Child Care Policy and Informal Care

Joel McMurry *

November 5, 2021

Abstract

Early child care experiences vary widely across the distribution of socio-economic status (SES), and sizeable skill gaps open up before children enter publicly-provided schooling. SES gradients in the quality of informal, relative-provided, care are particularly large. To understand how variation in the availability and quality of informal care contributes to skill inequality, I estimate a model of child care, mother labor supply, and child skill development, allowing for unequal access to informal care. I exploit the timing of grandmother deaths relative to a child's birth to identify substitution patterns between informal, formal, and mother-provided child care. I quantify the effect of having access to informal care on child development and mother labor supply, and I estimate that, for a substantial fraction of less-advantaged children, the availability of informal care is detrimental to skill development. I ex ante analyze the effects of policies such as universal public daycare, subsidies for formal care, and cash transfers, and show that accounting for heterogeneity in the availability and quality of informal care is quantitatively important for estimating the effect that such policies might have on skill inequality at the point of entry into K-12 schooling.

JEL: J13, J18

Keywords: Female labor supply, child care, human capital

*University of Wisconsin-Madison. Email: mcmurry2@wisc.edu.

1 Introduction

In this paper, I develop and estimate a model of skill production, child care choice, and mother labor supply during early childhood, explicitly modeling the choice to use informal care provided by relatives and accounting for heterogeneity in the quality of care within types of arrangements. Skill gaps between advantaged and disadvantaged groups in the US open up before children enter pre-kindergarten [Heckman, 2008], and one of the first sources of inequality that children confront is variation in the quality of early childhood care. Informal care is common in the US, with roughly 20% of families of young children using relative-provided child care. The use and quality of relative care varies significantly across the socio-economic spectrum. During the first three years of life, children of high school dropouts receive over 75% of their non-parental care from relatives, compared with under 40% for children of college-educated mothers. Gaps in quality between these groups are similarly stark. Children of lower-educated mothers receive informal care that is, on average, over 85% standard deviations lower quality than relative care experienced by their more-advantaged peers. By the same measure of care quality, this gap in formal center-based care is approximately 55% [Flood et al., 2021].

I quantify the importance of child care choice in generating skill gaps at the point of entry into K-12 schooling and estimate behavioral responses to frequently discussed but untried counterfactual policies, such as universal child care, on inequality in human capital, mother labor supply, and expenditures on formal care during a crucial period of child development. The effect on children of policies that aim to alter the early care they receive crucially depends on the counterfactual arrangements from which families endogenously switch [Kline and Walters, 2016]. Therefore, when developing my empirical model, I pay particular attention to the joint distribution of mother care qualty, relative care quality, and care arrangement elasticities. Put simply, children who receive poor quality care in the status quo will benefit from moving to higher-quality care. The empirical question is whether the parents of such children can be induced to change their choice of child care arrangements. Of course, the choice of how to care for one's child is made jointly with the labor supply decisions that themselves shape the environment in which a child grows. My paper offers a structural framework in which to study these decisions and their sensititvity to policy.

This paper contributes to our understanding of the early childhood origins of inequality and the sensitivity of the status quo to child care policy. I directly account for heterogeneity in the quality and availability of informal, child care provided by relatives. The existing structural literature has studied many important aspects of child care choice and child development, including the demand for paid care quality [Blau and Hagy, 1998], the developmental implications of maternal vs non-maternal care [Bernal, 2008], the productivity of parental time allocation

[Del Boca et al., 2014], and the subsitutability of home- and market-provided child care Caucutt and Lochner [2020]. Two closely-related papers model the choice to take up subsidized center-based care programs in the form of Head Start [**?**] and the Infant Health and Development Program [Chaparro et al., 2020]. My paper contributes to this literature by modeling the decision to use relative care as an option distinct from maternal or formal, center-based care, whereas existing studies include informal care in aggregated groups of either non-maternal or non-program care.[1] The large socioeconomic status (SES) gradient in the usage and quality of relative care suggests that accounting for dispersion in relative care quality is crucial for understanding inequality in childrens' care experiences.

Indeed, I find that our conclusions regarding the efficacy of policy hinge crucially on the modeling treatment of informal care access. I show that the estimated effects of large-scale policies such as formal care subsidies or an expanded public option are very sensitive to the joint distribution of relative availability and family preferences. The intuition is that policy intended to improve child outcomes requires that families who use lower-quality care arrangements actually do switch to higher-quality formal care when treated by a subsidy or offer. If families who are not currently using formal care have a large distaste for it, few will actually switch.[2] Estimating the distribution of preferences given choice data requires taking a stand on the choice set that a family faces and rationalizing observed child care care choices.

A significant econometric challenge in studying the choice to use informal care is that the availability of such care is not randomly assigned, and relatives who offer to provide child care likely systematically differ from those who do not. To address this endogeneity concern, I use grandmother mortality as an instrument for informal care usage, explicitly incorporating the instrument into my structural estimation. Previous design-based work in economics has used an IV strategy to study grandparents as a child care option. One strand of this literature uses geographic proximity to a grandparent as an instrument for the use of grandparent child care [Dimova and Wolff, 2008, Compton and Pollak, 2014, Boca et al., 2018]. Another set of papers uses, as I do, grandparent mortality [Posadas and Vidal-Fernandez, 2013, García-Morán and Kuehn, 2017], and or age relative to a pension cut-off [Zamarro, 2020] and geograhic variation in child care subsidies [Truskinovsky, 2020]. My approach exploits this same type of variation, but for a different purpose. Instead of seeking to uncover the effect of being cared for by

[1] One paper that does consider relative care as a separate option is the linear IV study of [Bernal and Keane, 2010]. Consistent with that paper, I find that relative care used by lower-SES mothes is of poorer quality, however a central point in my paper is the large dispersion in relative quality that is masked in design-based studies.

[2] It is important to note that a "distaste" for formal care may be a primitive preference or some other unobservable not captured by the model, such as lack of access to paid care. See Pilarz et al. [2019] for evidence that non-standard working hours among low-income mothers hinders usage of formal child care.

relatives, as these papers do, I use the information contained in responses to grandparent death to help identify the effects of child care policy. I estimate that roughly 40% of families have access to informal care, of which 42% actually take it up. Higher-educated mothers are about 10 percentage points less likely to have access to relative care.

I find that access to relative care has a small positive effect on the skills of more-advantaged children and no effect on those of lesser-advantaged children (with socio-economic advantage measured by mother education). This result is driven by differing quality of relatives across the SES spectrum and varying patterns of substitution that mothers of different means make between mother time, relative time, and formal care of different quality. On the other hand, access to informal care does have a large impact on mother labor supply, increasing participation of lower-educated mothers by 18% and that of higher-educated mothers by 23%. Family expenditures on paid care are also reduced by 20-30% just by having access to informal care. Thus the primary effect of informal care access is to alter family resources as opposed to shifting child skills during early life.

With the estimated model in hand, I evaluate several child care policies: a formal care subsidy, a direct cash transfer to families, and the public provision of a high-quality care program. I estimate that the public provision of high-quality care is most effective at reducing skill inequality, while cash transfers have no effect. However, subsidies for paid care have the greatest effect on mother labor supply, reflecting the distribution of income and substitution effects characterized by the model. I find that a 50% paid care subsidy would increase the labor force participation of mothers with access to informal care by 5.6%.

In my model, children are born into a family characterized by family structure, a mother wage offer, non-labor income, and the quality and availability of informal care. Each period, the family realizes the mother's wage offer and chooses maternal labor supply and the allocation of a fixed number of required child care hours across mother care, paid care, and (if available) relative care. Child skills accumulate according to a production function that takes as an input the time-weighted care qualities of each arrangement used.

I start by introducing the model in Section 2. Section 3 describes the data, and Section 4 discusses econometric challenges in identifying and estimating the model with the available data. In Section 5, I discuss identification and describe my estimation strategy. Section 7 studies several counterfactual policies, and Section 8 discusses the importance of identifying relative availability for policy analysis. Section 9 concludes.

2 A Model of Child Care

The empirical model that I outline in this section is designed for two related purposes. First, I aim to characterize the extent to which access to informal child care shapes child development during early life. This requires identifying the population that indeed does have access to informal care and estimating how mother choice (and thus child development) shifts in response to the membership of informal care in a mother's choice set. Second, I use the model as a lens through which to study the sensitivity of policy conclusions with respect to assumptions made about access to informal care.

I consider the dynamic discrete choice problem of a mother with a newborn child. She[3] must jointly decide whether or not to work and how the child will be cared for. Families differ at birth by mother care quality, mother wage offer, non-labor income (including labor income of a father), presence of a father, and the availability and quality of relative care. Initial conditions also include additional exogenous characteristics that drive the stochastic processes in the model. Beyond initial conditions, inequality at the point of starting K-12 schooling is induced by the interaction of optimal decision rules with shocks to family structure and resources.

2.1 Environment

The model takes place over two periods: ages 0-2 and 3-4. At age 5, children in the United States begin to enter K-12 schooling, which I take as a policy-relevant terminal point. An agent in the model is a unitary family, which consists of at least one child and one mother, but whose structure may further vary in terms of the presence of a father, other children, or a grandparent/relative. Markets are incomplete: the family has no access to state contingent claims, and it is subject to a no-borrowing and no-saving constraint, so a flow budget constraint binds in each period. The central friction in this environment is missing markets for mother quality and informal care. That is, mothers cannot improve their own care quality and they cannot buy either better informal care or access to informal care if they lack it. It is this variation in the menu of child care options that distorts family responses to child care policy, relative to a world in which everyone faces the same choice set.

