From Cradle to Consequence: The Dynamic Impact of Parental Behaviors on Teenage Outcomes Among Fragile Families

Danny Turkson†

University of North Carolina, Chapel Hill

Click here for the latest version

Abstract

How do parental behaviors at different stages of childhood causally shape a child's teenage outcomes? This paper estimates the causal impacts of observed parental behaviors (marriage/cohabitation, employment, welfare receipt, substance use, and criminal activity) on cognitive ability and on teenage educational attainment, substance use, and criminal behavior among fragile families. I jointly estimate a dynamic model of parental behaviors, child cognitive skill production, and teenage outcome production using data from the Fragile Families and Child Wellbeing Study. The model accounts for theoretically relevant and empirically tested exclusion restrictions, and permanent and time-varying family unobserved heterogeneity. A key finding is that parental behaviors shape teenage outcomes primarily through direct behavioral pathways rather than through cognitive skill formation, suggesting that policies aimed solely at boosting cognitive test scores are unlikely to undo the long-run damage of adverse parental trajectories: relative to more stable trajectories, sustained marital instability and economic disadvantage each increase adverse outcomes (school suspension, police contact, and early sexual initiation) by at least 8 percentage points. Three policy experiments, involving a sin tax, father-marriage promotion, and peer-environment upgrades, illustrate that sustained, multi-dimensional interventions in early childhood and adolescence offer the greatest potential for improving long-run outcomes in fragile families. Consistent with Heckman's critical-window hypothesis, sustaining paternal co-residence through age five reduces grade failure and school suspension by at least 3.4 percentage points, an effect that vanishes once the same intervention begins only after this window has closed.

Keywords: Parental behavior, child development, fragile families, dynamic structural model.
JEL Codes: J10, I30, K42, I24, I14

---

†I am grateful to my dissertation committee (Donna Gilleskie, Jane Fruehwirth, Luca Flabbi, Désiré Kédagni, and Jacob Kohlhepp) for their guidance and feedback. I also thank participants at the UNC Applied Micro Workshop for helpful comments and suggestions. Access to the Fragile Families and Child Wellbeing Study (FFCWS) data was generously provided by the Center for Research on Child and Family Wellbeing at Princeton University.
PhD Candidate, Department of Economics, University of North Carolina at Chapel Hill. dturkson@unc.edu

1 Introduction

Family structure in the United States has transformed over the past four decades. A notable change is the increase in the number of children born to unwed parents. This trend has resulted in many children being born to disadvantaged women (i.e., women who are racial minorities, and/or poor, and/or lower educated), thereby giving rise to the concept of 'Fragile Families'. Children born into such families often live with either a single mother or a mother who is living with a partner to whom she is not married or is not the biological father of the child. In the 1980s fewer than 20 percent of children born each year in the United States were born to unmarried women. However, since 2007, this figure has more than doubled, with the trend remaining consistently above 40 percent (Craigie et al., 2012). Fragile families are mostly characterized by family instability, parental incarceration, substance use disorders, and fewer parental resources compared to married parents (Osborne and Berger, 2009; Wildeman and Western, 2010; Waldfogel et al., 2010). Approximately 50 percent of fathers in a fragile family have at least one of the following risk factors: physical abuse, substance use disorder and incarceration (Waller and Swisher, 2006).

Parental involvement in a child's life is foundational to healthy development (Cunha and Heckman, 2007). Family inputs have been proven to contribute significantly to improving educational and cognitive outcomes and reducing incidence of criminality for minor offenses (Todd and Wolpin, 2003; Akee et al., 2010). Yet the effects of parental inputs are not uniform across childhood. Recent evidence shows that the timing and sequencing of parental input matter as much as their magnitude or level, and the parental behavioral trajectories sustained across the developmental period is what ultimately shapes children's schooling, earnings, and wellbeing (Carneiro et al., 2021; Chang et al., 2025). More broadly, differences in cognitive and behavioral outcomes among children are largely shaped by the environments children are exposed to (Heckman and Mosso, 2014). Parental behavior is among the most immediate of those environmental forces, but it does not operate in isolation: peer influences and the policy and institutional landscape children grow up within also leave lasting marks. These parental inputs represent optimal allocations made jointly with decisions regarding marriage/cohabitation, employment, welfare receipt, substance use, and criminal behavior over time. Yet these inputs are hard to find and measure in longitudinal data (Cunha et al., 2010).

It is difficult to predict how a child will fare when a parent is marrying, divorcing, and remarrying; in and out of a job; on and off welfare; in and out of rehabilitation; intermittently or continually incarcerated throughout their childhood. Unbiased and non-spurious measurement of these impacts on child outcomes is elusive. Most of the research has been correlational without establishing any causal impact, and the few that attempt to do so neglect the dynamic and jointly determined nature of parental behaviors during the most sensitive stages of the child's life. The studies that address the dynamic nature of parental behaviors have modeled the impact of employment and welfare and/or substance use on child outcomes (mostly educational outcomes), or the impact of parental incarceration on child outcomes. No work, to my knowledge, has studied jointly the causal impact of important correlated parental behaviors (marriage/cohabitation, employment/welfare, substance use, and criminal behavior) on children's outcomes. Moreover, the previous research that controls for child or family heterogeneity does not account for the joint nature of these parental behaviors, and estimates the effects of such behaviors on child outcomes separately (ignoring both correlation through family unobservables as well as across child outcomes).

The objective of this paper is to estimate the causal impacts of observed parental behaviors (marriage/cohabitation, employment, social services assistance, substance use, and criminal behavior) at different ages of child development on the child's teenage outcomes, specifically, cognitive ability, educational attainment, substance use, and criminal outcomes. Using data from the Fragile Families and Child Wellbeing Study (FFCWS) (now known as Future of Families and Child Wellbeing Study), I jointly estimate children's teenage outcomes with parents' age-specific behaviors, accounting for both permanent and time-varying unobserved family heterogeneity, and allowing for the effects of parental behavioral trajectories to operate through the child's cognitive skill development, providing an indirect channel of influence, as well as allowing a direct effect, which may indicate alternative channels.

This paper makes several contributions to the existing literature. The first and most fundamental contribution is the joint estimation of multiple parental behaviors and child outcomes within a single structural framework. Unlike previous studies that estimate the effects of individual parental behaviors in isolation, I model marriage/cohabitation, employment, welfare receipt, substance use, and criminal behavior as jointly determined choices that evolve dynamically over the child's developmental lifecycle. Since parents do not make decisions independently of one another, studies that ignore this joint nature risk introducing omitted variable and endogeneity biases. I address these biases by estimating all parental behaviors jointly with child outcomes using Full Information Maximum Likelihood (FIML), theoretically relevant instrumental variables for identification, and Discrete Factor Random Effects (DFRE), which accounts for both permanent and time-varying unobserved family heterogeneity.

A second contribution is the decomposition of the total causal effect of each parental behavior on teenage outcomes into two distinct channels: (1) an indirect cognitive effect, which captures the impact that operates through the parental behavior's cumulative influence on the child's cognitive skill development over the pre-teenage years; and (2) a direct behavioral effect, which captures any additional impact of parental behavior on teenage outcomes holding the child's cognitive skill fixed. This decomposition sheds new light on the mechanisms through which parental decisions shape long-run child outcomes, distinguishing between behaviors that primarily affect teenagers through channels outside cognitive development and those that operate primarily by shaping cognitive trajectories during the critical early developmental window.

Third, this study is unique in its many-to-many design: multiple parental behaviors are estimated jointly alongside multiple child outcomes spanning educational attainment, substance use, and deviant behavior. Unlike Attanasio et al. (2020), who consider observed parental investment as a single time-varying input, I model the much more commonly observed parental behaviors that are made jointly with unobserved parental investments. Finally, the use of the FFCWS contributes to our understanding of this process because most of the economics literature studying the dynamic nature of the parent-child relationship has relied on the NLSY, PSID, Add Health, or SIPP. Since the FFCWS was designed specifically to follow at-risk families, it is particularly well-suited for studying parental behavior and child outcomes in fragile families.

The findings reveal that the timing, persistence, and joint nature of parental behaviors matter for teenage outcomes. Stable family structure during childhood (i.e., the continuous presence of the biological father) is among the strongest predictors of favorable outcomes: relative to an unstable marital trajectory, sustained marriage to the biological father significantly improves college graduation prospects, reduces school suspension, and reduces police contact by at least 8.0 percentage points (p.p.). Conversely, persistent welfare dependence and maternal unemployment generate the largest detrimental effects among the outcomes studied, significantly increasing class failure by 11.7 p.p., police contact by 26.1 p.p., and early sexual activity by 8.1 p.p.. Parental substance use and criminal record, when varied jointly, show only one statistically significant effect: teenagers under the most adverse conditions are 28.8 p.p. less likely to be stopped by police than under the most favorable counterfactual conditions, a counterintuitive result consistent with a deterrence- or incapacitation-type channel. The decomposition of total effects into direct behavioral and indirect cognitive channels reveals that most of the impact of parental behavior on teenage outcomes operates through direct pathways rather than through cognitive skill formation. The indirect cognitive channel, while consistently present, accounts for only a modest share of the total effect across all behaviors and outcomes. This finding suggests that parental behaviors shape adolescent outcomes through mechanisms beyond cognitive development alone, that includes non-cognitive development and peer exposure, among others. Therefore, policies targeting cognitive skill formation alone are unlikely to fully offset the downstream consequences of adverse parental behavioral trajectories.

Three policy experiments further illustrate these findings. A Sin-to-Skill policy that combines excise taxes on beer and cigarettes with reinvestment into county-level education significantly reduces welfare receipt by up to 10.9 p.p. and curbs parental smoking and marijuana use and reduces teenage drinking by up to 3.6 p.p.. A Father-to-Foundation policy that sustains biological-father co-residence and non-incarceration operates as a dose-response: a one-year intervention generates no statistically significant gains, but sustaining co-residence through age five significantly reduces class failure by 3.9 p.p., school suspension by 3.4 p.p., and marijuana use by 2.1 p.p.. A direct placebo-style test confirms this timing story: imposing the identical intervention exclusively after age five, once Heckman's critical window has closed, generates no significant improvements and, in several cases, reverses in sign. A Peer-to-Promise policy that upgrades a teenager's peer environment shows a similar dose-response, with modest shifts significantly reducing teenage drinking alone, and larger shifts additionally and significantly reducing marijuana use, police contact, sexual activity, school suspension, and college graduation expectations by at least 3.1 p.p.. The policy results underscore that improving child outcomes in fragile families requires sustained, multi-dimensional interventions rather than isolated policy changes targeting a single behavior at a single point in time.

2 Related Literature

A large body of research has examined how parental behaviors affect children's outcomes, but the literature has largely studied these behaviors in isolation, that is one behavior at a time, and rarely in a structural framework that accounts for their joint and dynamic nature.[1] This paper contributes to four related strands of this literature: parental relationship dynamics, parental employment, parental incarceration, and parental substance use.

The link between maternal employment and child cognitive outcomes was studied fairly extensively. Bernal (2008), using NLSY data, developed and estimated a dynamic structural model of single mother's employment and childcare decisions to evaluate the effects on child cognitive ability. They concluded that children of mothers working full time and using formal childcare had lower test score ability, and therefore advocated for longer parental leave policies. Liu et al. (2010) similarly used the NLSY to estimate the effect of maternal employment and migration on child educational outcomes. They found a significant negative relationship between maternal employment and child outcomes after controlling for endogeneity, but a positive relationship when endogeneity was not addressed; highlighting why causal identification mattered in this literature. Schaller and Zerpa (2019) studied the impact of parental job loss on children's health using a child fixed effects specification and found that involuntary job loss was detrimental to children's physical and mental health, with the effects being worse for lower socioeconomic status families. Agostinelli and Sorrenti (2021) explored the EITC expansions to separately identify the income and substitution effects of maternal labor supply on children's cognitive and behavioral development. They found that while higher family income improved cognitive scores, increased maternal working hours negatively impacted noncognitive outcomes. The negative labor supply effect was concentrated among disadvantaged mothers, precisely the population that faced the most constrained access to quality alternative childcare. Also, Attanasio et al. (2020) endogenized parental inputs at different ages of

[1] See the summary of literature of parental behaviors on child outcomes in Appendix Table A.27

the child's development and concluded that the degree of persistence of parental inputs changed over the child's life, underscoring that the timing of parental behavior mattered and not just the behavior level.

The effects of parental incarceration on children's outcomes have been studied more in the sociology and public policy literatures than in economics, and few studies address unobserved heterogeneity directly. Cho (2009b) examined maternal imprisonment effects on cognitive outcomes among elementary school children in Chicago and, after accounting for unobserved child heterogeneity, found no significant association with reading or math test scores. Cho (2009a) extended this work to grade retention using a difference-in-difference approach with propensity score matched controls, and found that children of incarcerated mothers are actually less likely to experience grade retention in the immediate years of the jail term. Wildeman (2010) used OLS and fixed effects models and found that paternal incarceration significantly increases physical aggression among boys but not girls. More recently, a series of studies has used the random assignment of defendants to judges with different incarceration tendencies as an instrument to establish causal effects. Norris et al. (2021), using Ohio administrative data, found that parental and sibling incarceration substantially reduces criminal activity before age 25 but has no meaningful effect on academic outcomes. Similarly, Dobbie et al. (2018) used the same judge IV approach in Sweden and found that parental incarceration increases teenage pregnancy and crime and reduces early employment, with the largest effects among the most disadvantaged families. Arteaga (2023) applied the same IV in Colombia and found that parental incarceration increases children's educational attainment by approximately one year. Kailaheimo-Lönnqvist et al. (2022) used a cousin fixed effects approach and found a negative association between parental criminality and children's educational attainment at all stages, with the magnitude varying by whether the crime was violent or non-violent.

The literature on parental substance use documents significant effects on children across a range of outcomes. Bailey et al. (2018) used fixed effects estimation and found that having a substance-using family member increases the probability of child substance use, both concurrently and one year later. Schmidt and Tauchmann (2011) documented substantial heterogeneity in the intergenerational transmission of alcohol consumption, showing that estimated effects of parental drinking at the 95th percentile of the distribution are more than three times larger than at the median. Cunningham and Finlay (2013) used methamphetamine prices as an instrument and concluded that children with meth-using parents have a higher probability of entering foster care and experiencing child neglect and abuse. On long-run effects, Balsa (2008) found that having a problem-drinking parent is associated with longer periods out of the labor force, higher unemployment, and lower wages in adulthood, with worse outcomes for male children. Chatterji and Markowitz (2001) used mother-child and family fixed effects and found that maternal cocaine and marijuana use is positively associated with children's behavioral problems, while the effect of alcohol use is less consistent. Moreover, Ali and Dwyer (2009) found that parents and peers significantly influence adolescent smoking, with peer effects generally larger than parental effects.

A smaller but directly relevant set of studies uses the FFCWS, the same data used in this paper. Geller et al. (2009) found that paternal incarceration had no significant association with age-3 cognitive test scores but was positively associated with externalizing behavior patterns like temper tantrums. Osborne and Berger (2009) found that parental substance use is associated with behavioral problems among 3-year-olds, with associations being larger when both parents use substances and with father's substance use having a larger association than mother's. Craigie et al. (2012) found that stable marriage is associated with the highest age-5 cognitive scores, lower obesity, and fewer asthmatic episodes relative to step parent families. Geller (2013) found that paternal incarceration at any child-age significantly decreases father-child contact during and after jail time. Turney (2017) used propensity score matching to examine age-9 outcomes and found no significant difference in cognitive skills between children with and without incarcerated parents, but found persistent behavioral differences. Cummings and Gearhardt (2020) found that maternal illicit drug use significantly predicted increased sugary food and beverage intake among teenagers at age 15.

The key gap in the literature is that parental behaviors are studied one at a time and at a single point in time, even though parents make marital, employment, welfare, substance use, and criminal decisions jointly and simultaneously. Studies that do not account for this joint nature risk omitted variable and endogeneity biases. Additionally, treating parental behavior as a snapshot misses the importance of behavioral timing, sequencing, and persistence across the child's development. Recent work has begun to formalize this intuition. Carneiro et al. (2021), using administrative data from Norway, showed that the timing of parental income over childhood matters for children's schooling and earnings, with early and late childhood income both outperforming middle-childhood income. Chang et al. (2025) extended this approach by modeling both parental income and family structure as full developmental trajectories, finding that their effects on adult income are strongest during middle childhood and adolescence. These studies make a compelling case that what happens across the entire developmental period, not just at any single point, is what shapes long-run child outcomes.

Carneiro et al. (2021) and Chang et al. (2025) focus on income and family structure in their frameworks applied to general or high-income populations, and neither models the full range of behavioral choices that characterize fragile families. Using the FFCWS, Chan and Liu (2025) develop a quasi-structural framework that jointly models relationship transitions, labor supply, welfare participation, and time investment by both biological and social fathers, using a control function approach to identify the causal impact of these inputs on children's cognitive and noncognitive outcomes at age five. Lessem and Sanders (2025) take a structural approach on the FFCWS but narrow their scope to relationship and employment decisions, treating the father as a fixed latent type rather than an active behavioral agent and holding substance use, welfare, and criminal behavior outside the model. None of these papers accounts for the joint determination of multiple parental behaviors across the full span of childhood, from birth through adolescence, which is central to the fragile family experience where incarceration, substance use disorders, and family instability tend to reinforce one another over time. This paper jointly models parental behaviors (marriage/cohabitation, employment, welfare receipt, substance use, and criminal behavior) as endogenously determined dynamic choices, with child cognitive achievement and teenage outcomes; accounts for permanent and time-varying unobserved heterogeneity; and allows for both direct effects of behavioral trajectories and indirect effects operating through the cognitive development channel.

3 Data

3.1 Data Source

The data used in this study is from the Future of Families and Child Wellbeing Study (FFCWS) formally known as the Fragile Families and Child Wellbeing Study (FFCWS, 2023). These data are collected by the Office of Population Research at Princeton University. The FFCWS is a panel study of a multistage sample: first, U.S. cities with population over 200,000 in the year 1998 were sampled, then sampling of hospitals within cities, and finally sampling births within hospitals. Figure 3 shows the cities and the proportion of sample allocated to each city.[2] The resulting sample was 4898 children born with an oversampling ratio of 3 to 1 of births to unmarried mothers, resulting in a considerable number of Black, Hispanic and low-income families. Table 1 describes the distribution of US data (from 1998) and baseline FFCWS data (1998-2000), showing the over sample of disadvantaged families in the survey.

The first wave interviews were conducted between 1998-2000 in the hospitals where the mothers delivered. The subsequent waves were conducted between 1999-2001, 2001-2003, 2003-2006, 2007-2010, and 2014-2017 when the children were expected to be approximately age 1, 3, 5, 9 and 15, respectively. Across the 6 waves, mother, father, primary care giver (PCG) if not parent, child and teacher were interviewed, and in-home assessments were conducted as shown in Figure 4. Final analysis sample for this study is 1615 families who were present in all the waves, with children of the same age at the time of interview and non-missing child outcome variables.[3]

[2] The cities were shortlisted using intensity of three dimensions (welfare, child support, and labor market indicators) and the final cities in each strata were selected randomly, with the selection probability for each city proportional to its population Selection with probabilities proportional to size assures the representativeness of the sample.

[3] See Appendix A1.1 for full details on sample selection.

3.2 Variable Description

Endogenous Parental Behaviors and Teenage Outcome Variables

The time-varying endogenous variables in this study are marriage/cohabitation, employment, welfare receipt, substance use, criminal behaviors and child cognitive achievement. Table 2 shows the distribution of these endogenous variables across the different waves. Over 30 percent of mothers were on welfare when the child was born, but this percentage falls as the child ages. The presence of biological dads in the life of the child decreases as the child gets older, because of the changes in the marital status of the mothers and the prevailing incarceration rate among the dads.

I also observe a measure of child's cognitive ability, specifically performance on the Peabody Picture Vocabulary Test (PPVT). The PPVT examines the children on verbal ability and assessed receptive vocabulary knowledge and comprehension of spoken English. Dunn and Dunn (1965) notes that the test also helps in establishing rapport, testing preschool children, and screening for both giftedness and mental retardation. The PPVT score is in the range of $41 - 185$ with an average 112 and standard deviation 20.07. However, because different versions of the test are administered at different ages, I standardize the scores across all ages to have a mean of 80 and a standard deviation of 10.

Table 3 presents the summary statistics of the modeled teenage outcomes. The outcome variables for this study are categorized in three different groups: educational, substance use and deviant outcomes. The educational outcomes include ever failed a class, suspended from school and the child's subjective likelihood of graduating college. Close to 1 in 2 of these children have ever failed a class and approximately 1 in 3 have faced some form of suspension from school. Thirteen percent have drunk an alcoholic drink more than two or three times when they were not with their parents, and 18 percent have used marijuana. Finally, 1 in 4 of these children have been stopped by the police and 15 percent have ever had sex.

3.3 Descriptive Motivation from FFCWS

Table 4 presents unconditional differences in mean teenage outcomes by parental behaviors observed, highlighting substantial disparities across educational, substance use, and deviant outcomes. Across nearly all outcomes, children exposed to more stable and less risky parental behavior environments exhibit markedly better outcomes during adolescence. Starting with mother's marital status, children whose mothers are always married to the biological dad consistently perform better across all outcomes relative to any other family structure. Differences by mother's employment status are also pronounced, though somewhat smaller in magnitude. Children whose mothers are always employed are significantly less likely to fail a class, be unlikely to graduate college, or be suspended from school. In contrast, differences in alcohol and marijuana use are relatively small and statistically insignificant, suggesting that employment stability primarily contributes to academic and institutional outcomes rather than substance use.

The largest disparities in the table arise from mother's welfare status. Children whose mothers are always on welfare experience worse outcomes compared to children whose mothers are not always on welfare. These large gaps suggest that persistent welfare receipt is a strong indicator for sustained socioeconomic disadvantage. Turning to maternal substance use, drinking shows relatively modest differences, with only a weakly significant effect on college graduation prospects. In contrast, smoking and especially marijuana use are strongly associated with worse adolescent outcomes. Maternal smoking is linked to significantly higher probabilities of class failure, suspension, marijuana use, police contact, and early sexual activity. Maternal marijuana use is associated with some of the largest differences in the table, including a 26.6 percentage point increase in school suspension and a 15.6 percentage point increase in teen marijuana use. Finally, parental criminal involvement, both maternal convictions and paternal incarceration, are strongly correlated with adverse outcomes. Children whose mothers have a criminal record or whose dads have been incarcerated are more likely to fail classes, be suspended, engage in substance use, encounter police, and initiate sex early.

While these patterns in Table 4 are descriptive, they motivate the need for a dynamic structural framework that jointly models parental behaviors and accounts for unobserved heterogeneity. The consistency and magnitude of these differences suggest that parental behaviors co-evolve and jointly shape the environments in which children develop, reinforcing the importance of modeling them jointly rather than in isolation.

4 Empirical Framework

The empirical model is grounded in a structural framework where forward-looking parents optimally select their behaviors (marital status, employment, welfare receipt, substance use and criminal record) each period to maximize lifetime utility. Parents internalize not only the current cognitive development of the child but also the child's future teenage outcomes, which are realized at the end of the final period. The jointly chosen parental behaviors influence the child's pre-teenage cognitive outcomes over time and, through that cognitive trajectory, shape the child's teenage outcomes in education, substance use and delinquency.

4.1 Model

4.1.1 Model Timing and Per-Period Behavioral Alternatives

A family enters each period t with an observed information set, Ω_t, which consist of mother's histories of marriage/cohabitation, M_{t-1}; employment, E_{t-1}; welfare receipt, W_{t-1}; substance use,

$\mathbf{S}_{t-1}$ (which includes alcohol use, smoking, marijuana use); criminal record, R_{t-1}; dad's history of imprisonment (jail time), J_{t-1}; and the focal child's cognitive achievement from the end of the previous period, θ_{t-1}. This information set also includes exogenous characteristics of mother, dad and child (that is $\mathbf{X}_t = \{X_t^m, X_t^d, X_t^c\}$); and the state or county-level policies and supply side conditions, $\mathbf{Z_t}$. This vector of information known at the beginning of period t is denoted by $\Omega_t = (M_{t-1}, E_{t-1}, W_{t-1}, \mathbf{S}_{t-1}, R_{t-1}, J_{t-1}, \theta_{t-1}, \mathbf{X}_t, \mathbf{Z}_t)$. The state and local policy environment is denoted by the vector $\mathbf{Z}_t = (Z_t^M, Z_t^E, Z_t^W, \boldsymbol{Z}_t^{S_l}, Z_t^R, Z_t^J, Z_t^\theta, Z_t^p)$. The superscripts indicate the variables in $\mathbf{Z_t}$ that relate to particular endogenous parental behaviors and child cognitive achievement; namely, cost or supply-side variables that affect marital/cohabitation status, Z_t^M; employment, Z_t^E; welfare receipt, Z_t^W; substance use, $\boldsymbol{Z}_t^{S_l}$; criminal record, Z_t^R; dad's serving jail time, Z_t^J; child cognitive achievement, Z_t^θ; and peer characteristics affecting teenage outcomes, Z_t^p.

