consider the current period surplus, but also weigh how separation affects their child's development and the value of having the option to continue the relationship in the future. If δ_{t+1}^H and ξ_{t+1}^H are positive, then separation only lowers future values of θ_t^H if

$$\theta_t^R > -\frac{\delta_{t+1}^H}{\xi_{t+1}^H}$$

Thus when parental relationship quality falls below a certain threshold, parents may actually face an incentive to separate to improve their child's outcomes. It is also interesting that θ_t^H does not enter the choice equation, even though parents care about their child's development and take into account the effect separation has on the evolution of θ_t^H. Since ρ does not enter the optimal policy function, there is no need to model how it is determined. That the sharing rule fails to alter separation decisions is a well-known result from transferable utility models (Becker (1993)).

4.6 Vector-valued child human capital

Family structure and parental relationship quality may have different effects on cognitive and non-cognitive skill development in children. Furthermore, even if parental separation only has a direct effect on the development of one type of skill, it may indirectly influence the evolution of the other skill through dynamic complementarities. To incorporate these possibilities, replace the scalar θ_t^H

in the above model with the vector $\boldsymbol{\theta}_t^H = (\theta_t^N, \theta_t^C)'$, where θ_t^N and θ_t^C denote the child's non-cognitive and cognitive skills. The law of motion for $\boldsymbol{\theta}_t^H$ can be written as a vectorized version of (4.2):

$$\begin{pmatrix} \theta_{t+1}^N \\ \theta_{t+1}^C \end{pmatrix} = \begin{pmatrix} \alpha_{t+1}^N \\ \alpha_{t+1}^C \end{pmatrix} + \begin{pmatrix} \beta_{t+1}^N \\ \beta_{t+1}^C \end{pmatrix} R_t + \begin{pmatrix} \gamma_{t+1}^{N,R} & \gamma_{t+1}^{N,N} & \gamma_{t+1}^{N,C} \\ \gamma_{t+1}^{C,R} & \gamma_{t+1}^{C,N} & \gamma_{t+1}^{C,C} \end{pmatrix} \begin{pmatrix} \theta_t^R \\ \theta_t^N \\ \theta_t^C \end{pmatrix} + \begin{pmatrix} \delta_{t+1}^N \\ \delta_{t+1}^C \end{pmatrix} R_{t+1} + \begin{pmatrix} \xi_{t+1}^N \\ \xi_{t+1}^C \end{pmatrix} R_{t+1} \cdot \theta_t^R + \begin{pmatrix} \epsilon_{t+1}^N \\ \epsilon_{t+1}^C \end{pmatrix} \quad (4.15)$$

The analysis of the model in Section 4.5 remains unchanged.

CHAPTER 5

ESTIMATION

5.1 Model equations

The FFCWS data were used to estimate the law of motion for parental relationship quality, the laws of motion governing child human capital development, and the choice equations describing the parents' optimal separation decisions. The choice equations were estimated by taking a linear approximation to the $p_t(\cdot)$ functions in (4.14):

$$p_t(\boldsymbol{\theta}_t) = \alpha_t^P + \gamma_t^P \cdot \theta_t^R \tag{5.1}$$

No attempt was made to recover the structural preference parameters from the choice equations, as these were not needed to construct the counterfactuals of interest in this dissertation.

The set of estimating equations is given by

$$\begin{aligned}
\theta_{t+1}^R &= R_{t+1} \cdot \left(\mathbf{x}' \boldsymbol{\alpha}_{t+1}^R + \gamma_{t+1}^{R,R} \cdot \theta_t^R + \epsilon_{t+1}^R \right) \\
\theta_{t+1}^N &= \mathbf{x}' \boldsymbol{\alpha}_{t+1}^N + \beta_{t+1}^N \cdot R_t + \boldsymbol{\theta}_t{}' \boldsymbol{\gamma}_{t+1}^N + \left(\delta_{t+1}^N + \xi_{t+1}^N \cdot \theta_t^R \right) \cdot R_{t+1} + \epsilon_{t+1}^N \\
\theta_{t+1}^C &= \mathbf{x}' \boldsymbol{\alpha}_{t+1}^C + \beta_{t+1}^N \cdot R_t + \boldsymbol{\theta}_t{}' \boldsymbol{\gamma}_{t+1}^C + \left(\delta_{t+1}^C + \xi_{t+1}^C \cdot \theta_t^R \right) \cdot R_{t+1} + \epsilon_{t+1}^C \\
R_{t+1} &= R_t \cdot \mathbb{1} \left(\mathbf{x}' \boldsymbol{\alpha}_{t+1}^P + \gamma_{t+1}^P \theta_t^R - \eta_{t+1}^P > 0 \right)
\end{aligned} \tag{5.2}$$

where $\boldsymbol{\gamma}_{t+1}^N = \left(\gamma_{t+1}^{N,R}, \gamma_{t+1}^{N,N}, \gamma_{t+1}^{N,C} \right)'$ and $\boldsymbol{\gamma}_{t+1}^C = \left(\gamma_{t+1}^{C,R}, \gamma_{t+1}^{C,N}, \gamma_{t+1}^{C,C} \right)'$. Since the data did not contain measurements of non-cognitive and cognitive skills until periods 1 and 2, respectively, the estimated skill sequences started whenever the relevant measurements become available. The model in Section 4 can easily be adjusted to accommodate this data structure.

5.2 Measurement system

Although relationship quality and child human capital are difficult to quantify, the data used for estimation measure these concepts with a rich array of survey questions (see Section 3 and Tables B.1–B.3 in the Appendix). While it is possible to construct proxies by simply averaging the available measurements, this approach is unsatisfactory for several reasons:

- It ignores attenuation bias from measurement error and fails to account for missing measurements in a structured way
- Since some survey items likely contain more information than others, it should be possible to construct more accurate proxies by weighting the measurements asymmetrically
- Measurement averages often have a heavily skewed empirical distribution that can be difficult to fit parameterically

This dissertation addresses these problems by estimating an explicit measurement system for relationship quality and child skills. Let $M^j_{t,k}$ denote measurement k of θ^j_t, where $j \in \{R, N, C\}$. Let K^j_t denote the number of measurements of θ^j_t that are available in the data. Define the index

$$U^j_{t,k} = \gamma^j_{t,k} \cdot \theta^j_t + \eta^j_{t,k} \tag{5.3}$$

where $\gamma^j_{t,k}$ is a factor-loading parameter and $\eta^j_{t,k}$ is a stochastic measurement error term. The measurement error terms were assumed to be mutually independent, independent of the predetermined regressors, and independent of the error terms in (5.2). If $M^j_{t,k}$ took values in the finite set $\{1, \cdots, n\}$, it was assumed to be determined by the ordered threshold-crossing model

$$M_{t,k}^j = \begin{cases} 1 & \text{if } U_{t,k}^j \leq c_{1,t,k}^j \\ l & \text{if } c_{l-1,t,k}^j < U_{t,k}^j \leq c_{l,t,k}^j, \quad l = 2, \cdots, n-1 \\ n & \text{if } U_{t,k}^j > c_{n-1,t,k}^j \end{cases} \tag{5.4}$$

where $\left\{c_{1,t,k}^j, \cdots, c_{n-1,t,k}^j\right\}$ are threshold parameters. If $M_{t,k}^j$ was continuous, then it was assumed to be determined by

$$M_{t,k}^j = \mu_{t,k}^j + U_{t,k}^j \tag{5.5}$$

where $\mu_{t,k}^j$ denotes the mean of the measurement.

Since the measurement system only measures relationship quality and child skills indirectly, normalizations were needed to set the sign, location, and scale of each latent variable. The sign of θ_t^j was set by assuming

$$\gamma_{t,k}^j \geq 0 \tag{5.6}$$

for all t, k, j. Although a weaker assumption could have been used, (5.6) is easy to interpret, led to good computational performance, and was never binding in practice. The location and scale of θ_t^j were set by assuming

$$\mathbb{E}\left[\theta_t^j\right] = 0, \qquad \text{var}\left[\theta_t^j\right] = 1$$

5.3 Outcomes

Let O_j denote the j-th discrete outcome (see Section 3.5 and Tables C.1 and C.2 in the Appendix). O_j was modelled using a threshold-crossing model with latent index given by

$$U_j^O = \mathbf{x}'\boldsymbol{\alpha}_j^O + \gamma_j^{O,N} \cdot \theta_4^N + \gamma_j^{O,C} \cdot \theta_4^C + \eta_j^O \tag{5.7}$$

Unlike the dedicated measurements in Section 5.2, the outcomes can depend on the vector of covariates and on multiple latent variables. The signs of $\gamma_j^{O,N}$ and $\gamma_j^{O,C}$ were also left unrestricted. The η_j^O error terms were assumed to be mutually independent, independent of the pre-determined regressors, and independent of all other error terms in the model.

5.4 Parameteric assumptions

The specification of the data generating process was completed by imposing parametric assumptions on the unobserved terms. The unobserved terms in the model equations in (5.2) were drawn according to

$$\eta_t^P \sim N(0,1), \qquad \begin{pmatrix} \epsilon_t^R \\ \epsilon_t^N \\ \epsilon_t^C \end{pmatrix} \sim MVN\left(\mathbf{0}, \mathbf{\Sigma}_t\right), \tag{5.8}$$

where $N(\cdot)$ and $MVN(\cdot)$ represent normal and multivariate normal distributions, respectively, and $\mathbf{\Sigma}_t$ was a covariance matrix to be estimated.

If $M_{t,k}^j$ was a discrete measurement, then $\eta_{t,k}^j$ was drawn from a standard logistic:

$$\eta_{t,k}^j \sim \text{Logistic}(0,1) \tag{5.9}$$

If $M_{t,k}^j$ was continuous, then $\eta_{t,k}^j$ was drawn according to

$$\eta_{t,k}^j \sim N(0, \sigma_{t,k}^j) \tag{5.10}$$

where $\sigma_{t,k}^j$ denotes the standard deviation of the measurement error. The outcome errors were drawn according to

$$\eta_j^O \sim \text{Logistic}(0,1) \tag{5.11}$$

5.5 Identification

Although the estimates were derived from a parameteric model, it is instructive to analyze how the effects of interest might be identified non-parametrically. If $\boldsymbol{\theta}_t$ were observed, the effect of a separation could be recovered by simply matching parents by $(\mathbf{x}, \boldsymbol{\theta}_t)$ and comparing outcomes for parents who made different separation choices. This identification strategy relies solely on the following non-parametric assumptions:

- The vector $(\mathbf{x}, \boldsymbol{\theta}_t)$ captures all information parents have about future shocks
- The probability of parental separation is 0 or 1 for any value in the support of $(\mathbf{x}, \boldsymbol{\theta}_t)$

The first assumption follows from the information structure in Section 4.4 and the second assumption is satisfied as long as the support of the parental outside options is unbounded.