Each period, the mother chooses family consumption, mother labor supply, and child care

[3] I abstract away from any strategic interactions between parents, so the mother's problem is identical to that of a family planner. For simplicity, I study only those households with a present mother, abstracting away from households headed by single fathers, same-sex couples, or other arrangement. This restriction captures the vast majority of children in the data. As every family in this model contains at least a mother, for convenience I will refer to the mother as decisionmaker.

arrangements. All children require the same amount (time) of care, but the family can meet this care requirement by splitting care hours among three options: mother care, paid care, or relative care. Paid care entails a pecuniary cost and relative care is an option only for families with a relative available. That is, if no relative is available, relative care hours are restricted to be zero. The quality of mother and relative care is exogenous, but the quality of paid care is a choice variable. Families face an upward-sloping supply curve of paid care, and so the price of a time unit of care increases in care quality. The government generates revenue via progressive taxation and makes means-tested transfers.

2.2 Child Care

All children require care for $\bar{\tau}$ units of time per week, and the family allocates care across all options available to it. Mothers optimize over the space of average weekly time, so all care arrangements (and labor supply decisions) are written in terms of weekly time and restricted to be homogenous throughout the model period. Letting $\tau_{mt}, \tau_{pt}, \tau_{rt}$ denote mother, paid, and relative care time, the following care hours constraints must hold every period:

$$\bar{\tau} \equiv \tau_{mt} + \tau_{pt} + \tau_{rt} \tag{1}$$

$$\tau_{kt} \geq 0 \tag{2}$$

Central to this study is the idea that relative care is not an option for all families. Whether due to relative preferences, geography, or mortality, it is likely that there are families in which no relative offers to provide child care. In this case the choice set is restricted so that $\tau_{rt} = 0$. I model the availability of relative care as a latent, Normally-distributed random variable

$$\phi_i^* = X_i'\delta_0 + Z_i'\gamma_0 + \epsilon_i^{\phi} \tag{3}$$

where Z_i is a vector of instruments, including the death of a grandmother before birth and $\phi_i = 1$ if $\phi_i^* \geq 0$.

Along with deciding how to allocate care time, mothers choose the quality of paid care, facing an upward-sloping supply curve of care quality. Normalizing the price of the consumption good to be 1, I denote the relative price of a quality-unit of paid care to be p. Total family expenditure on paid care is then

$$\text{Paid care expenditures}_t \equiv p \times \tau_{pt} \times q_{pt} \tag{4}$$

Mother and relative care hours entail no direct pecuniary cost, but mothers have preferences over the time they spend with the child (see equation 13). Mother care time is also associated with an opportunity cost that varies with her wage offer. I model the quality of mother and relative care as functions of observables and a random component:

$$\ln q_{mi} = X_i' \beta^m + \nu_i^m \tag{5}$$

$$\ln q_{ri} = X_i' \beta^r + \nu_i^r \tag{6}$$

Equations 5 - 6 specifiy that the menu of quality faced by the family is time-invariant. In principle, allowing the quality of each arrangement to stochasticallly evolve is straightforward but would increase computational expense. Beyond issues of tractability, it is unlikely that there is enough variation in the available data to identify and estimate such a model. Over the relatively short time horizon of early childhood, any evolution of the quality of care is not likely to be quantitatively important.

2.3 Skill Production

The quality of care is understood in relation to the production of child cognitive skills. I adopt an "exposure" concept of care services received, in which the quality of a given care arrangement enters a skill production function in proportion to the total time spent in that arrangement:

$$\ln \theta_{t+1} = A_t + \gamma_{1t} \ln \theta_t + \gamma_{2t} \sum_k \frac{\tau_{kt}}{\bar{\tau}} \ln q_{kt} + \eta_t \tag{7}$$

where η_t is an innovation realized after decisions are made in period t. The implicit assumption in the structure of equation 7 is that child care quality is perfectly substitutable across care types. This should be thought of as a statement about the definition of the concept "quality" as opposed to any claim about the technology of skill production. By this definition, the care quality of an arrangement is the thing that produces skills according to the fraction of time the child spends in that particular arrangement. Indeed, if the "quality" of two types of care arrangements were not perfectly substitutable, then they would not be measuring the same input. However, as discussed in Section 5, taking the model to the data requires putting restrictions on how measures in the data map into the model object "quality."

The key technological restrictions in 7 is that the elasticity of substitution between current skills and inputs is unity and that there is no dynamic complementarity. Ruling out dynamic complementarity is done for simplicity given that this study is focused on how constraints that mothers face shape the *inputs* into child skill production, as opposed to studying the dynamics of skill production as is the focus of Cunha and Heckman [2007], Cunha et al. [2010], Agostinelli and Wiswall [2020]. These papers use measures of the home environment as measures of care inputs but do not account for non-parental care providers. Other studies of skill development and child care make a distinction between parental and non-parental care but do not account for quality differences within types [Bernal, 2008, Del Boca et al., 2014, Kline and Walters, 2016, Caucutt and Lochner, 2020]. Closest to my treatment are Griffen [2019], Chaparro et al. [2020] who also model inputs as a time-weighted average of the quality of each care type to which the child is exposed. Relative to these two papers, the primary innovation of my model is that I explicitly model variation in a family's choice set and show that this is important for estimating the effect of untried policies.

2.4 Wage Offers and Exogenous Income

I model mother wage offers and non-mother-wage income as linear functions of observable characteristics and permanent, unobservable hetereogeneity:

$$\ln w_{it} = X_i'\beta^w + \iota d_{it} + \xi_i^w \tag{8}$$

$$\ln y_{it} = X_{it}'\beta^y + \xi_i^y \tag{9}$$

ξ_i^w, ξ_i^y are permanent, reflecting unobservable heterogeneity at the birth of the child, such as parent human capital. I do not allow for innovations to ξ_i^w, ξ_i^y, and so I rule out any possibility of mothers making labor supply decisions based on unobservable shocks to either their own wage offer or that of their partner. d_{it} is the history of the mother's labor supply decisions starting from the first period, capturing any human capital depreciation occurring during the period. The only other temporal variation in resources occurs through X_{it}, as I allow exogenous income to depend on the presence of a father, who may exogenously enter or exit the family.

2.5 Initial Conditions, Exogenous Transitions, and Discretization

In addition to exogenous care quality, relative availability, and initial values of the mother's wage offer and other income, initial conditions include initial child skill, modeled as

$$\ln \theta_{i0} = X'_{i0}\beta^{\theta} + \nu_i^{\theta} \tag{10}$$

Given the functional form assumptions 3, 8 - 9, and 10, the family's period-0 state can be written as a vector of observable characteristics and idiosyncratic shocks:

$$\Omega_0 = (\xi_i^w, \xi_i^y, \nu_i^m, \nu_i^r, \nu_i^{\theta}, \epsilon_i^{\phi}, \Pi_0, X_0) \tag{11}$$

I assume that the six shocks in 11 are each independent Normal random variables with variances of the first five to be estimated (the variance of ϵ_i^{ϕ} is normalized to 1). In order to cast the model in a discrete choice framework, I approximate the marginal distributions of $(\xi_i^w, \xi_i^y, \nu_i^m, \nu_i^r, \nu_i^{\theta})$ via uniform distributions with $n = 3$ points of support, following the method of Kennan [2006].[4] States X transition according to a first-order Markov process represented by the matrix Π_0.

2.6 Discrete Choice Set

At each state, mothers must choose labor supply h, child care time allocation $\{\tau_k\}$, and paid care quality $\ln q_p$. Budget constraints for family resources and mother time imply that the choice $(h, \tau_m, \tau_r, \tau_p, \ln q_p)$ determine mother leisure and family consumption. I discretize the choice set, allowing mothers three choices of labor supply and three choices each of paid care, relative care (if available), and paid care quality. Feasible values of labor supply are not-working, part-time work (30 hours per week), or full-time work (60 hours per week). Expressed as fractions of the total time that children must be cared for in a week, valid choices of non-parental care arrangements include 0, $\frac{1}{4}$, or $\frac{1}{2}$ of total required time. However, I do not allow mothers to completely outsource care time, so $\tau_r = \tau_p = \frac{1}{2}$ is not a feasible choice (and indeed, as seen below, mother preferences imply this is never an optimal choice). Care quality may be purchased at four levels, expressed in standard deviations: -$\frac{1}{2}$, 0, $\frac{1}{2}$, or 1. The units of paid

[4] Kennan [2006] shows the counterintuitive result that the best discrete approximation of a continous distribution is uniform. Each gridpoint occurs with equal probability; the variance of the distribution is reflected in the location of the support points.

quality are anchored to a standardized measure in the data, discussed below. So families in the model may purchase paid care quality that is of the average value of all (paid and informal) non-parental care observed in the data, one-half standard deviations above or below average, or one full standard deviation over average.[5] Care quality is supplied via an affine market supply curve.

To fix notation, let $\mathcal{D}_t(\phi)$ denote the feasible choice set at time t, conditional on relative availability ϕ, and let D_{tj} denote the j^{th} member of that set.

2.7 Preferences and Unobserved Hetereogeneity

The mother is an expected utility maximizer with per-period preferences over consumption c, leisure ℓ, labor supply h, time spent caring for the child τ_m, and the aggregate quality of non-parental care received by the child, Q_n. The payoff enjoyed from choosing choice D_{tj} is the sum of a utility function $\tilde{u}$ describing preferences over consumption, leisure, care time, and non-parent care and an idiosyncratic, additive choice-specific payoff, ϵ_D:

$$u(D_{tj}; X, \zeta) = \tilde{u}(c, \ell, h, Q_n; X, \zeta) + \epsilon_D(D_{tj}) \tag{12}$$

The deterministic component is the sum of five terms:

$$\begin{aligned}\tilde{u}(c, \ell, h, Q_n; X, \zeta) = &\ln c + \Psi_\ell(X, \zeta) \times \ln \ell + \Psi_m(X, \zeta) \times \ln \tau_m &(13)\\ &+ \Psi_q(X, \zeta) \times Q_n + \Psi_p(X, \zeta) \times (\tau_p > 0) &(14)\end{aligned}$$

$\Psi_x > 0$ are taste shifters that are functions of observables X and a vector of permanent unobservable components, ζ. Marginal utility shifters take the functional form of

$$\Psi_x = \exp(Z_x' \psi_x) \tag{15}$$

where Z_x may include observables X and unobservables.[6] Recall that the model is written relative to the birth of a particular child, and so mother leisure includes *all* uses of mother time that are neither child care for *that particular child* or market work. This includes home

[5] I rule out the possibility of mothers buying care quality of -1, as the model lacks sufficient detail with respect to modeling credit constraints to rationalize such low quality and still match average values in the data.