In each period t a parent chooses among the alternatives associated with each behavior, m_t, e_t, w_t, $\left\{s_t^l\right\}_{l=1}^3$, r_t and j_t. Specifically:

- Marital/cohabitation, $m_t \in \{0, 1, 2\}$: single (no dad present); married/cohabiting with biological father (biological dad present); married/cohabiting with another man (another dad present)

- Employment, $e_t \in \{0, 1\}$: not employment; employed

- Welfare receipt, $w_t \in \{0, 1\}$: do not receive welfare; receive welfare

- Substance use, $s_t^l \in \{0, 1\}$: do not use the l^{th} substance; use the l^{th} substance. Where $l \in \{\text{alcohol, cigarettes, marijuana}\}$

- Conviction, $r_t \in \{0, 1\}$: no conviction; have conviction

- Dad's incarceration, $j_t \in \{0, 1\}$: not in-jail; in-jail

The model spans six periods corresponding to key developmental stages: t_1 (Birth, age 0), t_2 (Age 1), t_3 (Age 3), t_4 (Age 5), t_5 (Age 9), and t_6 (Age 15). Periods t_1 through t_5 constitute the pre-teenage developmental stage, during which parental behaviors jointly shape the child's cognitive skill. Period t_6 is the teenage period, in which parents continue to make behavioral choices and the child's teenage outcomes are observed at the close of the period. Prior to the onset of period t_1, each family draws a permanent type μ, which captures time-invariant unobserved heterogeneity in preferences, endowments, and circumstances that persist across the entire lifecycle. At the beginning of each period t, the family additionally receives an idiosyncratic shock ν_t, capturing time-varying unobserved heterogeneity that evolves over the developmental process. For each period t, the family enters with an observed information set Ω_t. Conditioning on their permanent type μ and the current period shock ν_t, parents observe the current state and make their behavioral choices, after which the child's cognitive skill is updated. Figure 1 illustrates the within-period timing of events during the pre-teenage stage. In the terminal period t_6 (teenage period), the child's teenage outcomes are realized as a function of the full history of accumulated cognitive skill and parental behaviors over the lifecycle. Figure 2 illustrates the terminal-period timing of events.[4]

Figure 1: **Pre-Teenage Period: Timing of Behavioral Choices**

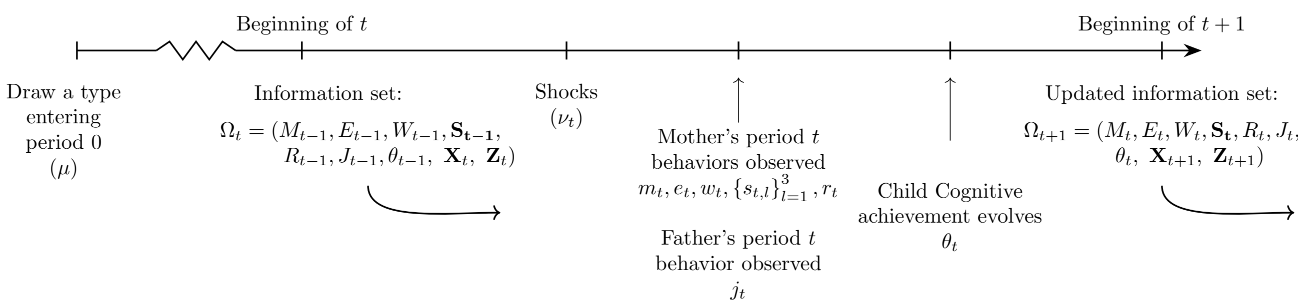
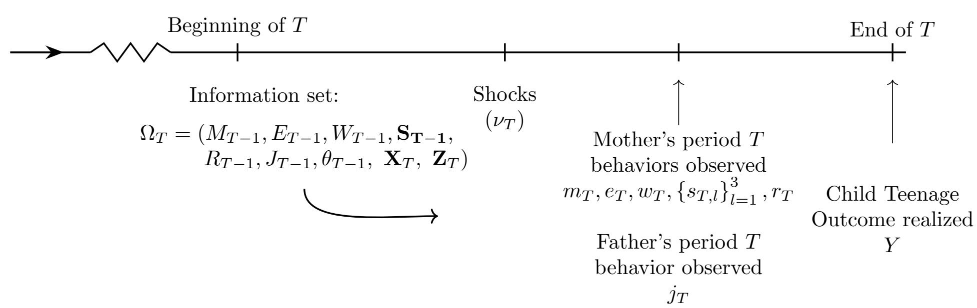


Figure 2: **Teenage (Terminal) Period: Timing of Behavioral Choices**



4.1.2 Parental Behavior Function

In this model, parents choose their behaviors (B_t) in each period after observing their child's cognitive achievement (θ_{t-1}) from the end of the last period and aware of their own lagged behavior (B_{t-1}). Let $\boldsymbol{X}_t^f$ denote a vector of parent/family characteristics and ε_t^b denote unobserved determinants of parental behavior. The parental behaviors may be correlated over time through both observed and unobserved variables. Hence, the inclusion of the lagged behavior as determinant requires dependence on exogenous supply side shifters ($\boldsymbol{Z}_t$) that vary over time across state or county

[4] A formal derivation of the parental utility function, budget constraint, value function, and the dynamic optimization problem underlying these behavioral choices is provided in Appendix A2.

level. The $\boldsymbol{Z}_t^b$ contains exogenous variables that are theoretically justified as impacting mom's or dad's behavior but do not affect the child's cognitive production function conditional on her or his behaviors. Disaggregating the parental behaviors into mother's ($B^m \in \{M, E, W, S, R\}$) and dad's ($B^d \in \{J\}$) behaviors, the parental behavior (or demand) function at age t is modeled as:

$$P(B_t^k = j) = f^{B^k}(\boldsymbol{B}_{t-1}^k, \theta_{t-1}, \boldsymbol{X}_t^f, \boldsymbol{Z}_t^k, \varepsilon^{B^k}) \tag{1}$$

where $j = 0, 1, \ldots, L^b$, with L^b being the total number of alternatives for behavior b. Also $k \in \{m, d\}$ and $\boldsymbol{B}_{t-1}$ is a vector of most recent lag of all the observed parental behaviors (i.e. $\boldsymbol{B}_{t-1}^m = \{M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}\}$ and $\boldsymbol{B}_{t-1}^d = \{J_{t-1}\}$).

4.1.3 Cognitive Achievement Production Function

End of period child cognitive achievement (or production) is a function of per-period parental behaviors, family and child observed characteristics, and unobserved determinants of cognitive achievement. Similarly, there is the need for an exogenous supply side shifter (Z_t^θ) given the inclusion of the lagged cognitive achievement as a determinant. Under the assumption of a value-added skill production function, the evolution of child cognitive skills can be defined by:

$$\theta_t = f^\theta(\theta_{t-1}, \boldsymbol{B}_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \varepsilon_t^\theta) \tag{2}$$

The value-added structure of equation (2) implies that only the contemporaneous parental behavior $\boldsymbol{B}_t$ enters each period's cognitive production function. Past parental behaviors $\boldsymbol{B}_{t-1}, \boldsymbol{B}_{t-2}, \ldots$ affect θ_t only indirectly through the accumulated skill path $\theta_{t-1} \to \theta_{t-2} \to \cdots$. This structure plays a key role in the decomposition of direct and indirect effects in Appendix section A3.

4.1.4 Teenage Outcome Production Function

In this section, I begin by asking the reader to consider a simple model of child's teenage outcomes. Let's conceive of these outcomes as dynamic, meaning they depend on previous values of the outcome and contemporaneous parental inputs (Todd and Wolpin, 2003). Let Y_t denote the child's outcome when the child is age t, and F_t represent the family input into the child outcome production function. Unobserved determinants of the child's outcome are denoted by ε_t^c and may include time varying and/or non-time varying unobserved heterogeneity. The teenager outcome production function at age $t+1$ is defined by:

$$Y_{t+1} = h(Y_t, F_t, \theta_t, X_t^c, \varepsilon_{t+1}^c) \qquad t = 1, \ldots, T \tag{3}$$

Assuming that the initial ($t = 1$) child outcome is given by $Y_1 = g(F_0)$, I substitute out Y_t in equation (3) with its determinants such that:

$$\begin{aligned}
Y_{t+1} &= h(Y_t, F_t, \theta_t, X_t^c, \varepsilon_{t+1}^c) \\
&= h(Y_{t-1}, F_t, F_{t-1}, \theta_t, \theta_{t-1}, X_t^c, X_{t-1}^c, \varepsilon_{t+1}^c, \varepsilon_t^c) \\
&\vdots \\
&= h\left(F_t, F_{t-1}, \ldots, F_0,\ \theta_t, \theta_{t-1}, \ldots, \theta_1,\ \underbrace{X_t^c, \ldots, X_1^c}_{\boldsymbol{X}_t^c},\ \underbrace{\varepsilon_{t+1}^c, \ldots, \varepsilon_1^c}_{\boldsymbol{\varepsilon}_{t+1}^c} \right) \\
&= h(F_t, F_{t-1}, \ldots, F_0,\ \theta_t, \theta_{t-1}, \ldots, \theta_1,\ \boldsymbol{X}_t^c,\ \boldsymbol{\varepsilon}_{t+1}^c) \qquad (4)
\end{aligned}$$

Family input (F_t) represents an optimal allocation of parental resources and time. Those allocations depend on jointly chosen parental behaviors such as marriage/cohabitation, employment and welfare receipt, parental consumption including substance use, or criminal behavior, among other things. Parents choose behaviors (e.g., employment, substance use, etc.), while making money and/or time allocation to child inputs. The key assumption here is that between choosing behaviors and allocating family inputs, parents receive additional information (ε_t^f) about their circumstance. For example, when parents choose to work, they do so based on a wage offer, but the realized wage or job amenities may differ. However, after choosing their behavior, new information (ε_t^f) is revealed that may affect how they allocate resources to their child. Similarly, a parent might anticipate receiving a particular level of public assistance (TANF, cash transfers, etc.), but the realized benefit may be different, hence impacting the family input allocation. Therefore, family inputs can be expressed as a function of parental observed behavior, family characteristics and unobserved determinants of family inputs, such that:

$$F_t = f(B_t, X_t^f, \varepsilon_t^f) \qquad (5)$$

Now substituting for F_t in equation (4):

$$\begin{aligned}
Y_{t+1} &= g\left(B_t, B_{t-1}, \ldots, B_0,\ \theta_t, \theta_{t-1}, \ldots, \theta_1,\ \underbrace{X_t^f, \ldots, X_1^f}_{\boldsymbol{X}_t^f},\ \boldsymbol{X}_t^c,\ \underbrace{\varepsilon_t^f, \ldots, \varepsilon_1^f}_{\boldsymbol{\varepsilon}_t^f},\ \boldsymbol{\varepsilon}_{t+1}^c \right) \\
Y_{t+1} &= g\left(B_t, B_{t-1}, \ldots, B_0,\ \underbrace{\theta_t, \theta_{t-1}, \ldots, \theta_1}_{\boldsymbol{\theta}_t},\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \underbrace{\boldsymbol{\varepsilon}_t^f,\ \boldsymbol{\varepsilon}_{t+1}^c}_{\boldsymbol{\varepsilon}_{t+1}^y} \right) \\
Y_{t+1} &= g(B_t, B_{t-1}, \ldots, B_0,\ \boldsymbol{\theta}_t,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \boldsymbol{\varepsilon}_{t+1}^y) \qquad (6) \\
Y_{t+1} &= g(B_t, B_{t-1}, \ldots, B_0,\ \theta_t,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \boldsymbol{\varepsilon}_{t+1}^y) \qquad (7)
\end{aligned}$$

In moving from equation (6) to equation (7), I adopt a Markovian assumption in which the child's current cognitive achievement θ_t captures all past skills and serves as a sufficient measure for the history of past skill levels $\{\theta_{t-1}, \theta_{t-2}, \ldots, \theta_1\}$. This assumption is standard in dynamic models of child development and human capital formation literature (e.g., Agostinelli and Wiswall, 2025; Cunha and Heckman, 2007; Todd and Wolpin, 2003), where skill evolves recursively and accumulates the effects of past investments and shocks. Under this framework, θ_t summarizes all relevant cognitive achievement needed to predict child teenage outcomes, rendering the inclusion of lagged cognitive achievement variables redundant conditional on θ_t.[5]

The unobserved family inputs are correlated over time, that is $\text{corr}(\varepsilon_t^f, \varepsilon_{t-1}^f) \neq 0$, implying that the entire history of family shocks does matter in the child's outcome production function. To isolate exogenous variation in behavioral pathways, I introduce a vector of supply-side shifters (Z_t^θ, $\boldsymbol{Z}_t^p$) that operates as an instrumental variable, shifting parental decisions while remaining orthogonal to ε_{t+1}^y, hence:

$$Y_{t+1} = g(B_t, B_{t-1}, \ldots, B_0,\ \theta_t,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ Z_t^\theta,\ \boldsymbol{Z}_t^p, \varepsilon_{t+1}^y) \tag{8}$$

Note that the specification in equation (8) assumes that the whole history of parental behavior affects the teenage outcomes. Inclusion of variables representing these period-specific behaviors introduces potential multicollinearity. Alternatively, I can model the sequence of period-specific parental behaviors $\{B_t, B_{t-1}, B_{t-2}, \ldots, B_0\}$ with a behavioral string or trajectory variable $\overrightarrow{B}_t$ (e.g., employment – "always employed", "never employed", "intermittent employed"; for marital status – "always single", "always married", "unstable" etc.). One could argue that the individual separate B_t terms are "redundant" because their direct, time-specific effects are already captured in the construction of the child's cognitive skill path θ_t. In particular, the cognitive achievement production function, estimated dynamically across the preteenage period, explicitly captures how each parental behavior at time t influences the formation of cognitive skill θ_{t+1}.[6] However, these parental behaviors could affect non-cognitive development, which is omitted in this specification. Hence, the past behaviors may still matter. To avoid multicollinearity concerns, the alternative specification assumes that any residual effect of parental behavior on the teenage outcome (i.e., not through cognitive development) operates through broader patterns or long-run trajectories, such as persistent disadvantage or consistent positive investment, which is better captured by behavioral strings than by isolated indicators. Hence, the alternative (to equation 8) child's teenage production function

[5] $P(Y_{t+1} \mid \theta_t, \bullet) = P(Y_{t+1} \mid \theta_t, \theta_{t-1}, \theta_{t-2} \ldots, \theta_1, \bullet)$. Additionally, this Markov property is also invoked in the decomposition of direct and indirect effects in Section A3, where it justifies expressing the full indirect cognitive channel solely through the terminal-period skill θ_T.

[6] As a result, the lagged effects of isolated behaviors are absorbed into the child's cumulative cognitive measure. In estimating the teenage outcome function, θ_t acts as a sufficient statistic summarizing the *quality and timing* of preteenage parental investment, as well as exposure to risky or protective behaviors.

is:

$$Y_{t+1} = g(\overrightarrow{B}_t,\ \theta_t,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ Z_t^\theta,\ \boldsymbol{Z}_t^p,\ \varepsilon_{t+1}^y) \tag{9}$$

The use of the behavioral string or trajectory specification enhances econometric tractability, dimensionality reduction, and mitigates multicollinearity.[7] The approach is consistent with human capital and developmental psychology theories of sustained exposure and risk accumulation (Chang et al., 2025; Cunha and Heckman, 2007; Sampson and Laub, 1995). Additionally, the behavioral strings are more interpretable and policy-relevant, for instance, categories like "always employed", "never on welfare" etc. Table 5 defines the behavioral string categories used in estimation for each parental behavior. For each behavior, strings are constructed from the sequence of binary (or multinomial) period-specific choices across the six developmental periods t_1 through t_6.

4.2 The Set of Structural Equations

Parents choose their behaviors conditional on the child's observed cognitive achievement at the beginning of the period (among other variables). The child's cognitive achievement is then updated at the end of the period (and is carried over into the next period), and this cycle continues dynamically over time.

Pre-teenage Period:

$$\begin{aligned}
\text{Parental Behavior:}\quad & P(B_t = j) = f^b(\boldsymbol{B}_{t-1}, \theta_{t-1}, \boldsymbol{X}_t^f, \boldsymbol{Z}_t, \varepsilon_t^B) && t = 1, \dots, 6 \\
\text{Child Cognitive Achievement:}\quad & \theta_t = f^\theta(\theta_{t-1}, \boldsymbol{B}_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \varepsilon_t^\theta) && t = 1, \dots, 5
\end{aligned}$$

Teenage Period:

$$\text{Child Teenage Outcome:}\quad Y_{t+1} = g(\overrightarrow{\boldsymbol{B}}_t,\ \theta_t,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ Z_t^\theta,\ \boldsymbol{Z}_t^p,\ \varepsilon_{t+1}^y) \qquad t = 6$$

The joint system of estimated equations is specified below, where some of these explanatory variables act as exclusion restrictions that aid in identifying the system of equations to be estimated. The probability of being married or cohabiting to the father of the focal child ($m = 1$) or another man ($m = 2$) in period t relative to being single (no father present, $m = 0$) in log odds is given by:

$$\ln\left[\frac{P(M_t = m)}{P(M_t = 0)}\right] = f^m(M_{t-1}, E_{t-1}, W_{t-1}, S_{t-1}^1, S_{t-1}^2, S_{t-1}^3, R_{t-1}, \theta_{t-1}, \boldsymbol{X}_t^f, \boldsymbol{Z}_t, \varepsilon_{mt}^M)$$
$$m = 1, 2.$$

[7] I also estimate two other alternative models: (1) Summation of parental behavior model; and (2) Contemporaneous parental behavior model. See section 8 for details.

The probability of mother being employed ($E_t = 1$) in period t relative to being unemployed ($E_t = 0$) in log odds is given by:

$$\ln\left[\frac{P(E_t = 1)}{P(E_t = 0)}\right] = f^e(M_{t-1}, E_{t-1}, W_{t-1}, S^1_{t-1}, S^2_{t-1}, S^3_{t-1}, R_{t-1}, \theta_{t-1}, \boldsymbol{X}^f_t, \boldsymbol{Z}_t, \varepsilon^E_t) \tag{10}$$

The probability of mother being on welfare ($W_t = 1$) in period t relative to not being on welfare ($W_t = 0$) in log odds is given by:

$$\ln\left[\frac{P(W_t = 1)}{P(W_t = 0)}\right] = f^w(M_{t-1}, E_{t-1}, W_{t-1}, S^1_{t-1}, S^2_{t-1}, S^3_{t-1}, R_{t-1}, \theta_{t-1}, \boldsymbol{X}^f_t, \boldsymbol{Z}_t, \varepsilon^W_t) \tag{11}$$

In estimating the substance use of the mother, I consider three main substances (i.e., alcohol, cigarettes, and marijuana). More formally $l \in \{$*alcohol, cigarettes, marijuana*$\}$, and each substance has its own equation with its own coefficient vector. The probability of mother using substance l ($S^l_t = 1$) in period t relative to not using ($S^l_t = 0$) in log odds is given by:

$$\ln\left[\frac{P(S^l_t = 1)}{P(S^l_t = 0)}\right] = f^{s^l}(M_{t-1}, E_{t-1}, W_{t-1}, S^1_{t-1}, S^2_{t-1}, S^3_{t-1}, R_{t-1}, \theta_{t-1}, \boldsymbol{X}^f_t, \boldsymbol{Z}_t, \varepsilon^{S^l}_t) \tag{12}$$
$$l \in \{1, 2, 3\}, \text{ estimated for each substance.}$$

The probability of mother being convicted of crime ($R_t = 1$) in period t relative to not ($R_t = 0$) in log odds is given by:

$$\ln\left[\frac{P(R_t = 1)}{P(R_t = 0)}\right] = f^r(M_{t-1}, E_{t-1}, W_{t-1}, S^1_{t-1}, S^2_{t-1}, S^3_{t-1}, R_{t-1}, \theta_{t-1}, \boldsymbol{X}^f_t, \boldsymbol{Z}_t, \varepsilon^R_t) \tag{13}$$

The probability of dad being in jail or convicted ($J_t = 1$) in period t relative to not ($J_t = 0$) in log odds is given by:

$$\ln\left[\frac{P(J_t = 1)}{P(J_t = 0)}\right] = f^j(J_{t-1}, \theta_{t-1}, \boldsymbol{X}^f_t, \boldsymbol{Z}_t, \varepsilon^J_t) \tag{14}$$

Child's cognitive achievement is given by:

$$\theta_t = f^\theta(M_t, E_t, W_t, S^1_t, S^2_t, S^3_t, R_t, J_t, \theta_{t-1}, \boldsymbol{X}_t, Z^\theta_t, \varepsilon^\theta_t) \tag{15}$$

The child's teenage outcome variables to be estimated are such that $Y^o \in \{$*Ever Failed a Class, Ever Suspended/Expelled from School, Likely to Graduate College, Ever Drunk Alcohol, Ever Used Marijuana, Ever Stopped by Police, Ever had Sex*$\}$. The probability of Y^o for a child expressed in log odds is given by:

$$\ln\left[\frac{P(Y^o_{T+1}=1)}{P(Y^o_{T+1}=0)}\right] = g^o(\overrightarrow{M}_t, \overrightarrow{E}_t, \overrightarrow{W}_t, \overrightarrow{S}^1_t, \overrightarrow{S}^2_t, \overrightarrow{S}^3_t, \overrightarrow{R}_t, \overrightarrow{J}_t, \theta_t, \boldsymbol{X}_t, Z^\theta_t, \boldsymbol{Z}^p_t, \varepsilon^R_t) \tag{16}$$

4.2.1 Error Structure

In addition to the observed permanent and time-varying characteristics of the mother, biological father, and child, unobserved heterogeneity undoubtedly influences the parental behaviors and child outcomes modeled in this paper. To account for the unobservables, I consider an unobserved family effect that contains both a permanent component and a time-varying component, each of which may influence every observed behavior and outcome in the system. Importantly, the estimated impact of this unobserved heterogeneity is modeled as non-linear, allowing its effect on each unit (i.e. mother, father, or child) to differ across behaviors and outcomes.

Formally, the error term in each equation q, ε^q_t, is decomposed into three components: a permanent component μ^q, a time-varying component ν^q_t, and an equation-specific idiosyncratic component u^q_t; that is,

$$\varepsilon^q_t = \mu^q + \nu^q_t + u^q_t \tag{17}$$

The permanent component μ^q captures time-invariant unobserved family heterogeneity that induces correlation across all equations and over time. For instance, some parents may have inherent tendencies towards a specific parenting style, like authoritarian, authoritative or permissive, which affects how they set boundaries, discipline and respond to their child's needs regardless of their chosen behaviors. Also, parents vary in terms of how forward-looking they are (i.e., a long term planner or someone more oriented to immediate gratification). These intrinsic characteristics may affect child's cognitive skill development regardless of behaviors like employment and welfare receipt. The time-varying component ν^q_t allows for an additional source of correlation across equations within a given period, capturing time-varying unobserved family characteristics over the developmental lifecycle such as a health shock to the child that simultaneously affects cognitive development and multiple parental behavioral decisions. The idiosyncratic component u^q_t is assumed to be i.i.d. across individuals and over time.

4.2.2 Initial Conditions

Since the initially-observed behaviors and cognitive achievement are endogenous, they must be modeled (rather than treated as exogenous) and correlated with unobserved determinants that may affect subsequent behaviors. Conditioning on predetermined outcomes would otherwise generate an initial conditions bias, such as, families entering the sample with different pre-sample histories would appear identical, conflating state dependence with unobserved heterogeneity. I address this issue by modeling each initial condition as a function of exogenous family characteristics, $\boldsymbol{X_t}$, appropriate exclusion restrictions, $\boldsymbol{Z_t}$, and the permanent family unobserved heterogeneity, μ. All of the initial condition equations are jointly-estimated with the set of structural dynamic parental behavior and child cognitive equations, and teenage outcome equations.

For each initially-observed behaviors $B_1 \in \{M_1, E_1, W_1, S_1^1, S_1^2, J_1\}$, the probability of a parent choosing the behavior ($B_1 = 1$) in period 1 relative to not ($B_1 = 0$) in log odds is given by:

$$\ln\left[\frac{P(B_1 = 1)}{P(B_1 = 0)}\right] = f^I(\boldsymbol{X}_t^f, \boldsymbol{Z}_t, \varepsilon_1^I) \tag{18}$$

The initially-observed cognitive achievement is given by:

$$\theta_1 = f^{\theta_1}(M_t, E_t, W_t, S_t^1, S_t^2, R_t, J_t, \boldsymbol{X}_t^f, Z_t^\theta, \varepsilon_1^\theta) \tag{19}$$

Table 6 presents the full system of 23 jointly estimated equations (8 parental behaviors, 1 cognitive outcome, 7 teenage outcomes, and 7 initial conditions) alongside their respective determinants. The dependent variable for each equation appears in the first column, while columns 2–4 detail the corresponding regressors, comprising pre-determined endogenous variables, exogenous variables, and unobserved heterogeneity terms. The resulting probabilities from each equation enter directly into the likelihood function specified below.

5 Estimation Method

I jointly estimate the system of equations using the discrete factor random effects (DFRE) method. The DFRE method approximates two components of the decomposed error terms ($\boldsymbol{\varepsilon_t}$) as discrete distributions: the permanent family heterogeneity ($\boldsymbol{\mu}$) and the time-varying family heterogeneity ($\boldsymbol{\nu}_t$), together with an independently and identically distributed idiosyncratic term ($\boldsymbol{u_t}$). The permanent heterogeneity, $\boldsymbol{\mu}$, is correlated across equations and over time and captures information on the family that is not observed by the econometrician. The time-varying unobserved heterogeneity, $\boldsymbol{\nu}_t$, is correlated across equations within a time period and captures family unobservables that vary over time. The equation-specific idiosyncratic error terms follow a Type-I Extreme Value distribution for discrete parental behaviors (u_t^B where $B \in \{M, E, W, S^1, S^2, S^3, R, J\}$; and $u_1^{B_1}$ where $B_1 \in \{M_1, E_1, W_1, S_1^1, S_1^2, J_1\}$) and teenage outcomes ($u_t^o$ where $o = 1, 2, \ldots, 7$), yielding logit and multinomial logit probabilities that contribute to the likelihood function. For the continuous variable, child cognitive score, the idiosyncratic term u_t^θ is assumed to be an i.i.d. normal distribution. The DFRE method assumes that the distributions of the heterogeneity terms can be approximated by a discrete step function, estimated with the number of support points determined by the model's best fit of the data (Mroz, 1999). I estimate the discrete distribution for μ, with each mass point

μ_k assigned an estimated probability τ_k; a family draws its permanent type once, and that draw is held fixed across all periods. I estimate a single discrete distribution for ν, with each mass point ν_g assigned an estimated probability π_g; unlike μ, a new draw from this same distribution is made independently in each period t.