Although this argument provides useful intuition, it is not directly applicable because $\boldsymbol{\theta}_t$ is unobserved in the data. Because these latent variables enter non-linear choice and measurement equations, the parameters cannot be non-parameterically point identified for a fixed number of time periods and measurements (Honoré and Tamer (2006)). There is still a sense in which this non-parametric argument holds asymptotically, since the true values of $\boldsymbol{\theta}_t$ are revealed as the number of time periods grows large (Arellano and Bonhomme (2011), Williams (2018)).

5.6 Likelihood and posterior

Let $\mathbf{M}_t^j$ denote the vector of measurements of θ_t^j and let $\mathbf{O}$ denote the vector of outcomes. The full vector of modelled data is given by

$$\mathbf{y} = \left(R_1, \cdots, R_5, \mathbf{M}_0^R, \cdots, \mathbf{M}_4^R, \mathbf{M}_1^N, \cdots, \mathbf{M}_4^N, \mathbf{M}_2^C, \cdots \mathbf{M}_4^C, \mathbf{O}\right) \tag{5.12}$$

Let $\boldsymbol{\epsilon}$ denote a vector containing all the laws of motion error terms from the equations in (5.2) and let $\boldsymbol{\psi}$ denote a vector containing all the model parameters. The joint distribution of $\mathbf{y}$ conditional on $\boldsymbol{\epsilon}$ and $\mathbf{x}$ can then be written as

$$\begin{aligned} & f(\mathbf{y} \mid \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}) = \\ & \left(\prod_{t=0}^{4} f\left(R_{t+1} \mid R_t, \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}\right)\right) \cdot \\ & \left(\prod_{t=0}^{4} \prod_{k=1}^{K_t^R} f\left(M_{t,k}^R \mid R_1, \cdots, R_t, \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}\right)\right) \cdot \left(\prod_{t=1}^{4} \prod_{k=1}^{K_t^N} f\left(M_{t,k}^N \mid R_1, \cdots, R_t, \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}\right)\right) \cdot \\ & \left(\prod_{t=2}^{4} \prod_{k=1}^{K_t^C} f\left(M_{t,k}^C \mid R_1, \cdots, R_t, \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}\right)\right) \cdot \left(\prod_{j=1}^{J} f\left(O_j \mid R_1, \cdots, R_4, \boldsymbol{\epsilon}, \mathbf{x}; \boldsymbol{\psi}\right)\right) \end{aligned} \tag{5.13}$$

The model equations in (5.2), along with the parametric assumptions in (5.8)–(5.11), fully specify the conditional densities on the right-hand side of equation (5.13). Indexing a sample of independently drawn observations by $i = 1, \cdots, n$, we can write the integrated log-likelihood function as

$$\ell(\boldsymbol{\psi}) = \sum_{i=1}^{n} \log\left(\int f(\mathbf{y}_i \mid \boldsymbol{\epsilon}, \mathbf{x}_i; \boldsymbol{\psi}) \cdot f(\boldsymbol{\epsilon} \mid \boldsymbol{\psi}) \, d\boldsymbol{\epsilon}\right) \tag{5.14}$$

where $f(\boldsymbol{\epsilon} \mid \boldsymbol{\psi})$ is determined by the parametric forms in (5.8). The random-effects maximum likelihood estimator (MLE) is the value of $\boldsymbol{\psi}$ that maximizes $\ell(\boldsymbol{\psi})$.

The high-dimensional integral in (5.14) makes locating the MLE computationally difficult. This dissertation instead adopts a Bayesian estimation strategy. Let $f(\boldsymbol{\psi})$ denote a prior distribution, which is fully specified in Section E in the Appendix. The kernel of the posterior distribution of can be written as

$$f\left(\boldsymbol{\psi}, \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_n \mid \mathbf{y}_1, \cdots, \mathbf{y}_n, \mathbf{x}_1, \cdots, \mathbf{x}_n\right) \propto f(\boldsymbol{\psi}) \cdot \prod_{i=1}^{n} f(\mathbf{y}_i \mid \boldsymbol{\epsilon}, \mathbf{x}_i, \boldsymbol{\psi}) \cdot f(\boldsymbol{\epsilon}_i \mid \boldsymbol{\psi}) \tag{5.15}$$

As long as $f(\boldsymbol{\psi})$ is not dogmatic, the Bayesian posterior mean of $\boldsymbol{\psi}$ is consistent and asymptotically normal (Koop et al. (2007)). If the distributions of $\boldsymbol{\epsilon}$ are well-specified, functions of $\boldsymbol{\psi}$ and $\boldsymbol{\epsilon}$ (e.g. average partial effects, counter-factual means) are also consistent. Even if these distributions are misspecified, functions of $\boldsymbol{\psi}$ and $\boldsymbol{\epsilon}$ are consistent as the number of measurements grows (Arellano and Bonhomme (2009)). This result follows because knowledge about $\boldsymbol{\epsilon}$ accumulates via Bayesian updating, a property that does not hold for the MLE (Arellano and Bonhomme (2011)).

5.7 Missing data

Since the data were collected over several years, many measurements, separation decisions, and outcomes were missing due to sample attrition and item non-response. To account for missing information, partition the vector in (5.12) as

$$\mathbf{y} = (\mathbf{y}^{obs}, \mathbf{y}^{mis})'$$

where $\mathbf{y}^{obs}$ and $\mathbf{y}^{mis}$ correspond to the observed and missing data, respectively. Let $\mathbf{m}$ denote a vector of indicators that specifies which components of $\mathbf{y}$ are missing. Suppose the distribution of $\mathbf{m}$ depends on a finite dimensional vector of parameters $\boldsymbol{\varphi}$, and let $f(\boldsymbol{\psi}, \boldsymbol{\varphi})$ denote a joint prior distribution for the model and missing process parameters. The joint distribution of $\mathbf{y}$ and $\mathbf{m}$ can be factored as

$$f\left(\mathbf{y}^{obs}, \mathbf{y}^{mis}, \mathbf{m} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}\right) = f\left(\mathbf{m} \mid \mathbf{y}^{obs}, \mathbf{y}^{mis}, \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}\right) \cdot f\left(\mathbf{y}^{obs}, \mathbf{y}^{mis} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}\right) \quad (5.16)$$

The following missing-at-random (MAR) assumptions significantly simplify the handling of missing data:

$$\begin{aligned} f(\mathbf{m} \mid \mathbf{y}, \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}) &= f(\mathbf{m} \mid \mathbf{x}, \boldsymbol{\varphi}) \\ f(\mathbf{y}^{obs}, \mathbf{y}^{mis} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}) &= f(\mathbf{y}^{obs}, \mathbf{y}^{mis} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}) \\ f(\boldsymbol{\psi}, \boldsymbol{\varphi}) &= f(\boldsymbol{\psi}) \cdot f(\boldsymbol{\varphi}) \end{aligned} \tag{5.17}$$

The first and most demanding assumption in (5.17) requires the missing process to only depend on the observed vector of covariates. This assumption allows the probability of attrition and item non-response rates to differ across observed demographic groups, but rules out the possibility that the latent variables or the observed measurements and choices affect these probabilities. The second equation in (5.17) excludes the missing process parameters from the model equations. The third equation requires the model and missing process parameters have independent priors.

Under the assumptions in (5.17), the expression in (5.16) simplifies to

$$f\left(\mathbf{y}^{obs}, \mathbf{y}^{mis}, \mathbf{m} \mid \boldsymbol{\varphi}, \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}\right) = f\left(\mathbf{m} \mid \boldsymbol{\varphi}, \mathbf{x}\right) \cdot f\left(\mathbf{y}^{obs}, \mathbf{y}^{mis} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}\right)$$

It is now possible to simply integrate out the missing data:

$$f\left(\mathbf{y}^{obs}, \mathbf{m} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}, \boldsymbol{\varphi}\right) = f\left(\mathbf{m} \mid \boldsymbol{\varphi}, \mathbf{x}\right) \cdot f\left(\mathbf{y}^{obs} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}\right) \tag{5.18}$$

The kernel of the posterior distribution can now be written as

$$\begin{aligned}
&f\left(\boldsymbol{\psi}, \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_n \mid \mathbf{y}_1^{obs}, \cdots, \mathbf{y}_n^{obs}, \mathbf{m}_1, \cdots, \mathbf{m}_n, \mathbf{x}_1, \cdots, \mathbf{x}_n\right) \\
&\qquad\qquad\propto f(\boldsymbol{\psi}) \cdot f(\boldsymbol{\varphi}) \cdot \prod_{i=1}^{n} f(\mathbf{y}_i^{obs}, \mathbf{m}_i \mid \boldsymbol{\epsilon}, \mathbf{x}_i, \boldsymbol{\psi}, \boldsymbol{\varphi}) \cdot f(\boldsymbol{\epsilon}_i \mid \boldsymbol{\psi}) \\
&\qquad\qquad\propto f(\boldsymbol{\psi}) \cdot f(\boldsymbol{\varphi}) \cdot \prod_{i=1}^{n} f\left(\mathbf{m} \mid \boldsymbol{\varphi}, \mathbf{x}\right) \cdot f(\mathbf{y}^{obs} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}\big) \cdot f(\boldsymbol{\epsilon}_i \mid \boldsymbol{\psi}) \\
&\qquad\qquad\propto f(\boldsymbol{\psi}) \cdot \prod_{i=1}^{n} f(\mathbf{y}^{obs} \mid \boldsymbol{\varepsilon}, \mathbf{x}, \boldsymbol{\psi}) \cdot f(\boldsymbol{\epsilon}_i \mid \boldsymbol{\psi})
\end{aligned}$$

where the second line follows from substituting (5.18) and the last line follows from eliminating terms that do not influence the kernel. This expression shows that the missing process can be left unmodelled because it does not affect the posterior of the model parameters.

5.8 Computation

Draws from the posterior distribution were obtained using a Hamiltonian Monte Carlo (HMC) algorithm (Carpenter et al. (2017)). Neal (2011) provides a useful review of this approach. HMC performs much better in high-dimensional spaces than the Metropolis-Hastings algorithm and does not require conjugate prior distributions like the Gibbs sampler.