[6] See Appendix A for particular functional forms

production, caring for other children or family members, and sleep. To capture the fact that co-habitating mothers and single mothers likely receive different levels of assistance with home production and child care, the taste shifter on leisure may vary with the presence of a spouse.

I choose to model preferences over non-parental child care quality instead of child skills for two reasons: one practical and one economic. The former is that by choosing not to condition future value functions on realizations of child skills, I avoid introducing a continuous state variable into a discrete choice problem. More importantly, by modeling preferences over non-parental quality instead of skills, I avoid making any assumptions about mother beliefs over the skill production process.[7] However, the cost of robustness to incorrect beliefs is that I rule out mother responses to skill shocks. That is, a mother cannot observe a bad shock over ages 0-2 and then endogenously compensate or reinforce that shock with child care choices at ages 3-4. I view this restriction as relatively innocuous, as there is little evidence that mothers engage in such behavior [**?**]. I also omit mother care quality from preferences (although it is an input into child skill development). This is done for simplicity and the fact that without writing the utility function as a function of skill production outputs, preferences over the mother input are not separately identified from correlation between preferences and mother quality.

ζ_i is a vector of idiosyncratic taste shocks, fixed throughout the model duration. I introduce correlation in the components of ζ_i by allowing multiple dimensions of ζ_i to enter the same utility shifter. I approximate each marginal distribution as a uniform distribution with three points of support and estimate the variance. $\epsilon_D(D_{tj})$ is an idiosyncratic payoff enjoyed if the choice $D_{tj} = (h, \{\tau_k\}, q_p)'_j$ is made. I assume ϵ_D is distributed according to a Type I Extreme Value distribution [McFadden, 1974, Rust, 1987]. The model ends on the child's fifth birthday, and terminal states (mother work history) are valued according to the parametric function $V_T(\Omega_T)$

$$V_T(\Omega_T) = \exp(X_T'\psi_T) \times d_T \tag{16}$$

leading to the period-0 problem:

$$V_0(\Omega_0) = \max_D E_D\left[\sum_{t=0}^{1} \beta^t u_t(\cdot, D_t)|\Omega_0\right] \tag{17}$$

where D_t is the optimal choice at time t. Every period, the following constraints must hold:

[7] See Cunha et al. [2013] for evidence that mothers do not have correct beliefs about the skill production process.

$$c_t + p\tau_{pt}q_{pt} = w_t h_t + y_t + g(w_t h_t, y_t) \quad (18)$$

$$\sum \tau_{kt} = \bar{\tau} \quad (19)$$

$$\tau_{mt}\bar{\tau} + \ell_t + h_t = 1 \quad (20)$$

$$\tau_{kt} \geq 0, \tau_{rt} = 0 \text{ if } \phi = 0 \quad (21)$$

$$\tau_{mt} > 0 \quad (22)$$

The government taxes income and distributes transfers according to g. The consumption floor is set so that that no family's after-tax income falls below 13.5% of median household earnings.[8] As households in my model make up only a fraction of the total tax base, I do not require that tax revenues from these families equal transfer payments. This reflects the substantial redistribution across households that occurs in the United States.

The expectation is taken with respect to transitions between observable states and realizations of choice shocks ϵ_D. Notice that the mother's time constraint 20 is normalized to 1, which represents 168 hours in a week. Child care allocation choices, however, are expressed in fractions of the total required care time $\bar{\tau}$. Therefore, the fraction of the *mother's* time spent caring for the child is given by $\tau_{mt}\bar{\tau}$, where $\bar{\tau} \leq 1$. In practice, I assume children must be cared for 91 hours per week (I do not count child sleep time as time requiring care), so $\bar{\tau} = \frac{91}{168} \approx 0.54$. Therefore, a mother who provides 100% of the required care time ($\tau_{mt} = 1$) spends $\tau_{mt} \times \bar{\tau} = 1 \times \bar{\tau} \approx 0.54$ of her *total* time in child care for the particular child whose childhood is studied in the model.

2.8 Instruments

An additional component of the data-generating process is a vector of instruments Z_i. By assumption, Z_i are uncorrelated with all unobservables in the model and only affect decisions through their effect on state variables relevant to the mother's decision problem. Thereore, Z_i may be easily incorporated into model simulations by considering the projections

$$z_i = X_i'\beta + \epsilon_i \quad (23)$$

as part of the data-generating process, assuming Normality of ϵ_i to allow for simulation.

[8] This follows Ashman and Neumuller [2020].

2.9 Model Solution

We may define the conditional value function associated with choice j at time t:

$$v_t(D_{tj}; X_t, \zeta_i) = u(D_{tj}; X_t, \zeta) - \epsilon_D(D_{tj}) + \beta \int V_{t+1}(X_{t+1}, \epsilon_D(D_{tj}))dF(X_{t+1}|X_t, D_{tj})dG(\epsilon_D) \tag{24}$$

where the integral is taken over future realizations of choice shocks ϵ_D and other stochastic processes such as family transitions. The distributional assumption on ϵ_D generates conditional choice probabilities of the familiar form

$$P(D_t = D_{tj}|X_t, \zeta) = \frac{\exp(v_t(D_{tj}; X_t, \zeta))}{1 + \sum_k \exp(v_t(D_{tk}; X_t, \zeta))} \tag{25}$$

I solve the model via backward induction, computing terminal period conditional value functions V_T and then integrating over state variable transitions in periods $t < T$. The resulting choice probabilities for a given parameter vector form the basis of my simulation-based estimator.

3 Data

I use the structure of the model to combine information from three data sources: the Panel Study of Income Dynamics (PSID), the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), and the American Time Use Survey (ATUS).

The primary data source that I use is the Panel Study of Income Dynamics (PSID) core family files and supplemental instrument, the Child Development Supplement (CDS). These data allow me to observe the joint distribution of child skills, care arrangements used, and family and grandparent characteristics before and after birth. For moments of the distribution of care quality of different arrangements, I use the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B). I also use the American Time Use Survey (ATUS) for additional information on the time allocation of mothers. In this section, I discuss each data source in turn and then highlight crucial measurement concepts and variable definitions.

3.1 PSID

Studying the extended family requires a dataset that can link family members across multiple generations. The PSID is ideal, as it is a long-running longitudinal survey of households that follows the children of PSID respondents (and thus the children of those children, etc), allowing us to view three generations of a family. Beginning in 1968 with 5,000 nationally-representative US families, the PSID tracks education, labor supply, income, wealth, health, and family structure. The main study provides demographic information every year between 1968 and 1996 and every two years between 1997 and 2017. The main study also includes a wealth supplement that collects extensive information on household balance sheets every five years between 1984 and 1999, and every two years thereafter.

In addition to the main study family-level files, I draw upon the Child Development Supplement (CDS). The original CDS study was initiated in 1997 (CDS-I), documenting detailed measures of child cognitive and non-cognitive skills, health, and relationship with parents for 3,501 children. The second and third waves (CDS-II and CDS-III) of the original study aimed to follow up on the CDS-I children as they aged. Carried out in 2002-2003 and 2007, respectively, CDS-II included 2,907 respondent children and CDS-III re-interviewed 1,608. As the original cohort aged out of childhood, the CDS began following a new cohort of 4,311 children in 2014. The CDS also includes a time diary, measures of expenditures on children (including education), and an extensive questionnaire on current and past child care arrangements. A limitation of the CDS is that the sampling interval is 5 years. Therefore, I observe a particular child between ages 0 and 7 at most twice. This is particularly restrictive when studying the formation of skills during the first few years of life.

Crucially, the PSID provides a parent identification file that allows one to link parents and children (and thus grandparents, parents, and children) across all studies. This allows me to observe the correlation between the care arrangements that a child experiences and whether the child's grandmother died shortly before birth, forming the basis for my identification strategy.

3.2 ECLS-B

As the PSID-CDS lacks a useful panel dimension during early childhood, I use the ECLS-B as the main source of information on the joint distribution of child skills and care quality. The ECLS-B is a longitudinal study following a nationally-representative cohort of roughly 14,000 children born in 2001. The first survey wave was carried out when the youngest children were 9 months old, and follow-up waves then collected information when these children were 2 years old, 4 years old, and at entry into kindergarten. For my purposes, the most valuable components in this study are child skill measures and observer-based assessments of the quality of non-parental care arrangements.

3.3 ATUS

Although the PSID-CDS contains information on mother labor supply, I use the ATUS for more detailed information on mother time and a much larger sample size. The ATUS is an ongoing, nationally-representative time diary, in which one respondent from each sample household is selected to complete a time diary, documenting their activities during a 24-hour period. Many existing papers use the ATUS to study mother caregiving, which I do not do here. The reason is that my model is child-centric, dealing with care time allocated to a particular child. The ATUS, on the other hand, collects information about the time use of mothers (as well as fathers) and thus can only measure caregiving allocating to *all* children in the family.

3.4 Child Skills

Child skills are measured in both the ECLS-B and the PSID-CDS, although I primarily use those assessments available in the former, as the latter lacks a useful panel dimension. The ECLS-B contains several measures of child skills, although the measures differ across ages. Therefore, I construct standardized Z-scores within survey wave and year of age for each measure and take a child's average score as child skills.[9] Therefore, child skills at any age are expressed in units of the standard deviation of child skills at that age.