5.1 Likelihood Function

The unconditional likelihood function for family (parent with focal child) i is given by:

$$
\begin{aligned}
\mathcal{L}_{ic}(\Theta) = \prod_{i=1}^{N}\sum_{k=1}^{K}\tau_k\Bigg\{ & \prod_{m=0}^{1}\mathbb{P}(M_{i1}=m \mid \mu_k)^{\mathbb{1}[M_{i1}=m]} \bullet \prod_{e=0}^{1}\mathbb{P}(E_{i1}=e \mid \mu_k)^{\mathbb{1}[E_{i1}=e]} \\
& \bullet \prod_{w=0}^{1}\mathbb{P}(W_{i1}=w \mid \mu_k)^{\mathbb{1}[W_{i1}=w]} \bullet \prod_{l=1}^{2}\prod_{s^l=0}^{1}\mathbb{P}(S_{i1}^{l}=s^{l} \mid \mu_k)^{\mathbb{1}[S_{i1}^{l}=s^{l}]} \\
& \bullet \prod_{r=0}^{1}\mathbb{P}(R_{i1}=r \mid \mu_k)^{\mathbb{1}[R_{i1}=r]} \bullet \prod_{j=0}^{1}\mathbb{P}(J_{i1}=j \mid \mu_k)^{\mathbb{1}[J_{i1}=j]} \bullet \phi^{\theta_1}(\theta_{i1} \mid \mu_k) \\
& \bullet \prod_{t=2}^{T}\Bigg[\sum_{g=1}^{G}\pi_g\Bigg(\prod_{m=0}^{1}\mathbb{P}(M_{it}=m \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[M_{it}=m]} \bullet \prod_{e=0}^{1}\mathbb{P}(E_{it}=e \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[E_{it}=e]} \\
& \qquad \bullet \prod_{w=0}^{1}\mathbb{P}(W_{it}=w \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[W_{it}=w]} \bullet \prod_{l=1}^{3}\prod_{s^l=0}^{1}\mathbb{P}(S_{it}^{l}=s^{l} \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[S_{it}^{l}=s^{l}]} \\
& \qquad \bullet \prod_{r=0}^{1}\mathbb{P}(R_{it}=r \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[R_{it}=r]} \bullet \prod_{j=0}^{1}\mathbb{P}(J_{it}=j \mid \Omega_{it},\mu_k,\nu_{tg})^{\mathbb{1}[J_{it}=j]} \\
& \qquad \bullet \phi^{\theta_t}(\theta_{it} \mid \Omega_{it},\mu_k,\nu_{tg})\Bigg)\Bigg] \\
& \bullet \prod_{o=1}^{7}\Bigg[\sum_{g=1}^{G}\pi_g\Bigg(\prod_{y=0}^{1}\mathbb{P}(Y_{i\ T+1}^{o}=y \mid \Omega_{it},\mu_k,\nu_{Tg})^{\mathbb{1}[Y_{i\ T+1}^{o}=y]}\Bigg)\Bigg]\Bigg\}
\end{aligned} \tag{20}
$$

where family i faces type k permanent heterogeneity, μ_k, with probability τ_k throughout all periods, and also experiences each of the type g time-varying types, ν_{tg}, each period t with probability π_g. $\mathbb{P}(x)$ is the probability of a behavior or outcome if x is discrete and $\phi(x)$ represents the density of a behavior or outcome when x is continuous. The parameters of the likelihood function are estimated with full information maximum likelihood (FIML).

5.2 Identification

Credibly identifying the causal effects of parental behavior on child cognitive achievement demands exogenous variables that predict the joint behaviors yet exert no direct influence on child achievement beyond their role as determinants of those behavioral inputs. The exclusion restriction ($\boldsymbol{Z_t}$)

is a vector of instrumental variables that: (1) affect the mother's or father's behaviors but do not affect the child's achievement; (2) are independent of all the independently and identically distributed error terms ($\boldsymbol{u_t}$). These exclusion restrictions enter the parent behavioral equations but do not enter the child's cognitive achievement equation and serve as shifters in the estimation of endogenous variables to identify the system of equations.

Another source of identification comes from the dynamic structure of the model itself. Each parental behavior and cognitive achievement equation at period t includes the lagged value of the relevant outcome at period $t-1$ as a predetermined regressor (i.e., B_{t-1} or θ_{t-1} entering the period-t equation). Similar to Arellano and Bond (1991), these lagged endogenous variables serve as internal instruments: since each B_{t-1} or θ_{t-1} has itself been instrumented by the supply-side exclusion restrictions $\boldsymbol{Z}_{t-1}$ or Z^{θ}_{t-1} in the previous period, and its variation at $t-1$ is orthogonal to the idiosyncratic error u^b_t at period t. The inclusion of these lags serve as a sequence of exclusion restrictions that contribute to identification across all pre-teenage periods, supplementing the external supply-side instruments. In effect, the lagged instrumented behaviors act as predetermined instruments for current behaviors in each period, exploiting the panel structure of the data to identify the dynamic path of parental decisions and cognitive skill formation.

Identification of the teenage outcome equations requires a separate argument, as both the behavioral strings $\overrightarrow{B}_t$ and the terminal cognitive score θ_t are endogenous regressors in that equation. Identification is achieved through three channels. First, the same supply-side instruments $\boldsymbol{Z}_t$ that shift parental behaviors in the pre-teenage periods propagate forward into the behavioral string classification; by construction, instruments that shift behavior in any single period can shift the string category, creating variation in $\overrightarrow{B}_t$ that is orthogonal to the teenage outcome idiosyncratic shock u^o_{t+1}. Second, the county-level income-adjusted test score Z^{θ}_t shifts θ_t without independently affecting teenage outcomes conditional on the behavioral strings and other exogenous controls. Third, peer characteristics $\boldsymbol{Z}^p_t$ (including peer drinking alcohol, smoking cigarettes, marijuana use, study habits, and college ambitions) serve as additional instruments that enter the teenage outcome equation directly. This approach follows Fruehwirth et al. (2019), who similarly use peer characteristics as an instrument to identify the causal effect of adolescent religiosity on depression. These peer variables affect the child's adolescent outcomes through social influence and peer norms, but are excludable from the parental behavior and cognitive achievement equations, where peer effects do not operate as direct determinants. The permanent and time-varying unobserved heterogeneity terms μ and ν_t are also shared across the behavior, cognitive, and outcome equations, providing within-family covariance restrictions that jointly help identify the full system.

Identifying variables associated with marital behavior are county level fraction of people 15 or older who are married and not separated. This measure is from the Fragile Families' county intergenerational mobility data appendage from Chetty and Hendren (2018) opportunity insights index computed based on de-identified tax records on more than five million people in US counties. The variable associated with employment behavior is unemployment rate in the mother's city of residence in year of interview. This measures the ease and likelihood of the mother getting a job but does not affect any of the child's outcome variables. With regards to welfare, there have been a few instruments used in previous studies like state policies and distance from the survey respondent's residence to the nearest TANF program office (Ratcliffe et al., 2011; Yen et al., 2012; McKernan et al., 2021). Given that I do not have the GPS location of respondent in my data, I do not use distance as an instrument, but rather county-level percentage of households on public assistance. Associated with substance abuse, are state level alcohol excise tax on beer and state level tax per pack of cigarettes in the year of interview. Williams et al. (2011) investigated why people want cannabis to be legalized and what contributes to cannabis usage. Cannabis prices and decriminalization were found to be strong instruments in influencing non-medical usage. Therefore, I use the variation in the timing of marijuana legalization across states. The county violent crime rate (per 100,000 people) in the year of interview for mom and dad identifies mom's conviction and dad's jail-time equations.

At the child's level, the county-level income-adjusted test score serves as the instrument for the child's cognitive equation. This variable shifts county-level average cognitive performance through variation in local school quality and resources, while remaining excludable from the teenage outcome equation conditional on parental behaviors and household characteristics. Finally, peer characteristics $\boldsymbol{Z}_t^p$ (i.e. peers' drinking, smoking, marijuana use, study habits, and likelihood of graduating college) identify the teenage outcome equations. These variables capture the social environment the child is embedded in during adolescence and directly shape adolescent risk behaviors and educational ambitions, but are excludable from the parental behavior and cognitive production equations, where peer influences do not operate as direct determinants of parental choices or pre-teenage cognition. To assess instrument relevance, Appendix Table A.24 reports Wald F-tests for the joint significance of excluded instruments across all equations; the instruments are jointly significant providing empirical support for the relevance condition. Table 7 provides a summary of all the supply side variables categorized by their related behavior.

6 Results

6.1 Model Fit

I use the dynamic multiple equation model and the estimated coefficients to simulate behaviors and outcomes overtime in order to assess how well the model reproduces the patterns present in the actual data. Using initial period data, I generate parental behaviors and child cognitive achievement, at age 1. The simulated endogenous variables are carried forward to determine behaviors and cognitive scores in each subsequent period. In the terminal period, teenage outcomes are then generated as a function of the full history of simulated behaviors and cognitive achievement accumulated over the lifecycle. Each family, N from the sample size, is simulated R times (i.e., replications). For each replication, a permanent family type is assigned once at the outset by drawing from the estimated unobserved heterogeneity distribution, while a time-varying family type is drawn every period. Table 8 reports a comparison of summary statistics across observed and model-generated behaviors, cognitive scores, and teenage outcomes. The average deviation between simulated and observed means is only 0.005, indicating the model matches behaviors and outcomes to within less than one percentage point on average. Additionally, the dynamic behavior probabilities and average cognitive score are depicted in Figure 5 and suggests that the estimated model fits the data well.

Prior to estimating my main model, I also estimated a naive model, that is, estimating the structural equations separately (not jointly) without accounting for unobserved heterogeneity. This model treats the endogenous behaviors and cognitive achievement as exogenous in the equations they enter. A likelihood ratio test was conducted to evaluate whether incorporating permanent and time-varying unobserved heterogeneity yields a statistically meaningful improvement in model fit, as gauged by the log-likelihood value. The technical details of this test are provided in Appendix A4.1.1. The test conclusively favors the specification that accommodates unobserved heterogeneity over the restricted version that omits it, reinforcing the necessity of jointly accounting for endogeneity and simultaneity in recovering the parameters of interest. The results provide strong statistical evidence in favor of the richer specification with unobserved heterogeneity, justifying its use in the analysis and in the computation of marginal effects and counterfactuals.

Among the specifications considered, the one achieving the best statistical fit supports four mass points for the permanent heterogeneity distribution and four mass points for the time-varying heterogeneity distribution. Table 9 presents the associated probability weights and mass points, with columns 2 and 3 reporting estimates of the probabilities and mass points for the permanent heterogeneity distribution and columns 4 and 5 reporting the corresponding quantities for the time-varying distribution. Standard errors are shown in brackets throughout.

6.2 Results from the Dynamic Multiple-equation Model

Estimation results for the parental behavior equations are reported in Online Appendix Tables OA.3–OA.10, where endogenous regressors appear first, followed by exogenous variables. Across the parental behaviors, the results consistently show that lagged parental behaviors (i.e., dynamic behavioral persistence), and also child cognitive performance, and policy relevant county or state characteristics exert meaningful and statistically significant effects on parental behaviors. The comparison between models estimated without unobserved heterogeneity (no-UH) and those that incorporate unobserved heterogeneity (UH) is particularly revealing. The inclusion of UH plays a critical role in refining estimates, often attenuating inflated coefficients and in several cases reveal relationships that were obscured in the no-UH specifications. This finding highlights the presence of persistent and time-varying family-specific unobserved heterogeneity and characteristics that confound estimates unless explicitly modeled.

All lagged behaviors exhibit statistically significant and substantively large effects across all estimation of the dynamic behaviors. These lagged effects remain robust even under UH, although they typically decrease in magnitude, confirming that part of the persistence observed in naive model (no-UH model) arises from unobserved traits such as risk tolerance, parenting style, child's intrinsic nature, dealing with trauma or stress, intergenerational behavioral patterns, among others. Additionally, some lagged cross-behaviors are significant predictors of the dynamic parental behaviors. For example, being married/cohabiting with biological dad decreases, and using marijuana increases, the probability of receiving welfare. With regard to criminal behavior: marriage/cohabitation decreases, and being on welfare increases, the probability of being convicted. The cross-behavior effects show how the behaviors are intertwined, reinforcing the jointly estimated model specification.

The state-, county- and city-level variables function as critical supply-side shifters, and retain significance even after UH is included implying that their effects are not absorbed by unobserved family traits. The persistence of their effects in UH specifications strengthens the argument that they serve as valid sources of exogenous variation in the dynamic parental behavior environment and child achievement. Other demographic and household characteristics mostly show expected signs, although their magnitudes typically decrease in the UH models. Age and age-squared capture life-cycle dynamics, and higher education improves employment and marital stability while reducing welfare dependence and smoking. The coefficients on the race aligns with well known racial and ethnic disparities. Also poor health increase the probability of welfare receipt and risky behavior like drinking, smoking and marijuana use, implying that although the mother's health stock may be depleting, the utility she gains from using these substances enforces this cycle to continue leading addiction.

The estimation of children's cognitive achievement in Online Appendix Table OA.11 reveals a strong degree of persistence over time, with the lagged cognitive score emerging as the most powerful predictor of current performance, which is consistent with the value added literature. Parental behaviors also play an important role in child cognitive development. Presence of a father figure increases child cognitive scores, while welfare receipt and criminal conviction decline cognitive score, each of which becomes more pronounced under UH model.

7 Counterfactuals and Total Effects Decomposition

Before turning to the results, I first describe how the total effect of a counterfactual change in parental behavior is decomposed into a direct and an indirect component. For each counterfactual comparison, I simulate the model forward under Counterfactual (1) and separately under Counterfactual (2), and take the difference in the resulting predicted teenage outcome; this difference is the total effect of the counterfactual. To isolate the direct effect, that is, the portion of the total effect that does not operate through the child's cognitive development, I re-run both simulations a second time, but in each case I replace the child's cognitive score, which would otherwise evolve endogenously along the counterfactual behavioral path, with the child's cognitive score as actually observed in the data. Because cognitive achievement is now held fixed at the same observed value under both counterfactuals, any remaining difference between the two simulated outcomes cannot be attributed to the cognitive channel, isolating the direct effect. The indirect cognitive effect is then recovered as a residual, the total effect less the direct effect. I apply this procedure to every counterfactual comparison below; the formal derivation is provided in Appendix A3.

Tables 10 to 12 present these counterfactual decompositions across a range of teenage outcomes, with positive values indicating worse outcomes under the more disadvantaged or high-risk scenario. In all the results presented, the total effect is driven almost entirely by the direct effect, with indirect effects through child cognitive achievement being small in magnitude and mostly negative, since we expect a higher cognitive score to reduce the likelihood of these teenage outcomes.

Marital, employment and welfare behaviors

Table 10 evaluates the role of parental marital stability by contrasting an unstable marital trajectory, that is, cycling between single status, cohabitation with the biological father, and cohabitation with another man (Counterfactual 1), against a benchmark of continuous marriage to the biological father (Counterfactual 2). Marital instability significantly worsens two educational outcomes: the probability of being unlikely to graduate college rises by 8.7 percentage points (p.p.), and the probability of ever being suspended from school rises by 8.0 p.p., under the unstable marital counterfactual. Marital instability also significantly raises the probability of being stopped by police, by 8.1 p.p.. Table 11 turns to maternal economic behavior, contrasting a counterfactual in which the mother is unemployed and persistently on welfare (Counterfactual 1) with one in which she is employed and never receives welfare (Counterfactual 2). Disadvantaged economic conditions significantly raise the probability of ever failing a class by 11.7 p.p. and the probability of having had sex by 8.1 p.p.. The most striking result in Table 11 is that the probability of ever being stopped by police rises by 26.1 p.p.. Across both tables, these significant effects are overwhelmingly driven by the direct behavioral component, with indirect cognitive effects close to zero.

Substance use and criminal behaviors

Table 12 presents a joint counterfactual in which maternal substance use and parental criminal involvement are varied simultaneously: Counterfactual (1) is the most adverse profile, in which the mother drinks alcohol, smokes cigarettes, and uses marijuana in every period, is convicted in every period, and the father is incarcerated in every period; Counterfactual (2) is the most favorable profile, in which the mother abstains from all substances and is never convicted, and the father is never incarcerated. The only statistically significant result is on police contact: the probability of ever being stopped by police is 28.8 p.p. lower under the adverse profile than under the favorable one, an effect driven almost entirely by the direct behavioral component. This counterintuitive direction is consistent with a deterrence-type channel; for instance, a father's continuous incarceration may reduce the teenager's exposure to the situations, peers, or environments that would otherwise result in police contact.

Finally, the marginal effects in Appendix Table A.26 do not equal the sum of effects from Tables 10 – 12. This non-additivity is central to the interpretation of the model. When behaviors are altered jointly, their effects interact in nonlinear ways that amplify overall disadvantage. For instance, unstable marital status, unemployment, welfare dependence, and substance use jointly shape parent's daily routines, time allocations, supervision quality, and stress levels in ways that cannot be captured by summing isolated marginal effects. The presence of criminal involvement and paternal incarceration further reinforces these mechanisms by disrupting household stability. This finding provides strong empirical support for jointly modeling parental behaviors, rather than treating them as independent inputs into child development. Family structure emerges endogenously from a bundle of interrelated parental behaviors, and teenage outcomes respond to the nexus of behaviors rather than to any single dimension in isolation.

8 Model Comparison: Alternative Behavioral Transmission Mechanisms

8.1 Summation of Parental Behavior Model

Following Chan and Liu (2025), I assume that parental behaviors enter the child's teenage outcome production function not as isolated period-specific shocks, nor as categorical strings, but rather as cumulative behavioral exposure over the child's development window. Formally, consider the teenage outcome equation:

$$Y_{t+1} = g\big(B_t, B_{t-1}, B_{t-2}, \ldots, B_0, \theta_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \boldsymbol{Z}_t^p, \mu^y, \nu_t^y, u_{t+1}^y\big)$$

To capture the cumulative structure implied by behavioral persistence, I rewrite the model such that the direct behavioral entry is summarized through the cumulative behavioral score:

$$Y_{t+1} = g\big(\breve{B}_t, \theta_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \boldsymbol{Z}_t^p, \mu^y, \nu_t^y, u_{t+1}^y\big) \tag{21}$$

where $\breve{B}_t$ denotes the total number of periods in which the parent engaged in behavior B.[8] The key difference between the baseline (string-based) model and the behavioral summation model is that, the baseline model emphasizes pattern recognition, but the summation model emphasizes exposure volume allowing the outcome equation to respond proportionally rather than categorically. Table 13 presents the estimated marginal effects of summed parental behavioral histories and child cognitive achievement on teenage outcomes.[9] Unlike the pattern-based specification, in which behavioral timing and sequencing matter, this specification treats each parental behavior as a continuous measure and interprets as an accumulated exposure. Hence, a coefficient represents the change in the probability of a teenage outcome given one additional period of parental involvement in that specific behavior.

One additional period spent living with the biological father significantly reduces the likelihood that the child will ever fail a class by 1.6 percentage points (p.p.) and the probability of being unlikely to graduate college by 1.9 p.p.. These magnitudes demonstrate meaningful long-run significant effects of stable paternal presence on academic readiness and schooling attainment. For welfare receipt, cumulative exposure significantly increases school suspension risk by 2.2 p.p., police stops by 2.6 p.p., and early sexual initiation by 1.8 p.p.. The magnitude of the police stop effect is particularly notable because it exceeds all other effects of mother's behavior. Also, cumulative exposure to mother's smoking increases the probability of teenage marijuana use by 2.4 p.p.. Similarly, paternal criminal history demonstrates some of the largest effects in the Table. An additional period of father's jail time increases teenage drinking by 1.4, marijuana use by 2.1 and police stops by 2.2 p.p., but surprisingly decreases the probability of not having college ambitions by 2.5 p.p.. The presence of repeated paternal incarceration appears to normalize contact with formal sanctions, possibly lowering perceived stigma or increasing adolescent interaction with law enforcement systems. Finally, a 10-point increase in the cognitive score reduces the likelihood of class failure by 4.0 p.p., respectively.

In summary, the summed parental behavior specification provides a complementary interpretive layer to the pattern-based findings. Here, parental behaviors translate not through timing effects but through linear accumulation.

[8] More precisely, $\breve{B}_t = B_0 + B_1 + \cdots + B_t$. For example, if a mother is Employed at $t = 2$ and $t = 5$, with periods defined as ages 0, 1, 3, 5, 9, 15, then: $\{E_{age0}, E_{age1}, E_{age3}, E_{age5}, E_{age9}, E_{age15}\} = \{0, 1, 0, 0, 1, 0\} \quad \Rightarrow \quad \breve{E}_t = 2$

[9] Full model-fit diagnostics for this specification are reported in Online Appendix Tables OA.19–OA.20 and Figure OA.1.

8.2 Contemporaneous Parental Behavior Model

Additionally, as a robustness check, this section evaluates an alternative behavioral transmission mechanism in which teenage outcomes depend only on the most recent parental behavior, rather than on the full history of behaviors. Again, consider the teenage outcome equation:

$$Y_{t+1} = g\big(B_t, B_{t-1}, B_{t-2}, \dots, B_0, \theta_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \boldsymbol{Z}_t^p, \mu^y, \nu_t^y, u_{t+1}^y\big)$$

Now, instead of allowing all past parental behaviors to enter the teenage outcome production function through a pattern or cumulatively, I impose a value-added structure similar to the assumption used for child cognitive development. Under this assumption, the teenage outcome is:

$$Y_{t+1} = g\big(B_t, \theta_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \boldsymbol{Z}_t^p, \mu^y, \nu_t^y, u_{t+1}^y\big) \tag{22}$$

where only contemporaneous parental behavior at time t directly affects adolescent outcomes, while earlier behaviors matter only insofar as they have shaped current cognition.[10] This specification contrasts with the baseline model in which the entire behavioral history $\{B_t, B_{t-1}, \dots, B_0\}$ enters directly into the teenage outcome equation.

Table 14 reports marginal effects under the contemporaneous parental behavior specification.[11] Mother's marital status continues to play a significant role even under this restrictive assumption. Living with another man, as opposed to the biological father, significantly increases the probability of teenage school suspension and marijuana use by 5.3 and 3.8 percentage points (p.p.), respectively. In contrast, children of mothers married with the biological father reduces the likelihood of failing a class and unlikely to graduate college by 9.3 and 7.8 p.p. compared to single mothers. Also, maternal employment remains the only behavior which is negative across all outcomes. Employed mother's significantly reduces the probability of marijuana use by 5.4 p.p. and being stopped by police by 7.4 p.p.. These effects suggest that contemporaneous mother's labor-force participation may improve household economic structure or serve as role modeling in ways that immediately translate into lower adolescent risk exposure.

Welfare participation increases the likelihood of failing a class by over 13.2 p.p., this is a large effect even by the standards of the baseline model. This magnitude remains sizable even when past welfare exposure is excluded, highlighting the immediacy of economic conditions in shaping schooling outcomes. Current smoking is associated with significant increases in marijuana use by 6.2 p.p., and early sexual activity by 7.2 p.p.. A contemporaneous maternal conviction increases the probability of failing a class by 18.6 p.p.. This magnitude underscores how active criminal involvement can

[10] $P(Y_{t+1}|B_t, \bullet) = P(Y_{t+1}|B_t, B_{t-1}, B_{t-2} \dots, B_0, \bullet)$

[11] Full model-fit diagnostics for this specification are reported in Online Appendix Tables OA.21–OA.22 and Figure OA.2.

expose children to instability in ways that have immediate educational consequences. Finally, child cognitive achievement remains a powerful predictive factor even under the contemporaneous parental behavior framework. A ten-point increase in cognitive score significantly reduces the likelihood of failing a class and of early sexual activity.

The results suggest that modeling parental behavior as a dynamic pattern over the pre-teenage period is empirically and conceptually superior to a contemporaneous parental behavior specification that conditions only on the most recent behavior. While the contemporaneous parental behavior model captures short-run impacts, it understates the effects that arise from persistent exposure to parental behaviors. The baseline model (pattern-based approach) more accurately reflects how children experience parental behaviors, not as an isolated one-period shocks, but as sustained regimes of stability or instability that shape teenage outcomes. Empirically, the pattern specification generates more consistent effects on teenage educational attainment, substance use, and deviant outcomes.[12]

9 Policy Experiments

Having established the dynamic impacts of parental behaviors on teenage outcomes, I use the model to assess the extent to which policy interventions aimed at shifting parental behaviors along different dimensions can generate meaningful improvements in teenage outcomes. Specifically, the first experiment imposes an excise tax to improve parental behavior and teenage outcomes, the second promotes sustained co-residence with the biological father, and the third upgrades a teenager's peer environment.

9.1 Experiment 1: From Sin-to-Skill Policy

In the Sin-to-Skill policy, I consider a counterfactual policy that imposes an excise tax on beer and cigarettes (sin tax), with all revenues invested in child education that generate a 0.1 standard deviation (SD) increase in test score, hence the name Sin-to-Skill policy. Specifically, the policy considers three increments in tax levels (i.e., \$0.5, \$1, and \$1.5 per gallon of beer and pack of cigarette). I assume that the associated revenues are reinvested through child development programs that collectively raise county-level public school test scores by 0.1 standard deviation (SD), which is an increment consistent with the teacher quality/value-added literature and Project STAR (Krueger, 1999; Chetty et al., 2011). These investments may include programs like early education

[12] This ranking is confirmed formally by a Likelihood Ratio test comparing the behavior pattern (baseline) model against both the behavior summation model and the contemporaneous behavior model. (see Appendix A4.1.2 for details).

and parental support programs, and direct family-oriented educational spending targeting needy families.

I simulate this policy experiment by imposing an increase in beer and cigarette tax by \$0.5 all periods. The tax increase and resulting test-score gain are imposed as exogenous shifts to the relevant behavioral and production equations, rather than as a policy that parents or children anticipate and re-optimize against in a forward-looking manner. To account for administrative lag in public spending and the time required for investment in education to affect children's cognitive achievement, I model the associated increase in county-level test score as occurring only after the first two periods. Also, rather than imposing a uniform increase in test score, I allow the policy effect to vary across counties by drawing the test score increment from a normal distribution, $\mathcal{N}(0.1 * \text{SD}, 1)$, reflecting differences in school quality and program implementation. I repeat this experiment for the case when the tax is increased by \$1 and \$1.5. Table 15 reports the effects of the sin tax policy on parental behaviors and teenage outcomes. The table presents predicted outcomes under each tax level, a no-intervention baseline, and the difference between intervention and baseline outcomes.

Among parental behaviors, only welfare receipt responds significantly to the policy, and it does so at every tax level: it falls by 6.0 percentage points (p.p.) under the \$0.5 tax, 9.4 p.p. under the \$1 tax, and 10.9 p.p. under the \$1.5 tax. Parental smoking declines significantly only at the lowest tax level, by 4.5 p.p., and parental marijuana use declines significantly at the \$0.5 and \$1 tax levels, by 1.3 and 1.9 p.p. respectively. These results suggest that the sin taxes, combined with educational investment, shift families away from public assistance toward labor market attachment, and is effective in curbing substance use. Among teenage outcomes, only alcohol use responds significantly to the policy: the probability of ever having drunk alcohol falls by 2.9, 3.3, and 3.6 p.p. at the \$0.5, \$1, and \$1.5 tax levels, respectively.

A key insight from Table 15 emerges when comparing marginal changes as the tax increases from \$0.5 to \$1 and from \$1 to \$1.5. For parental behaviors, each additional \$0.5 increase in the tax generates sizable improvements, particularly for employment, welfare receipt and smoking. In contrast, for teenage outcomes, the marginal gains from increasing the tax diminish as the tax rises. This distinction has important policy implications. If the government's primary objective is to curb harmful parental behaviors, such as substance use or over-reliance on welfare, then higher and progressively increasing sin taxes appear highly effective. However, if the policy goal extends to improving downstream teenage outcomes, then most of the gains appear to be realized at relatively modest tax levels, with diminishing returns to larger tax increases.