Let $\boldsymbol{\psi}_t^j$ denote the parameters that determine the distribution of $\mathbf{M}_t^j$. To reduce computation time, $\boldsymbol{\psi}_t^j$ were computed in a first step by drawing from

$$f(\boldsymbol{\psi}_t^j, \theta_t^j \mid \mathbf{M}_t^j) \propto \prod_{k=1}^{K_t^j} f(M_{t,k}^j \mid \theta_t^j; \boldsymbol{\psi}_t^j) \cdot \phi(\theta_t^j) \cdot f(\boldsymbol{\psi}_t^j)$$

where $\phi(\theta_t^j)$ denotes the standard normal distribution. $\boldsymbol{\psi}_t^j$ was then fixed at its posterior mean for the estimation of the full model. This two-step procedure has no effect on the asymptotic consistency of the estimator, but the standard errors reported below do not account for the error in the estimation of ψ_t^j. Since the posteriors of these parameters were highly concentrated relative to the other parameters in the model, this estimation error had a negligible impact on the results.

The standard errors reported in the tables below reflect symmetric 95% credible sets estimated from 5,000 posterior draws. The draws were obtained from eight chains with a warm-up period of 1,000 iterations per chain. Parameters were initialized by transforming them to have unrestricted support and then drawing from a uniform distribution over $[-.5, .5]$. Convergence was checked using the estimated potential scale reduction statistic $\hat{R}$ (Carpenter et al. (2017)). $\hat{R}$ was less than 1.05 for all parameters, indicating that the chains converged to the same stationary distribution. The estimated effective sample size for each parameter exceeded 300.

CHAPTER 6

RESULTS

6.1 Measurement system

Rather than report the raw factor loadings, Tables B.1–B.3 in the Appendix display the fraction of each measurement's latent index variance that corresponds to a signal of the latent variable:

$$\text{Signal } \% = \frac{\left(\gamma_{t,k}^{j}\right)^2}{\left(\gamma_{t,k}^{j}\right)^2 + \text{var}\left(\eta_{t,k}^{j}\right)} \times 100$$

The signal percentage differs considerably across survey items, suggesting that some measurements were more informative about the latent variables than others.

The signal percentages for the parental relationship quality measurements are displayed in Table B.1. Positive questions, such as "He expresses affection or love for you" and "He encourages or helps you to do things that are important to you", contained more information than negative questions like "He insults or criticizes you and your ideas". Questions that probed whether the couple had considered breaking up in the past year were particularly informative about relationship quality.

The signal percentages for non-cognitive skill measurements are displayed in Table B.2. The mother's report of age 1 behavior was more informative about non-cognitive skills than the father's report. At older ages, CBCL items that were tangentially related to aggression, such as "(He/She) wants a lot of attention" or "(He/She) talks too much", contained less information than items like "(He/She) physically attacks people" or "(He/She) gets in many fights".

The signal percentages for cognitive skill measurements are displayed in Table B.3. In addi-

Table 6.1: Parameter estimates for the relationship quality law of motion

	θ_1^R	θ_2^R	θ_3^R	θ_4^R
θ_{t-1}^R	.347***	.347***	.319***	.297***
	(.034)	(.037)	(.044)	(.047)
R^2	.541	.506	.472	.449

Relationship quality law of motion: $\theta_t^R = R_t \cdot (\mathbf{x}'\alpha_t^R + \gamma_t^R\theta_{t-1}^R + \epsilon_t^R)$. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1 - \frac{\alpha}{2})$-th posterior quantile.

* $p < .1$ ** $p < .05$ *** $p < .01$

tion to the child's test scores, parental test scores were also used to measure the child's baseline cognitive ability. The digit span test score, which asked children to repeat a number read by an interviewer backwards, contained less information about the child's cognitive skills than the other standardized test scores.

Figure B.1 in the Appendix illustrates the model fit by plotting observed versus simulated measurement averages for each latent variable. Measurements were simulated for each posterior draw and then each observation's draws were averaged. To ensure the measurements were on the same scale, each measurement was normalized to have a zero mean and unit standard deviation. The simulated averages closely fit most of the observed averages, even in the tails of the distributions. The child's non-cognitive skill at age 1 and the child's cognitive skill at age 3 were the only latent variables fit somewhat poorly. These discrepancies may reflect the difficulty of obtaining a reliable set of measurements for skills at very young ages.

6.2 Parameter estimates

Table 6.1 contains parameter estimates for the relationship quality law of motion. The dynamics of relationship quality appear fairly stable over time. The R^2 results suggest that a large fraction of the variance in relationship quality remains unexplained each period. This suggests parents

Table 6.2: Parameter estimates for the non-cognitive skill law of motion

	θ_1^N	θ_2^N	θ_3^N	θ_4^N
R_{t-1}		.079	-.060	-.047
		(.087)	(.077)	(.095)
R_t	-.141	-.122*	.075	.001
	(.099)	(.073)	(.074)	(.094)
θ_{t-1}^R	.125	.027	-.068	-.001
	(.085)	(.058)	(.057)	(.068)
θ_{t-1}^N		.257***	.544***	.470***
		(.033)	(.032)	(.035)
θ_{t-1}^C			.027	.104**
			(.047)	(.042)
$R_t \cdot \theta_{t-1}^R$	-.003	.163***	.123**	.134*
	(.086)	(.066)	(.062)	(.078)
R^2	.160	.249	.478	.430

Non-cognitive skill law of motion: $\theta_t^N = \mathbf{x}'\boldsymbol{\alpha}_t^N + \beta_t^N R_{t-1} + \boldsymbol{\theta}_{t-1}'\boldsymbol{\gamma}_t^N + (\delta_t^N + \xi_t^N \cdot \theta_{t-1}^R) \cdot R_t + \epsilon_t^N$. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1 - \frac{\alpha}{2})$-th posterior quantile.
* $p < .1$ ** $p < .05$ *** $p < .01$

face substantial uncertainty about how their relationship quality will evolve, even over short time frames.

Table 6.2 contains parameter estimates for the non-cognitive skill laws of motion. The coefficients on parental relationship status at the start and end of each period, reported in the first two rows of Table 6.2, are typically modest and statistically insignificant. The coefficients on parental relationship quality are also small and insignificant. By contrast, the interaction coefficients in the last row of the table are large in periods 3, 4 and 5. This coefficient is statistically significant at conventional levels for periods 3 and 4 and marginally significant in the last period. The results suggest that a parental separation can potentially protect children from a low relationship quality shock. The effects of a healthy or unhealthy parental relationship accumulate while the parents stay together.

Table 6.3: Parameter estimates for the cognitive skill law of motion

	θ_3^C	θ_4^C	θ_5^C
R_{t-1}	-.268**	.083	-.023
	(.123)	(.085)	(.069)
R_t	.177*	-.016	.022
	(.105)	(.081)	(.065)
θ_{t-1}^R	-.097	.052	-.020
	(.084)	(.066)	(.050)
θ_{t-1}^N	.065	.090**	.017
	(.048)	(.039)	(.027)
θ_{t-1}^C		.471***	.492***
		(.099)	(.115)
$R_t \cdot \theta_{t-1}^R$	.057	-.059	.014
	(.098)	(.075)	(.056)
R^2	.979	.989	.993

Cognitive skill law of motion: $\theta_t^C = \mathbf{x}'\boldsymbol{\alpha}_t^C + \beta_t^C R_{t-1} + \boldsymbol{\theta}_{t-1}'\boldsymbol{\gamma}_t^C + (\delta_t^C + \xi_t^C \cdot \theta_{t-1}^R) \cdot R_t + \epsilon_t^C$. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.
* $p < .1$ ** $p < .05$ *** $p < .01$

Table 6.3 contains parameter estimates for the cognitive skill laws of motion. As was the case with non-cognitive skills, the coefficients on parental relationship status indicators and parental relationship quality are small and statistically insignificant except in the second period. But unlike the non-cognitive skill law of motion, the interaction coefficients in the last row of the table are insignificant. The results indicate that parental relationship quality does not have a large direct effect on cognitive skill development, regardless of whether the parents stay together or separate. Cognitive skills exhibit much greater persistence than non-cognitive skills, consistent with evidence that suggests these skills are determined at very early ages (Cunha et al. (2006)).

Table 6.4 contains estimates of the choice functions. The first row reports estimates of the coefficient on relationship quality while the second row reports the estimated average partial effect of relationship quality on the probability of staying together. Parental relationship quality exhibits a

Table 6.4: Estimated optimal separation policy functions

	R_1	R_2	R_3	R_4	R_5
θ^R_{t-1}	.710***	1.03***	.980***	.962***	5.99***
	(.100)	(.087)	(.090)	(.101)	(1.69)
APE	.100***	.208***	.198***	.200***	.131***
	(.009)	(.008)	(.011)	(.016)	(.027)

Choice equations: $R_{t+1} = R_t \cdot \mathbb{1}\left(\mathbf{x}'\boldsymbol{\alpha}^p_t + \gamma^p_{1,t}\theta^R_t - \eta^S_t > 0\right)$. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.
* $p < .1$ ** $p < .05$ *** $p < .01$

strong effect on the separation probability, with a standard deviation decline in relationship quality increasing the probability of separation by 10 to 20 percentage points.

Table 6.5 reports estimates of the binary outcome equations. The first three outcomes involve the child's relative performance in different subjects. These outcomes are very responsive to differences in cognitive skills, but do not respond much to variation in non-cognitive skills. The next three outcomes involve whether the child received different special services at school. These outcomes respond to both skills, with receipt of an individualized education plan responding more to cognitive skills and receipt of a behavior intervention plan responding more to non-cognitive skills. The last three outcomes represent different behaviors reported by the child during his or her age 9 interview. These outcomes respond strongly to differences in the non-cognitive skill, but do not depend on child cognitive skills.