3.5 Care Arrangements

Although each of these three datasets contains information on caregiving, I use retrospective histories in the PSID-CDS primary caregiver interview as my main source of information on the care arrangements used for children of different demographic cells. The first two waves of the CDS, carried out in 1997 and 2002-2003, included a battery of questions designed to retroactively elicit how a child was cared for before entering kindergarten. These retrospective reports are only available in CDS I-II.[10] Pre-K care in the 2014 CDS wave is measured as

[9] An alternative way to reduce the dimension of multiple measures is to explicitly model the measurement system [Cunha et al., 2010, Agostinelli and Wiswall, 2020]. However, doing so requires sufficient measures at each age (generally 3, with a linear measurement system) or sufficient measures at a single age paired with an assumption of age-invariance of the measure [Agostinelli and Wiswall, 2020]. The ECLS-B lacks sufficient measures and so this route is infeasible.

[10] The youngest child at the time of the CDS-II survey was 5 years old and so no pre-K care was recorded between CDS II and CDS III.

regular child care arrangements used in the last four weeks and reports of current child care arrangements in CDS I-III are limited to children of school age. Therefore, the simplest way to construct a full history of how a given respondent child was cared for during the first five years of life is to use the histories in CDS I-II.

Non-parental care arrangements in the CDS consist of relative care (in the child's home or the relative's home), daycare, Head Start, before/after school programs, and non-relative care in the child's home or the non-relative's home.[11] I aggregate these categories into three non-parental arrangements: relative care, center care (daycare, Head Start, before/after school), and home-based care (non-relative in the child's home or in the non-relative's home).[12] Tables 1-2 display the extensive and intensive margins, respectively, of child care by age. Table 1 shows the frequency with which children experienced each care arrangement and, conditional upon using any non-parental care, the number of arrangements used. We see that average relative care is declining (at the extensive margin) as the child ages, while center care increases until children reach school age. Usage of home-based care is non-monotone but broadly decreasing with child age. An important point is that by age 6, almost 88% of children use no care arrangement other than their parent. This reflects the fact that kindergarten absorbs many weekly care hours starting at that age. Table 2 presents average hours in each non-parental arrangement, both unconditional and conditional upon use.

A key motivating idea of this paper is that child care choices vary by socio-economic status (SES). Figure 2 plots average (unconditional) hours in each of the three non-parental care categories by child age and mother education. The fourth plot shows the frequency of a child receiving no non-parental care. We see stark differences in hours[13] spent in each non-parental care type, but these gradients mostly shrink to zero by ages 5-6, when children start entering kindergarten. At all ages, children of less-educated mothers are more likely to receive no non-parental care, and the use of paid care is broadly monotonic in mother education. It is noteworthy that the use of relative care is non-monotone in mother education. Children of the highest- and lowest-educated mothers receive the least relative care. Due to time constraints, maternal labor supply must be correlated with non-parental care usage. Figures 3 - 4 show the same moments computed while conditioning on mother labor force participation. Indeed we see that non-parental care hours are much lower for children of non-working moms at all ages and levels of maternal education. However, the education gradients do not disappear when

[11]Other arrangements include the child looking after him or herself or "other." Neither category is relevant for pre-K children.

[12]I combine center and home-based care when estimating the model; both are paid child care.

[13]Unconditional hours muddies the distinction between the intensive and extensive margins, but it is the object that most naturally maps into the mother's choice set in the model. Further, hours received (whether zero or positive) in a care arrangement are likely what matters for child development.

conditioning on labor supply. Interestingly, comparing Figure 2 and Figure 4 shows that relative care is an important source of care hours for children of working mothers with less than a high school diploma. The lower unconditional average is driven by weak labor force attachment of this group. These facts are consistent with relative care being an important source of child care for working mothers at all education levels.

3.6 Maternal Care Quality

As a measure of the quality of maternal care time, I use self-reports of the frequency with which the mother reads to the child, as reported in the ECLS-B. This commonly-used measure of mother quality is a component of the Home Observation for Measurement of the Environment (HOME) Inventory [Caldwell and Bradley, 1984], an instrument designed to measure the quality of the environment in which a child develops.[14] To capture the notion of quality per unit of maternal time, I divide the reported frequency of reading per week by the number of hours per week the child spends with his or her mother. I then standardize this reading-per-hour score by child age and take the resulting Z-score as a measure of maternal quality.

3.7 Non-parental Care Quality

Measures of non-maternal (either paid or relative care) are also taken from the ECLS-B. In particular, I use the Arnett Scale of Caregiver Behavior. The Arnett score is an observer-based measure of the quality of interactions between a non-parental caregiver and a child. It is designed to be comparable across different types of non-parental care, and so I can compare the quality of non-parental care time received in a formal care setting with that from a relative caregiver. As with the frequency of mother reading per hour, I standardize the Arnett score by child age.

4 Econometric Issues

Estimating the model of Section 2 is complicated by several different flavors of missing data. First, as is clear from the fact I use multiple data sources, no single dataset contains the

[14]The PSID-CDS contains a superset of the HOME items reported in the ECLS-B, as well as detailed child time diaries that may be used to construct measure of parental caregiving quality. However, the PSID-CDS provides little information on the quality of non-parental time, unlike the ECLS-B. As the joint distribution of mother and non-parent quality is a key object in my model, I rely only on the measures of mother time quality in the ECLS-B. However, maternal education gradients in measures of mother time quality in the PSID-CDS match those constructed from the ECLS-B.

joint distribution of child skills, care arrangements, care quality, mother labor supply, and characteristics (crucially, mortality) of a child's grandparents. I use the structure of the model to integrate over the gaps in any given dataset, exploiting observables common to each dataset. Second, the familiar selection-on-unobservables problem in the choice to work and to use relative care bite, as I do not see relative quality or mother wage offers for families in which no relative care is used and the mother does not work, respectively. Without an instrument for either, I rely on the selection model implicit in my model of mother decision-making to recover the latent distribution of wage offers and relative quality.

The third, and most fundamental, issue of missing data is that I do not observe relative availability. That is, when a child in the data is not cared for by a relative, I do not know if this was because no informal care was offered or if it was offered and the mother chose not to use it. As shown in Section 8, data on the choices that different families make is insufficient to separately identify relative availability and preferences. Towards a resolution to this under-identification problem, I exploit the genealogical design of the PSID. In particular, I study how the death of a PSID grandmother shortly before birth shifts care arrangements and mother labor supply choices. The operative assumption is that the death of a grandmother reduces the likelihood of being offered informal care, as the grandmother who died might have made such an offer. Therefore, the choices that families make after the death of a grandmother contain information about substitution patterns between different care types, as some families face a restricted choice set induced by grandmother mortality. The intuition is that families who would use grandmother care if offered are made up of two types characterized by the choice they would make if grandmother care is not available: those who would use mother care only and those who would purchase formal care. The composition of these two groups is of first-order importance for child care policy, and yet multiple mixtures are consistent with data on observed choices.

5 Identification and Estimation

I estimate the model of Section 2 in two steps. First, I estimate exogenous transitions directly from the data and calibrate the formal care quality supply curve based on the joint distribution of quality measures and cost per hour observed. I then estimate the remaining parameters using the method of simulated moments (SMM) [Gourieroux et al., 1993]. Although it is general preferable to use the Maximum Likelihood Estimator (MLE) when estimating discrete choice problems [Eisenhauer et al., 2015], the fact that I use information from three different datasets complicates forming a likelihood function. In this section, I discuss identification and estimation at each step. In the results presented below, I simplify the degree of preference heterogeneity that I allow. I restrict the preference shock parameter vector ζ to be a scalar that loads into the taste shifter for using paid care, $\Psi_p(X, \zeta)$. Allowing for more flexible preference heterogeneity is the subject of future work.

5.1 Calibrated Parameters

I set the parameters of the paid quality supply curve to match the average hourly cost of non-parental care paid for two different levels of non-parental quality, measured as standardized Arnett scores taken from the ECLS-B. Table 3 reports the slopes and intercepts of these two supply curves, and we see that an hour of average quality care during ages 0-2 costs \$3.42 an hour, with every additional standard deviation costing \$6.66 per hour. The analogous numbers for ages 3-4 are \$3.38 and \$9.3. Mothers may purchase at most 40.5 hours per week of care, and so at 48 weeks of available care per year, these supply curves translate into maximum expenditures of \$6,648 and \$6,570 for average quality care at each age. Maximum expenditures for care of $\frac{1}{2}$ standard deviation of quality at each age are \$13,122 and \$15,610, reflecting an increasing marginal cost of quality as children age.

I calibrate the tax schedule following Chatterjee and Eyigungor [2015], modeling five tax brackets over pre-tax income per resident parent. Therefore, total taxable income for a household with a resident father is half that of a single mother household with identical total income. An average tax rate is applied to total taxable income within each bracket. Table 4 lists tax brackets and average tax rates, with taxable income presented in as fractions of median household income.

5.2 Family Transitions

I assume that transitions in family structure are exogenous conditional on current family state and mother education. Letting $\breve{X}_{it}$ denote the subset of observables that characterize family structure, the exogeneity assumption is that

$$P(\breve{X}_{it+1}|\Omega_{it}, \epsilon_D) = P(\breve{X}_{it+1}|\breve{X}_{it}, \text{maternal education}) \tag{26}$$

The right-hand side of 26 is non-parametrically identified from the distribution of $\breve{X}_{it+1}$ conditional on $\breve{X}_{it}$ and mother education, and so a cell estimator consistently estimates family transition probabilities. The estimated transition process is given in Table 5 with each column

j presenting the probability of moving to state j at time $t+1$ conditional on mother education and being in state i at time t.