9.2 Experiment 2: Father-to-Foundation Policy

In 2025 the Office of Family Assistance (OFA) launched the FORGE Fatherhood program, which supports responsible fatherhood by promoting and sustaining marriage, responsible parenting, and economic stability. This section evaluates a related counterfactual policy experiment that incentivizes biological fathers to cohabit with or marry the mother and remain present and non-incarcerated during the child's early years. Conceptually, the policy can be interpreted as a tax credit, cash transfer, or economic training program conditioned on biological parents residing together during targeted developmental periods in fragile families.

To assess whether the duration of co-residence matters for downstream teenage outcomes, I simulate the experiment across three incremental intervention windows: from birth through Age 1, from birth through Age 3, and from birth through Age 5. In each window, I impose that all mothers are married to or cohabiting with the biological father, and that the biological father is simultaneously present and non-incarcerated, for the duration of the intervention. Following each window, parents resume their endogenously determined lifecycle behaviors. Table 16 reports predicted teenage outcomes under each intervention duration alongside the no-intervention baseline.

The results reveal a clear dose-response pattern: no effect reaches statistical significance under the shortest, one-year intervention window, but effects strengthen and become statistically significant as the window extends toward Age 5. Extending co-residence through Age 3 makes the reduction in marijuana use significant, falling by 2.2 p.p.. Sustaining co-residence through Age 5 yields the largest estimated effects: the probability of ever failing a class falls by 3.9 p.p., school suspension falls by 3.4 p.p., and marijuana use falls by 2.1 p.p.. These findings are consistent with Heckman's technology of skill formation (Heckman, 2006; Cunha and Heckman, 2007), which emphasizes that the period from birth through age five is the most critical window for early childhood investment, since cognitive and non-cognitive skills are most malleable and early inputs are most productive. The model results suggest that paternal co-residence generates statistically detectable returns only when sustained through this full early-childhood window; interventions confined to the first year of life are too short to produce measurable improvements in teenage outcomes.

Consistent with this interpretation, the effects that do emerge likely operate through improved household resources and reduced maternal stress feeding into the cognitive production function during the period when it is most responsive, with these early cognitive gains propagating into the observed improvements in educational and behavioral outcomes at age 15. For the full-lifecycle comparison, in which the mother is always married to or cohabiting with the biological father throughout childhood, the reader is referred to Table 10, where effects are larger, underscoring that it is the sustained behavioral trajectory that drives the strongest outcomes.

9.2.1 The Heckman Critical Window Test

To directly test the critical window hypothesis, I conduct an additional experiment that imposes the same co-residence requirement (biological father married to or cohabiting with the mother, present and non-incarcerated) exclusively to the period after Age 5, when Heckman's critical window has closed. Table 17 presents the results side by side: the Birth–Age 5 intervention in the left column and the Above Age 5 intervention in the right column. Within the critical window (Birth–Age 5), co-residence reduces grade failure by 3.9 p.p., school suspension by 3.4 p.p., and marijuana use by 2.1 p.p.. Beyond the critical window (Above Age 5), the effects shrink and are insignificant, and in several cases reverse in sign: the probability of ever drinking alcohol, marijuana use, and having sex all rise. This within-versus-beyond comparison constitutes a direct empirical validation of Heckman's technology of skill formation within a dynamic model. Figure 6 traces this same pattern directly through the child's cognitive score: the simulated gain rises monotonically from 0.044 of a baseline SD under the Birth–Age 1 window, through to 0.072 under the full Birth–Age 5 window, compared to 0.017 SD gain under the Above Age 5 placebo. The same policy intervention (Father-to-Foundation) generates significant improvements in teenage outcomes when applied during the 0–5 critical period, and essentially no to less improvement when applied after it. The result is driven by when the parental behavior change occurs in the developmental lifecycle. This finding reinforces the message: for children in fragile families, the timing of paternal investment matters as much as its presence, and policies that fail to reach families during the early childhood critical window are unlikely to generate lasting improvements in teenage outcomes.

9.3 Experiment 3: Peer-to-Promise Policy

Adolescent peer networks are a well-documented channel through which risk behaviors and educational aspirations spread among teenagers (Ali and Dwyer, 2009). Policies that reshape a child's social environment, such as school choice and magnet-school access, or neighborhood mobility programs that relocate families to higher-opportunity areas (Chetty and Hendren, 2018), operate through exactly this channel by changing who a teenager's peers are, rather than by changing the teenager's own household directly. This section evaluates a counterfactual Peer-to-Promise policy that upgrades a teenager's peer environment: shifting peers away from substance use (drinking, smoking, and marijuana use) and toward positive peer behaviors (studying and college aspirations), without altering any of the parental behaviors modeled elsewhere in the paper.

I implement this experiment by reassigning teenagers' peer characteristics at three incremental magnitudes, $\lambda = 0.1$, 0.2, and 0.3 standard deviations (SD). For the substance-use peer variables (peer drinking, peer smoking, peer marijuana use), a fraction of peers is randomly reassigned from users to non-users such that the peer group mean falls by $\lambda\times$ SD; for the positive peer variables (peer study habits, peer likelihood of graduating college), a fraction of peers is randomly reassigned upward such that the peer group mean rises by $\lambda\times$ SD. All peer variables remain binary throughout the simulation. This policy can be interpreted as a stylized school reassignment, magnet school access, or neighborhood mobility program that shifts the social environment a teenager is embedded in, holding the family environment fixed. Table 18 reports the resulting teenage outcomes under each magnitude of intervention relative to a no-intervention baseline reflecting the observed peer environment in the FFCWS sample.

The policy exhibits a clear dose-response pattern: as the peer-environment shift grows from $\lambda = 0.1$ to $\lambda = 0.3$ SD, both the magnitude and the number of statistically significant effects increase. At the smallest shift ($\lambda = 0.1$), only teenage drinking responds significantly, falling by 2.2 p.p.. At the intermediate shift ($\lambda = 0.2$), the reduction in drinking grows to 3.9 p.p., and three additional outcomes become significant: marijuana use falls by 3.6 p.p., police contact falls by 2.8 p.p., and having had sex falls by 2.6 p.p.. At the largest shift ($\lambda = 0.3$), all four of these effects strengthen further (drinking falls by 5.3 p.p., marijuana use by 5.7 p.p., police contact by 4.3 p.p., and having had sex by 3.8 p.p.), and two educational outcomes become statistical significance: the probability of being unlikely to graduate college falls by 3.3 p.p., and the probability of ever being suspended from school falls by 3.1 p.p.. The probability of ever failing a class declines modestly at every magnitude but is not statistically significant at any level.

These results reveal that the peer channel operates with a threshold-like structure: modest peer-environment improvements meaningfully curb substance use, but noticeably larger shifts are needed before the policy registers detectable gains in school suspension and college aspirations, and failing a class never responds significantly within the range of magnitudes considered. This pattern suggests that peer-based interventions, such as school assignment or neighborhood mobility programs, may need to induce fairly substantial changes in a teenager's social environment before generating measurable academic benefits, even though more modest shifts are sufficient to reduce substance use and risk-taking behavior. Relative to the Sin-to-Skill policy, which achieves its most robust effects on parental welfare receipt and a narrow set of downstream teenage outcomes, the Peer-to-Promise policy reaches a broader set of statistically significant teenage outcomes at its largest magnitude, underscoring that adolescent peer environments are a potent lever for shaping teenage outcomes.

10 Conclusion

The paper estimates the causal impacts of observed parental behaviors (marriage/cohabitation, employment, social services assistance, substance use, and criminal behavior) at different ages of child development on the child's teenage outcomes, specifically, cognitive ability, educational attainment, substance use, and criminal outcomes. Using data from the Fragile Families and Child Wellbeing Study (FFCWS), I jointly estimate children's teenage outcomes with parents' age-specific behaviors, accounting for both permanent and time-varying unobserved family heterogeneity, and allowing for the effects of parental behavioral trajectories to operate through both direct pathways and an indirect channel running through the child's cognitive skill development. Beyond accounting for unobserved heterogeneity, the model jointly estimates five parental behaviors, marriage/cohabitation, employment, welfare receipt, substance use, and criminal behavior, as dynamically and jointly determined choices rather than as isolated, single-period decisions, and maps their evolving trajectories onto multiple teenage outcomes spanning education, substance use, and delinquency within one internally consistent framework. Endogeneity is addressed throughout using theoretically motivated instruments and the dynamic panel structure of the data, rather than relying on the single-behavior, single-period reduced-form approaches common in the existing literature.

I find that the timing, persistence, and joint nature of parental behaviors matter for teenage outcomes. Sustained marital stability and consistent employment generate the largest gains in educational attainment, while criminal record and substance use effects are more heterogeneous across outcomes. The joint modeling of these behaviors proves essential: the combined effect of adverse behavioral patterns exceeds the sum of their isolated marginal contributions, reflecting the nonlinear and mutually reinforcing nature of disadvantage in fragile families. The decomposition of behavioral effects into direct behavioral and indirect cognitive channels reveals that parental behaviors shape adolescent outcomes through mechanisms beyond cognitive development alone, that includes non-cognitive development, and role modeling.

Three policy experiments illustrate what the model's estimates imply for intervention design. A Sin-to-Skill policy combining excise taxes on beer and cigarettes with reinvestment into county-level education significantly reduces welfare dependence and, to a lesser degree, parental smoking, marijuana use, and teenage drinking. A Father-to-Foundation policy that sustains biological-father co-residence generates statistically detectable gains in educational and substance-use outcomes only when maintained through age five, underscoring that the benefits of early family stability accumulate over developmental time rather than materializing from short-run interventions. A Peer-to-Promise policy that upgrades a teenager's peer environment shows a similar dose-response, with larger shifts in peer composition generating significant improvements across a broad set of educational, substance-use, and deviant-behavior outcomes. The policy results underscore that improving child outcomes in fragile families requires sustained, multi-dimensional interventions rather than isolated policy changes targeting a single behavior at a single point in time.

References

Agostinelli, F. and Sorrenti, G. (2021). Money vs. time: family income, maternal labor supply, and child development. Technical Report 273, University of Zurich, Department of Economics.

Agostinelli, F. and Wiswall, M. (2025). Estimating the technology of children's skill formation. *Journal of Political Economy*, 133(3):846–887.

Akee, R. K., Copeland, W. E., Keeler, G., Angold, A., and Costello, E. J. (2010). Parents' incomes and children's outcomes: a quasi-experiment using transfer payments from casino profits. *American Economic Journal: Applied Economics*, 2(1):86–115.

Ali, M. M. and Dwyer, D. S. (2009). Estimating peer effects in adolescent smoking behavior: a longitudinal analysis. *Journal of Adolescent Health*, 45(4):402–408.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Arteaga, C. (2023). Parental incarceration and children's educational attainment. *Review of Economics and Statistics*, 105(6):1394–1410.

Attanasio, O., Bernal, R., Giannola, M., and Nores, M. (2020). Child development in the early years: Parental investment and the changing dynamics of different dimensions. Technical report, National Bureau of Economic Research.

Bailey, J. A., Epstein, M., Steeger, C. M., and Hill, K. G. (2018). Concurrent and prospective associations between substance-specific parenting practices and child cigarette, alcohol, and marijuana use. *Journal of Adolescent Health*, 62(6):681–687.

Balsa, A. I. (2008). Parental problem-drinking and adult children's labor market outcomes. *Journal of Human Resources*, 43(2):454–486.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4):1173–1209.

Carneiro, P., García, I. L., Salvanes, K. G., and Tominey, E. (2021). Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, 129(3):757–788.

Chan, M. K. and Liu, K. (2025). Changing families: Family relationships, parental decisions, and child development. *Journal of Labor Economics*, 43(S1):S399–S444.

Chang, Y., Durlauf, S. N., Hu, B., and Park, J. Y. (2025). The effects of parental income and family structure on intergenerational mobility: A trajectories-based approach. *Journal of Applied Econometrics*.

Chatterji, P. and Markowitz, S. (2001). The impact of maternal alcohol and illicit drug use on children's behavior problems: evidence from the children of the national longitudinal survey of youth. *Journal of health economics*, 20(5):703–731.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics*, 126(4):1593–1660.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Cho, R. M. (2009a). Impact of maternal imprisonment on children's probability of grade retention. *Journal of Urban Economics*, 65(1):11–23.

Cho, R. M. (2009b). The impact of maternal imprisonment on children's educational achievement: Results from children in chicago public schools. *Journal of Human Resources*, 44(3):772–797.

Craigie, T.-A. L., Brooks-Gunn, J., and Waldfogel, J. (2012). Family structure, family stability and outcomes of five-year-old children. *Families, Relationships and Societies*, 1(1):43–61.

Cummings, J. R. and Gearhardt, A. N. (2020). Prospective associations between parental substance use and youth highly processed food intake in the fragile families and child wellbeing study. *Drug and alcohol dependence*, 211:107965.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American economic review*, 97(2):31–47.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Cunningham, S. and Finlay, K. (2013). Parental substance use and foster care: Evidence from two methamphetamine supply shocks. *Economic Inquiry*, 51(1):764–782.

Dobbie, W., Grönqvist, H., Niknami, S., Palme, M., and Priks, M. (2018). The intergenerational effects of parental incarceration. Technical report, National Bureau of Economic Research.

Dunn, L. M. and Dunn, L. M. (1965). *Peabody Picture Vocabulary Test.* American Guidance Service, Circle Pines, MN.

FFCWS (2023). Future of families child welfare survey. *Princeton University.*

Fruehwirth, J. C., Iyer, S., and Zhang, A. (2019). Religion and depression in adolescence. *Journal of Political Economy*, 127(3):1178–1209.

Geller, A. (2013). Paternal incarceration and father–child contact in fragile families. *Journal of Marriage and Family*, 75(5):1288–1303.

Geller, A., Garfinkel, I., Cooper, C. E., and Mincy, R. B. (2009). Parental incarceration and child well-being: implications for urban families. *Social science quarterly*, 90(5):1186–1202.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902.

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, 6(1):689–733.

Kailaheimo-Lönnqvist, S., Kuja-Halkola, R., Larsson, H., Lichtenstein, P., and Latvala, A. (2022). Parental criminality and children's educational attainment: A population-based extended family study. *Journal of criminal justice*, 81:101920.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, 105(3):473–522.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, 114(2):497–532.

Lessem, R. and Sanders, C. (2025). The effects of family structure on children's outcomes. Preliminary manuscript, July 2025.

Liu, H., Mroz, T. A., and Van der Klaauw, W. (2010). Maternal employment, migration, and child development. *Journal of Econometrics*, 156(1):212–228.

McKernan, S.-M., Ratcliffe, C., and Braga, B. (2021). The effect of the us safety net on material hardship over two decades. *Journal of Public Economics*, 197:104403.

Mroz, T. A. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics*, 92(2):233–274.

Norris, S., Pecenco, M., and Weaver, J. (2021). The effects of parental and sibling incarceration: Evidence from ohio. *American Economic Review*, 111(9):2926–2963.

Osborne, C. and Berger, L. M. (2009). Parental substance abuse and child well-being: A consideration of parents' gender and coresidence. *Journal of Family Issues*, 30(3):341–370.

Ratcliffe, C., McKernan, S.-M., and Zhang, S. (2011). How much does the supplemental nutrition assistance program reduce food insecurity? *American journal of agricultural economics*, 93(4):1082–1098.

Sampson, R. J. and Laub, J. H. (1995). *Crime in the making: Pathways and turning points through life*. Harvard University Press.

Schaller, J. and Zerpa, M. (2019). Short-run effects of parental job loss on child health. *American Journal of Health Economics*, 5(1):8–41.

Schmidt, C. M. and Tauchmann, H. (2011). Heterogeneity in the intergenerational transmission of alcohol consumption: A quantile regression approach. *Journal of Health Economics*, 30(1):33–42.

Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485):F3–F33.

Turney, K. (2017). The unequal consequences of mass incarceration for children. *Demography*, 54(1):361–389.

Waldfogel, J., Craigie, T.-A., and Brooks-Gunn, J. (2010). Fragile families and child wellbeing. *The Future of children/Center for the Future of Children, the David and Lucile Packard Foundation*, 20(2):87.

Waller, M. R. and Swisher, R. (2006). Fathers' risk factors in fragile families: Implications for "healthy" relationships and father involvement. *Social Problems*, 53(3):392–420.

Wildeman, C. (2010). Paternal incarceration and children's physically aggressive behaviors: Evidence from the fragile families and child wellbeing study. *Social Forces*, 89(1):285–309.

Wildeman, C. and Western, B. (2010). Incarceration in fragile families. *The future of children*, pages 157–177.

Williams, J., Van Ours, J. C., and Grossman, M. (2011). Why do some people want to legalize cannabis use? Technical report, National Bureau of Economic Research.

Yen, S. T., Bruce, D. J., and Jahns, L. (2012). Supplemental nutrition assistance program participation and health: Evidence from low-income individuals in tennessee. *Contemporary Economic Policy*, 30(1):1–12.

Tables

Table 1: **FFCWS Sample Compared to the 1998 US Population**

		FFCWS data	**1998 US Population**	
		Women in baseline sample	Single women giving birth	All women
Race	White non-Hispanic	21%	40%	60%
	Black non-Hispanic	48%	32%	15%
	Hispanic	27%	24%	19%
	Other	4%	4%	6%
Unmarried		76%	100%	44%
US–Born		83%	83%	88%
Education	Less than High School	35%		22%
	High School	30%		33%
	Some College	24%		22%
	College or Graduate	11%		23%

Notes: This table compares the demographic characteristics of mothers in the Fragile Families and Child Wellbeing Study (FFCWS) baseline sample to two reference groups drawn from the 1998 US population: single women who recently gave birth, and all women of childbearing age. The FFCWS is a birth cohort study that surveyed families at the time of a child's birth in large US cities (population over 200,000) between 1998 and 2000, with a deliberate oversample of unmarried parents. As a result, the FFCWS baseline sample is disproportionately composed of unmarried, minority, and less-educated mothers relative to the general US population, but is broadly comparable to the national population of single women giving birth. Education statistics for single women giving birth are not available for the 1998 US population comparison and are therefore left blank. Percentages may not sum to 100 due to rounding.

Table 2: **Summary Statistics of Endogenous Variables across Waves (N = 1,615)**

	Wave:	**1**	**2**	**3**	**4**	**5**	**6**
Period t Behaviors	**Child's age:**	**Birth**	**Year 1**	**Year 3**	**Year 5**	**Year 9**	**Year 15**
Married/Cohabitation (Dad Present)							
Single (No Dad)		39%	41%	45%	49%	54%	42%
With biological father		61%	58%	54%	47%	38%	32%
With another man			1%	1%	4%	8%	26%
Mom Employed		81%	77%	81%	80%	79%	73%
Mom on Welfare		34%	24%	20%	17%	17%	9%
Mom Substance Use							
Alcohol		10%	35%	55%	58%	65%	64%
Smoking		18%	26%	21%	30%	28%	24%
Marijuana			1%	3%	2%	4%	5%
Mom Convicted		1%	1%	1%	3%	2%	3%
Dad in Jail		4%	5%	7%	7%	7%	16%
Cognitive Score (average)			80.000	80.017	80.039	80.003	

Notes: The table reports the distribution of endogenous parental behaviors and child cognitive achievement across the six waves of the Fragile Families and Child Wellbeing Study (FFCWS), corresponding to the child's birth, age 1, age 3, age 5, age 9, and age 15. The sample comprises 1,615 families who are observed across all six waves. All behavioral variables are binary indicators reported as proportions, with the exception of marital/cohabitation status, which is a three-category variable distinguishing between single (no father present), cohabiting or married with the biological father, and cohabiting or married with another man. Alcohol use at Wave 1 (Birth) reflects drinking reported during pregnancy and is therefore substantially lower than subsequent waves. Marijuana use and conviction rates are blank at Birth as these variables are not observed at that wave. The cognitive score is a standardized measure of the child's cognitive achievement normalized to a mean of 80 within each wave. The sharp increase in father's incarceration at Age 15 and the rise in cohabitation with another man over time are consistent with the broader literature on relationship instability in fragile families.

Table 3: **Summary Statistics of Teenage Outcome Variables (N = 1,615)**

	Wave	**6**
	Child's age	**Year 15**
Education	Ever Failed a Class	46.2%
	Likely Not to Graduate College	32.5%
	Ever Suspended/Expelled from School	27.6%
Substance Use	Ever Drunk Alcohol	13.0%
	Ever Used Marijuana	17.8%
Deviant Outcomes	Ever Stopped by Police	25.7%
	Ever Had Sex	15.0%

Notes: The table reports the seven teenage outcome variables measured at Wave 6 of the Fragile Families and Child Wellbeing Study (FFCWS), when the focal child is 15 years of age. All outcomes are binary indicators reported as proportions. *Education* outcomes capture academic performance and school behavior, including whether the child has ever failed a class, likely not of graduating college, and whether the child has ever been suspended or expelled from school. *Substance use* outcomes capture whether the child has ever consumed alcohol or used marijuana by age 15. *Deviant outcomes* capture early engagement in risky behaviors, including whether the child has ever been stopped by the police and whether the child has ever had sex.

Table 4: **Difference in Means of Teenage Outcomes by Parental Behaviors**

	Teenage Outcomes						
	Ever Failed Class	Unlikely to Graduate College	Ever Suspended from Sch	Ever Drunk Alcohol	Ever Used Marijuana	Ever Stopped by Police	Ever Had Sex
Mom's Marital Status							
Always Married to Dad	0.239	0.209	0.086	0.108	0.056	0.157	0.026
Not Always Married to Dad	0.506	0.348	0.313	0.134	0.202	0.277	0.175
Difference	−0.268***	−0.139***	−0.227***	−0.026	−0.146***	−0.120***	−0.149***
Mom's Employment Status							
Always Employed	0.412	0.273	0.249	0.133	0.171	0.231	0.140
Not Always Employed	0.502	0.366	0.296	0.128	0.183	0.277	0.159
Difference	−0.090***	−0.093***	−0.047**	0.005	−0.012	−0.046**	−0.019
Mom's Welfare Status							
Not Always on Welfare	0.459	0.324	0.271	0.129	0.176	0.251	0.147
Always on Welfare	0.700	0.450	0.650	0.200	0.300	0.700	0.450
Difference	−0.241**	−0.126	−0.379***	−0.071	−0.124	−0.449***	−0.303***
Mom's Drinking Behavior							
Never Drinks Alcohol	0.467	0.275	0.263	0.100	0.146	0.233	0.138
Drinks Alcohol	0.461	0.334	0.278	0.135	0.183	0.261	0.153
Difference	0.006	−0.059*	−0.015	−0.035	−0.037	−0.028	−0.015
Mom's Smoking Behavior							
Never Smokes Cigarettes	0.412	0.303	0.231	0.121	0.140	0.232	0.124
Smokes Cigarettes	0.542	0.359	0.347	0.144	0.238	0.297	0.192
Difference	−0.130***	−0.056**	−0.116***	−0.022	−0.098***	−0.066***	−0.067***
Mom's Marijuana Use							
Never Uses Marijuana	0.444	0.314	0.251	0.125	0.163	0.247	0.138
Uses Marijuana	0.639	0.435	0.517	0.177	0.320	0.354	0.272
Difference	−0.195***	−0.121***	−0.266***	−0.052*	−0.156***	−0.106***	−0.134***
Mom's Criminal Behavior							
No Criminal Record	0.441	0.310	0.256	0.124	0.163	0.246	0.142
Criminal Record	0.630	0.448	0.431	0.177	0.293	0.343	0.221
Difference	−0.189***	−0.138***	−0.175***	−0.053**	−0.130***	−0.096***	−0.079***
Dad's Criminal Behavior							
Never In Jail	0.377	0.272	0.171	0.105	0.111	0.201	0.094
Been In Jail	0.543	0.376	0.376	0.154	0.241	0.310	0.205
Difference	−0.166***	−0.104***	−0.205***	−0.049***	−0.130***	−0.109***	−0.111***

Notes: The table reports mean teenage outcomes at age 15 separately by parental behavior group, along with the difference in means between groups, for each of the eight parental behaviors modeled in this paper. All teenage outcomes and parental behavior indicators are binary variables; values represent proportions. For each parental behavior, families are classified based on whether the mother (or father) ever exhibited that behavior across the five pre-teenage developmental periods (Birth, Age 1, Age 3, Age 5, and Age 9). Marital status is classified based on whether the mother was always married to or cohabiting with the biological father across all periods. Sample size is N = 1,615 families. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: **Parental Behavior Patterns**

Mom Marital Status ($\overrightarrow{M}_t$)	***Mom Employment*** ($\overrightarrow{E}_t$)	***Mom Welfare Receipt*** ($\overrightarrow{W}_t$)
always single	always unemployed	never on welfare
with biodad always	always employed	always on welfare
single early	ever emp. early & always unemp. late	ever on wel. early & never on wel. late
with biodad early	always unemp. early & ever emp. late	never on wel. early & ever on wel. late
unstable relationship	intermittent employment	intermittent welfare
other marital patterns		
Mom Drinking ($\overrightarrow{S}_t^1$)	***Mom Smoking*** ($\overrightarrow{S}_t^2$)	***Mom Marijuana Use*** ($\overrightarrow{S}_t^3$)
never drinks	never smokes	never use marijuana
always drink	always smoke	ever use marijuana
start drinking early & quit late	start smoking early & quit late	
never drunk early & start drinking late	never smoked early & start smoking late	
intermittent drinking	intermittent smoking	
Mom Conviction ($\overrightarrow{R}_t$)	***Dad in Jail*** ($\overrightarrow{J}_t$)	
never convicted	never in jail	
first-ever conviction was early	first-ever jail was early	
first-ever conviction was late	first-ever jail was late	

Notes: The table summarizes the parental behavior trajectory patterns. Rather than using single-period behavioral snapshots, each pattern captures the full lifecycle history of a parental behavior, denoted by the arrow notation $\overrightarrow{\cdot}$, accumulated across the six developmental periods: Birth, Age 1, Age 3, Age 5, Age 9, and Age 15. "Early" refers to behaviors occurring during the first three periods (Birth, Age 1, and Age 3), while "Late" corresponds to behaviors observed during the latter three periods (Age 5, Age 9, and Age 15). Marital status patterns distinguish between stability with the biological father, instability, and transitions involving other partners. Employment and welfare patterns capture persistence, entry, exit, and intermittency over the lifecycle. Substance use patterns for drinking and smoking additionally distinguish between early initiation with later cessation, late initiation, and intermittent use. Due to its lower prevalence in the sample, marijuana use is captured by a binary indicator of whether the mother ever used marijuana over the lifecycle. Similarly, conviction and incarceration patterns distinguish between never occurring, first occurring early, and first occurring late in childhood. All patterns are constructed from the FFCWS data and reflect the observed behavioral heterogeneity among fragile families.