6.3 Counter-factual analysis

Tables 6.6 and 6.7 contain counter-factual estimates for how age 9 skills would change if parents separated at different dates. Parents are grouped by the period in which they separated in the survey, with the first row corresponding to parents that separated before the child's birth, the second row corresponding to parents that separated between the child's birth and his or her first birthday, and

Table 6.5: Estimated binary outcome equations

	θ_4^N		θ_4^C	
	Factor loading	APE	Factor loading	APE
Below average language and literacy skills	-.089 (.109)	-.008 (.010)	-4.28*** (.322)	-.406*** (.010)
Below average science and social studies skills	-.047 (.115)	-.004 (.009)	-4.19*** (.335)	-.343*** (.011)
Below average mathematical skills	.008 (.095)	.001 (.010)	-3.48*** (.228)	-.378*** (.011)
Special education services through an IEP	-.350*** (.100)	-.028*** (.008)	-2.11*** (.153)	-.168*** (.010)
Received Behavior Intervention Plan	-.884*** (.117)	-.046*** (.006)	-.543*** (.115)	-.028*** (.006)
Received services because of ADD/ADHD	-.595*** (.133)	-.024*** (.006)	-.697*** (.129)	-.028*** (.005)
Purposely damaged or destroyed property	-.491*** (.066)	-.054*** (.007)	.015 (.073)	.002 (.008)
Had a fist fight with another person	-.480*** (.053)	-.084*** (.009)	.031 (.059)	.005 (.010)
Suspended or expelled from school	-.705*** (.064)	-.079*** (.007)	.007 (.067)	.001 (.008)

Outcome equations: $O_j = \mathbb{1}\left\{\mathbf{x}'\boldsymbol{\alpha}_j^O + \gamma_j^{O,N}\theta_4^N + \gamma_j^{O,C}\theta_4^C + \eta_j^O > 0\right\}$. APE columns report partial effects averaged over the sample. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.

* $p < .1$ ** $p < .05$ *** $p < .01$

so forth. The left panels tabulate the number of observations in each group, as well as each group's average child skill. Children whose parents did not separate have significantly higher non-cognitive and cognitive skills than other children in the sample. However, children whose parents separated later do not have more skills than children whose parents separated earlier.

The right panels in Tables 6.6 and 6.7 contain estimates for how each group's average skills

Table 6.6: Observed and counter-factual mean non-cognitive skills at age 9

Observed			**Counter-factual**				
Separation period	N	Mean of θ_5^N	Change in mean of θ_5^N				
			0	1	2	3	≥ 4
0	468	-.098*** (.026)		-.051 (.030)	-.153*** (.054)	-.347*** (.119)	-.481*** (.137)
1	892	-.141*** (.018)	.061* (.030)		-.112** (.046)	-.298** (.114)	-.425*** (.126)
2	568	-.081*** (.023)	.085* (.047)	.068 (.046)		-.186 (.107)	-.307*** (.122)
3	418	-.060* (.034)	.197 (.107)	.190 (.106)	.153 (.108)		-.137* (.078)
4	278	-.196*** (.029)	.080 (.062)	.075 (.058)	.043 (.061)	.015 (.089)	
Did not separate	1,126	.221*** (.016)	-.125 (.083)	-.116 (.077)	-.124 (.080)	-.060 (.118)	

Left panel tabulates average non-cognitive outcome at age 9 by observed separation period. Right panel displays estimates for how much mean outcomes would change if parents separated in different periods. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.
* $p < .1$ ** $p < .05$ *** $p < .01$

would change if parents chose to separate in different periods. Entries above (below) the main diagonal of this panel represent separating later (earlier) than the separation period observed in the data. Except in the last row, all the entries in Table 6.6 above the diagonal are negative and all the entries below the diagonal are positive. This pattern indicates children of separated parents would have had significantly lower non-cognitive skills if their parents had stayed together, and in fact may have had higher non-cognitive skills if their parents separated earlier. On the other hand, the negative entries in the last row indicate that children of parents who stayed together during the survey would have had lower non-cognitive skills had their parents separated.

In contrast to the entries in Table 6.6, most of the entries on the right panel of Table 6.7 are

Table 6.7: Observed and counter-factual mean cognitive skills at age 9

Observed			**Counter-factual**				
Separation period	N	Mean of θ_4^C	Change in mean of θ_4^C				
			0	1	2	3	≥ 4
0	468	-.137*** (.024)		-.080** (.047)	-.024 (.061)	-.028 (.097)	-.022 (.115)
1	892	-.231*** (.015)	.081** (.047)		.050 (.047)	.045 (.092)	.051 (.107)
2	568	-.142*** (.022)	-.000 (.053)	-.080 (.052)		-.003 (.083)	.004 (.099)
3	418	-.118*** (.029)	.026 (.084)	-.053 (.086)	.026 (.086)		.005 (.065)
4	278	-.076*** (.026)	.017 (.063)	-.062 (.061)	.017 (.056)	-.020 (.063)	
Did not separate	1,126	.332*** (.012)	.015 (.085)	-.064 (.080)	.022 (.070)	-.030 (.082)	

Left panel tabulates average cognitive outcome at age 9 by observed separation period. Right panel displays estimates for how much mean outcomes would change if parents separated in different periods. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.
* $p < .1$ ** $p < .05$ *** $p < .01$

small and insignificant. These estimates are consistent with the parameters in Table 6.3 and suggest that changing the timing of parental separation has little effect on age 9 cognitive skills.

Although the latent variables are normalized to have unit standard deviations, the magnitudes of the entries in Tables 6.6 and 6.7 do not have an obvious interpretation in terms of observable outcomes. To anchor the results, Table 6.5 displays estimates for how the age 9 outcomes would change if parents who separated during the sample instead chose to stay together. The Δ and $RR(\%)$ columns in Table 6.8 denote predicted percentage point and relative changes in the probability of each outcome's occurrence, respectively. Choosing not to separate has almost no effect on the probabilities of the first three outcomes. On the other hand, choosing not to separate increases

Table 6.8: Counter-factual age 9 outcomes for children of separated parents

	Counter-factual: did not separate		
	N	Δ	RR (%)
Below average language and literacy skills	2346	-.003 (.038)	1.03 (.223)
Below average science and social studies skills	2346	-.003 (.034)	1.03 (.261)
Below average mathematical skills	2346	-.006 (.036)	0.99 (.163)
Special education services through an IEP	2346	.010 (.019)	1.10 (.177)
Received Behavior Intervention Plan	2346	$.024^{**}$ (.011)	1.34^{***} (.164)
Received services because of ADD/ADHD	2346	$.011^{*}$ (.007)	1.21^{**} (.127)
Purposely damaged or destroyed property	2346	$.022^{***}$ (.008)	1.15^{***} (.056)
Had a fist fight with another person	2346	$.030^{***}$ (.010)	1.11^{***} (.037)
Suspended or expelled from school	2346	$.034^{***}$ (.012)	1.21^{***} .077

IEP = Individualized Education Program, ADD/ADHD = Attention Deficit/Hyperactivity Disorder. Δ and RR(%) columns denote the percentage point and relative change in probability of each outcome if the couples who separated in the sample instead chose to stay together. Posterior standard deviations reported in parentheses. Tests of significance level $\alpha\%$ performed by determining whether zero (Δ column) or one (RR(%) column) fell below the $\frac{\alpha}{2}$-th posterior quantile or above the $(1-\frac{\alpha}{2})$-th posterior quantile.

* $p < .1$ ** $p < .05$ *** $p < .01$

the probability of receiving a behavior intervention plan in school by 34 percent and increases the likelihood of receiving special services because of ADD or ADHD by 21 percent. Not separating also increases the probability of purposely damaging property, having a fist fight, and getting suspend or expelled from school by 15 percent, 11 percent, and 21 percent, respectively.

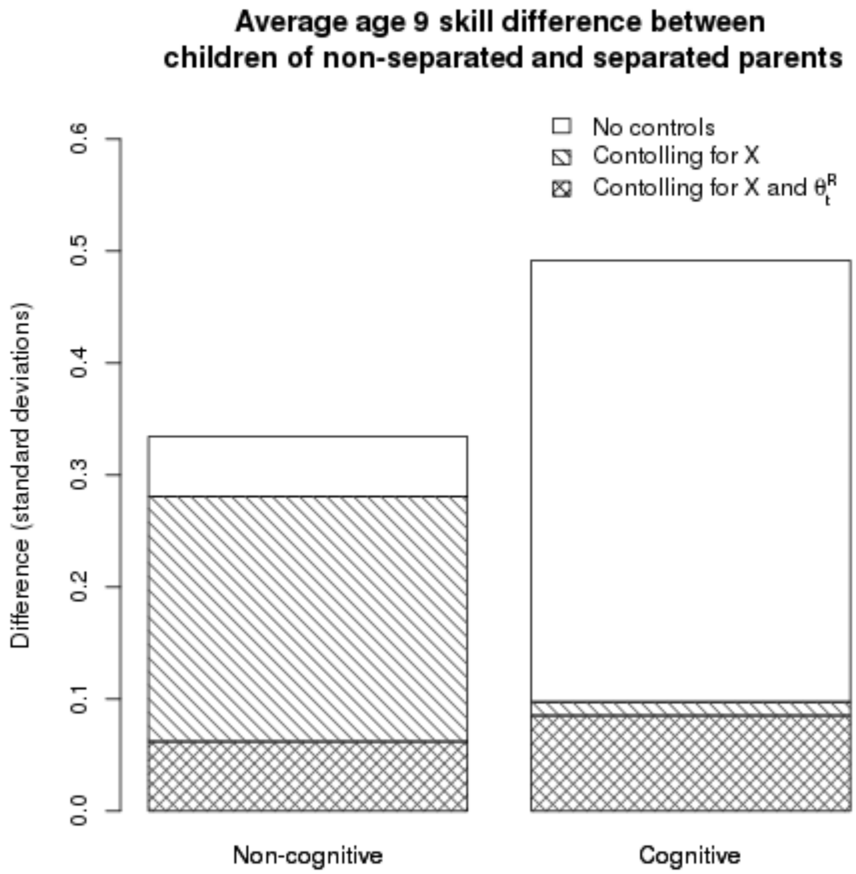


Figure 6.1: Mediation analysis

6.4 Mediation analysis

Figure 6.1 illustrates how much differences in the predetermined covariate vector and parental relationship quality can account for the skill gap between children of non-separated and separated parents. The two bars in the figure correspond to the average age 9 skill difference between these two groups of children. On average, children whose parents did not separate had a third of a standard deviation higher non-cognitive skills and half a standard deviation higher cognitive skills than children whose parents separated sometime during the survey. As a comparison, the skill gap between children of college-educated mothers and the rest of the sample is about .2 standard deviations for non-cognitive skills and 1.1 standard deviations for cognitive skills.