5.3 Skill Production

Identifying the parameters of the skill production function 7 requires restrictions on the measurement of care inputs and outputs. I assume that the averaged age-standardized skill measures in the ECLS-B perfectly measure child skills. As discussed in Section 3, I have access to several instruments that aim to measure the quality of care provided by mothers, formal caregivers, and relatives. The quality of formal and informal non-parental caregivers are measured using the same measure, but mother caregiving is not assessed according to that same instrument. Let Z_m denote the available measure of mother care and Z_{nx} a measure of non-parental care in care arrangement type $x \in \{p, r\}$, either paid care or that provided by a relative. I treat the model object $\ln q_k$ as measured without error in the same units of Z_{nx}, so that Z_{np} and Z_{nr} are perfectly substitutable. To transform units of the mother quality measure Z_m into those of $\ln q_m$, I assume that Z_m differs from Z_{nx} in scale and but not location. So we can write $Z_m = \lambda \ln q_m$.

Let $\tilde{Q}_{mt} = \tau_{mt} Z_{mt}$ and $\tilde{Q}_{nt} = \tau_{pt} Z_{npt} + \tau_{rt} Z_{nrt}$ denote the total effective quality units of mother and non-parental care, respectively, that the child enjoys at time t. Normalizing the total care time requirement $\bar{\tau}$ to 1 and substituting into the production function yields a linear equation characterizing the relationship between the structural parameters, skill outcomes, and measured inputs:

$$\ln \theta_{t+1} = \tilde{\beta}_{0t} + \tilde{\beta}_{1t} \ln \theta_t + \tilde{\beta}_{2t} \tilde{Q}_{mt} + \tilde{\beta}_{3t} \tilde{Q}_{nt} + \tilde{\eta}_t \tag{27}$$

where

$$\tilde{\beta}_{0t} = A_t \tag{28}$$

$$\tilde{\beta}_{1t} = \gamma_{1t} \tag{29}$$

$$\tilde{\beta}_{2t} = \frac{\gamma_{2t}}{\lambda} \tag{30}$$

$$\tilde{\beta}_{3t} = \gamma_{2t} \tag{31}$$

I take skill production to be exogenous, conditional upon inputs, and thus the "reduced form" parameter vector $\tilde{\beta}$ is recovered from a projection of $t+1$ skills onto the right-hand side inputs in 27. I view this simplification as an upper bound on the productivity of care quality inputs, as it is likely that any correlation between the skill shock η_t and endogenously-chosen care quality is positive. Clearly this endogeneity issue is concerning if our primary goal is to estimate the technology of skill production, but here I am primarily interested in studying how our conclusions regarding policy change when we take into account variation in access to informal care. To the extent that my estimates of the elasticity of skills with respect to care quality inputs are upwardly-biased, the estimated effects of policy on skill inequality should also be treated as upper bounds. From the standpoint of a risk-averse planner who considers whether to implement a costly large-scale program with uncertain benefits but certain costs, producing an upper bound on the effect of policy can be considered a conservative estimate.

The ECLS-B contains observations from the joint distribution of current skills, parental (almost always mother) time and quality, and the time spent in each non-parental arrangement. However, not all non-parental arrangements used are measured with the Arnett score of caregiver quality. Only the primary non-parental arrangement is selected for quality evaluation, and only those primary arrangements used for at least 10 hours per week are measured. Therefore, time spend with a secondary non-parental care is an omitted variable. In practice, few children spend significant time in more than one non-parental care arrangements at any given age.

The right-hand-sides of equations 28 - 31 are functions of the data, and the aggregate measured quality $\tilde{Q}_{xt}$ is observable. By assuming that η_t is exogenous, the OLS estimand recovers $\tilde{\beta}$ and then we have λ as well as the structural parameters. In practice I estimate 27 on pooled samples, pooling ages 0-2 and 3-4 to match the period frequency of the model. Table 6 displays estimates of the production process and the mother measure scaling λ. As is consistent with other studies, skills become more persistent as the child ages: the self-productivity of skills at ages 3-4 is twice that of its value ages 0-2.

5.4 Care Quality

With the mapping between measures of mother quality in the data to quality in the model in hand, the fact that all mothers provide some care to their children allows me to identify the parameters of the mother quality equation 5 straight from the data. Put another way, participation in maternal caregiving is 100%, and so there is no selection problem on the extensive margin. As I observe the entire distribution of mother quality measures and know the scaling between measured mother quality Z_m and $\ln q$, I can compute the maternal education gradient by simply plugging λZ_m into equation 5:

$$Z_{mi} = X_i' \tilde{\beta}^m + \tilde{\nu}_i^m \tag{32}$$

where

$$\tilde{\beta}^m = \frac{\beta^m}{\lambda} \tag{33}$$

and the variance of the structural quality shock ν_i^m equals $\frac{1}{\lambda^2}\sigma_{\tilde{\nu}}^2$. Given that I assume ν is exogenous, the parameters of 32 may be consistently estimated via OLS and we can then recover β^m, σ_ν given our estimate of λ.

5.5 Preferences and Relative Availability: Death of a Grandmother

Given the assumption of additively-separable log utility, preference shifters are identified by mother time allocation, conditional on state variables. The extensive margin of mother labor supply (and how it changes between income groups) identify the terminal value of mother experience, as the accumulation of experience is assumed to be independent of the number of hours worked. By assumption, instruments for relative availability are exogenous, and so the parameters of the Probit characterizing the probability of grandmother death may be estimated with MLE using the genealogical dimension of the PSID.

5.6 SMM Algorithm

Conditional on calibrated parameters and those estimated outside of the model, I form a simulation estimator and search over a vector of structural parameters to match moments in the data [Gourieroux et al., 1993]. With distributional assumptions on all stochastic terms, I can solve the model for a given parameter vector and simulate a synthetic panel of individuals. The estimator is defined as

$$\hat{\Upsilon} \equiv \text{argmin}_{\Upsilon}(\hat{\beta}(\Upsilon) - \hat{\beta}(\text{Data}))' W(\text{Data})(\hat{\beta}(\Upsilon) - \hat{\beta}(\text{Data})) \tag{34}$$

where $\hat{\beta}(\Upsilon)$ is a set of moments implied by a given parameter vector Υ and $\hat{\beta}(\text{Data})$ is the set of analogous moments computed from the PSID data. $W(\text{Data})$ is some positive-definite weighting matrix with potential dependence on the data. I use a diagonal matrix whose elements are the inverse of the squared variances of $\hat{\beta}$(Data). The choice of moments is a degree of freedom left to the analyst, and in this case I use a set of OLS regression coefficients of choices on states and outcomes on choices and states as well as as well as differences in the conditional means laborsupply and care choices conditioning on the event of grandmother death. The estimation algorithm proceeds as follows:

0. Estimate auxiliary models on data: $\hat{\beta}(\text{Data})$

1. Guess structural parameter vector Υ and solve discrete choice problem via backward induction, generating conditional choice probabilities

2. Draw N simulated families from distribution of initial conditions

3. For each $n \in N$, draw S paths of shocks from stochastic components of model, including discrete choice shocks, simulating choices and state transitions for each shock path s

4. For each simulated panel, estimate auxiliary models on simulated data: $\hat{\beta}_s(\Upsilon)$

5. Form $\hat{\beta}(\Upsilon) = \frac{1}{S}\sum_s \hat{\beta}_s(\Upsilon)$ and compute the objective function

$$(\hat{\beta}(\Upsilon) - \hat{\beta}(\text{Data}))' W(\text{Data})(\hat{\beta}(\Upsilon) - \hat{\beta}(\text{Data})) \tag{35}$$

I iterate over steps 1 - 5 according to the "TikTak" procedure described in Arnoud et al. [2019]. This is a global optimzation routine that begins by pre-testing N points $\Upsilon_1, ...\Upsilon_N$ in a $\#(\Upsilon)$-dimensional hypercube defined by a pseudo-random Sobol sequence. The TikTak algorithm then selects the $M < N$ seed points $s_1, ..., s_M$ with the lowest objective value. Without loss of generality, label these M points in ascending order according to their image in 35, $f(s_1) \leq f(s_2) \leq ... \leq f(s_M)$. The next step of the algorithm is to carry out M local search routines where the starting point for the first local search is s_1 and that for the m^{th} search is given by

$$\tilde{s}_m = (1 - \theta_m)s_{m+1} + \theta_m s_m^* \tag{36}$$

where s_m^* is the best minimizer for all searches $1...m$. θ_m is a mixing weight that approaches 1 as $m \to M$. In practice I use the Nelder-Mead simplex algorithm as the local optimizer.

6 The Importance of Informal Care

Skill inequality is partly determined by variation in the circumstances of birth. Such variation takes the form of initial skills, family resources, mother preferences, and caregiver quality. An additional source of endowment inequaliy is variation in the *availability* of informal care. That is, in some families, relatives or other informal caregivers offer to provide child care. From the mother's perspective, the availability of informal care is unambiguously good, as it simply represents an expansion of the choice set. However, insofar as mother have preferences over anything other than child skills, it is not obvious that the availability of informal care is beneficial for any given child's development. In this section, I use the model to estimate the effect of having access to informal care on mother choices and child skill outcomes. That is, if we could manipulate whether or not a relative offered to provide care, would a child be better or worse off in terms of skill development and family resources?