Table 6: **Specification Summary: Jointly Estimated Behaviors and Outcomes**

Dependent Variables	**Explanatory Variables** Endogenous	Exogenous	**Unobservables** ε_t	**Likelihood Contribution**
Parental Behaviors				
Marital Status M_t	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^M,\ \nu_t^M,\ u_t^M$	$\mathbb{P}(M_t = m),\ m = 0, 1, 2$
Employment E_t	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^E,\ \nu_t^E,\ u_t^E$	$\mathbb{P}(E_t = e),\ e = 0, 1$
Welfare W_t	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^W,\ \nu_t^W,\ u_t^W$	$\mathbb{P}(W_t = w),\ w = 0, 1$
Drinking S_t^1	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{S^1},\ \nu_t^{S^1},\ u_t^{S^1}$	$\mathbb{P}(S_t^1 = s^1),\ s^1 = 0, 1$
Smoking S_t^2	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{S^2},\ \nu_t^{S^2},\ u_t^{S^2}$	$\mathbb{P}(S_t^2 = s^2),\ s^2 = 0, 1$
Marijuana use S_t^3	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{S^3},\ \nu_t^{S^3},\ u_t^{S^3}$	$\mathbb{P}(S_t^3 = s^3),\ s^3 = 0, 1$
Conviction R_t	$M_{t-1}, E_{t-1}, W_{t-1}, \boldsymbol{S}_{t-1}, R_{t-1}, \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^R,\ \nu_t^R,\ u_t^R$	$\mathbb{P}(R_t = r),\ r = 0, 1$
Jail J_t	$J_{t-1},\ \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^d,\ \boldsymbol{Z}_t$	$\mu^J,\ \nu_t^J,\ u_t^J$	$\mathbb{P}(J_t = j),\ j = 0, 1$
Child				
Cognitive score θ_t	$M_t,\ E_t,\ W_t,\ \boldsymbol{S}_t,\ R_t,\ J_t,\ \theta_{t-1}$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{X}_t^d, Z_t^\theta$	$\mu^\theta,\ \nu_t^\theta,\ u_t^\theta$	$\phi^{\theta_{t+1}}(\cdot)$
Teenage outcome Y_{T+1}^o	$\overrightarrow{M}_t,\ \overrightarrow{E}_t,\ \overrightarrow{W}_t,\ \overrightarrow{\boldsymbol{S}}_t,\ \overrightarrow{R}_t,\ \overrightarrow{J}_t,\ \theta_t$	$\boldsymbol{X}_t^c, \boldsymbol{X}_t^m, \boldsymbol{X}_t^d, Z_t^\theta, \boldsymbol{Z}_t^p$	$\mu^o,\ \nu_t^o,\ u_{t+1}^o$	$\mathbb{P}(Y_{T+1}^o = y),\ y = 0, 1$ $o = 1, 2, \ldots, 7$
Initial Conditions				
Initial marital status M_1		$\boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{M_1},\ u_1^{M_1}$	$\mathbb{P}(M_1 = m),\ m = 0, 1$
Initial employment E_1		$\boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{E_1},\ u_1^{E_1}$	$\mathbb{P}(E_1 = e),\ e = 0, 1$
Initial welfare W_1		$\boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{W_1},\ u_1^{W_1}$	$\mathbb{P}(W_1 = w),\ w = 0, 1$
Initial drinking S_1^1		$\boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{S_1^1},\ u_1^{S_1^1}$	$\mathbb{P}(S_1^1 = s^1),\ s^1 = 0, 1$
Initial smoking S_1^2		$\boldsymbol{X}_t^m, \boldsymbol{Z}_t$	$\mu^{S_1^2},\ u_1^{S_1^2}$	$\mathbb{P}(S_1^2 = s^2),\ s^2 = 0, 1$
Initial jail J_1		$\boldsymbol{X}_t^d,\ \boldsymbol{Z}_t$	$\mu^{J_1},\ u_1^{J_1}$	$\mathbb{P}(J_1 = j),\ j = 0, 1$
Initial cognitive score θ_1	$M_t,\ E_t,\ W_t,\ \boldsymbol{S}_t,\ R_t$	$\boldsymbol{X}_t^m, \boldsymbol{X}_t^d, Z_t^\theta$	$\mu^{\theta_1},\ u_1^{\theta_1}$	$\phi^{\theta_1}(\cdot)$

Notes: The table summarizes the full system of 23 jointly estimated equations comprising 8 parental behavior equations, 1 child cognitive score equation, 7 teenage outcome equations, and 7 initial conditions equations. $\boldsymbol{X}_t^c$, $\boldsymbol{X}_t^m$, and $\boldsymbol{X}_t^d$ denote vectors of time-varying exogenous characteristics of the child, mother, and father, respectively. The mother's policy and supply-side environment is $\boldsymbol{Z}_t = (Z_t^M, Z_t^E, Z_t^W, Z_t^{S^l}, Z_t^R)$, comprising supply-side variables related to marital status, employment, welfare receipt, substance use, and criminal conviction, respectively. For the father's jail equation, $\boldsymbol{Z}_t = \{Z_t^J\}$ denotes the supply-side environment governing incarceration. Z_t^θ denotes county-level variables affecting child cognitive achievement, and $\boldsymbol{Z}_t^p$ denotes peer characteristics that enter the teenage outcome equations. μ and ν_t denote permanent and time-varying unobserved family heterogeneity, respectively, while u_t is an equation-specific idiosyncratic error term. I estimate the joint distribution of the μ's and ν_t's. The arrow notation $\overrightarrow{\cdot}$ denotes the full history of a variable accumulated over the developmental lifecycle. Discrete equations contribute logit or multinomial logit probabilities to the likelihood, while continuous equations contribute normal densities $\phi(\cdot)$.

Table 7: **Supply-Side Variables Categorized by Related Behavior ($\boldsymbol{Z}_t$)**

Related Behavior	**Instrumental Variable**
Marital Status (M)	County-level fraction of people aged 15 or older who are married and not separated
Employment (E)	City-level unemployment rate
Welfare (W)	County-level percentage of households on public assistance
Substance Use (S)	
Alcohol (S^1)	State-level alcohol tax per gallon of beer
Smoking (S^2)	State tax per pack of cigarettes
Marijuana (S^3)	State marijuana decriminalization laws
Crime (R and J)	County-level violent crime rate
Cognitive Score (θ)	County-level income-adjusted test score
Teenage Outcomes (Y^o)	Peer drinking, peer smoking, peer marijuana use, peer study habits, peer likelihood of graduating college

Notes: All supply-side instruments are measured in the calendar year corresponding to the observed outcome B_t or θ_t in which they serve as exclusion restrictions. In addition to these external instruments, lagged endogenous behaviors B_{t-1} and lagged cognitive score θ_{t-1} serve as internal instruments in the dynamic equations, exploiting the sequential structure of the model. Marital status, employment, welfare, crime and cognitive instruments are drawn from county- and city-level administrative data from the Bureau of Labor Statistics and Opportunity Insight. Alcohol and smoking instruments are drawn from Tax Policy Center and CDC (Orzechowski and Walker) respective. Marijuana instrument data are from Marijuana Policy Project and MJBizDaily. Peer characteristics are drawn from the FFCWS and capture the local peer environment faced by the focal child at age 15.

Table 8: **Summary Statistics for Model Fit**

	Observed		Simulated	
	Mean	Std. Dev.	Mean	Std. Dev.
Parental Behaviors				
Marriage/cohabitation status				
single	0.450	0.498	0.453	0.498
with biodad	0.484	0.500	0.480	0.500
with another man	0.066	0.249	0.068	0.251
Employed	0.782	0.413	0.782	0.413
Welfare Receipt	0.203	0.402	0.198	0.398
Substance Use				
Drinking	0.476	0.499	0.476	0.499
Smoking	0.243	0.429	0.232	0.422
Marijuana use	0.025	0.157	0.024	0.154
Conviction	0.016	0.125	0.015	0.121
Jail	0.076	0.266	0.073	0.261
Child				
Cognitive score	80.012	11.189	79.991	11.082
Teenage Outcomes				
Ever failed class	0.462	0.499	0.473	0.499
Unlikely to graduate college	0.325	0.468	0.331	0.471
Ever suspended from school	0.276	0.447	0.287	0.452
Ever drunk alcohol	0.130	0.336	0.133	0.340
Ever used marijuana	0.178	0.382	0.178	0.383
Ever stopped by police	0.257	0.437	0.260	0.439
Ever had sex	0.150	0.358	0.157	0.364

Notes: The table assesses the in-sample fit of the jointly estimated model by comparing observed and simulated means and standard deviations for all endogenous variables in the system. Observed values are computed directly from the FFCWS sample of N = 1,615 families. Simulated values are generated by drawing 1000 replications from the estimated model, assigning each family a permanent type μ and a time-varying shock ν_t each period from the estimated discrete factor random effects distributions, and simulating the full lifecycle of parental behaviors, child cognitive skill, and teenage outcomes forward from the initial period. All parental behavior and teenage outcome variables are binary indicators; reported means therefore represent proportions. Overall model fit is summarized by the Mean Absolute Deviation (MAD) between simulated and observed means, defined as MAD = $\frac{1}{K}\sum_{k=1}^{K} |\text{Simulated}_k - \text{Observed}_k| = 0.005$, indicating that the model reproduces the observed data moments with a high degree of accuracy.

Table 9: **Estimation Results: Correlated Unobserved Heterogeneity – Baseline Model**

	Married to / Cohabiting with Biodad		**Married to / Cohabiting with Another Man**		**Employment**		**Welfare**		**Prob. Weight**
Mass Point	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	
Permanent									
2	−0.900	0.544	−0.030	0.222	0.770***	0.260	−0.169	0.300	0.341
3	0.310	0.559	0.017	0.256	−0.488**	0.195	0.539***	0.195	0.191
4	−1.035	0.923	0.287	0.390	0.694***	0.266	−0.029	0.294	0.133
Time-varying									
2	−50.654***	0.395	−0.855	1.193	−0.599***	0.196	0.624***	0.192	0.538
3	−51.931***	0.513	−1.034	1.289	1.053	0.700	−15.494***	0.000	0.136
4	−75.975***	0.656	−1.222	1.227	−0.021	0.223	1.375***	0.213	0.174

	Drinking		**Smoking**		**Marijuana Use**		**Conviction**		**Prob. Weight**
Mass Point	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	
Permanent									
2	2.251***	0.189	2.678***	0.738	2.745***	0.696	−0.245	0.542	0.341
3	1.207***	0.316	5.705***	0.646	2.238***	0.642	0.297	0.459	0.191
4	3.165***	0.398	6.246***	0.723	3.873***	0.638	0.800	0.520	0.133
Time-varying									
2	−0.167	0.180	−0.417	0.271	−0.199	0.446	−0.387	0.754	0.538
3	−0.846**	0.330	−1.354***	0.482	−0.919	0.665	−56.060***	0.000	0.136
4	−0.164	0.183	0.406	0.285	0.647	0.406	1.162**	0.530	0.174

	Jail		**Ever Failed Class**		**Unlikely to Graduate College**		**Ever Suspended from School**		**Prob. Weight**
Mass Point	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	
Permanent									
2	0.664***	0.236	0.302	0.501	−0.252	0.400	0.243	0.543	0.341
3	0.469*	0.254	−0.553	0.750	−0.937	0.787	0.241	0.848	0.191
4	0.436	0.402	0.289	1.011	−0.442	0.994	1.671	1.096	0.133
Time-varying									
2	2.557***	0.478	−1.146***	0.397	−0.869**	0.442	−0.733	0.495	0.538
3	2.555***	0.707	1.076	0.898	0.732	0.772	1.693**	0.796	0.136
4	4.311***	0.517	−0.450	0.439	−0.423	0.534	0.631	0.527	0.174

	Ever Drunk Alcohol		**Ever Used Marijuana**		**Ever Stopped by Police**		**Ever Had Sex**		**Prob. Weight**
Mass Point	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	
Permanent									
2	0.658	0.693	1.552*	0.914	0.740	0.521	1.568**	0.763	0.341
3	0.043	1.072	2.082*	1.232	1.181	0.909	1.111	1.112	0.191
4	−0.149	1.315	3.054*	1.654	2.237*	1.280	1.326	1.295	0.133
Time-varying									
2	−2.385***	0.707	−2.173***	0.695	−1.308**	0.520	−1.536**	0.678	0.538
3	0.523	0.925	2.164*	1.260	−0.351	0.767	1.244	0.881	0.136
4	0.794	0.632	3.782***	0.967	0.118	0.518	2.273***	0.789	0.174

Notes: Permanent and time-varying UH mass points are relative to type 1, which is normalized to zero with probability weights of 0.335 and 0.152, respectively. Coefficients for the marital status equation are log odds relative to the reference category Single (No Dad Present). All other behavioral and outcome equations report log odds coefficients. All types are estimated jointly via FIML. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table 10: **Marital Status: Counterfactual and Total Effects Decomposition**

	Counterfactual (1)	**Counterfactual (2)**	**Total Effect**	**Direct Behavioral Effect**	**Indirect Cognitive Effect**
Ever failed class	0.473	0.412	0.060	0.055	0.005
			(0.048)		
Unlikely to graduate college	0.374	0.286	**0.087****	0.090	−0.003
			(0.043)		
Ever suspended from school	0.327	0.246	**0.080****	0.079	0.001
			(0.040)		
Ever drunk alcohol	0.145	0.164	−0.018	−0.020	0.001
			(0.034)		
Ever used marijuana	0.181	0.177	0.004	0.002	0.002
			(0.035)		
Ever stopped by police	0.316	0.236	**0.081***	0.082	−0.001
			(0.043)		
Ever had sex	0.150	0.143	0.006	0.005	0.001
			(0.030)		

Notes: The table reports counterfactual predicted teenage outcomes and a decomposition of the total effect of maternal marital status trajectories on teenage outcomes, simulated using the estimated dynamic model. Two counterfactual marital status trajectories are compared: *Counterfactual (1) reflects an unstable marital history in which the mother transitions through the sequence single → with biological father → single → with another man → single. Counterfactual (2) reflects a stable marital history in which the mother is always married to or cohabiting with the biological father.* Total Effect = Counterfactual (1) − Counterfactual (2). Total Effect is decomposed into a Direct Behavioral Effect, and an Indirect Cognitive Effect. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 11: **Employment and Welfare: Counterfactual and Total Effects Decomposition**

	Counterfactual (1)	Counterfactual (2)	Total Effect	Direct Behavioral Effect	Indirect Cognitive Effect
Ever failed class	0.546	0.429	**0.117*** (0.070)	0.102	0.015
Unlikely to graduate college	0.345	0.296	0.049 (0.065)	0.051	−0.002
Ever suspended from school	0.365	0.258	0.107 (0.066)	0.100	0.007
Ever drunk alcohol	0.092	0.128	−0.036 (0.037)	−0.036	0.000
Ever used marijuana	0.155	0.165	−0.009 (0.044)	−0.013	0.004
Ever stopped by police	0.479	0.218	**0.261***** (0.072)	0.258	0.003
Ever had sex	0.220	0.139	**0.081*** (0.048)	0.072	0.009

Notes: The table reports counterfactual predicted teenage outcomes and a decomposition of the total effect of maternal employment and welfare receipt trajectories on teenage outcomes, simulated using the estimated dynamic model. Two counterfactual scenarios are compared jointly along both the employment and welfare dimensions: *Counterfactual (1) reflects the most disadvantaged economic trajectory, in which the mother is persistently unemployed and continuously receives welfare* across all developmental periods. *Counterfactual (2) reflects the most advantaged economic trajectory, in which the mother is consistently employed and never receives welfare* across all developmental periods. Total Effect = Counterfactual (1) − Counterfactual (2). Total Effect is decomposed into a Direct Behavioral Effect, and an Indirect Cognitive Effect. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 12: **Substance Use and Criminal Record: Counterfactual and Total Effects Decomposition**

	Counterfactual (1)	**(2)**	**Total Effect**	**Direct Behavioral Effect**	**Indirect Cognitive Effect**
Ever failed class	0.598	0.452	0.147	0.126	0.020
			(0.214)		
Unlikely to graduate college	0.385	0.234	0.151	0.155	−0.004
			(0.195)		
Ever suspended from school	0.195	0.293	−0.099	−0.107	0.008
			(0.186)		
Ever drunk alcohol	0.188	0.105	0.083	0.082	0.002
			(0.153)		
Ever used marijuana	0.312	0.212	0.099	0.094	0.006
			(0.165)		
Ever stopped by police	0.062	0.351	**−0.288****	−0.291	0.003
			(0.135)		
Ever had sex	0.267	0.216	0.052	0.044	0.008
			(0.165)		

Notes: The table reports counterfactual predicted teenage outcomes and a decomposition of the total effect of maternal substance use and parental criminal involvement on teenage outcomes, simulated using the estimated dynamic model. Both dimensions are varied simultaneously across counterfactuals: *Counterfactual (1) reflects the most adverse profile, in which the mother drinks alcohol, smokes cigarettes, and uses marijuana in every developmental period while also being convicted in every period, and the father is incarcerated in every period. Counterfactual (2) reflects the most favorable profile, in which the mother abstains from all substances and is never convicted, and the father is never incarcerated, across all developmental periods.* Total Effect = Counterfactual (1) − Counterfactual (2). Total Effect is decomposed into a Direct Behavioral Effect, and an Indirect Cognitive Effect. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 13: **Marginal Effects of Summed Parental Behavior and Cognitive Achievement on Teenage Outcomes**

Summed Behavior	**Ever Failed Class**	**Unlikely to Graduate College**	**Ever Suspended from Sch**	**Ever Drank Alcohol**	**Ever Use Marijuana**	**Ever Stopped by Police**	**Ever had Sex**
Biodad	−0.016**	−0.019***	0.006	−0.001	−0.005	0.005	0.001
	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)
Another Man	−0.014	−0.007	0.015	0.014	0.013	0.005	0.014
	(0.018)	(0.017)	(0.015)	(0.013)	(0.014)	(0.015)	(0.012)
Employment	−0.008	−0.008	0.012	0.016	0.017	−0.012	0.010
	(0.013)	(0.014)	(0.012)	(0.014)	(0.017)	(0.012)	(0.012)
Welfare	0.015	0.007	0.022**	0.007	0.010	0.026***	0.018***
	(0.011)	(0.010)	(0.009)	(0.006)	(0.006)	(0.009)	(0.006)
Drink	0.000	0.006	−0.018	0.021	0.026	0.009	0.006
	(0.028)	(0.032)	(0.029)	(0.025)	(0.036)	(0.024)	(0.026)
Smoke	0.007	−0.004	−0.010	0.010	0.024**	−0.003	0.008
	(0.016)	(0.017)	(0.013)	(0.009)	(0.010)	(0.019)	(0.013)
Marijuana	−0.007	0.027	0.012	0.029	0.035	−0.015	0.008
	(0.029)	(0.033)	(0.022)	(0.031)	(0.032)	(0.022)	(0.018)
Conviction	0.044	−0.008	0.004	−0.021	−0.002	−0.025	−0.033
	(0.037)	(0.032)	(0.032)	(0.019)	(0.022)	(0.028)	(0.022)
Jail	0.002	−0.025**	0.016	0.014*	0.021**	0.022**	0.009
	(0.014)	(0.012)	(0.011)	(0.008)	(0.009)	(0.010)	(0.009)
Cognitive (10-point scale)	−0.040**	0.010	−0.018	−0.004	−0.010	−0.006	−0.017
	(0.017)	(0.016)	(0.014)	(0.012)	(0.012)	(0.014)	(0.011)

Notes: The table reports marginal effects from the behavioral summation alternative specification of the teenage outcome equation, estimated as part of the alternative specifications analysis. This specification replaces the pattern-based behavioral history used in the baseline model with a cumulative behavioral exposure measure, $\breve{B}_t = B_0 + B_1 + \cdots + B_t$, which counts the total number of developmental periods in which the parent engaged in each behavior. Under this specification, each reported coefficient represents the change in the predicted probability of a teenage outcome associated with one additional period of parental exposure to that behavior, holding all other behaviors and child characteristics at their observed values. This contrasts with the baseline model, in which the timing and sequencing of parental behaviors matter, and with the contemporaneous specification reported in Table 14, in which only the most recent period's behavior enters the teenage outcome equation directly. The summation specification treats behavioral exposure as continuous and proportional, emphasizing cumulative volume of exposure rather than the pattern or recency of behaviors. Marital status is decomposed into two separate cumulative exposure measures: the number of periods cohabiting with the biological father (Biodad) and the number of periods cohabiting with another man (Another Man), both relative to periods spent as a single mother. The cognitive score is entered on a 10-point scale, so that the reported coefficient represents the change in the probability of a teenage outcome associated with a 10-point increase in the child's accumulated cognitive score. The model is estimated by Full Information Maximum Likelihood (FIML) with Discrete Factor Random Effects (DFRE) on the FFCWS sample of N = 1,615 families. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 14: **Marginal Effects of Contemporaneous Parental Behavior and Cognitive Achievement on Teenage Outcomes**

	Ever Failed Class	**Unlikely to Graduate College**	**Ever Suspended from Sch**	**Ever Drank Alcohol**	**Ever Use Marijuana**	**Ever Stopped by Police**	**Ever had Sex**
Biodad	−0.093**	−0.078**	0.022	0.010	0.005	0.003	0.005
	(0.047)	(0.036)	(0.035)	(0.035)	(0.036)	(0.035)	(0.036)
Another Man	0.014	0.043	0.053**	0.024	0.038*	0.037	0.034
	(0.033)	(0.031)	(0.025)	(0.021)	(0.022)	(0.026)	(0.021)
Employment	−0.046	−0.037	−0.052	−0.027	−0.054**	−0.074***	−0.020
	(0.044)	(0.036)	(0.034)	(0.026)	(0.025)	(0.029)	(0.026)
Welfare	0.132***	0.052	0.054	0.070	0.039	0.050	0.028
	(0.048)	(0.041)	(0.033)	(0.052)	(0.047)	(0.039)	(0.044)
Drink	0.005	0.033	0.019	0.008	0.008	0.035	0.011
	(0.043)	(0.037)	(0.040)	(0.040)	(0.042)	(0.043)	(0.039)
Smoke	0.029	−0.036	−0.038	0.013	0.062*	0.077*	0.072*
	(0.052)	(0.046)	(0.037)	(0.033)	(0.038)	(0.046)	(0.039)
Marijuana	0.017	0.039	0.062	0.007	0.032	−0.042	0.040
	(0.060)	(0.056)	(0.051)	(0.052)	(0.048)	(0.056)	(0.043)
Conviction	0.186**	0.069	0.080	−0.037	−0.010	−0.019	−0.011
	(0.081)	(0.066)	(0.055)	(0.090)	(0.083)	(0.079)	(0.080)
Jail	0.017	−0.030	0.014	0.001	0.018	−0.005	−0.030
	(0.052)	(0.040)	(0.041)	(0.035)	(0.035)	(0.034)	(0.037)
Cognitive (10-point scale)	−0.042**	0.010	−0.020	−0.003	−0.011	−0.010	−0.021*
	(0.017)	(0.016)	(0.014)	(0.011)	(0.011)	(0.015)	(0.012)

Notes: The table reports marginal effects from the contemporaneous parental behavior alternative specification of the teenage outcome equation, estimated as part of the alternative specifications analysis. Under this specification, teenage outcomes depend only on parental behaviors in the most recent developmental period t, rather than on the full history of behaviors as in the baseline model. Formally, the teenage outcome equation is specified as $Y_{t+1} = g(B_t, \theta_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, \boldsymbol{Z}_t, \mu^y, \nu_t^y, u_{t+1}^y)$, where only contemporaneous parental behavior B_t enters directly, while earlier behaviors matter only insofar as they have shaped the child's accumulated cognitive skill θ_t over the developmental lifecycle. This specification imposes a value-added structure on the teenage outcome equation, analogous to the assumption used for the child cognitive skill production function, and contrasts with both the baseline pattern-based model — in which the entire behavioral history enters directly through categorical trajectory indicators — and the summation model reported in Table 13, in which cumulative behavioral exposure enters as a continuous count variable. Each reported coefficient represents the change in the predicted probability of a teenage outcome associated with the presence of a given parental behavior in the terminal developmental period, holding all other contemporaneous behaviors and child characteristics at their observed values. Marital status is decomposed into two binary indicators: cohabiting with the biological father (Biodad) and cohabiting with another man (Another Man), both relative to being single. The cognitive score is entered on a 10-point scale, so that the reported coefficient represents the change in the probability of a teenage outcome associated with a 10-point increase in the child's cognitive score at the terminal period. The model is estimated by Full Information Maximum Likelihood (FIML) with Discrete Factor Random Effects (DFRE) on the FFCWS sample of $N = 1{,}615$ families. Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 15: **Policy Experiment 1: Effects of Sin Tax on Parental Behaviors and Teenage Outcomes**

	Intervention			No Intervention	Difference		
	Tax = \$0.5	Tax = \$1	Tax = \$1.5		Tax = \$0.5	Tax = \$1	Tax = \$1.5
Parental Behaviors							
Married/cohabit status							
single	0.435	0.420	0.409	0.453	−0.017	−0.032	−0.043
					(0.048)	(0.064)	(0.081)
with biodad	0.489	0.494	0.493	0.480	0.010	0.015	0.013
					(0.052)	(0.071)	(0.092)
with another man	0.076	0.085	0.098	0.068	0.008	0.018	0.030
					(0.011)	(0.021)	(0.035)
Employed	0.793	0.802	0.810	0.782	0.011	0.020	0.028
					(0.036)	(0.059)	(0.082)
Welfare receipt	0.138	0.104	0.089	0.198	**−0.060****	**−0.094****	**−0.109*****
					(0.029)	(0.037)	(0.042)
Substance use							
Drinking	0.465	0.458	0.452	0.476	−0.011	−0.019	−0.024
					(0.045)	(0.065)	(0.087)
Smoking	0.187	0.167	0.153	0.232	**−0.045***	−0.064	−0.079
					(0.024)	(0.039)	(0.056)
Marijuana use	0.011	0.006	0.003	0.024	**−0.013***	**−0.019***	−0.021
					(0.007)	(0.011)	(0.014)
Conviction	0.013	0.011	0.009	0.015	−0.002	−0.004	−0.005
					(0.005)	(0.010)	(0.018)
Jail	0.072	0.072	0.072	0.073	−0.001	−0.001	−0.002
					(0.010)	(0.010)	(0.011)
Teenage Outcomes							
Ever failed class	0.465	0.458	0.453	0.473	−0.007	−0.014	−0.019
					(0.025)	(0.031)	(0.037)
Unlikely to grad. college	0.334	0.325	0.321	0.331	0.002	−0.006	−0.010
					(0.023)	(0.027)	(0.031)
Ever suspended from sch.	0.261	0.259	0.258	0.287	−0.026	−0.028	−0.029
					(0.021)	(0.023)	(0.027)
Ever drunk alcohol	0.104	0.100	0.097	0.133	**−0.029****	**−0.033****	**−0.036****
					(0.014)	(0.015)	(0.017)
Ever used marijuana	0.163	0.159	0.156	0.178	−0.015	−0.019	−0.022
					(0.017)	(0.019)	(0.021)
Ever stopped by police	0.245	0.248	0.252	0.260	−0.015	−0.012	−0.008
					(0.021)	(0.026)	(0.031)
Ever had sex	0.149	0.146	0.143	0.157	−0.008	−0.012	−0.014
					(0.017)	(0.019)	(0.021)