The single-hatched regions indicate how much of the gap remains after controlling for the predetermined variables in $\mathbf{x}$. These gaps were calculated by assigning separated parents a covariate vector drawn randomly from the covariate vectors of non-separated parents. A counterfactual evolution of each child's development was then simulated. The process was repeated 500 times to reduce simulation noise. Controlling for $\mathbf{x}$ in this fashion reduced the gap in cognitive skills by 80 percent to less .1 standard deviations. However, controlling for $\mathbf{x}$ only reduced the gap in non-cognitive skills by 16 percent to .28 standard deviations.

The cross-hatched regions in the figure indicate how much of the gap remains after additionally controlling for differences in relationship quality. This gap was calculated by replacing the observed relationship quality sequence with a sequence drawn randomly from the set of non-separated parents. Controlling for relationship quality only reduced the cognitive skill gap by 2.4 percent or .012 standard deviations. However, controlling for parental relationship quality reduced the gap in non-cognitive skills by almost two thirds the size of the original gap. The results indicates that differences in parental relationship quality largely explain why children of separated parents have worse non-cognitive outcomes than children whose parents remain together.

CHAPTER 7

CONCLUSION

This dissertation estimated a dynamic model of parental relationship quality, family structure, and child outcomes using rich longitudinal data. The estimation procedure controlled for observed demographics, unobserved heterogeneity, and measurement error. The results indicated that parental separation and relationship quality have little impact on cognitive skills, but have a large impact on non-cognitive skills. The effect of parental separation on non-cognitive skills was found to positively depend on the quality of the parental relationship. Since parents who separate tend to have poor relationships, the model predicted their children would have had worse behavioral outcomes had they chosen to stay together. Differences in parental relationship quality explained a large fraction of the non-cognitive skill gap between children of separated and non-separated parents, but explained little of the gap in cognitive skills.

In the future, the model developed in this dissertation could be extended to distinguish between different types of parental relationships (e.g. cohabitation versus marriage). Post-separation decisions like the decision to start a relationship with a non-biological parent could also be included. Modelling parental choices like investment in children and labor supply could provide more information about mechanisms. Allowing for reverse causality (e.g. an unhappy child causing poor relationship quality) would also be a useful extension.

The results also have important policy implications. First, policies that discourage parental separations may actually worsen child behavioral outcomes because relationships that dissolve are typically poor environments for children. Second, the results indicate that improving parental relationship quality is critical to closing the non-cognitive skill gap between children of non-separated and separated parents. How policy can be designed to improve parental relationship quality should be an area of future research.

REFERENCES

Thomas Achenbach. *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Department of Psychiatry, University of Vermont, Burlington, VT, 1991.

Thomas Achenbach and Leslie Rescorla. *Manual for the ASEBA school-age forms & profiles*. University of Vermont, Research Center for Children, Youth, and Families, Burlington, VT, 2001.

Anna Aizer and Sara McLanahan. The impact of child support enforcement on fertility, parental investments, and child well-being. *Journal of Human Resources*, 41(1):28–45, 2006.

Mathilde Almlund, Angela Lee Duckworth, James Heckman, and Tim Kautz. Personality psychology and economics. In Ludger Woessmann Eric Hanushek, Stephen Machin, editor, *Handbook of the Economics of Education*, volume 4, pages 1–181. Elsevier, 2011.

Paul Amato. Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family*, 72(3):650–666, 2010.

Paul Amato and Christopher Anthony. Estimating the effects of parental divorce and death with fixed effects models. *Journal of Marriage and Family*, 76(2):370–386, 2014.

Paul Amato and Juliana Sobolewski. The effects of divorce and marital discord on adult children's psychological well-being. *American Sociological Review*, pages 900–921, 2001.

Manuel Arellano and Stéphane Bonhomme. Robust priors in nonlinear panel data models. *Econometrica*, 77(2):489–536, 2009.

Manuel Arellano and Stéphane Bonhomme. Nonlinear panel data analysis. *Annual Reviews of Economics*, 3:395–424, 2011.

Jeremy Arkes. The temporal effects of divorces and separations on children's academic achievement and problem behavior. *Journal of Divorce & Remarriage*, 56(1):25–42, 2015.

Joëlle Barthassat. Positive and negative effects of parental conflicts on childrens condition and behaviour. *Journal of European Psychology Students*, 5(1):10–18, 2014.

Jennifer Baxter, Ruth Weston, and Lixia Qu. Family structure, co-parental relationship quality, post-separation paternal involvement and childrens emotional wellbeing. *Journal of Family Studies*, 17(2):86–109, 2011.

Gary Becker. *A Treatise on the Family*. Harvard University Press, Cambridge, MA, 1993.

Michael Brien, Lee Lillard, and Steven Stern. Cohabitation, marriage, and divorce in a model of match quality. *International Economic Review*, 47(2):451–494, 2006.

Susan Brown. Marriage and child well-being: Research and policy perspectives. *Journal of Marriage and Family*, 72(5):1059–1077, 2010.

Martin Browning, Pierre-André Chiappori, and Yoram Weiss. *Economics of the Family*. Cambridge University Press, New York, NY, 2014.

Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 2017.

Anne Carroll, Hope Corman, Kelly Noonan, and Nancy Reichman. Why do poor children lose health insurance in the SCHIP era? The role of family health. *The American Economic Review*, 97(2):398–401, 2007.

Andrew Cherlin, Frank Furstenberg, Lindsay Chase-Lansdale, Kathleen Kiernan, Philip Robins, Donna Ruane Morrison, and Julien Teitler. Longitudinal studies of effects of divorce on children in Great Britain and the United States. *Science*, 252(5011):1386–1389, 1991.

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007.

Flavio Cunha and James Heckman. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4):738–782, 2008.

Flavio Cunha, James Heckman, Lance Lochner, and Dimitriy Masterov. Interpreting the evidence on life cycle skill formation. In Ludger Woessmann Eric Hanushek, Stephen Machin, editor, *Handbook of the Economics of Education*, volume 1, pages 697–812. Elsevier, San Diego, CA, 2006.

Flavio Cunha, James Heckman, and Susanne Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

Janet Currie, Valentina Duque, and Irwin Garfinkel. The great recession and mothers' health. *The Economic Journal*, 125(588):F311–F346, 2015.

Patrick Davies and E Mark Cummings. Marital conflict and child adjustment: An emotional security hypothesis. *Psychological bulletin*, 116(3):387–411, 1994.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. *The Review of Economic Studies*, 81(1):137–185, 2014.

Robert Emery. Interparental conflict and the children of discord and divorce. *Psychological Bulletin*, 92(2):310–330, 1982.

Erin Fletcher. Match quality and maternal investments in children. *Review of Economics of the Household*, 14(1):83–102, 2016.

Libertad González and Tarja Viitanen. The long-term effects of legalizing divorce on children. *Oxford Bulletin of Economics and Statistics*, 80(2):327–357, 2018.

Jonathan Gruber. Is making divorce easier bad for children? The long-run implications of unilateral divorce. *Journal of Labor Economics*, 22(4):799–833, 2004.

John Grych and Frank Fincham. Marital conflict and children's adjustment: a cognitive-contextual framework. *Psychological Bulletin*, 108(2):267–290, 1990.

Juho Härkönen, Fabrizio Bernardi, and Diederik Boertien. Family dynamics and child outcomes: An overview of research and open questions. *European Journal of Population*, 33(2):163–184, 2017.

James Heckman and Stefano Mosso. The economics of human development and social mobility. *Annual Review of Economics*, 6:689–733, 2014.

James Heckman, Rodrigo Pinto, and Peter Savelyev. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *The American Economic Review*, 103(6):1–35, 2013.

Bo Honoré and Elie Tamer. Bounds on parameters in panel dynamic discrete choice models. *Econometrica*, 74(3):611–629, 2006.

Susan Jekielek. Parental conflict, marital disruption and children's emotional well-being. *Social Forces*, 76(3):905–936, 1998.

Gary Koop, Dale Poirier, and Justin Tobias. *Bayesian Econometric Methods*. Cambridge University Press, New York, NY, 2007.

Daniel Lewandowski, Dorota Kurowicka, and Harry Joe. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001, 2009.

Shirley Liu and Frank Heiland. Should we get married? The effect of parents' marriage on out-of-wedlock children. *Economic Inquiry*, 50(1):17–38, 2012.

Erika Lundquist, JoAnn Hsueh, Amy E. Lowenstein, Kristen Faucetta, Daniel Gubits, Charles Michalopoulos, and Virginia Knox. A family-strengthening program for low-income families: Final impacts from the supporting healthy marriage evaluation. Technical report, Office of Planning, Research and Evaluation, 2014.

Katherine Magnuson and Lawrence Berger. Family structure states and transitions: associations with children's well-being during middle childhood. *Journal of Marriage and Family*, 71(3): 575–591, 2009.

Joyce Martin, Brady Hamilton, Michelle Osterman, Anne Driscoll, and TJ Mathews. Births: Final data for 2015. *National Vital Statistics Reports*, 66(1), 2017.

Sara McLanahan and Gary Sandefur. *Growing Up with a Single Parent. What Hurts, What Helps.* Harvard University Press, Cambridge, MA, 1994.

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

Kelly Musick and Ann Meier. Are both parents always better than one? Parental conflict and young adult well-being. *Social Science Research*, 39(5):814–830, 2010.

Radford Neal. MCMC using Hamiltonian dynamics. In Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng, editors, *Handbook of Markov Chain Monte Carlo*, chapter 5, pages 113–162. CRC Press, 2011.

Patricia Pendry and Emma Adam. Child-related interparental conflict in infancy predicts child cognitive functioning in a nationally representative sample. *Journal of Child and Family Studies*, 22(4):502–515, 2013.

Thomas Piketty. The impact of divorce on school performance: Evidence from France, 1968-2002. CEPR Discussion Paper No. 4146, 2003.

Nancy Reichman, Julien Teitler, Irwin Garfinkel, and Sara McLanahan. Fragile Families: Sample and design. *Children and Youth Services Review*, 23(4):303–326, 2001.

David Ribar. What do social scientists know about the benefits of marriage? A review of quantitative methodologies. IZA Discussion Paper No. 998, 2004.

Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0*. Minnesota Population Center, Minneapolis, MN, 2010.

Carmen Solomon-Fears. Fatherhood initiatives: Connecting fathers to their children. CRS Report RL31025, Congressional Research Service, 2015.

Melissa Tartari. Divorce and the cognitive achievement of children. *International Economic Review*, 56(2):597–645, 2015.

United States Census Bureau. Historical living arrangements of children. https://www.census.gov/data/tables/time-series/demo/families/children.html, 2018. Accessed: 2018-07-21.