To do so, I first simulate a panel of individuals under the estimated data-generating process $\widehat{DGP}$. Each individual is characterized by initial conditions and a series of shocks, with conditional choice probabilities given by $\widehat{DGP}$. Let $P_i = (\{X_{it}\}, \{\epsilon_{it}\}, \hat{\sigma}_i)$ denote the "paths" of such a simulated individual. For each simulated individual, I compute two counterfactual paths $P_{i,1}, P_{i,0}$ in which I hold constant individual i's sequence of shocks ϵ_i and initial conditions, but I compute $P_{i,1}$ under the assumption that informal care is available for all families and $P_{i,0}$ under the assumption that relative care is never available. Therefore, any differences between $P_{i,1}$ and $P_{i,0}$ stem *only* from the total effect of the manipulated choice set, taking into account a mother's endogenous response to a change in her menu of care options. For any outcome X_{it}, the treatment effect of having a relative available is simply given by the difference in that value between the two paths:

$$\Delta X_{it} \equiv X_{it}(P_{i,1}) - X_{it}(P_{i,0}) \tag{37}$$

Outcomes of interest are terminal child skills as well as changes in mother labor supply and expenditures on paid care, which are informative about how the option of relative care shifts family resources during early childhood. Table 13 displays averages of 37 for several different outcomes. For each outcome, columns 1 and 2 show conditional average values for children whose mothers have no college education and those with higher-educated mothers, respectively, under a scenario in which no family has access to relatives. Columns 3 and 4 show the respective differences in these averages under a scenario in which all families have access to relatives. Columns 5 and 6 display this effect of relative access as a percentage of the no relative figures.

I estimate that access to informal care increases the skills of more-advantaged children and reduces those of the less-advantaged, although the magnitudes of these changes are small. In particular, Column 5 shows that the reduction in skills for children of lower-educated mothers is 0.7%, or 0.001 standard deviations, of the average value for this group if no relative care were available. The positive effect of relative access for children of higher-educated mothers is 2% of the no-relative value. These near-zero changes in skill reflect the substitution patterns of each group when relative care is not a member of the choice set. We see from Table 13 that access to relative care causes substantial changes in mother labor supply and care choices: having relative care as an option increases lower-educated mother labor force participation by 18.6% and higher-educated participation by 23%. The corresponding percentage-point increases (4.5% and 9.1%, respectively) are in line with other studies of the relationship between mother labor supply and relative-provided child care: Compton and Pollak [2014] estimate a 4-10 percentage-point increase for married women living near in-laws, Dimova and Wolff [2011] find that grandparent care increases labor force participation by 11.5 percentage points, and García-Morán and Kuehn [2017] estimate a 6 percentage point increase. Comparing these quantities across studies is not entirely straightforward as the counterfactual in each study is somewhat different, but the magnitudes are broadly similar and support the conclusion that access to informal care is a quantitatively large shifter of mother labor supply.

The remaining outcomes in Table 13 illustrate how mothers alter care decisions when given access to informal care. As is mechanical given the finite number of hours in a day, informal care reduces the time that children spend with mothers and in paid care. Lower-educated mothers reduce time with children by 10.8 percentage points (roughly 9 hours per week) and higher-educated mothers respond similarly. Time in paid care falls by 1.5 percentage points and 2.8 percentage points, respectively, representing a 24% decrease for lower-educated mothers and a 30% decrease for higher-educated mothers. These numbers show that mothers primarily trade own-care time for relative time, and substitute much less between formal and informal care.

Finally, the total sum of non-parental quality inputs into child skill production falls for less-advantaged children and increases for more-advantaged children when relatives are available. As non-parental quality inputs are a function of both time in and quality of non-parental arrangements, these numbers (6 percentage point decrease lower-educated families and 7.2 percentage point increase for higher-educated families) reflect how mothers alter the total amount of non-parental care used as well as the difference in quality between the relative care hours that replace paid care when offered informal care. The fact that total time with mothers falls more than total paid care time implies that total non-parental care time *increases* when offered informal care. Therefore, it must be the case that lower-educated mothers are willing to take up low quality relative care, while higher-educated mothers select into relative care of higher quality than that purchased in the market when no informal care is available. These different selection patterns reflect education differences in endowments of relative care quality, preferences for non-parental care quality, and the marginal valuation of consumption.

7 Child Care Policy

The efficacy of policies aimed at reducing skill inequality primarily depends on the joint distribution of treatment effects and take-up propensities in the population. As seen in Table 13, access to relative care shifts the care and labor supply choices that mothers make. To study how informal care interacts with the effects of child care policy, I simulate the effects of three types of policies: a subsidy for formal care, a direct cash transfer, and public provision of a relatively high-quality center-based care option. Each simulated policy effect involves recomputing choices and state transitions under a baseline scenario and a counterfactual scenario in which the relevant policy object is manipulated. Shocks to choices and states are held constant, and I compare average outcomes between the baseline and the counterfactual for children of low- and high-educated mothers. Crucially, I compute average policy effects for the population as well as for the subgroups with and without access to informal care.

7.1 Formal Care Subsidy

Table 14 displays estimates of the effect of a 50% formal care subsidy. That is, the hourly cost of a quality unit of formal care is reduced by 50% for any amount of care and quality chosen. We see that this subsidy reduces total time with mothers by 1 and 0.7 percentage points for each respective education group and increases paid care time by 1.1 and 0.8 percentage points, respectively. As most of a child's time is spent with the mother, these absolute numbers reflect a 19.5% increase in paid care time in lower-educated families and a 9.5% increase in higher-educated families (roughly 1 hour and 0.75 hours, respectively). Expenditure on formal care drops by 30 and 40% ($370 and $1,230) for each education group. Families in which relative care is available are slightly more responsive with respect to paid care time (1.3 percentage point increase for both education levels), but the effect on skills is smaller than in families with no offer of informal care. In fact, skills for children of higher-educated mothers fall by 1.1%.

7.2 Direct Cash Transfer

An alternative to subsidizing formal care is to simply transfer money to families. In my model, this is unambiguously good for mothers, as they are free to decide how to allocate the fungible resources. However, the effect on children is not certain. Table 15 present estimates of the effect of a \$1,700 yearly transfer to families and shows very little response of skills and all choices but for mother labor force participation, which falls by 1.8% for lower-educated mothers and 1% for their higher-educated counterparts. This illustrates that 1-2% of mothers are very weakly attached to the labor force and are very responsive to income effects. Among the lower-educated group, the most responsive mothers have no access to informal care, representing a population that works primarily to finance formal child care. In contrast, labor supply responsiveness for higher-educated mothers varies less (0.9pp vs 1.3pp) by the availability of relatives.

In the pooled population, neither higher- nor lower-educated mothers increase expenditures on average in response to the direct transfer. However, the null effect for lower-educated mothers masks a 5.5% increase in paid expenditures for families with relatives available and a 1.6% decrease in expenditures for those without. The increase in paid care expenditure for families with relatives available is explained by the fraction of that group with very high marginal utility of wealth and poor-quality relatives. A transfer relaxes their budget constraint enough for them to substitute away from poor quality relatives and into paid care (an increase of 10.6%). Those without relatives, on the other hand, respond to the income effect of a transfer by reducing working hours and expenditures on paid care.

The null effect on child skills is worth discussing, particularly given evidence from expansion in the Earned Income Tax Credit (EITC) that child skills respond to an increase in income [Dahl and Lochner, 2012].[15] The first relevant distinction between my exercise and the EITC evidence is that I simulate a one-shot transfer and not an increase in permanent income.[16] In fact, the magnitudes that I find are in line with the effect of income on child skills estimated cross-sectionally in Dahl and Lochner [2012]. Second, the only mechanism through which I allow dollars to increase scores is through child care. In reality, there are more channels through which resources alter child outcomes. Finally, as I only simulate a single period, I likely understate the effect of transfers on child skills, as there is no scope for continued self-productivity of skills.

[15] That paper estimated a 6% standard deviation increase in test scores in response to a \$1,000 increase in income. A revised manuscript updated that effect to 3.8% of a standard deviation.

[16] The degree to which an expansion in the EITC reflects changes in permanent income is a function of how long families expect to remain eligible. But it is reasonable to think that such an expansion in public benefits is much more akin to a shift in permanent income than is a single transfer.

7.3 Publicly-Provided Care

Finally, I simulate the effect of a publicly-provided care option whose quality is equal to one-half of a standard deviation of non-parental quality. This policy has a larger effect on skill development for less-advantaged children (a 6.5% increase) and very little effect for children of higher-educated mothers (0.6% increase). The positive effect is larger for children in families with no access to relatives. Hours in paid care response meaningfully to such an offer. Average hours in paid care increase from roughly 5 hours per week (5.5% of total care time) to 8.3 hours per week. Therefore, total non-parental quality received increases dramatically, however the relatively modest effect on skills reflects the fact that, for young children, non-parental care consists of a relatively small fraction of their total care.

8 What Do We Learn from Care Choices Alone?

The effectiveness of child care policy differs for children with and without access to informal care. Since such access is not observed, it is worth considering the sensitivity of our views on policy to the identification of this unobserved quantity. As argued above, the effects of grandmother death on choices are key moments in identifying relative availability and preferences. In this section, I ask how estimates of the effectiveness of the 50% formal care subsidy change if we were to ignore this information and exploit only the choices that mothers make unconditional on any plausible shifter in choice sets. Thus I reintroduce the identification problem discussed above: we do not know whether a mother who does not choose relative care had access to such care. The parameter space is now set-identified: there is a locus of relative availability parameters and tastes for paid care that rationalize the choices we observe in the data. The substitution patterns revealed by grandmother death help locate the true data generating process on this locus, and here I ask how much this mattrers for our evaluation of policy.

To do so, I carry out a simulation exercise in which I take my baseline estimated values as the true data-generating process and simulate a panel of individuals. I then fix the variance of unobserved tastes of paid care, σ^2_ζ and estimate the parameters determining relative availability using only no information on grandmother death. This generates a vector of estimates $(\widehat{\ln\sigma_\zeta}, \hat{\delta}_0)$ that, along with all other parameters held constant at the baseline estimates, characterize a “valid” data-generating process $\widehat{\text{DGP}}$ that rationalizes observed choices, where the scare quotes reflect the fact that $\widehat{\text{DGP}}$ is under-identified. For each such estimated DGP, I simulate the effect of a 50% subsidy and plot the estimate response of skills, non-parental quality, mother labor force participation, and time in paid care for each education group. The distribution of $\widehat{\text{DGP}}$ generates a distribution of policy effects, and the quantitative question at hand is whether the support of this distribution is sufficiently large to alter our views on the efficacy of formal care subsidies.