Notes: The table reports the results of a Sin-to-Skill policy experiment simulated using the estimated dynamic model. The policy combines an excise tax on beer and cigarettes at three incremental levels (\$0.5, \$1.00, and \$1.50 per unit) with reinvestment of the associated tax revenues into county-level education programs. To account for administrative lags in public spending, the associated improvement in county-level test scores (modeled as a 0.1 standard deviation increase, consistent with estimates from the teacher quality and Project STAR literatures) is assumed to take effect only after the first two developmental periods. The test score increment is drawn from a normal distribution $\mathcal{N}(0.1 \times \text{s.d.}, 1)$ across counties to reflect heterogeneity in school quality and program implementation. The *Intervention* columns report predicted parental behaviors and teenage outcomes under each tax level. The *No Intervention* column reports predicted outcomes under the baseline environment reflecting observed parental behaviors in the FFCWS sample. *Difference* = *With Intervention* − *No Intervention*; so that a negative difference indicates an improvement in the outcome, with the exception of employment and marriage with the biological father where a positive difference indicates improvement. Parental behavior outcomes are reported as proportions across all pre-teenage developmental periods, while teenage outcomes are binary indicators reported as proportions at age 15. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 16: **Policy Experiment 2: Effects of Co-residence with Biological Father on Teenage Outcomes**

	Intervention			No Intervention	Difference		
	Birth–Age 1	Birth–Age 3	Birth–Age 5		Birth–Age 1	Birth–Age 3	Birth–Age 5
Ever failed class	0.448	0.441	0.434	0.473	−0.025 (0.021)	−0.032 (0.024)	**−0.039*** (0.023)
Unlikely to grad. college	0.311	0.306	0.304	0.331	−0.020 (0.018)	−0.025 (0.021)	−0.027 (0.019)
Ever suspended from sch.	0.267	0.257	0.253	0.287	−0.020 (0.017)	−0.030 (0.021)	**−0.034*** (0.020)
Ever drunk alcohol	0.123	0.118	0.121	0.133	−0.010 (0.012)	−0.015 (0.014)	−0.012 (0.013)
Ever used marijuana	0.162	0.156	0.157	0.178	−0.016 (0.012)	**−0.022*** (0.013)	**−0.021*** (0.012)
Ever stopped by police	0.266	0.265	0.259	0.260	0.006 (0.017)	0.005 (0.020)	−0.001 (0.019)
Ever had sex	0.150	0.145	0.145	0.157	−0.007 (0.015)	−0.012 (0.017)	−0.012 (0.016)

Notes: The table reports the results of a co-residence promotion policy experiment simulated using the estimated dynamic model. The intervention requires all mothers to be married to or cohabiting with the biological father, and simultaneously requires the biological father to be present and not incarcerated, across three incremental durations: from birth through Age 1, from birth through Age 3, and from birth through Age 5. Following each intervention window, parents resume their endogenously determined lifecycle behaviors. The *Intervention* columns report predicted teenage outcomes under each policy duration. The *No Intervention* column reports predicted outcomes under the baseline environment reflecting observed parental behaviors in the FFCWS sample. *Difference* = *With Intervention* − *No Intervention*; a negative difference indicates an improvement in the outcome. All outcomes are binary indicators reported as proportions. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 17: **Co-residence with Biological Father: Within vs. Beyond the Critical Window**

	Intervention		No Intervention	Difference	
	Birth–Age 5	Above Age 5		Birth–Age 5	Above Age 5
Ever failed class	0.434	0.463	0.473	**−0.039*** (0.023)	−0.010 (0.021)
Unlikely to grad. college	0.304	0.330	0.331	−0.027 (0.019)	−0.001 (0.020)
Ever suspended from sch.	0.253	0.284	0.287	**−0.034*** (0.020)	−0.003 (0.019)
Ever drunk alcohol	0.121	0.152	0.133	−0.012 (0.013)	0.019 (0.015)
Ever used marijuana	0.157	0.188	0.178	**−0.021*** (0.012)	0.010 (0.014)
Ever stopped by police	0.259	0.245	0.260	−0.001 (0.019)	−0.015 (0.017)
Ever had sex	0.145	0.173	0.157	−0.012 (0.016)	0.016 (0.016)

Notes: The table compares two co-residence promotion policy experiments to assess the role of Heckman's (2006) early childhood critical window. *Birth–Age 5* requires all mothers to be married to or cohabiting with the biological father, with the father present and non-incarcerated, from birth through Age 5, spanning the full early childhood critical period. *Above Age 5* imposes the same co-residence requirement only after Age 5, when the critical window has closed. In both cases, parents resume their endogenously determined lifecycle behaviors following the intervention window. The *No Intervention* column reports predicted outcomes under the baseline environment reflecting observed parental behaviors in the FFCWS sample. *Difference* = *With Intervention* − *No Intervention*; a negative difference indicates an improvement in the outcome. All outcomes are binary indicators reported as proportions. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 18: **Policy Experiment 3: Effects of Peer Environment Upgrade on Teenage Outcomes**

	Intervention			No Intervention	Difference		
	λ = 0.1 SD	λ = 0.2 SD	λ = 0.3 SD		λ = 0.1 SD	λ = 0.2 SD	λ = 0.3 SD
Ever failed class	0.465	0.458	0.450	0.473	−0.008	−0.015	−0.023
					(0.019)	(0.019)	(0.020)
Unlikely to grad. college	0.322	0.309	0.298	0.331	−0.009	−0.022	**−0.033***
					(0.016)	(0.017)	(0.017)
Ever suspended from sch.	0.277	0.266	0.256	0.287	−0.010	−0.021	**−0.031****
					(0.015)	(0.015)	(0.016)
Ever drunk alcohol	0.111	0.094	0.081	0.133	**−0.022****	**−0.039*****	**−0.053*****
					(0.011)	(0.011)	(0.011)
Ever used marijuana	0.160	0.142	0.122	0.178	−0.018	**−0.036*****	**−0.057*****
					(0.013)	(0.013)	(0.013)
Ever stopped by police	0.248	0.232	0.218	0.260	−0.012	**−0.028***	**−0.043*****
					(0.015)	(0.015)	(0.015)
Ever had sex	0.144	0.131	0.119	0.157	−0.013	**−0.026****	**−0.038*****
					(0.013)	(0.013)	(0.013)

Notes: The table reports the results of a peer environment upgrade policy experiment simulated using the estimated dynamic model. The intervention improves the adolescent peer environment across three incremental magnitudes (λ = 0.1, 0.2, and 0.3 standard deviations). For substance-use peer variables (drinking alcohol, smoking cigarette, marijuana use), the intervention randomly reassigns a fraction of peers from substance users to non-users such that the mean falls by $\lambda\times$ SD. For positive peer variables (peer study habits, and likely to graduate college), the intervention randomly reassigns a fraction of peers upward such that the mean rises by $\lambda\times$ SD. All peer variables remain binary throughout. The policy can be interpreted as a school reassignment, magnet school access, or neighborhood mobility program that shifts the social environment the child is embedded in during adolescence. The *Intervention* columns report predicted teenage outcomes under each policy magnitude. The *No Intervention* column reports predicted outcomes under the baseline environment reflecting observed parental behaviors in the FFCWS sample. *Difference = With Intervention − No Intervention*; a negative difference indicates an improvement in the outcome. All outcomes are binary indicators reported as proportions. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figures

Figure 3: **US cities included in the survey (N = 4898 births)**

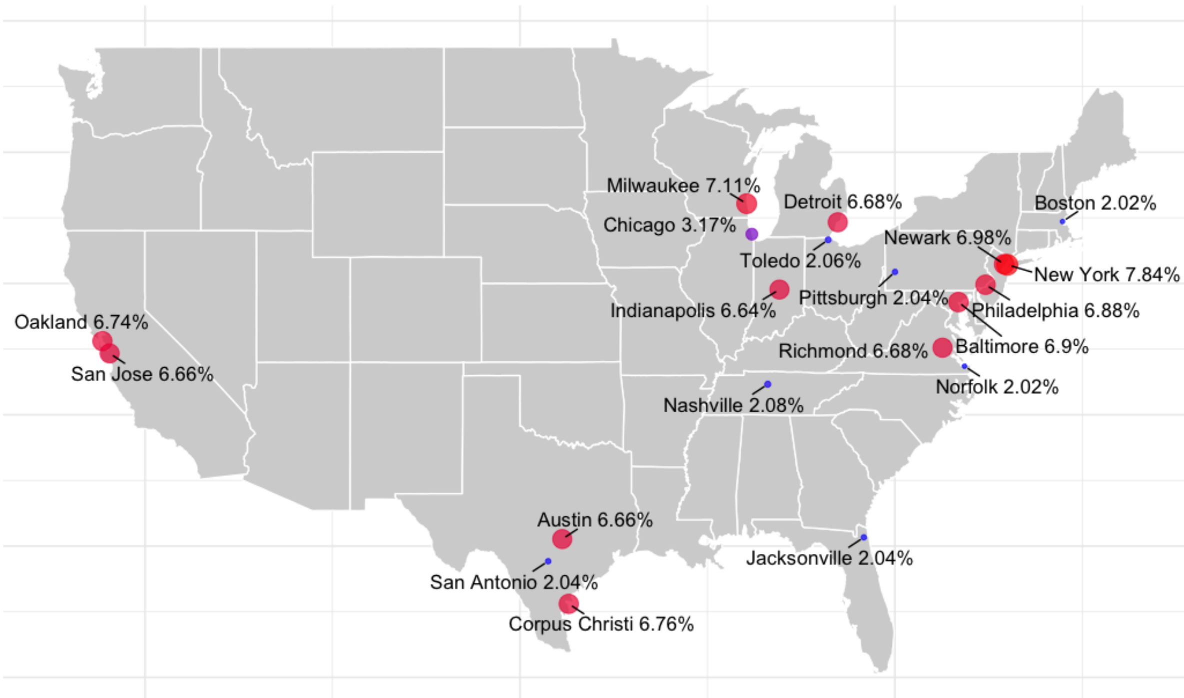
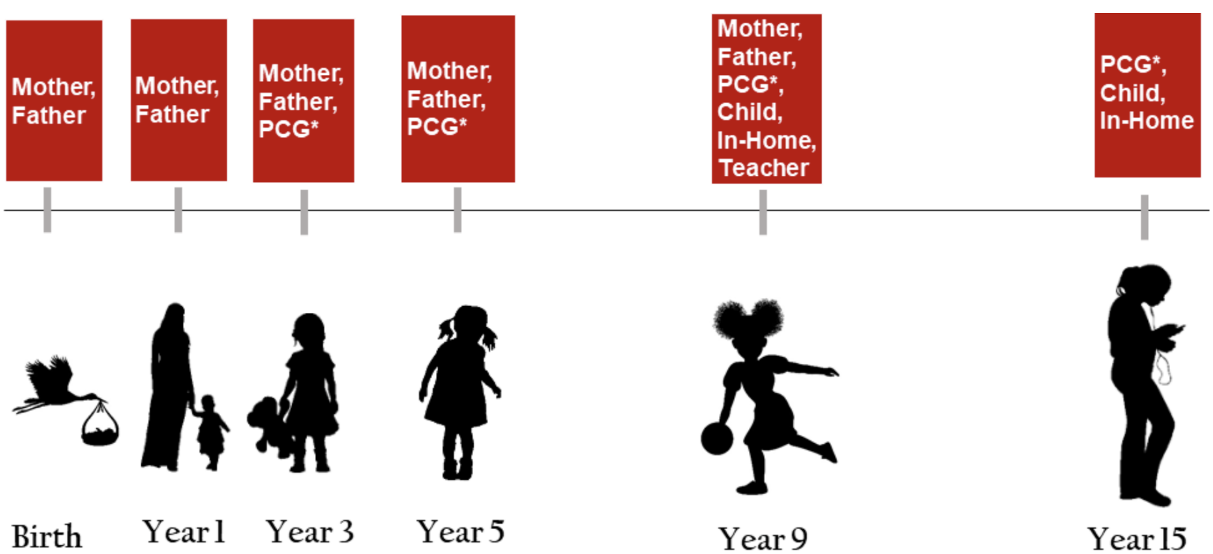


Figure 4: **FFCWS timeline and agents of survey**



Source: Future of Families and Child Wellbeing Study FFCWS (2023)

Figure 5: **Model Fit for Dynamic Behaviors and Cognitive Achievement**

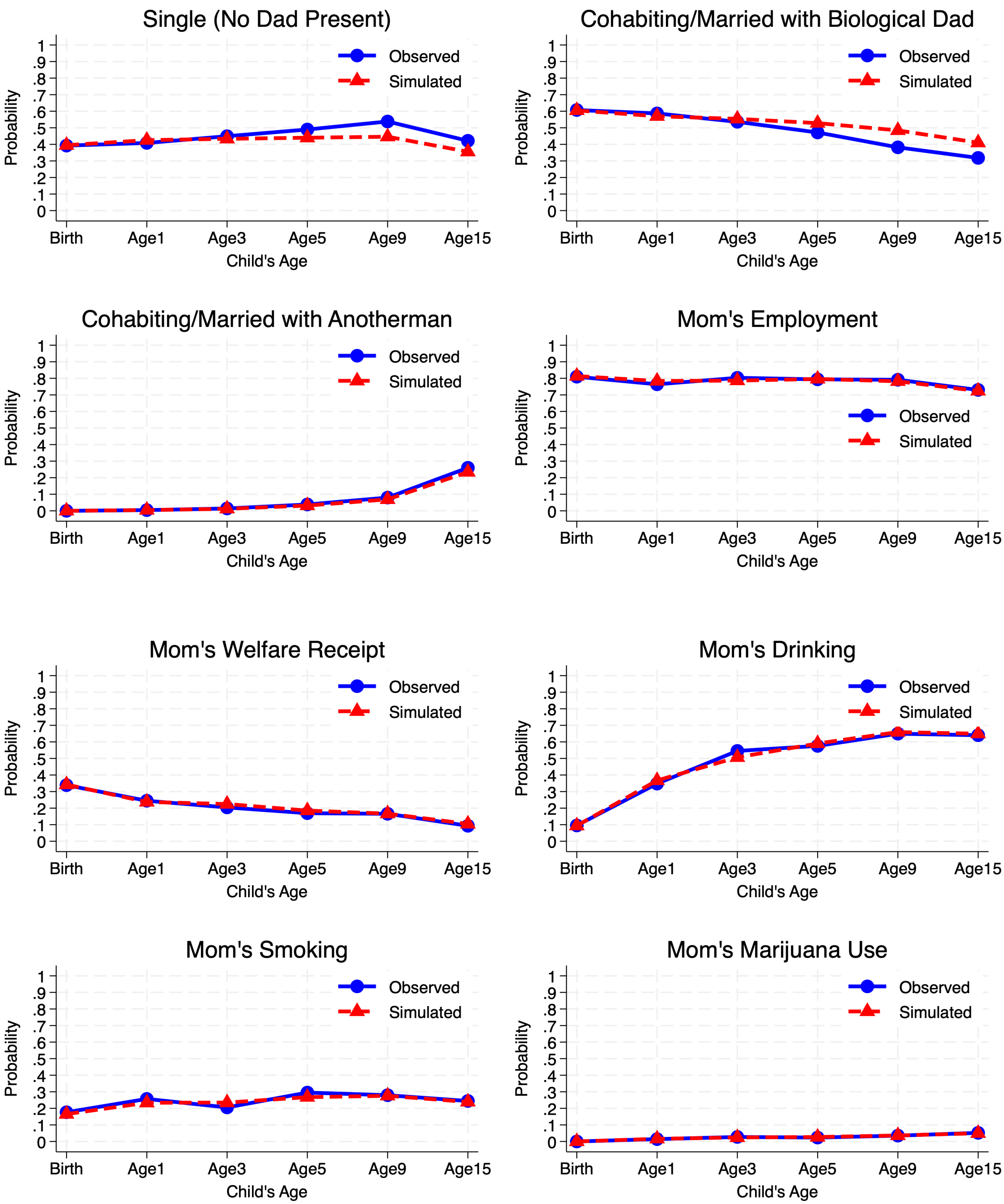


Figure 5: **Model Fit for Dynamic Behaviors and Cognitive Achievement (continued)**

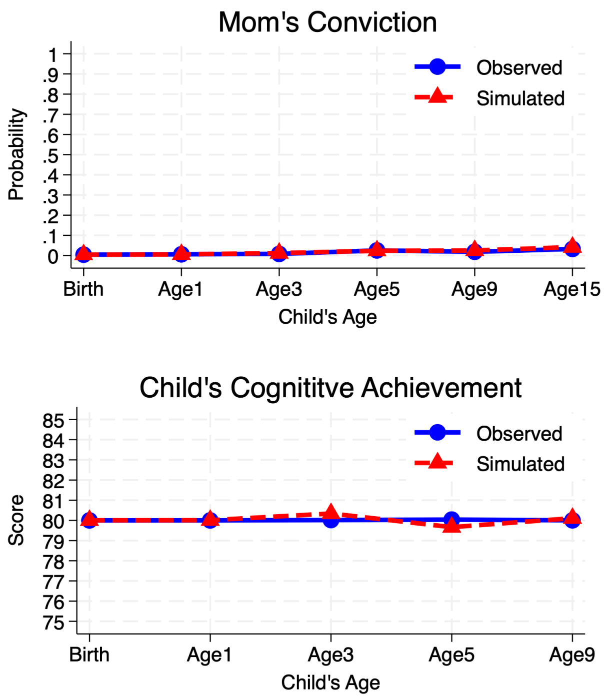
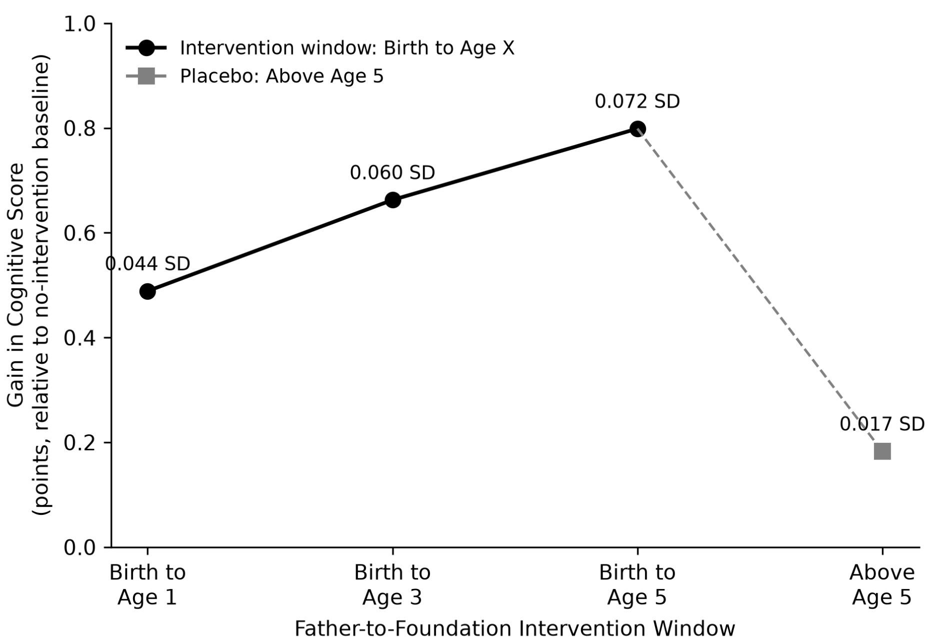


Figure 6: **Cognitive Score Gain Under the Father-to-Foundation Policy**[13]



[13] Gains are measured as the increase in the child's simulated cognitive score at the end of the pre-teenage period, expressed as a percentage of its baseline standard deviation, relative to the no-intervention baseline; simulated outcomes are averaged over 1,000 parametric bootstrap replications. The *Birth to Age X* windows impose co-residence with the biological father from birth through the indicated age. The *Above Age 5* bar is a placebo: it imposes the same co-residence requirement only after Age 5, once Heckman's (2006) early-childhood critical window has closed. Corresponding significance levels are reported in Table 17.

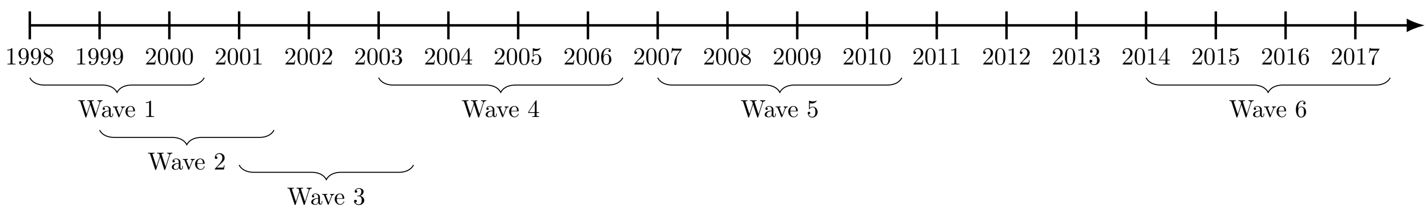
Appendix

A1 Appendix on Section 3: Data

A1.1 Sample Selection

The estimation sample for this study was selected based on three criteria and Table A.19 summarizes the resulting sample sizes. The first sample (labeled N1) criterion is that the mother must have been present in all the waves (waves 1 to 6). Here, attrition 1 is defined as absence of a mother from at least one of the waves. As of wave 6 (when the child is 15 years) there were 2530 mothers who were present in all the waves.

Figure A.7: **Timeline of survey for each wave**



The second sample (labeled N2) criterion was that included children should be the same age at the interview. Given the survey period for each wave was across multiple years and some waves also were overlapping as shown in Figure A.7, the children were of different ages at the time of interview. For instance, the raw data indicates the age of the children in waves 4 and 5 to be 57 – 72 months and 104 – 132 months. Therefore, at wave 4 where children were supposed to be 5 years, some children were already 6 years due to delay in administering the interview. I desire to streamline disparities across the child's age so we can have consistent results. Therefore, I indicate whether a child's age is between 3 months of the expected age of the wave and a month away from his/her next age in years. For example, for wave 2 (expected child's age is 1 year/12 months), a child is considered to have the correct age if their age is in the range 9 months – 23 months; for wave 3 (expected child's age is 3 years/36 months) the observed age is 33 months – 47 months; for wave 4 (expected child's age is 5 years/60 months) the observed age is 57 months – 71 months, etc. Attrition 2 is defined as mothers whose children were not the focal age when interviewed, conditional on mother being present in the all the waves (criterion 1).

The final criterion used in the sample selection was that the estimation sample (labeled N3) include children with no missing child outcomes among those to be studied, conditional on satisfying criteria 1 and 2. The attrition rate for the mothers is indicated in Table A.19. The sample size that satisfied all three criteria was 1618. However, there were 3 mothers with missing race, which I drop from the sample. Therefore, the final analysis sample was 1615 mothers who were present in all the waves, with children of the same age at the time of interview and non-missing child outcome variables. The larger attrition rates across wave 5 and wave 6 is likely attributed to the long span of time between survey waves. We can see from the descriptive statistics in Table A.20

Table A.19: **Progression of number of mothers in the sample**

Wave	**1**	**2**	**3**	**4**	**5**	**6**
Child's age	**Birth**	**Year 1**	**Year 3**	**Year 5**	**Year 9**	**Year 15**
N1 (mum in wave)	4897	4364	4009	3675	2986	2530
Attrition 1		533	355	334	689	456
Attrition Rate 1 (%)		10.9	8.1	8.3	18.7	15.3
N2 (focal age child)	4897	4249	3868	3541	2679	1809
Attrition		648	381	327	862	870
Attrition Rate 2 (%)		13.2	8.9	8.5	24.3	32.5
N3 (non-missing outcome)	2822	2601	2466	2358	2152	1615
Attrition		221	135	108	206	537
Attrition Rate 3 (%)		7.8	5.2	4.4	8.7	25

that the distribution of the core demographics across the waves have been preserved conditional on the sample that satisfies all the sample selection criteria (N3) despite the attrition, hence implying that the attrition is random.

Table A.20: **Demographics distribution over the waves**

Wave	**1**	**2**	**3**	**4**	**5**	**6**
Child's age	**Birth**	**Year 1**	**Year 3**	**Year 5**	**Year 9**	**Year 15**
N3	**2822**	**2601**	**2466**	**2358**	**2152**	**1615**
Mom Race						
White	20%	21%	22%	22%	22%	23%
Black	52%	51%	51%	51%	51%	51%
Hispanic	25%	25%	24%	24%	24%	23%
Other	3%	3%	3%	3%	3%	3%
Child's Sex						
Male	51%	51%	51%	51%	52%	51%
Female	49%	49%	49%	49%	48%	49%

A1.2 Endogenous variables

Furthermore, I depict the persistence in parental behaviors in Figure A.8, which shows the dynamic evolution of parental behaviors across the child's life. Panels (a) to (d) highlight substantial heterogeneity in early childhood years, pronounced behavioral persistence over time, and clear evidence that parental behaviors tend to consolidate into more stable patterns as children age. Panel (a) shows trajectories of mothers' marital status from birth through age 15. At birth and age 1, there is considerable dispersion across marital states, with a large share of mothers single, alongside significant proportions cohabiting with the biological dad or married. The figure shows a gradual decline in transitions and a growing concentration in stable states, either remaining single or remaining married to the biological dad. Panel (b) presents the evolution of mom's employment and welfare participation. Early in the child's life, employment and welfare combinations are highly fluid. Over time, however, employment patterns become increasingly stable. From age 5 onward, the share of mothers consistently employed rises, while persistent unemployment, particularly without welfare, also becomes more common for a subset of families.