Stephanie Ventura and Christine Bachrach. Nonmarital childbearing in the United States, 1940-99. *National Vital Statistics Reports*, 48(16), 2000.

Alessandra Voena. Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *The American Economic Review*, 105(8):2295–2332, 2015.

Elizabeth Washbrook, Carol Propper, and Kapil Sayal. Pre-school hyperactivity/attention problems and educational outcomes in adolescence: prospective longitudinal study. *The British Journal of Psychiatry*, 203(4):265–271, 2013.

Benjamin Williams. Identification of a nonseparable model under endogeneity using binary proxies for unobserved heterogeneity. Working paper, 2018.

APPENDIX

A First-differencing with heterogeneous effects

Let θ_0^H and θ_1^H denote a child's human capital level in periods 0 and 1, respectively, and let θ^R denote parental relationship quality. Suppose all parents are in a relationship in period 0 and that some choose to separate in period 1; let R denote an indicator identifying which parents stay together. Suppose child human capital is determined by the equations

$$\begin{aligned} \theta_0^H &= \alpha + \epsilon_0 \\ \theta_1^H &= \left(\beta + \xi \cdot \theta^R\right) \cdot R + \alpha + \epsilon_1 \\ &= \beta \cdot R + u \end{aligned} \tag{A.1}$$

where α is a fixed effect, ϵ_0 and ϵ_1 are exogenous error terms, and $u = \xi \cdot \theta^R \cdot R + \alpha + \epsilon_1$. β is a parameter capturing the average effect of separation on child human capital development and ξ is a parameter representing how the effect of separation varies with the quality of the parental relationship.

The correlation between R and u is given by

$$\begin{aligned} \mathrm{cov}(R, u) = (\mathbb{E}\left[\alpha \mid R = 1\right] - \mathbb{E}\left[\alpha\right]) \cdot \Pr(R = 1) + \\ \xi \cdot \mathbb{E}\left[\theta^R \mid R = 1\right] \cdot \Pr(R = 1) \cdot (1 - \Pr(R = 1)) \end{aligned}$$

The first term on the right arises because time-invariant unobserved determinants of the child's human capital (e.g. inherited genetic factors, parental human capital) may be correlated with the parental decision to separate. The second term appears because parents with lower quality relationships are more likely to separate, and these separations are less detrimental than average. Both terms are likely to be positive, causing the ordinary least squares (OLS) estimate of β to be biased upward.

The fixed effect approach attempts to eliminate the problematic correlation by differencing the equations in (A.1):

$$\begin{aligned}\theta_1^H - \theta_0^H &= (\beta \cdot + \xi \cdot \theta^R) \cdot D + \epsilon_1 - \epsilon_0 \\ &= \beta \cdot D + u_\Delta\end{aligned}$$

Although this procedure eliminates the first source of bias, the second source remains:

$$\operatorname{cov}(D, u_\Delta) = \xi \cdot \mathbb{E}\left[\theta^R \mid D = 1\right] \cdot \Pr(D = 1) \cdot (1 - \Pr(D = 1))$$

Thus a regression of differenced outcomes on parental separation will still overestimate the average effect β.

B Measurement tables

Table B.1: Parental relationship quality measurements

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
How often, if at all, do you have open disagreements about...					
Money?	26.9%				
Spending time together?	31.2				
Sex?	22.7				
The pregnancy?	33.4				

(*Continued on next page*)

Table B.1: Parental Relationship Quality Measurements (continued)

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
Drinking or drug use?	34.8				
Being faithful?	41.7				
Now, think about how (FATHER) behaves towards you. For each statement I read, please tell me how often he behaves this way.					
He is fair and willing to compromise when you have a disagreement	26.4	27.0	29.8	25.1	41.8
He expresses affection or love for you	33.0	50.3	52.8	43.6	52.0
He encourages or helps you to do things that are important to you	34.7	55.8	57.4	47.7	53.9
He listens to you when you need someone to talk to		68.4	62.7	55.5	55.0
He really understands your hurts and joys		66.6	63.4	54.7	63.1
He insults or criticizes you or your ideas	25.3	34.4	38.2	33.7	38.3
He tries to keep you from seeing or talking with your friends or family		34.1	37.5	39.7	49.6
He tries to prevent you from going to work or school		28.8	33.9	33.9	41.2
He withholds money, makes you ask for money, or takes your money		48.9	48.6	55.3	51.1
He withholds sex to try to control your behavior				43.1	54.3
He insults or criticizes you for not taking good enough care of the child or your home				41.8	46.8
For the next set of statements, please tell me how often each is true about your relationship with (FATHER) over the past year.					

(*Continued on next page*)

Table B.1: Parental Relationship Quality Measurements (continued)

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
You thought your relationship with (FATHER) might be in trouble?				63.2	69.5
You and (FATHER) discussed ending your relationship?				61.3	61.8
You talked to a close friend or relative about breaking up with (FATHER)?				70.3	62.8
In general, would you say that your relationship with (FATHER) is excellent, very good, good, fair, or poor?		50.2	52.7	57.5	54.7
No matter how well parents get along, they sometimes have arguments. How often do you and (FATHER) argue about things that are important to you?		31.7	26.4		
After I read each statement, please tell me whether or not you strongly disagree, disagree, neither agree nor disagree, agree, or strongly agree.					
My relationship with (FATHER) is more important to me than almost anything else in my life				14.0	
I may not want to be with (FATHER) a few years from now				26.5	
I like to think of (FATHER) and me more as a couple than as two separate people				41.4	
I want this relationship to stay strong no matter what rough times we may encounter				46.8	
I am happy with my sexual relationship with (FATHER)				55.2	
I can trust that (FATHER) will not cheat on me with other people.				47.0	

Numbers correspond to the estimated signal percentage $\%$ Signal $= \frac{\left(\gamma^j_{t,k}\right)^2}{\left(\gamma^j_{t,k}\right)^2 + \mathrm{var}\left(\eta^j_{t,k}\right)} \times 100.$

Table B.2: Non-cognitive skill measurements

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
On a scale from 1 (not at all like your child) to 5 (very much like your child), how well does each of the following statements describe your child?					
Mother's report:					
(He/She) often fusses and cries		28.7%			
(He/She) gets upset easily		46.7			
(He/She) reacts strongly when upset		32.7			
Father's report:					
(He/She) often fusses and cries		10.0			
(He/She) gets upset easily		11.6			
(He/She) reacts strongly when upset		9.9			
Is this statement not true, somewhat or sometimes true, very true or often true for (CHILD)?					
(He/She) can't stand waiting; (he/she) wants everything now			44.4		
(He/She) is defiant			37.0		
(His/Her) demands must be met immediately			46.3		
(He/She) destroys (his/her) own things				48.6	56.1
(He/She) destroys things belonging to (his/her) family or other children			38.6	55.1	58.6
(He/She) is disobedient			42.4		
(He/She) is disobedient at home				45.7	53.7
(He/She) is disobedient at school or in childcare				25.5	44.1

(*Continued on next page*)

Table B.2: Non-cognitive skill measurements (continued)

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
(He/She) doesn't seem to feel guilty after misbehaving			20.7		
(He/She) is easily frustrated			33.0		
(He/She) gets in many fights			36.6	58.1	56.1
(He/She) hits others			38.7		
(He/She) hurts animals or people without meaning to			24.2		
(He/She) has angry moods			44.3		
(He/She) physically attacks people			46.8	62.2	65.2
Punishment doesn't change (his/her) behavior			29.2		
(He/She) screams a lot			38.3	46.0	52.8
(He/She) is selfish or won't share			31.9		
(He/She) is stubborn, sullen, or irritable			41.5	28.3	57.2
(He/She) has temper tantrums or hot temper			54.5	35.7	66.7
(He/She) is uncooperative			42.3		
(He/She) wants a lot of attention			22.0	19.5	39.4
(He/She) argues a lot				39.9	41.4
(He/She) brags or boasts				17.2	
(He/She) is cruel, bullies and shows meanness to others				54.4	57.8
(He/She) is easily jealous				20.0	
(He/She) shows off or clowns around				25.7	
(He/She) has sudden changes in mood or feelings				27.5	51.7
(He/She) talks too much				20.0	

(*Continued on next page*)

Table B.2: Non-cognitive skill measurements (continued)

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
(He/She) teases a lot				35.0	45.9
(He/She) threatens people				54.0	71.9
(He/She) is unusually loud				34.7	44.0
(He/She) sulks a lot					40.4
(He/She) is suspicious					35.3

Figures correspond to the estimated signal percentage $\%$ Signal $= \frac{\left(\gamma^j_{t,k}\right)^2}{\left(\gamma^j_{t,k}\right)^2 + \mathrm{var}\left(\eta^j_{t,k}\right)} \times 100$.

Table B.3: Cognitive skill measurements

Survey item	Year 0	Year 1	Year 3	Year 5	Year 9
Peabody Picture Vocabulary Test					
Child's score			20.8%	40.6	52.7
Mother/caretaker's score			57.9		
Weschler Intelligence Scale Tests					
Mother's score			24.7		
Father's score			11.7		
Child's digit span test score					31.1
Woodcock Johnson Tests					
Letter-Word Recognition				52.8	
Passage Comprehension					70.8
Applied Problems					61.1
Kindergarten teacher skill assessment				47.7	

Figures correspond to the estimated signal percentage $\%$ Signal $= \frac{\left(\gamma^j_{t,k}\right)^2}{\left(\gamma^j_{t,k}\right)^2 + \mathrm{var}\left(\eta^j_{t,k}\right)} \times 100$.