Figure 1 displays the distribution of estimated policy effects for each set of parameters. Column A shows effects in lower-educated families and Column B presents those for higher-educated families. Each point represents a $(\widehat{\ln \sigma_\zeta}, \hat{\delta}_0)$ pair that rationalizes care choices, with the baseline estimate generated with the grandmother death instrument plotted for comparison. All effects are presented as percentage changes from baseline values, comparable with Columns 5 - 6 in Table 14. We see that estimated effects vary substantially. For example, the top row of Figure 1 shows that we could conclude that a subsidy increases skills of less-advantaged children by anywhere from 0.5% to 2.5%, and anywhere from a null efffect (as found in the baseline) to a *drop* in skills of the more-advantaged of over 2.5%. Similarly, we might conclude that mother labor supply responds to a formal care subsidy anywhere between 8% to -1%. These differences are economically large and highlight the policy importance of identifying the population that actually has access to informal care.

9 Conclusion

I estimate a model of maternal labor supply, child care choices, and skill development, allowing for heterogeneity in access to informal care provided by relatives. I find that predicted responses to child care policy vary by the availability of informal care and that failing to account for this variation in care choice sets leaves reponses to policy set identified. Indeed, the set of structural parameters that rationalize observed child care and labor supply choices allows for a distribution of policy responses with an economically large support.

References

Francesco Agostinelli and Matthew Wiswall. Estimating the Technology of Children's Skill Formation. *NBER Working Papers Series*, (February 2019), 2020. 8, 15

Antoine Arnoud, Fatih Guvenen, and Tatjana Kleineberg. Benchmarking Global Optimizers *. Technical report, 2019. URL http://ab-initio.mit.edu/wiki/index.php/NLopt. 23

Hero Ashman and Seth Neumuller. Can income differences explain the racial wealth gap? A quantitative analysis. *Review of Economic Dynamics*, 35:220–239, 2020. ISSN 10942025. doi: 10.1016/j.red.2019.06.004. URL https://doi.org/10.1016/j.red.2019.06.004. 12, 36

Raquel Bernal. The Effect of Maternal Employment and Child Care on Children's Cognitive Development. *International Economic Review*, 49(4):1173–1209, 2008. ISSN 00206598. doi: 10.1111/j.1468-2354.2008.00510.x. URL http://doi.wiley.com/10.1111/j.1468-2354.2008.00510.x. 2, 8

Raquel Bernal and Michael P. Keane. Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, 156(1):164–189, 2010. ISSN 03044076. doi: 10.1016/j.jeconom.2009.09.015. URL http://dx.doi.org/10.1016/j.jeconom.2009.09.015. 3

David M Blau and Alison P Hagy. The Demand for Quality in Child Care. Technical report, 1998. 2

Daniela Del Boca, Daniela Piazzalunga, and Chiara Pronzato. The role of grandparenting in early childcare and child outcomes. *Review of Economics of the Household*, pages 477–512, 2018. ISSN 1573-7152. doi: 10.1007/s11150-017-9379-8. URL http://dx.doi.org/10.1007/s11150-017-9379-8. 3

B.M Caldwell and R.H. Bradley. *Home Observation for Measurement of the Environment*. University of Arkansas at Little Rock, Little Rock, 1984. 17

Elizabeth M Caucutt and Lance Lochner. Early and Late Human Capital Investments, Borrowing Constraints, and the Family. Technical report, 2020. 3, 8

Juan Chaparro, Aaron Sojourner, and Matthew Wiswall. Early Childhood Care and Cognitive Development. Technical report, NBER Working Paper Series 26813, 2020. URL http://www.nber.org/papers/w26813. 3, 8

Satyajit Chatterjee and Burcu Eyigungor. A QUANTITATIVE ANALYSIS OF THE U.S. HOUSING AND MORTGAGE MARKETS AND THE FORECLOSURE CRISIS Satyajit Chatterjee Federal Reserve Bank of Philadelphia A Quantitative Analysis of the U.S. Housing and Mortgage Markets and the Foreclosure Crisis *. (15), 2015. URL www.philadelphiafed.org/research-and-data/publications/working-papers/. 19, 36

Janice Compton and Robert A. Pollak. Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics*, 79:72–90, jan 2014. ISSN 00941190. doi: 10.1016/j.jue.2013.03.007. 3, 25

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007. ISSN 00028282. doi: 10.1257/aer.97.2.31. 8

Flavio Cunha, James Heckman, and Susanne Schennach. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931, 2010. ISSN 0012-9682. doi: 10.3982/ecta6551. 8, 15

Flávio Cunha, Irma Elo, and Jennifer Culhane. Eliciting Maternal Expecations about the Technology of Cognitive Skill Formation. 2013. 11

Gordon B Dahl and Lance Lochner. The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, 2012. doi: 10.1257/aer.102.5.1927. URL http://dx.doi.org/10.1257/aer.102.5.1927. 27

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. *Review of Economic Studies*, 81(1):137–185, 2014. ISSN 00346527. doi: 10.1093/restud/rdt026. 3, 8

Ralitza Dimova and François Charles Wolff. Grandchild care transfers by ageing immigrants in france: Intra-household allocation and labour market implications. *European Journal of Population*, 24(3):315–340, sep 2008. ISSN 01686577. doi: 10.1007/s10680-007-9122-x. URL https://link.springer.com/article/10.1007/s10680-007-9122-x. 3

Ralitza Dimova and François Charles Wolff. Do downward private transfers enhance maternal labor supply? Evidence from around Europe. *Journal of Population Economics*, 24(3): 911–933, jul 2011. ISSN 09331433. doi: 10.1007/s00148-010-0305-0. URL https://link.springer.com/article/10.1007/s00148-010-0305-0. 25

Philipp Eisenhauer, James J Heckman, and Stefano Mosso. Estimation of Dynamic Discrete Choice Models by Maximum Likelihood and the Simulated Method of Moments. *International Economic Review*, 56(2):331–357, 2015. 18

Sarah Flood, Joel McMurry, Aaron Sojourner, and Matthew Wiswall. Inequality in Care Experience by U.S. Children. 2021. 2

Eva García-Morán and Zoë Kuehn. With strings attached: Grandparent-provided child care and female labor market outcomes. *Review of Economic Dynamics*, 23:80–98, 2017. ISSN 10942025. doi: 10.1016/j.red.2016.09.004. URL `http://dx.doi.org/10.1016/j.red.2016.09.004`. 3, 25

C Gourieroux, A Monfort, and E Renault. Indirect Inference C. Gourieroux; A. Monfort; E. Renault Journal of Applied Econometrics , Vol. 8, Supplement: Special Issue on Econometric Inference Using Simulation Techniques. (Dec., 1993), pp. S85-S118. *Journal of Applied Econometrics*, 8, 1993. 18, 22

Andrew S. Griffen. Evaluating the Effects of Childcare Policies on Children's Cognitive Development and Maternal Labor Supply. *Journal of Human Resources*, 54(3):604–655, 2019. ISSN 0022-166X. doi: 10.3368/jhr.54.3.0315.6988r1. 8

James J Heckman. Schools, Skills, and Synapses. *Economic inquiry*, 46(3):289, jun 2008. ISSN 0095-2583. URL `http://www.ncbi.nlm.nih.gov/pubmed/20119503http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2812935`. 2

John Kennan. A Note on Discrete Approximations of Continuous Distributions. *University of Wisconsin-Madison and NBER*, september(September), 2006. URL `http://www.ssc.wisc.edu/$\sim$jkennan/research/`. 9

Patrick Kline and Christopher R Walters. Evaluating public programs with close substitutes: The case of head start. *Quarterly Journal of Economics*, 131(4):1795–1848, 2016. ISSN 15314650. doi: 10.1093/qje/qjw027. URL `http://qje.oxfordjournals.org/`. 2, 8

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*. 1974. doi: 10.1080/07373937.2014.997882. 11

Alejandra Ros Pilarz, Ying Chun Lin, and Katherine A. Magnuson. Do parental work hours and nonstandard schedules explain income-based gaps in center-based early care and education participation? *Social Service Review*, 93(1):55–95, 2019. ISSN 00377961. doi: 10.1086/702685. 3

Josefina Posadas and Marian Vidal-Fernandez. Grandparents' Childcare and Female Labor Force Participation. *IZA Journal of Labor Policy*, 2(1):1–20, 2013. ISSN 21939004. doi: 10.1186/2193-9004-2-14. 3

J B Y Rust. Optimal Replacement of GMC Bus Engines : An Empirical Model of Harold Zurcher Author. *Econometrica*, 55(5):999–1033, 1987. 11

Yulya Truskinovsky. The Unintended Consequences of Informal Childcare Subsidies for Older Women's Retirement Security. *Journal of Human Resources*, 2020. 3

Gema Zamarro. Family labor participation and child care decisions: the role of grannies. *SERIEs*, 11(3):287–312, 2020. ISSN 18694195. doi: 10.1007/s13209-020-00213-5. URL `https://doi.org/10.1007/s13209-020-00213-5`. 3

10 Tables and Figures

Table 1: Care Arrangements by Child Age

Child Age	Parent Only	Relative	Center	Home-Based	N Arrange
0	0.59	0.16	0.07	0.22	1.11
1	0.54	0.16	0.11	0.23	1.10
2	0.51	0.15	0.15	0.22	1.09
3	0.46	0.14	0.24	0.20	1.11
4	0.48	0.12	0.27	0.16	1.09
5	0.69	0.08	0.14	0.10	1.05
6	0.88	0.05	0.03	0.04	1.04

Source: PSID-CDS. Non-parent care usage is not mutually exclusive, so frequency of use may not sum to 1. Number of non-parent arrangements conditional on using any non-parent care. Statistics weighted by CDS child-primary caregiver sample weight, normalized within survey year.