Panel (c) illustrates trajectories of maternal substance use, distinguishing between no use, low intensity (use of one substance), moderate intensity (use of two substances), and high intensity (use of all three: alcohol, cigarettes, and marijuana). Substance use exhibits strong state dependence. Mothers who report no use early in the child's life are highly likely to remain non-users over time. In contrast, mothers who engage in substance use, particularly moderate or high intensity, tend to persist in these behaviors, with relatively few transitions back to non-use. Panel (d) focuses on parental criminal behavior, distinguishing between no crime, dad ever in jail only, mother ever convicted only, and joint parental criminal involvement. At birth and age 1, the majority of families have no criminal involvement, but a sizable fraction already experiences paternal incarceration.

Importantly, these behaviors do not evolve in isolation. Marital instability often coincides with employment and welfare transitions, and also substance use and criminal involvement show strong persistence and likely interact with economic and relational dynamics. Together, the panels emphasize the need for a joint dynamic modeling framework that captures both the timing and accumulation of parental behaviors, rather than relying on static or single-period indicators.

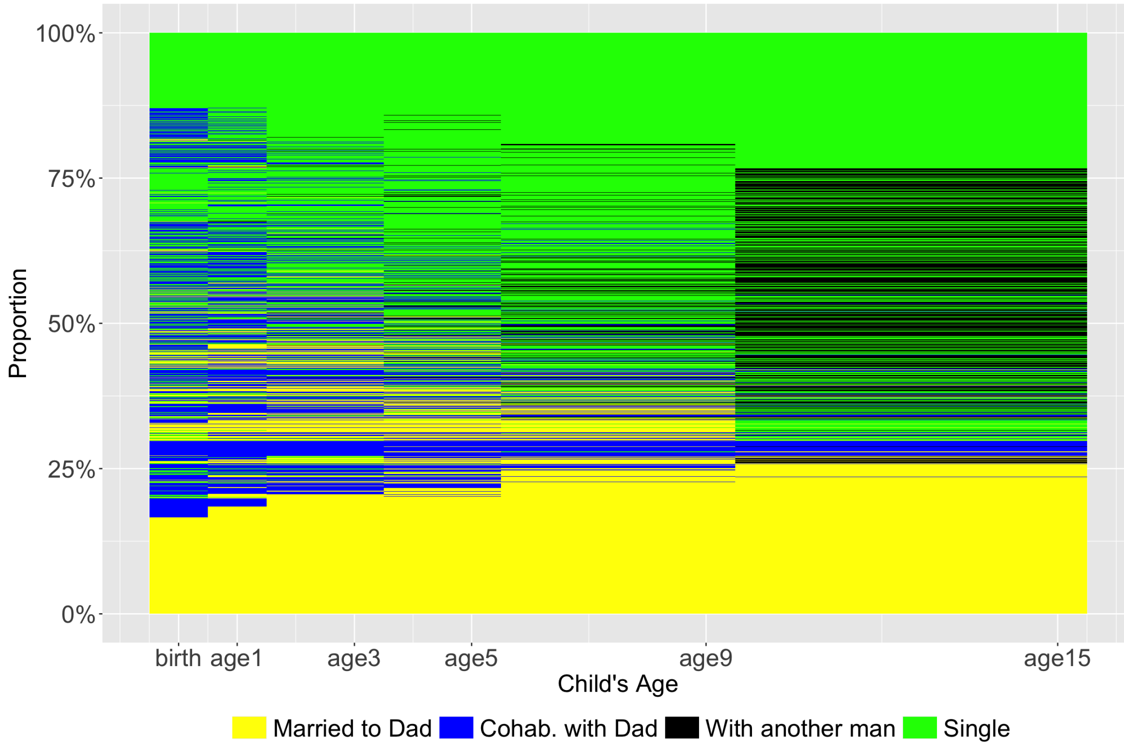
To capture the variation in states of behaviors across time within a family, I denote the various combinations across time in Online Appendix Table OA.1. For example, 44.2 percent of the mothers were employed throughout all the waves (111111). Approximately 3 percent of mothers were unemployed across all waves. Regarding alcohol, most mothers were not drinking when they were pregnant (in wave 1), but drank in the subsequent waves (011111). Smoking was similarly characterized by the two extremes of those who do not smoke and those who smoke throughout all waves. The main result in relation to dads is that, over 24 percent of biological dads had not been present in the life of the child throughout the entire 15 years of the child's life. However, given that others can step in as a father figure, as we see that only 13 percent of children did not have a father figure in their lives from birth to age 15.

A1.3 Exogenous variables

Summary statistics of the exogenous explanatory variables are provided in Online Appendix Table OA.2. The mother's educational level over the 15-year period shows a gradual increase in the distribution towards high levels of education. Educational attainment improves over time; that is, the share of mothers with less than a high school education declines from 28 percent at birth to 15 percent by age 15, while the fraction with some college or a college/graduate degree rises steadily. Given that these mothers were likely to be single, it is normal for them to work full-time rather than to enroll in school resulting in limited variation in the education variable. Living in public housing declines sharply after birth, suggesting improving housing conditions, while the presence of a grandparent in the household decreases through early childhood before rising substantially by age 15, potentially reflecting care-giving responsibility of grandparents in the sample. I define the mother's non-wage income as all income received from all other sources apart from wages while working as an employee. More formally I define it as $Y_t - w_t h_t$, where Y_t^m is mom's income and $w_t h_t$ is earned wages. I calculated the mother's non-wage income by accounting for all forms of income the mother received from entrepreneurial income (income from running own business) (K_t), income from returns on investment (I_t), illicit income (income from selling, drugs/prostitution, etc.) (Q_t) and income from benefits (b_t^j); where $j =$ Food stamps/EBT and other benefits. Mathematically as:

Figure A.8: **Parental Behavior Trajectories**

(a) Mother's Marital Status



(b) Mother's Employment and Welfare

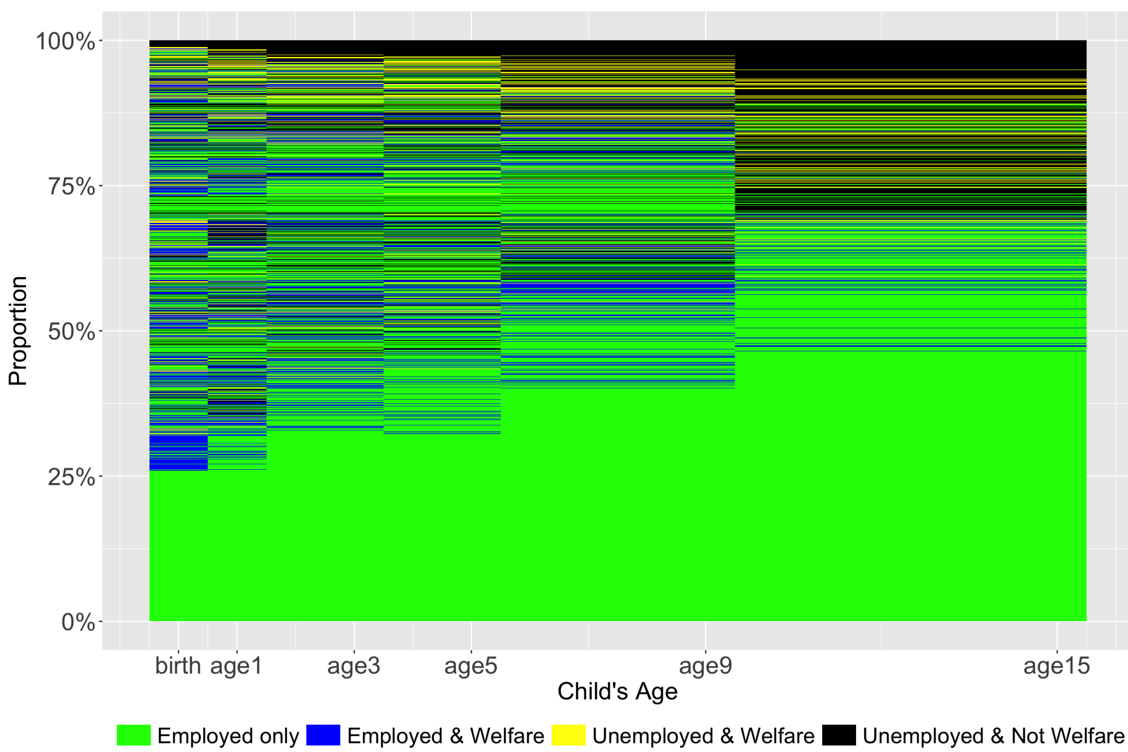
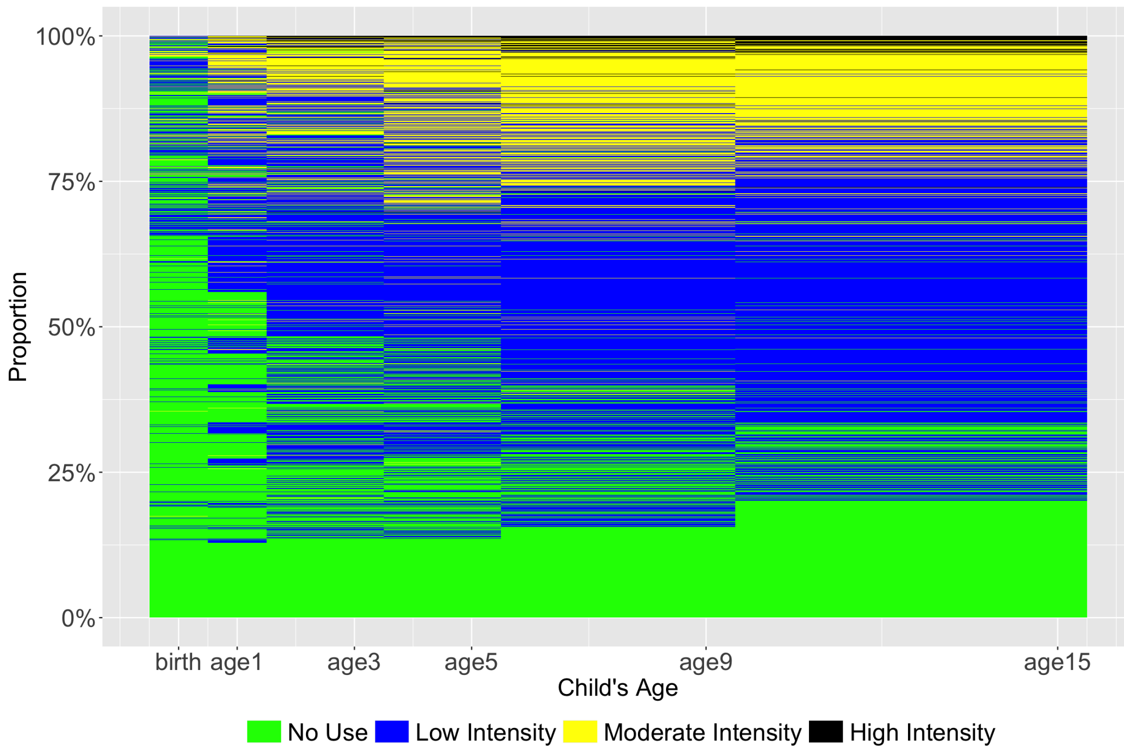
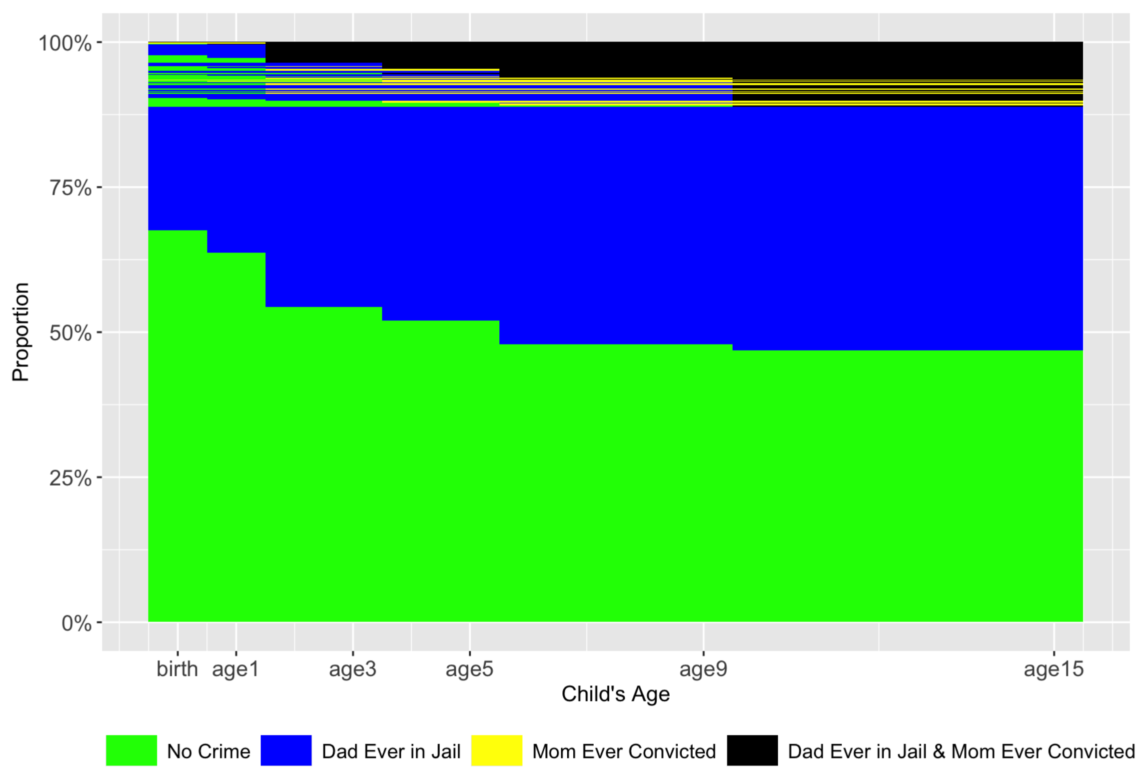


Figure A.8: **Parental Behavior Trajectories (continued)**

(c) Mother's Substance Use (Drinking, Smoking, and Marijuana Use)



(d) Parent's Criminal Record



$Y_t^m - w_t h_t = \mathbb{1}\{\text{selfemp}_t\} \cdot K_t + I_t + Q_t + \sum_{j=1}^J b_t^j$. Average non-wage income, reported in inflation-adjusted year-2000 dollars, fluctuates across waves but increases by ages 9 and 15, indicating greater access to transfers or non-labor income later in the child's life.

Child characteristics measured later in the panel indicate substantial heterogeneity in skills and behaviors. Average Woodcock–Johnson (WJ) scores in English and Math suggest moderate academic performance at age 9. Time use patterns show that most children engage infrequently in structured school or religious activities, with a majority reporting "never" or "once per month", though a nontrivial share participates weekly. Behavioral indicators reveal meaningful exposure to risk (i.e., 24 percent are reported as disobedient at school, 28 percent admit cheating, and 17 percent bully peers). Health indicators show that 11 percent are diagnosed with ADHD, and average BMI at age 15 is 23.8. Nearly half of peers are perceived as likely to graduate from college, and most peers study "most times". However, substantial proportions of peers smoke (17 percent), drink alcohol (26 percent), or use marijuana (32 percent), underscoring the importance of peer effects as a contextual influence during adolescence.

A2 Appendix for Section 4: Theoretical Motivation

A2.1 Parental Preferences, Value Function and Dynamic Problem

At the beginning of each period t, the family enters with an observed information set Ω_t. Conditioning on their permanent type μ and the current period shock ν_t, parents observe the current state and make their behavioral choices, after which the child's cognitive skill is updated. Figure 1 illustrates the within-period timing of events during the pre-teenage stage. In the terminal period t_6 (teenage period), the child's teenage outcomes are realized as a function of the full history of accumulated cognitive skill and parental behaviors over the lifecycle. Figure 2 illustrates the terminal-period timing of events.

Household Budget Constraint

The household's consumption at time t, denoted c_t, is determined by income from labor supply, welfare transfers, spousal earnings, child support and illicit income (i.e., income from selling drugs/prostitution etc.) minus the costs associated with risky behaviors:

$$c_t = e_t \cdot \omega_t + w_t \cdot \delta_t + \mathbb{1}[m_t \neq 0] \cdot y_t^s + \mathbb{1}[m_t \neq 1] \cdot y_t^{cs} + q_t - \sum_{l=1}^{3} p_t^l s_t^l - \kappa_r r_t - \kappa_j j_t - f_t r_{t-1} \tag{23}$$

Here, on the income side, ω_t is the realized wage, δ_t is the welfare benefit level, y_t^s is spousal income for mothers with a married/cohabiting spouse, y_t^{cs} is child support received by mothers who are not married/cohabiting with the child's biological dad, and q_t is illicit income. Illicit income q_t is treated as an exogenous endowment in a given period; it is not a separate choice variable. Finally, p_t^l is the price of the l^{th} substance, the κ terms capture monetary penalties associated with criminal behavior, and f_t captures any fines paid for being convicted in the previous period.

Parental Preferences

The parental behavior vector in period t is defined as $B_t = (m_t, e_t, w_t, s_t^1, s_t^2, s_t^3, r_t, j_t)$. The parent's per-period utility over consumption, behaviors, and child cognitive achievement is:

$$u(c_t, B_t, \theta_{t-1}, \mu, \nu_t) = u(c_t) + \gamma_\theta \theta_{t-1} + \gamma_m m_t + \gamma_e e_t + \gamma_w w_t + \sum_{l=1}^{3} \gamma_l s_t^l + \gamma_r r_t + \gamma_j j_t + \mu + \nu_t$$

where $u(c_t)$ is consumption utility and $\gamma_\theta \theta_{t-1}$ is utility the parent derives from the child's cognitive achievement, with γ_θ being the weight parents place on the child's cognitive achievement in their own utility. The remaining γ coefficients reflect behavioral utilities, disutility, stigma, or social value from choices. The term μ is a time-invariant, family-specific unobserved heterogeneity parameter capturing permanent differences in preferences, and ν_t is a family-level time-varying idiosyncratic shock. The utility function is additively separable across consumption, child cognition, and behavioral choices. This is a standard simplifying assumption in dynamic discrete choice models (see, e.g., Keane and Wolpin, 1997; Bernal, 2008), which facilitates identification and estimation while keeping the state space tractable. It implies that the marginal utility of each behavioral choice is independent of the levels of other choices. While interactions between behaviors are plausible, the joint dynamics of behaviors are captured through the shared unobserved heterogeneity terms μ and ν_t, as well as through the lagged behavioral state variables in Ω_t.

Laws of Motion

Child Cognitive Skill: The child's cognitive skill θ_{t+1} evolves according to the value-added production function such that: $\theta_{t+1} = f(\theta_t,\ B_t,\ X_t,\ \varepsilon_t)$, where $f(\cdot)$ captures the role of current cognitive skill, parental behaviors, household characteristics, and idiosyncratic shocks ε_t on next-period skill development.

Teenage Outcomes: At the close of the terminal period t_6, the child's teenage outcomes are realized as a vector: $\mathbf{Y} = (Y^1,\ Y^2,\ \ldots,\ Y^O)$, capturing outcomes in education, substance use, and delinquency. Each outcome Y^o is determined by the full history of accumulated cognitive skill and parental behaviors: $Y^o = g(\theta_{t_1}, \ldots, \theta_T;\ B_{t_1}, \ldots, B_T;\ X,\ \varepsilon_T)$

Dynamic Optimization Problem

The parent's problem is to choose a sequence of behaviors $\{B_t\}$ to maximize expected lifetime utility, anticipating the effects of current choices on their child's future cognitive skill and, ultimately, the teenage outcomes realized in the terminal period.

Pre-Teenage Periods (t_1 through t_5): In each pre-teenage period, the value function satisfies the Bellman equation:

$$V_t(\Omega_t) \;=\; \max_{B_t \in \mathcal{B}} \Big\{ \, u(c_t, B_t, \theta_{t-1}, \mu, \nu_t) \;+\; \beta \cdot \mathbb{E}\big[V_{t+1}(\Omega_{t+1}) \mid \Omega_t, B_t\big] \, \Big\}, \quad t = t_1, \ldots, t_5 \tag{24}$$

subject to the budget constraint (equation 23) and the law of motion $\theta_{t+1} = f(\theta_t, B_t, X_t, \varepsilon_t)$.

Terminal Period (t_6 = Age 15): In the terminal period t_6, the parent still makes behavioral choices B_T. At the close of this period, the teenage outcome vector $\mathbf{Y}$ is observed. The terminal value function is:

$$V_T(\Omega_T) \;=\; \max_{B_T \in \mathcal{B}} \Big\{ \, u(c_T, B_T, \theta_{T-1}, \mu, \nu_T) \;+\; \Psi(\mathbf{Y}, \, \Omega_T, \, B_T) \, \Big\} \tag{25}$$

where $\Psi(\cdot)$ is the terminal payoff capturing the parent's valuation of the child's teenage outcomes.

The full dynamic problem can thus be written recursively. Parents at birth (t_1) solve:

$$V_{t_1}(\Omega_{t_1}) \;=\; \max_{\{B_t\}} \; \mathbb{E}\left[\sum_{t=t_1}^{T-1} \beta^{t-t_1} \, u(c_t, B_t, \theta_{t-1}, \mu, \nu_t) \;+\; \beta^{T-t_1} \, \Psi(\mathbf{Y}, \, \Omega_T, \, B_T) \, \right] \tag{26}$$

where $\beta \in (0, 1)$ is the discount factor. Equation 26 ensures that teenage outcomes $\mathbf{Y}$ enter the parent's objective from the very first period, through the terminal payoff $\Psi(\cdot)$. Since the model has a finite horizon with no period beyond t_6, $\Psi(\cdot)$ plays the role of a bequest value, capturing the parent's long-run concern for the child's wellbeing beyond the final period of data observation. The specification of $\Psi(\cdot)$ is such that parents derive direct utility from each component of the child's teenage outcome vector: $\Psi(\mathbf{Y}) = \sum_o \lambda_o Y^o$, where λ_o is the parent's marginal utility weight on the o-th teenage outcome.

Fully solving the model requires specifying the evolution of parental beliefs in order to recover the functional forms of the utility and production functions, the demand functions, and the continuation value function. This would require computing the value function period-by-period across a vast number of state space combinations. The state space in this setting poses particular challenges for two reasons. First, it combines a continuously evolving variable, the child's cognitive skill θ_t, with a high-dimensional set of discrete parental behavioral choices, $B_t = (m_t, e_t, w_t, s_t^1, s_t^2, s_t^3, r_t, j_t)$, which together govern the child's developmental path.

Second, the sheer dimensionality of the joint behavioral choice set makes exact evaluation of the value function across all feasible state variable combinations computationally tasking. Approximating the decision rules offers a tractable path forward without sacrificing the model's essential structure. Working with flexible approximations to the parental behavioral demand functions allows me to preserve the dynamic nature of parental decision-making and its cumulative influence on child cognitive skill and teenage outcomes, without requiring the full solution of the model. A full solution would impose potentially strong restrictions on the utility function's functional form, on what parents are assumed to know about the cognitive skill production process, and on the specification of the terminal payoff $\Psi(\cdot)$.

To implement this approach, I represent each parental behavioral choice as a flexible function of the state space and whatever information is available to the parent prior to making that decision. Because all choices contained in B_t are made at the same point in time under a common information set Ω_t, every decision depends on the same set of variables. Specifically, I adopt a linear approximation for these decision rules, parameterizing each as a function of the state space Ω_t, the prevailing state and county-level policy and supply-side conditions Z_t, permanent unobserved heterogeneity μ, time-varying unobserved heterogeneity ν_t, and a behavior-specific idiosyncratic error term u_t. The permanent and time-varying heterogeneity components are handled through the discrete factor random effects approach (Mroz, 1999), under which families are drawn from a finite mixture of

k permanent types, one of which is assigned prior to period t_1, and g time-varying shock types, one of which is realized each period. The mass points μ_k and ν_{tg} and their associated probability weights are recovered jointly with all remaining structural parameters via full information maximum likelihood.

A3 Decomposition of Parental Behavioral Effect: Direct vs Indirect

In this section I outline the decomposition of the effect of parental behavior on teenage outcomes into direct and indirect effects. I present the decomposition under two cases to illustrate how the transmission channels differ across model specifications. The empirical model adopted in this paper is Case 2 (behavioral pattern), consistent with equation (9), however, Case 1 is presented as a theoretical baseline to clarify the role of the cognitive channel. Presenting Case 1 first allows a simpler specification in which behaviors enter the teenage outcome equation individually, the two-channel decomposition into direct and indirect cognitive effects can be introduced in its most basic form, and then extending the same logic to the baseline model which pattern-based specification in Case 2. Consider the full dynamic system of parental behavior, child cognitive achievement, and teenage outcomes:

$$\begin{aligned} P(B_t = j) &= f^b(\boldsymbol{B}_{t-1}, \theta_{t-1}, \boldsymbol{X}_t^f, \boldsymbol{Z}_t, \mu^b, \nu_t^b, u_t^b) \\ \theta_t &= f^\theta(\theta_{t-1}, \boldsymbol{B}_t, \boldsymbol{X}_t^f, \boldsymbol{X}_t^c, Z_t^\theta, \mu^\theta, \nu_t^\theta, u_t^\theta) \\ Y_{T+1} &= g\big(\cdot,\ \theta_T,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \boldsymbol{Z}_t,\ \mu^y,\ \nu_t^y,\ u_{t+1}^y\big) \end{aligned}$$

I decompose the effect of a behavioral choice B_k on the teenage outcome into direct and indirect components under two cases. Here k denotes the period in which the behavioral decision occurs, and the effect propagates forward through subsequent periods via the cognitive production technology.

A3.1 Case 1: Behaviors enter Teen Outcome Production Function Individually (No Pattern/String)

Assume that teenage outcomes depend on the full history of behaviors:

$$Y_{T+1} = g\big(B_T, B_{T-1}, \ldots, B_k, \ldots, B_0,\ \theta_T,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \boldsymbol{Z}_t,\ \mu^y,\ \nu_t^y,\ u_{t+1}^y\big)$$

Under the Markov assumption established in Section 4.1.4 (equation 7), θ_T is a sufficient statistic for the full history of past cognitive achievement. Moreover, since the value-added production function implies that θ_t depends on $\boldsymbol{B}_t$ only contemporaneously (i.e. past behavior $\boldsymbol{B}_k$ affects θ_T only through the accumulated skill chain $\theta_k \to \theta_{k+1} \to \cdots \to \theta_T$, and not directly on $\theta_T, \theta_{T-1}, \ldots, \theta_{k+1}$), the total derivative of Y_{T+1} with respect to B_k takes a two-channel form.