Figure B.1: Observed versus simulated measurement averages

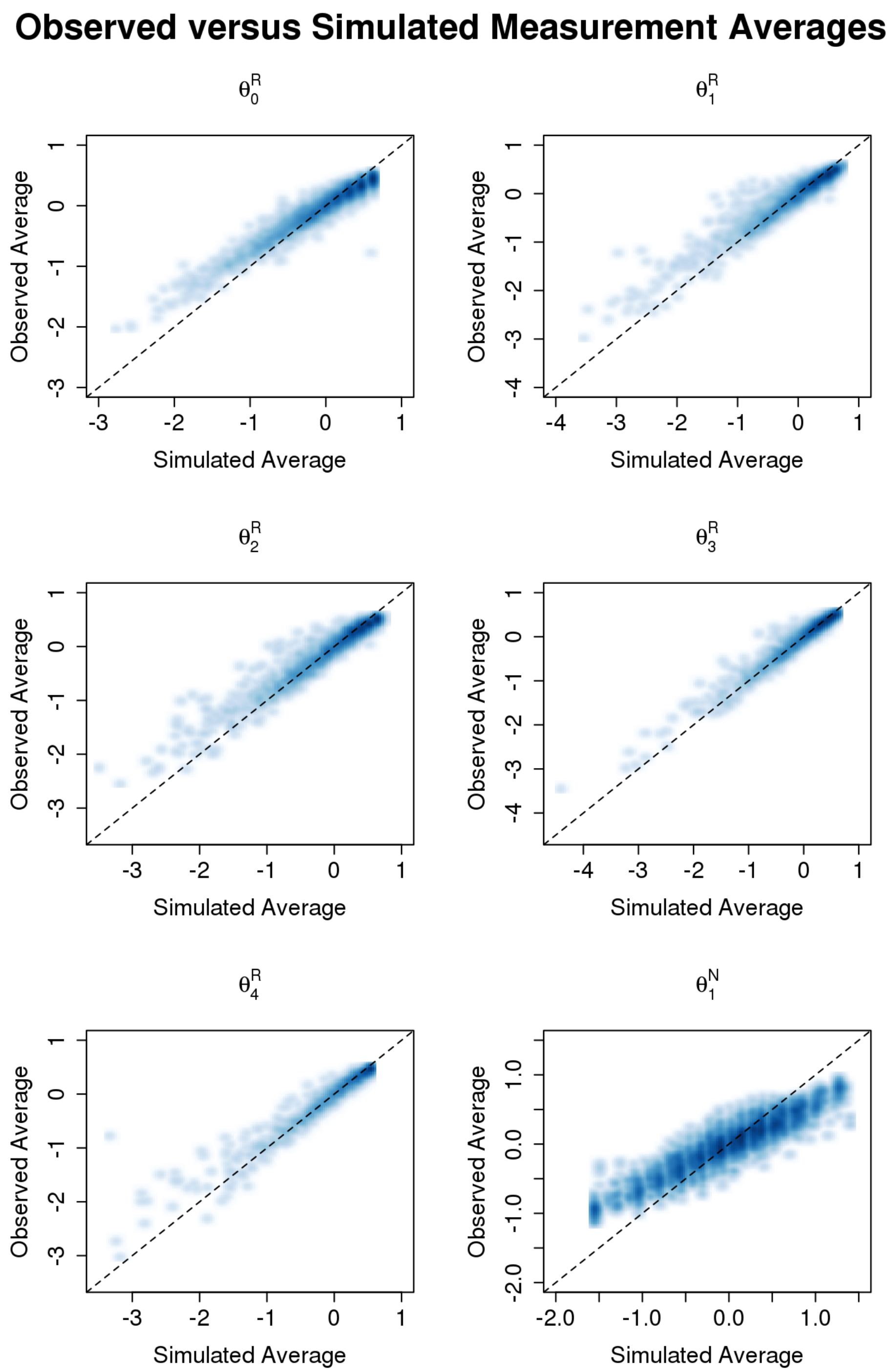
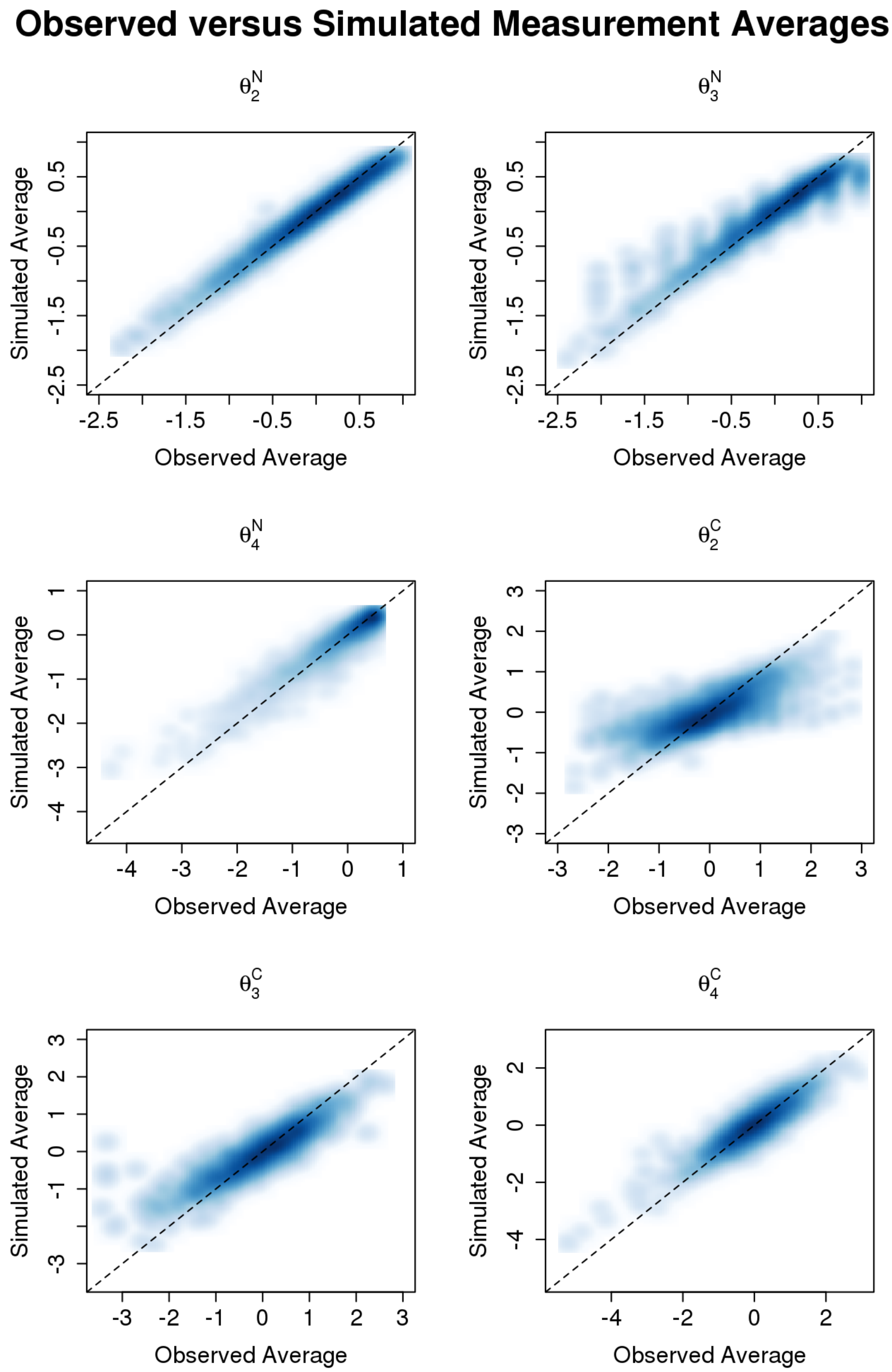


Figure B.1: Observed versus simulated measurement averages (continued)



Simulated measurement averages computed for each posterior draw and then averaged for each observation. Each measurement normalized to have a zero mean and unit standard deviation.

C Outcome tables

Table C.1: Year 9 outcomes

Outcome	$\gamma_j^{O,N}$	$\gamma_j^{O,C}$
Teacher rating of academic skills (5-point scale):		
Language and literacy skills	.056 (.070)	3.693 (.153)
Science and social studies	.062 (.071)	3.403 (.142)
Mathematics skills	.042 (.064)	3.101 (.124)
Special education services:		
Currently receiving special education services through an Individualized Education Program (IEP)*?	-.266 (.101)	-2.248 (.155)
Receiving any special education or related services because of Attention Deficit/Hyperactivity Disorder (ADD/ADHD)**	-.553 (.124)	-.748 (.127)
Received Behavior Intervention Plan (BIP), in or out of the classroom?***	-.874 (.116)	-.567 (.115)
Child response to early delinquency questions:		
Purposely damaged or destroyed property that wasn't yours?	-.632 (.068)	.056 (.074)
Taken or stolen something that didn't belong to you from another person or from a store?	-.702 (.073)	.063 (.083)
Taken some money at home that did not belong to you, like from your mothers' purse or from your parents' dresser?	-.788 (.086)	.078 (.096)
Cheated on a school test?	-.627 (.091)	-.185 (.100)
Had a fist fight with another person?	-.658 (.056)	.055 (.060)
Hurt an animal on purpose?	-.427 (.097)	-.132 (.112)

(*Continued on next page*)

Table C.1: Year 9 outcomes (continued)

Outcome	$\gamma_j^{O,N}$	$\gamma_j^{O,C}$
Gone into somebody's garden, backyard, house or garage when you were not supposed to be there?	-.595 (.086)	.063 (.099)
Run away from home?	-.866 (.138)	-.364 (.157)
Skipped school without an excuse?	-.614 (.140)	-.310 (.149)
Secretly taken a sip of wine, beer, or liquor?	-.597 (.110)	-.108 (.121)
Been suspended or expelled from school?	-.936 (.072)	.018(.074)
Written things or sprayed paint on walls or sidewalks or cars?	-.992 (.128)	.144 (.142)
Purposely set fire to a building, a car, or other property or tried to do so?	-.994 (.164)	-.020 (.185)
Avoided paying for things such as movies, bus or subway rides, or food?	-.282 (.094)	-.096 (.107)
Thrown rocks or bottles at people or cars?	-.886 (.104)	.083 (.116)

Outcome latent index equation: $U_j^O = \mathbf{x}'\boldsymbol{\alpha}_j^O + \gamma_j^{O,N} \cdot \theta_4^N + \gamma_j^{O,C} \cdot \theta_4^C + \eta_j^O$. Columns display posterior means of factor loadings (posterior standard deviations reported in parentheses).

*An Individualized Education Plan (IEP) is a federally mandated document that summarizes a disabled childs current level of performance and annual educational goals.

**Attention-Deficit/Hyperactivity Disorder (ADHD) is a brain disorder marked by an ongoing pattern of inattention and impulsivity that interferes with normal functioning.

***A Behavior Intervention Plan (BIP) is a list of steps teachers take to stop a child's problem behavior like disrupting the class, showing aggression toward the teacher or other children, or refusing to do classroom work.