The total discrete difference of the teenage outcome with respect to behavior at period k is:[14]

$$\frac{dY_{T+1}}{dB_k} = \underbrace{\frac{\partial g}{\partial B_k}}_{\text{Direct effect}} + \underbrace{\frac{\partial g}{\partial \theta_T} \cdot \frac{d\theta_T}{dB_k}}_{\text{Indirect cognitive effect}} \tag{27}$$

The cognitive transmission channel expands forward through the recursive learning technology:

$$\frac{d\theta_T}{dB_k} = \frac{\partial \theta_T}{\partial \theta_{T-1}} \frac{\partial \theta_{T-1}}{\partial \theta_{T-2}} \cdots \frac{\partial \theta_{k+1}}{\partial \theta_k} \frac{\partial \theta_k}{\partial B_k} \tag{28}$$

Therefore the total effect becomes:

$$\frac{dY_{T+1}}{dB_k} = \underbrace{\frac{\partial g}{\partial B_k}}_{\text{Direct effect}} + \underbrace{\frac{\partial g}{\partial \theta_T} \left[\frac{\partial \theta_T}{\partial \theta_{T-1}} \frac{\partial \theta_{T-1}}{\partial \theta_{T-2}} \cdots \frac{\partial \theta_{k+1}}{\partial \theta_k} \frac{\partial \theta_k}{\partial B_k} \right]}_{\text{Indirect cognitive effect}} \tag{29}$$

When behaviors enter the teenage outcome production function directly as individual period-specific inputs, the teenage outcome responds to a behavioral shock through two channels. The indirect cognitive channel implies parental behavior in the past affects the child's cognitive development, and since teenage outcomes depend on cognition, the dynamic accumulation of cognitive development over time generates an indirect transmission channel. A direct impact of lagged behaviors suggests that parental behavior affects teenage outcomes directly or possibly through omitted channels (such as non-cognitive development).

A3.2 Case 2: Behaviors enter Teen Outcome Production Function as a String/Pattern

Now assume the teenage outcome depends on the behavioral pattern (or string) rather than each period's behavior separately. Let $\overrightarrow{B} \equiv \mathcal{H}(B_0, B_1, \ldots, B_T)$ denote the full behavioral history mapped into a finite set of categories such as "always employed," "never employed," "early employment," etc. Then the teenage outcome satisfies:

$$Y_{T+1} = g\Big(\overrightarrow{B},\ \theta_T,\ \boldsymbol{X}_t^f,\ \boldsymbol{X}_t^c,\ \boldsymbol{Z}_t,\ \mu^y,\ \nu_t^y,\ u_{t+1}^y\Big)$$

Consider a unit change in behavior at some period $k \in \{1, 2, \ldots, T\}$. The total effect of B_k on the teenage outcome can be decomposed into a direct pattern channel and an indirect cognitive channel:

$$\frac{dY_{T+1}}{dB_k} = \underbrace{\frac{\partial g}{\partial \overrightarrow{B}} \cdot \frac{\Delta \overrightarrow{B}}{\Delta B_k}}_{\text{Direct pattern effect}} + \underbrace{\frac{\partial g}{\partial \theta_T} \cdot \frac{d\theta_T}{dB_k}}_{\text{Indirect cognitive effect}} \tag{30}$$

[14] Since B_k is a discrete choice variable, the "derivative" dY_{T+1}/dB_k should be understood as a discrete difference operator – the change in the expected outcome from a unit change in B_k, holding all other period behaviors fixed. The chain-rule notation is used for expositional clarity and corresponds to the marginal effect of the latent continuous index underlying the discrete choice.

where $\Delta\overrightarrow{B}/\Delta B_k$ denotes a discrete difference operator capturing whether a change in B_k causes a shift in the behavioral string category.[15] The cognitive component follows the same forward recursive learning technology as in Case 1:

$$\frac{d\theta_T}{dB_k} = \frac{\partial\theta_T}{\partial\theta_{T-1}}\frac{\partial\theta_{T-1}}{\partial\theta_{T-2}}\cdots\frac{\partial\theta_{k+1}}{\partial\theta_k}\frac{\partial\theta_k}{\partial B_k}$$

Therefore, the total effect can be written compactly as:

$$\frac{dY_{T+1}}{dB_k} = \underbrace{\frac{\partial g}{\partial\overrightarrow{B}}\cdot\frac{\Delta\overrightarrow{B}}{\Delta B_k}}_{\text{Direct pattern effect}} + \underbrace{\frac{\partial g}{\partial\theta_T}\left[\frac{\partial\theta_T}{\partial\theta_{T-1}}\frac{\partial\theta_{T-1}}{\partial\theta_{T-2}}\cdots\frac{\partial\theta_{k+1}}{\partial\theta_k}\frac{\partial\theta_k}{\partial B_k}\right]}_{\text{Indirect cognitive effect}} \tag{31}$$

When the teenage outcome depends on the pattern of parental behavior rather than specific period values, the direct channel reflects the child's exposure to long-run behavioral patterns. A change in behavior at time k affects the teenager if it causes the family to move to a different behavioral category (e.g., "always employed" vs. "intermittent employment"). The indirect channel through cognitive development remains fully dynamic, preserving the cumulative developmental consequences of parental behavior over the child's life.

A4 Appendix for Section 6: Results

A4.1 Likelihood Ratio Test

A4.1.1 Heterogeneity vs. No Heterogeneity

The Likelihood Ratio (LR) test assesses the relative goodness of fit of two nested models. In this context, the unrestricted specification, which incorporates unobserved heterogeneity is pitted against the restricted specification, which omits it, to determine whether the added flexibility of the former translates into a statistically meaningful gain in the log-likelihood value. The test statistic is constructed as:

$$LR = 2(\ell_\mathrm{n} - \ell_\mathrm{r}) \xrightarrow{d} \chi^2_{df} \tag{32}$$

where ℓ_n denotes the maximized log-likelihood of the unrestricted model and ℓ_r denotes the corresponding value for the restricted model. Under the null hypothesis that the restrictions are valid, this statistic follows a chi-squared asymptotic distribution, with degrees of freedom (df) given by the difference in the parameter count across the two specifications. Statistical significance is assessed by comparing the computed test statistic against the appropriate critical value from this distribution. The inputs used to construct the test statistic are summarized in Table A.21:

[15] The key assumption in this case is that the direct effect operates only through whether period-k behavior changes the pattern classification. Since B_k is discrete, the direct pattern effect is characterized via a discrete difference rather than a smooth partial derivative:

$$\frac{\Delta\overrightarrow{B}}{\Delta B_k} = \begin{cases} 1 & \text{if a unit change in } B_k \text{ changes the pattern category,} \\ 0 & \text{otherwise.} \end{cases}$$

Table A.21: Comparison of Likelihood Function Values: Heterogeneity vs. No Heterogeneity

Model	**Likelihood**	**No. of Parameters**	**LR Test Statistic**
No Heterogeneity	−41,127	837	
Heterogeneity	−40,078	987	2,098

Note: LR test statistic computed as $2(\ell_{\text{Het}} - \ell_{\text{No Het}})$.

At 150 degrees of freedom and a 1% significance level, the critical chi-squared value is $\chi^2_{0.01}(150) \approx 194.6$. Since the computed test statistic exceeds this threshold, the null hypothesis that the restrictions are valid is rejected, confirming that the specification accommodating unobserved heterogeneity achieves a statistically superior fit relative to the restricted model that excludes it. The results provide strong statistical evidence in favor of the richer specification with unobserved heterogeneity, justifying its use in the analysis and in the computation of marginal effects and counterfactuals.

A4.1.2 Behavior Pattern Model (Baseline Model) vs. Other Models

The critical value for chi-square with 119 degrees of freedom at the 5% significance level is $\chi^2_{0.95}(119) \approx 145.46$. Based on a comparison between the test statistic from the Tables A.22 and A.23, and this critical value, we reject the null hypothesis at $\alpha = 5\%$, indicating a statistically significant improvement in model fit for the Behavior Pattern Model (Baseline Model) relative to both the Behavior Summation Model and the Contemporaneous Behavior Model. These results provide strong statistical evidence in favor of the Behavior Pattern Model.

Table A.22: Comparison of Likelihood Function Values: Behavior Summation Model vs. Behavior Pattern Model (Baseline Model)

Model	**Likelihood**	**No. of Parameters**	**LR Test Statistic**
Behavior Summation Model	−40,156	868	
Behavior Pattern Model	−40,078	987	156

Note: LR test statistic computed as $2(\ell_{\text{pattern}} - \ell_{\text{summation}})$.

Table A.23: Comparison of Likelihood Function Values: Contemporaneous Behavior Model vs. Behavior Pattern Model (Baseline Model)

Model	**Likelihood**	**No. of Parameters**	**LR Test Statistic**
Contemporaneous Behavior Model	−40,155	868	
Behavior Pattern Model	−40,078	987	154

Note: LR test statistic computed as $2(\ell_{\text{pattern}} - \ell_{\text{contemporaneous}})$.

A4.2 Wald F-Tests for Instrument Relevance

Table A.24: **Wald F-Tests for Instrument Relevance**

	Wald Test			
Equation	# of restrictions (q)	F-statistic	χ^2	p-value
Mother's Behavioral Equations				
Marriage/Cohabitation	7	1.730	12.112	0.097^{*}
Employment	7	2.356	16.492	0.021^{**}
Welfare Receipt	7	9.591	67.138	$<0.001^{***}$
Alcohol Use	7	2.481	17.365	0.015^{**}
Smoking	7	1.962	13.735	0.056^{*}
Marijuana Use	7	2.142	14.997	0.036^{**}
Conviction	7	2.546	17.822	0.013^{**}
Father's Behavioral Equation				
Jail	2	7.269	14.539	0.001^{***}
Child Cognitive Skill Equation				
Cognitive Skill	1	5.899	5.899	0.015^{**}
Child Teenage Outcome Equations				
Ever failed class	6	3.909	23.454	0.001^{***}
Unlikely to graduate college	6	7.667	46.005	$<0.001^{***}$
Ever suspended from school	6	7.229	43.371	$<0.001^{***}$
Ever drunk alcohol	6	23.202	139.210	$<0.001^{***}$
Ever used marijuana	6	11.651	69.908	$<0.001^{***}$
Ever stopped by police	6	13.151	78.907	$<0.001^{***}$
Ever had sex	6	10.960	65.760	$<0.001^{***}$

Notes: Each row reports a Wald F-test for the joint significance of excluded instruments in the specified equation. q denotes the number of exclusion restrictions (instruments). The test statistic is $F = \frac{1}{q}\hat{\boldsymbol{\beta}}_z'[\mathbf{V}(\hat{\boldsymbol{\beta}}_z)]^{-1}\hat{\boldsymbol{\beta}}_z$, where $\hat{\boldsymbol{\beta}}_z$ is the vector of instrument coefficients and $\mathbf{V}(\hat{\boldsymbol{\beta}}_z)$ is the corresponding submatrix of the FIML variance-covariance matrix. The χ^2 statistic $= q \times F$. Instruments that enter mother's behavioral equations are state-level policy (fraction of married, unemployment rate, fraction of household on public assistance, beer tax, cigarette tax, marijuana laws, crime rate). Instruments that enter the father's incarceration equation are state-level crime rate and substance treatment facility availability. The instrument that enter the cognitive skill equation is the standardized state test score. Instruments that enter the teenage outcome equations are peer characteristics (peers' drinking, smoking, marijuana use, peer study habit, likelihood of graduating college) and standardized state test score. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

A4.3 Marginal Effects of Parental Behavior Patterns and Cognitive Achievement on Teenage Outcomes

Table A.25 reports the estimated marginal effects of parental behavioral patterns and child cognitive achievement, on teenage outcomes, which include school failure, college readiness, school suspension, alcohol and marijuana use, police contact, and early sexual activity.[16] The marginal effects presented are percentage points (p.p.) changes relative to the omitted baseline category in each behavioral dimension. Mothers who consistently reside with the biological father of the child reduce the probability of failing a class by 7.6 p.p. and decrease the probability of being unlikely to graduate college by 6.8 p.p., relatively to single mothers. Early singlehood and unstable relationship (i.e., a mother who has experienced all the types of marital status: singlehood, married to biobad, and married to another man) increases school suspension by 9.0 and 7.4 p.p., respectively. Teens whose mothers are in unstable relationships or had their parents divorce/separate in later period are likely to having a police contact.

Teens whose mothers were always employed are 14.6 p.p. more likely to have been suspended from school and 11.2 p.p. less likely to have been in contact with the police, relative to children of never-employed mothers. Mothers moving in and out of employment in the early and late periods also increase the probability of school suspension and teenage marijuana use by 19.9 and 11.8 p.p. respectively, relative to the baseline. These mixed effects are consistent with family time allocation tradeoff, in that regular employment may raise household income and stability but simultaneously reduce parental time available for monitoring. Persistent welfare use ("always on welfare") increases the likelihood of school suspension by 20.8 p.p., being stopped by the police by 34.6 p.p. and early sexual activity by 16.6 p.p., relatively to mothers who were never on welfare. These large magnitudes suggest that welfare participation may capture deep underlying disadvantage or chronic economic stress that interferes with early skill formation and later teenage outcomes. Mixed patterns, such as welfare use early but not late, and intermittent welfare use (moving in and out of welfare during early and late periods), shows sizeable worse effects on teenage class failure, and this highlights that welfare spells, even when temporary, are strongly predictive of adolescent educational difficulties.

On the contrary, early smoking followed by cessation ("start early & quit") decreases probability of risky sexual behavior. The reason may be that cessation represents a lifestyle shift within the household. When a mother quits smoking in the later period, the child is directly exposed to both the negative consequences of the earlier behavior and the intentional effort required to quit. This combination may create a powerful deterrent effect, which is the mother's cessation serves as an experiential lesson about health risks and long-term consequences. Similar deterrent effect also explains why teens whose mothers always drank alcohol are significantly less likely to be suspended from school or stopped by the police; although other irregular drinking patterns increases the likelihood of not graduating college and teenage drinking. Regarding paternal jail time, children whose father were jailed before by age 5 (first 3 periods) are 5.9 p.p. more likely to engaging in drinking during their teenage. Finally, a 10-point increase in the child's cognitive score reduces the probability of failing a class in their teenage by 3.9 p.p..

[16] See the full table of results in Online Appendix Tables OA.12 – OA.18

Table A.25: **Marginal Effects of Parental Behavior Patterns and Cognitive Achievement on Teenage Outcomes**

Dynamic Behavior Patterns	**Ever Failed Class**	**Unlikely to Graduate College**	**Ever Suspended from Sch**	**Ever Drank Alcohol**	**Ever Use Marijuana**	**Ever Stopped by Police**	**Ever had Sex**
Marital Status ($\overrightarrow{M}$)							
with biodad always	−0.076	−0.068	0.010	0.098***	0.019	0.022	−0.019
	(0.049)	(0.045)	(0.041)	(0.037)	(0.038)	(0.042)	(0.035)
single early	0.032	0.042	0.090**	0.087***	0.038	0.036	0.031
	(0.047)	(0.044)	(0.037)	(0.033)	(0.032)	(0.042)	(0.033)
with biodad early	−0.042	−0.037	0.021	0.030	−0.006	0.064*	0.015
	(0.045)	(0.042)	(0.037)	(0.031)	(0.028)	(0.037)	(0.030)
unstable relationship	−0.027	0.025	0.074**	0.077**	0.017	0.090**	0.022
	(0.046)	(0.042)	(0.036)	(0.033)	(0.032)	(0.039)	(0.033)
other marital patterns	−0.032	−0.047	0.042	0.067**	0.024	0.013	0.031
	(0.042)	(0.039)	(0.033)	(0.030)	(0.028)	(0.036)	(0.029)
Base: always single							
Mom Employment ($\overrightarrow{E}$)							
always employed	−0.039	0.007	0.146*	0.077	0.093	−0.112*	0.024
	(0.081)	(0.073)	(0.075)	(0.058)	(0.061)	(0.062)	(0.061)
ever emp. early & always unemp. late	−0.011	0.015	0.199**	0.008	0.069	−0.027	−0.011
	(0.087)	(0.079)	(0.080)	(0.066)	(0.066)	(0.076)	(0.070)
always unemp. early & ever emp. late	0.013	0.031	0.158*	0.067	0.077	0.025	−0.101
	(0.091)	(0.085)	(0.084)	(0.073)	(0.073)	(0.072)	(0.075)
intermittent employment	0.031	0.075	0.199***	0.078	0.118**	−0.062	0.037
	(0.074)	(0.066)	(0.071)	(0.056)	(0.060)	(0.058)	(0.056)
Base: always unemployed							
Mom Welfare Receipt ($\overrightarrow{W}$)							
always on welfare	0.176	0.074	0.208**	0.042	0.025	0.346***	0.166***
	(0.126)	(0.100)	(0.100)	(0.059)	(0.078)	(0.090)	(0.054)
ever on wel. early & always not on wel. late	0.071**	0.009	0.017	−0.010	−0.003	−0.003	0.029
	(0.033)	(0.031)	(0.027)	(0.022)	(0.022)	(0.028)	(0.024)
always not on wel. early & ever on wel. late	0.077	0.040	0.067*	0.033	0.023	0.044	−0.005
	(0.049)	(0.045)	(0.036)	(0.028)	(0.029)	(0.039)	(0.036)
intermittent welfare	0.068*	0.020	0.011	0.019	0.028	0.056*	0.061**
	(0.039)	(0.036)	(0.030)	(0.025)	(0.025)	(0.032)	(0.028)
Base: never on welfare							

Notes: The table reports marginal effects of parental behavioral trajectory patterns on seven teenage outcomes, estimated from the jointly estimated structural model via Full Information Maximum Likelihood (FIML) with Discrete Factor Random Effects (DFRE). Rather than relying on single-period snapshots of parental behavior, the marginal effects are computed with respect to the full lifecycle behavioral trajectory of each parent, denoted by the arrow notation $\overrightarrow{\cdot}$. "Early" refers to parental behaviors occurring during the first three developmental periods (Birth, Age 1, and Age 3), while "Late" corresponds to behaviors observed during the latter three periods (Age 5, Age 9, and Age 15). The table continues on the following page with marginal effects for substance use, criminal conviction, and father's incarceration trajectories, as well as for the child's accumulated cognitive skill. Standard errors are reported in parentheses and are obtained parametrically via bootstrap with N replications. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.25: **Marginal Effects of Parental Behavior Patterns and Cognitive Achievement on Teenage Outcomes (continued)**

Dynamic Behavior Patterns	**Ever Failed Class**	**Unlikely to Graduate College**	**Ever Suspended from Sch**	**Ever Drank Alcohol**	**Ever Use Marijuana**	**Ever Stopped by Police**	**Ever had Sex**
Mom Drinking ($\vec{S}^1$)							
always drink	−0.111	0.126	−0.145*	0.050	−0.009	−0.218**	−0.096
	(0.101)	(0.087)	(0.080)	(0.056)	(0.053)	(0.097)	(0.075)
start drinking early & quit drinking late	0.028	0.061	0.029	0.081*	0.014	−0.019	0.031
	(0.069)	(0.068)	(0.059)	(0.049)	(0.048)	(0.067)	(0.049)
never drinking early & start drinking late	0.011	0.098**	0.016	0.034	−0.010	−0.022	−0.028
	(0.045)	(0.042)	(0.040)	(0.030)	(0.031)	(0.042)	(0.036)
intermittent drinking	−0.053	0.099*	−0.032	0.003	−0.026	−0.045	−0.047
Base: never drinks	(0.063)	(0.055)	(0.052)	(0.038)	(0.035)	(0.057)	(0.045)
Mom Smoking ($\vec{S}^2$)							
always smoke	0.133	0.065	−0.027	0.060	−0.017	−0.185	−0.004
	(0.133)	(0.142)	(0.105)	(0.074)	(0.062)	(0.135)	(0.083)
start smoking early & quit smoking late	0.091	0.094	0.008	−0.027	−0.073	−0.089	−0.155*
	(0.076)	(0.071)	(0.060)	(0.048)	(0.057)	(0.082)	(0.085)
never smoking early & start smoking late	0.079	0.026	−0.016	−0.049	−0.031	−0.063	0.002
	(0.075)	(0.067)	(0.055)	(0.041)	(0.035)	(0.066)	(0.042)
intermittent smoking	0.108	0.062	−0.034	0.031	−0.050	−0.128	−0.009
Base: never smokes	(0.122)	(0.129)	(0.095)	(0.066)	(0.056)	(0.122)	(0.077)
Mom Marijuana Use ($\vec{S}^3$)							
ever use marijuana	0.038	0.052	0.056	0.003	0.008	−0.036	0.005
Base: never use marijuana	(0.050)	(0.044)	(0.039)	(0.028)	(0.027)	(0.041)	(0.035)
Mom Conviction ($\vec{R}$)							
first-ever conviction was early	0.081	−0.077	−0.035	−0.092	0.055	−0.053	0.086
	(0.139)	(0.119)	(0.118)	(0.069)	(0.086)	(0.097)	(0.185)
first-ever conviction was late	0.078	0.065	0.044	0.002	0.005	−0.020	−0.048
Base: never convicted	(0.051)	(0.046)	(0.036)	(0.031)	(0.028)	(0.044)	(0.031)
Dad in Jail ($\vec{J}$)							
first-ever jail was early	−0.041	−0.024	0.022	0.058**	0.029	0.041	0.029
	(0.041)	(0.038)	(0.033)	(0.024)	(0.024)	(0.034)	(0.037)
first-ever jail was late	0.005	−0.032	0.038	−0.002	−0.009	0.039	−0.066
Base: never in jail	(0.040)	(0.038)	(0.033)	(0.030)	(0.030)	(0.033)	(0.046)
Child Cognitive Achievement							
cognitive score (10-point scale)	−0.039**	0.013	−0.016	−0.005	−0.012	−0.005	−0.021
	(0.017)	(0.016)	(0.013)	(0.011)	(0.011)	(0.014)	(0.015)

Notes: The table reports marginal effects of parental behavioral trajectory patterns on seven teenage outcomes, estimated from the jointly estimated dynamic model via Full Information Maximum Likelihood (FIML) with Discrete Factor Random Effects (DFRE). Rather than relying on single-period snapshots of parental behavior, the marginal effects are computed with respect to the full lifecycle behavioral trajectory of each parent, denoted by the arrow notation $\vec{\cdot}$. "Early" refers to parental behaviors occurring during the first three developmental periods (Birth, Age 1, and Age 3), while "Late" corresponds to behaviors observed during the latter three periods (Age 5, Age 9, and Age 15). Standard errors are reported in parentheses and are obtained parametrically via bootstrap with N replications. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A4.4 Other Counterfactuals

Table A.26: **All Behaviors: Counterfactual and Total Effects Decomposition**

	Counterfactual (1)	**Counterfactual (2)**	**Total Effect**	**Direct Behavioral Effect**	**Indirect Cognitive Effect**
Ever failed class	0.662	0.389	0.273	0.242	0.030
			(0.198)		
Unlikely to graduate college	0.385	0.191	0.194	0.206	−0.012
			(0.183)		
Ever suspended from school	0.377	0.252	0.125	0.111	0.014
			(0.190)		
Ever drunk alcohol	0.093	0.128	−0.034	−0.038	0.003
			(0.111)		
Ever used marijuana	0.220	0.190	0.030	0.020	0.010
			(0.147)		
Ever stopped by police	0.363	0.267	0.096	0.092	0.004
			(0.190)		
Ever had sex	0.363	0.179	0.184	0.168	0.016
			(0.170)		

Notes: The table reports counterfactual predicted teenage outcomes and a decomposition of the total effect of the full vector of parental behaviors on teenage outcomes, simulated using the estimated dynamic model. This table represents the most comprehensive counterfactual comparison in the paper, simultaneously varying all parental behaviors for both the mother and father across the entire developmental lifecycle. *Counterfactual (1) reflects the most disadvantaged behavioral profile, in which the mother has an unstable marital history, is persistently unemployed, continuously receives welfare, drinks alcohol, smokes cigarettes, uses marijuana, and is convicted in every period, while the father is incarcerated* in every period. *Counterfactual (2) reflects the most advantaged behavioral profile, in which the mother is always married to or cohabiting with the biological father, is consistently employed, never receives welfare, abstains from all substances, and is never convicted, while the father is never incarcerated.* Total Effect = Counterfactual (1) − Counterfactual (2). Total Effect is decomposed into a Direct Behavioral Effect, and an Indirect Cognitive Effect. Simulated outcomes are averaged over 1,000 parametric bootstrap replications; standard errors are reported in parentheses. None of the total effects are statistically significant at conventional levels. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

A5 Appendix for Chapter 2

Table A.27: **Summary of Literature: Parental Behaviors on Child Outcomes**

Parental Behaviors:	**Employment/Welfare**	**Substance Abuse**	**Criminal Behavior**
Does not address Unobserved Heterogeneity	*Cognitive/Education Outcomes* Waldfogel et al. (2002) Han (2005) *Health Outcomes* Lawson et al. (2020)	*Health Outcomes* Waller & Swisher (2006) Cumming & Gearhardt (2020) Osborne & Berger (2009) Haghighi et al. (2013) Riedel et al. (2014) *Cognitive and Health Outcomes* Jeong et al. (2021)	*Cognitive/Education Outcomes* Haskins (2017) *Health Outcomes* Wildeman (2010) Murray et al. (2012) Washington (2018) *Welfare Outcomes* Geller (2013)
Address Unobserved Heterogeneity	*Cognitive/Education Outcomes* Agostinelli & Sorrenti (2021) Akee et al. (2010) Attenasio et al. (2020)* Bernal (2008) Bernal & Keane (2011) Bono et al. (2016) Dahl & Lochner (2012) Hsin & Felfe (2014) Løken et al. (2018) Lui et al. (2009)* Ruhm (2004) Carneiro & Rodrigues (2009) *Health Outcomes* Geleta (2023) Schaller & Zerpa (2019) *Criminal Outcomes* Akee et al. (2010)	*Cognitive and Non-cognitive Outcomes* Srivastava & Trinh (2021) Mangiavacchi et al. (2018) Frank & Meara (2009) *Health Outcomes* Chatterji & Markowitz (2001) Bella et al. (2023) *Labor Market Outcomes* Balsa (2008) *Welfare Outcomes* Cunningham & Finlay (2012)	*Education and Criminal Outcomes* Bhuller et al. (2018) Arteaga (2019) Norris et al. (2021) *Welfare Outcomes* Geller et al. (2011) Chung (2012) *Labor Market and Criminal Outcomes* Dobbie et al. (2018)

Note: Child outcomes are *italicized* within table cells. * These studies modelled parental behaviors/decisions jointly