Table C.2: Year 15 outcomes

Outcome	$\gamma_j^{O,N}$	$\gamma_j^{O,C}$
Youth's most recent grade in (4.0 scale) ...		
...English or language arts	.348 (.041)	.323 (.050)
...math	.315 (.041)	.315 (.048)
...history or social studies	.319 (.042)	.523 (.054)
...science	.269 (.040)	.365 (.050)
Youth's response to vandalism and violence questions:		
Paint graffiti or signs on someone else's property or in a public place	-.809 (.120)	.020 (.146)
Deliberately damage property that didn't belong to you	-.987 (.102)	.149 (.116)
Take something from a store without paying for it	-1.068 (.089)	.117 (.097)
Get into a serious physical fight	-.760 (.061)	.048 (.065)
Hurt someone badly enough to need bandages or care from a doctor or nurse	-.897 (.082)	.232 (.094)
Drive a car without its owner's permission	-.659 (.125)	.004 (.145)
Steal something worth more than $50	-1.031 (.148)	.134 (.165)
Go into a house or building to steal something	-.959 (.190)	.179 (.216)
Use or threaten to use a weapon to get something from someone	-.896 (.177)	-.185 (.201)
Sell marijuana or other drugs	-.968 (.138)	.372 (.168)
Steal something worth less than $50	-.980 (.084)	.378 (.096)
Take part in a fight where a group of your friends was against another group	-.740 (.070)	.069 (.081)
Were you loud, rowdy, or unruly in a public place	-.479 (.050)	.338 (.062)

(*Continued on next page*)

Table C.2: Year 15 Outcomes (continued)

Outcome	$\gamma_j^{O,N}$	$\gamma_j^{O,C}$
Questions on primary caregiver's survey:		
Has youth ever been arrested?	-.946 (.094)	-.215 (.098)
Does youth receive remedial math services?	-.345 (.065)	-1.126 (.083)
Does youth receive remedial English services?	-.237 (.067)	-1.256 (.089)
Does youth receive gifted and talented program services?	.163 (.061)	.918 (.080)
Does youth receive special education or related services?	-.515 (.079)	-2.232 (.126)
Has youth ever been suspended or expelled?	-1.033 (.064)	-.305 (.065)
Has youth repeated any grades?	-.406 (.072)	-.507 (.081)
Other questions on youth survey:		
Have you ever taken any honors courses in school?	.166 (.048)	1.084 (.067)
Do you ever skip school for a full day without an excuse?	-.569 (.064)	.076 (.077)
Have you been suspended or expelled from school in the past two years?	-.924 (.064)	-.196 (.066)
Have you ever smoked an entire cigarette?	-1.191 (.109)	.370 (.124)
Have you ever drank alcohol more than two times without parents?	-.656 (.062)	.372 (.076)
Have you ever been arrested or taken into custody by the police?	-1.031 (.116)	-.067 (.127)
Have you ever had sexual intercourse with anyone?	-.734 (.061)	.314 (.072)
Have you ever tried marijuana?	-.822 (.062)	.384 (.071)

Outcome latent index equation: $U_j^O = \mathbf{x}'\boldsymbol{\alpha}_j^O + \gamma_j^{O,N} \cdot \theta_4^N + \gamma_j^{O,C} \cdot \theta_4^C + \eta_j^O$ Columns display posterior means of factor loadings (posterior standard deviations reported in parentheses).

D General model

Suppose the period $t+1$ value functions can be written as

$$
\begin{aligned}
V_{t+1}^{M}(\varsigma_{t+1}) &= \sum_{t'=t+1}^{T} r^{t'-(t+1)}\alpha_{t'}^{M} + \eta_{t+1}^{M} + \rho V_{t+1}(\varsigma_{t+1}) \\
V_{t+1}^{F}(\varsigma_{t+1}) &= \sum_{t'=t+1}^{T} r^{t'-(t+1)}\alpha_{t'}^{F} + \eta_{t+1}^{F} + (1-\rho) V_{t+1}(\varsigma_{t+1})
\end{aligned} \tag{D.2}
$$

where the summation terms in each equation represent the expected present values of the parents' outside options and $V_{t+1}(\cdot)$ is a known total net value function. From (4.9), we see that (D.2) holds for $t = T-1$. If $R_t = 1$, the expected payoffs from choosing $R_{t+1} = 0$ are given by

$$
\begin{aligned}
&S_t^M + r\mathbb{E}\left(V_{t+1}^{M}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right) = \\
&\quad \sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^{M} + \eta_t^M + \rho\Bigg(g_t^S(\theta_t^H) + r\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right)\Bigg)
\end{aligned} \tag{D.3a}
$$

$$
\begin{aligned}
&S_t^F + r\mathbb{E}\left(V_{t+1}^{M}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right) = \\
&\quad \sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^{F} + \eta_t^F + (1-\rho)\Bigg(g_t^S(\theta_t^H) + r\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right)\Bigg)
\end{aligned} \tag{D.3b}
$$

while the expected payoffs from choosing $R_{t+1} = 1$ are given by

$$S_t^M + \rho(U_t - S_t) + \mathbb{E}\left(V_{t+1}^M(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right) =$$
$$\sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^F + \eta_t^M + \rho\Bigg(g_t^S(\theta_t^H) + U_t - S_t + r\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right)\Bigg) \tag{D.4a}$$

$$S_t^F + (1-\rho)(U_t - S_t) + r\mathbb{E}\left(V_{t+1}^F(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}) \mid \boldsymbol{\theta}_t, R_t\right) =$$
$$\sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^F + \eta_t^F + (1-\rho)\Bigg(g_t^S(\theta_t^H) + U_t - S_t + r\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right)\Bigg) \tag{D.4b}$$

By subtracting (D.3a) and (D.3b) from (D.4a) and (D.4b), respectively, we see that mother and father are both better off staying together through period t if and only if

$$U_t - S_t +$$
$$r\Bigg(\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right) - \mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right)\Bigg) > 0 \tag{D.5}$$

The parents' period t values functions can be written as

$$\begin{aligned} V_t^M(\varsigma_t) &= \sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^M + \eta_t^M + \rho V_t(\varsigma_t) \\ V_t^F(\varsigma_t) &= \sum_{t'=t}^{T} r^{t'-t}\alpha_{t'}^F + \eta_t^F + (1-\rho) V_t(\varsigma_t) \end{aligned} \tag{D.6}$$

where

$$V_t(\varsigma_t) = g_t^S(\theta_t^H) + r\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, R_t, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t, R_t\right) +$$
$$R_t \cdot \left(\begin{array}{l} \max\left\{ U_t - S_t + r\Bigg(\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right) - \right. \\ \qquad\qquad \left. \mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right) \Bigg), 0 \right\} \end{array} \right) \tag{D.7}$$

We conclude by induction that (D.6) holds for $t = 0, 1, \cdots, T$, where the total net value functions are defined recursively by (4.8) and (D.7). The sum of the first two terms in (D.7) represents the total net value in period t if the parents separate before period $t+1$, while the term enclosed by the large parentheses represents the option value of beginning period t in a relationship. Separation affects next period's total net value directly by eliminating the option to continue the relationship and indirectly by altering the evolution of the child's human capital.

The policy functions are given by

$$R_{t+1}(\varsigma_t) = R_t \cdot \mathbb{1}\left\{p_t(\boldsymbol{\theta}_t) - \eta_t^S > 0\right\} \tag{D.8}$$

where

$$p_t(\boldsymbol{\theta}_t) = g_t^U(\boldsymbol{\theta}_t) - \alpha_t^S - g_t^S(\theta_t^H) +$$
$$r\Bigg(\mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 1, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 1) \mid \boldsymbol{\theta}_t\right) - \mathbb{E}\left(V_{t+1}(\boldsymbol{\theta}_{t+1}(\boldsymbol{\theta}_t, 1, 0, \boldsymbol{\epsilon}_{t+1}), \boldsymbol{\eta}_{t+1}, 0) \mid \boldsymbol{\theta}_t\right) \Bigg) \tag{D.9}$$

E Priors

Independent $N(0,3)$ priors were used for all coefficients in (5.2).[1] A Beta$(1.5, 1.5)$ was used as the prior distribution for the standard deviations of the error terms in (5.2), while the Lewandowski-Kurowicka-Joe (LKJ) distribution was used as the prior on the error term correlation matrices (Lewandowski et al. (2009)). The kernel of the latter distribution is given by

$$\mathrm{LKJ}(\Sigma \mid \eta) \propto \det(\Sigma)^{\eta-1}$$

where Σ is a correlation matrix and η is a shape parameter. η was set so that the prior standard deviation of each correlation coefficient equals .4.

As mentioned in Section 5.2, the sign of at least one factor loading parameter per latent variable must be restricted. In practice, restricting the signs on all the factor loadings improved the algorithm's ability performance. Preliminary estimates of the measurement system were obtained using normalized averages of the observed measurements as proxies for the latent variables. The prior for factor loading $\gamma^j_{t,k}$ was then given by

$$\gamma^j_{t,k} \sim \mathrm{Gamma}\left(\left(\hat{\gamma}^j_{t,k}\right)^2, \hat{\gamma}^j_{t,k}\right)$$

where $\hat{\gamma}^j_{t,k}$ denotes the preliminary estimate.

The cutoff parameters in (5.4) were drawn from a truncated multivariate normal distribution with mean set equal to their preliminary estimates and covariance matrix equal to an identity matrix multiplied by 3. The support of the distribution was truncated to ensure the cutoffs were

[1] The priors for the coefficients on $\mathbf{x}$ were actually placed on a linear transformation of the parameters. Let

$$\mathbf{X} = (\mathbf{x}_1\ \mathbf{x}_2\ \cdots\ \mathbf{x}_n)'$$

denote the matrix of observed covariates. Let $\mathbf{X} = \mathbf{QR}$ denote the QR decomposition of $\mathbf{X}$. The rows of $\mathbf{Q}$ were used as the covariate vectors in estimation and priors were placed on vectors of the form $\mathbf{R}\boldsymbol{\alpha}$. This procedure ensures the priors have a reasonable scale, as the columns of $\mathbf{Q}$ all have unit standard deviation when a constant is included in $\mathbf{x}$.

ordered. The prior distribution for means for the continuous measurements in (5.5) were drawn from a normal distribution with standard deviation 3 and mean equal to the measurement's sample average. The measurement error standard deviations in (5.10) were drawn according to

$$\sigma_{t,k}^{j} \sim \mathrm{Gamma}\left(\left(\hat{\sigma}_{t,k}^{j}\right)^{2}, \hat{\sigma}_{t,k}^{j}\right)$$

where $\hat{\sigma}_{t,k}^{j}$ denotes the preliminary estimate of $\hat{\sigma}_{t,k}^{j}$.