Table 2: Non-Parental Hours by Child Age

Child Age	Relative	Relative Cond.	Center	Center Cond.	Home-Based	Home-Based Cond.
0	4.28	26.79	1.46	20.28	6.77	31.40
1	6.10	38.93	3.13	29.81	8.82	38.10
2	5.72	37.59	4.79	31.03	7.56	34.28
3	5.73	41.08	6.89	28.21	7.39	36.86
4	4.73	38.32	7.21	26.46	5.78	36.42
5	2.53	30.23	2.64	18.91	2.39	23.45
6	1.83	37.49	0.92	26.79	1.02	22.89

Source: PSID-CDS. Average weekly hours unconditional and conditional on use of arrangement. Statistics weighted by CDS child-primary caregiver sample weight, normalized within survey year Statistics weighted by CDS child-primary caregiver sample weight, normalized within survey year.

Table 3: Calibrated Parameters

Description	Estimate	Source
Care quality supply curve parameters (age 0 - 2)	\$3.42, \$6.66	ECLS-B
Care quality supply curve parameters (age 3 - 4)	\$3.38, \$9.3	ECLS-B
Annual hours full-time work	2,880	ATUS
Weekly child care time requirement	91	PSID-CDS Time Diaries
Consumption floor (fraction median income)	13.5 %	Ashman and Neumuller [2020]
Discount rate β	0.97	n/a

Table 4: Tax Schedule

Pre-Tax Income per Resident Parent	Average Tax Rate
0.00 - 0.37	0.15
0.37 - 0.88	0.28
0.88 - 1.34	0.31
1.34 - 2.40	0.36
$>$ 2.40	0.39

Notes: Tax schedule taken from Chatterjee and Eyigungor [2015]. Pre-tax income expressed as fraction of median earnings, set to \$65,000, taken from the PSID.

Table 5: Family Structure Transitions

	Only Child		Siblings	
	Two Parents	Single Mother	Two Parents	Single Mother
Mother College				
Two Parents, Only Child	0.669	0.035	0.291	0.006
Single Mother, Only Child	0.074	0.749	0.036	0.141
Both Parents, Mult. Child	0.006	0.001	0.952	0.041
Single Mother, Mult. Child	0.011	0.073	0.062	0.854
Mother No College				
Two Parents, Only Child	0.653	0.064	0.258	0.026
Single Mother, Only Child	0.031	0.717	0.033	0.219
Both Parents, Mult. Child	0.013	0.005	0.912	0.070
Single Mother, Mult. Child	0.006	0.055	0.050	0.888

Source: PSID linked family files. Sample limited to CDS children. Two-parent households are defined as households in which the respondent mother reports a residing spouse. Sibling indicator constructed from family roster indicating number of children within household. Statistics weighted by combined family sample weight, normalized within survey year.

Table 6: Parameter Estimates - Skill Production and Mother Quality Measurement

Description	Parameter	Estimate	SE
Ages 0-2			
TFP	A_t	0.002	(0.012)
Skill self-productivity	γ_{1t}	0.267	(0.011)
Input productivity	γ_{2t}	0.369	(0.017)
Mother quality measure scale	λ_t	0.400	(0.139)
Shock variance	σ_η^2	0.780	
Ages 3-4			
TFP	A_t	0.023	(0.02)
Skill self-productivity	γ_{1t}	0.615	(0.014)
Input productivity	γ_{2t}	0.127	(0.027)
Mother quality measure scale	λ_t	1.215	(0.733)
Shock variance	σ_η^2	0.525	

Source: ECLS-B. Standard errors computed via block bootstrap at the child/survey level.

Table 7: Parameter Estimates - Care Quality

Description	Parameter	Estimate	SE	Source
Mother Care Quality				
Coefficient - mom college	β^m	0.404		
Shock variance	$\sigma^2_{\nu,m}$	1		
Relative Care Quality				
Intercept	β^r_0	-0.57	(0.137)	SMM
Coefficient - mom college	β^r_1	0.742	(0.252)	SMM
Shock variance	$\sigma^2_{\nu,m}$	1.067	(0.11)	SMM

Table 8: Parameter Estimates - Preferences

Description	Parameter	Estimate	SE	Source
Leisure Shifter				
Intercept	$\psi_{\ell,0}$	1.85	(0.085)	SMM
Both parents	$\psi_{\ell,1}$	-0.26	(0.014)	SMM
Care Time Shifter				
Intercept	$\psi_{m,0}$	2.000	(0.035)	SMM
Care Quality Shifter				
Intercept	$\psi_{q,0}$	-0.705	(0.045)	SMM
Mom college	$\psi_{q,1}$	1.365	(0.438)	SMM
Psychic Cost of Paid Care				
Psychic cost for paid care variance	$\ln \sigma_{\zeta}$	0.500	(0.137)	SMM
Career Concern Shifter				
Intercept	$AT0$	-1.000	(0.4)	SMM
Mom college	$AT1$	1.300	(0.782)	SMM

Table 9: Parameter Estimates - Income Processes

Description	Parameter	Estimate	SE	Source
Mother Wage Offer				
Intercept	$\gamma_{w,0}$	10.006	(0.12)	SMM
Coefficient - mom college	$\gamma_{w,1}$	0.278	(0.057)	SMM
Shock variance	$\sigma^2_{\xi,w}$	0.287	(0.11)	SMM
Exogenous Income				
Intercept	$\gamma_{y,0}$	8.648	(0.021)	OLS
Coefficient - mom college	$\gamma_{y,1}$	0.107	(0.291)	OLS
Coefficient - both parents	$\gamma_{y,2}$	1.706	(0.025)	OLS
Coefficient - mom college $\times$ both parents	$\gamma_{y,3}$	0.381	(0.034)	OLS
Shock variance	$\sigma^2_{\xi,y}$	1.013		OLS residual variance

Table 10: Model Fit - Child Care Arrangements

	Model	Data	Data SE
Paid Care OLS Coefficients			
Usage - Intercept	0.193	0.229	0.015
Usage - Mom College	0.097	0.133	0.018
Hours - Intercept	5.030	5.036	0.468
Hours - Mom College	2.690	3.185	0.574
Relative Care OLS Coefficients			
Usage - Intercept	0.126	0.168	0.012
Usage - Mom College	−0.011	0.005	0.015
Hours - Intercept	3.336	3.812	0.344
Hours - Mom College	−0.119	−0.642	0.422
Maternal Education Quality Gradients			
Mom Care Quality	0.600	0.405	0.018
Paid Care Quality	0.190	0.185	0.086
Relative Care Quality	1.000	0.591	0.095

Table 11: Model Fit - Labor Supply and Family Resources

	Model	Data	Data SE
Mother Labor Supply OLS Coefficients			
MLFP - Intercept	0.255	0.284	0.018
MLFP - Mom College	0.160	0.182	0.023
Cond Hours Market Work - Intercept	31.299	45.854	0.676

Table 12: Model Fit - Grandmother Death "First Stage"

	Model	Data	Data SE
Relative Care Use - Extensive Margin			
Mom No College	−0.129	−0.118	0.044
Mom College	−0.117	−0.134	0.029
Mother Labor Force Participation			
Mom No College	−0.018	−0.060	0.048
Mom College	−0.047	−0.094	0.031
Mother Sole Caregiver			
Mom No College	0.138	0.070	0.054

Table 12: Model Fit - Grandmother Death "First Stage"

	Model	Data	Data SE
Mom College	0.036	0.125	0.042
Expenditures Paid Care			
Mom No College	-0.046	0.048	0.034
Mom College	0.127	0.208	0.054

Table 13: Average Relative Effect

	No Relative Avg		Avg Effect of Relative		Avg Effect \ |No Relative|	
	No College	College	No College	College	No College	College
Skills t = T	−0.116	0.079	−0.001	0.002	−0.007	0.020
Nonparent Quality	−0.004	0.012	−0.060	0.072	−14.948	5.803
Paid expenditures	0.131	0.326	−0.030	−0.102	−0.226	−0.313
MLFP	0.242	0.395	0.045	0.091	0.186	0.231
Time with mom	0.939	0.909	−0.108	−0.123	−0.115	−0.136
Time in paid care	0.061	0.091	−0.015	−0.028	−0.241	−0.307

Table 14: Average Effect of 50% Paid Care Subsidy

	Baseline Avg		Avg Effect		Avg Effect \ |Baseline|	
	No College	College	No College	College	No College	College
All Families						
Skills t = T	−0.116	0.079	0.002	0.000	0.015	0.001
Nonparent Quality	−0.022	0.029	0.003	0.004	0.135	0.139
Paid expenditures	0.121	0.301	−0.037	−0.123	−0.307	−0.410
MLFP	0.255	0.415	0.004	0.003	0.014	0.008
Time with mom	0.908	0.880	−0.010	−0.007	−0.011	−0.008
Time in paid care	0.055	0.085	0.011	0.008	0.195	0.095
No Relative Available						
Skills t = T	−0.109	0.082	0.002	0.000	0.017	0.003
Nonparent Quality	−0.004	0.013	0.003	0.005	0.899	0.423
Paid expenditures	0.131	0.331	−0.039	−0.136	−0.302	−0.412
MLFP	0.240	0.398	0.004	0.004	0.018	0.011
Time with mom	0.940	0.908	−0.011	−0.008	−0.012	−0.009
Time in paid care	0.060	0.092	0.011	0.008	0.183	0.085
Relative Available						
Skills t = T	−0.134	0.069	0.002	−0.001	0.012	−0.011
Nonparent Quality	−0.065	0.081	0.004	0.000	0.064	0.002
Paid expenditures	0.099	0.204	−0.028	−0.075	−0.280	−0.370
MLFP	0.290	0.469	0.016	0.001	0.056	0.003
Time with mom	0.831	0.791	−0.007	−0.007	−0.009	−0.009
Time in paid care	0.045	0.061	0.013	0.013	0.286	0.207

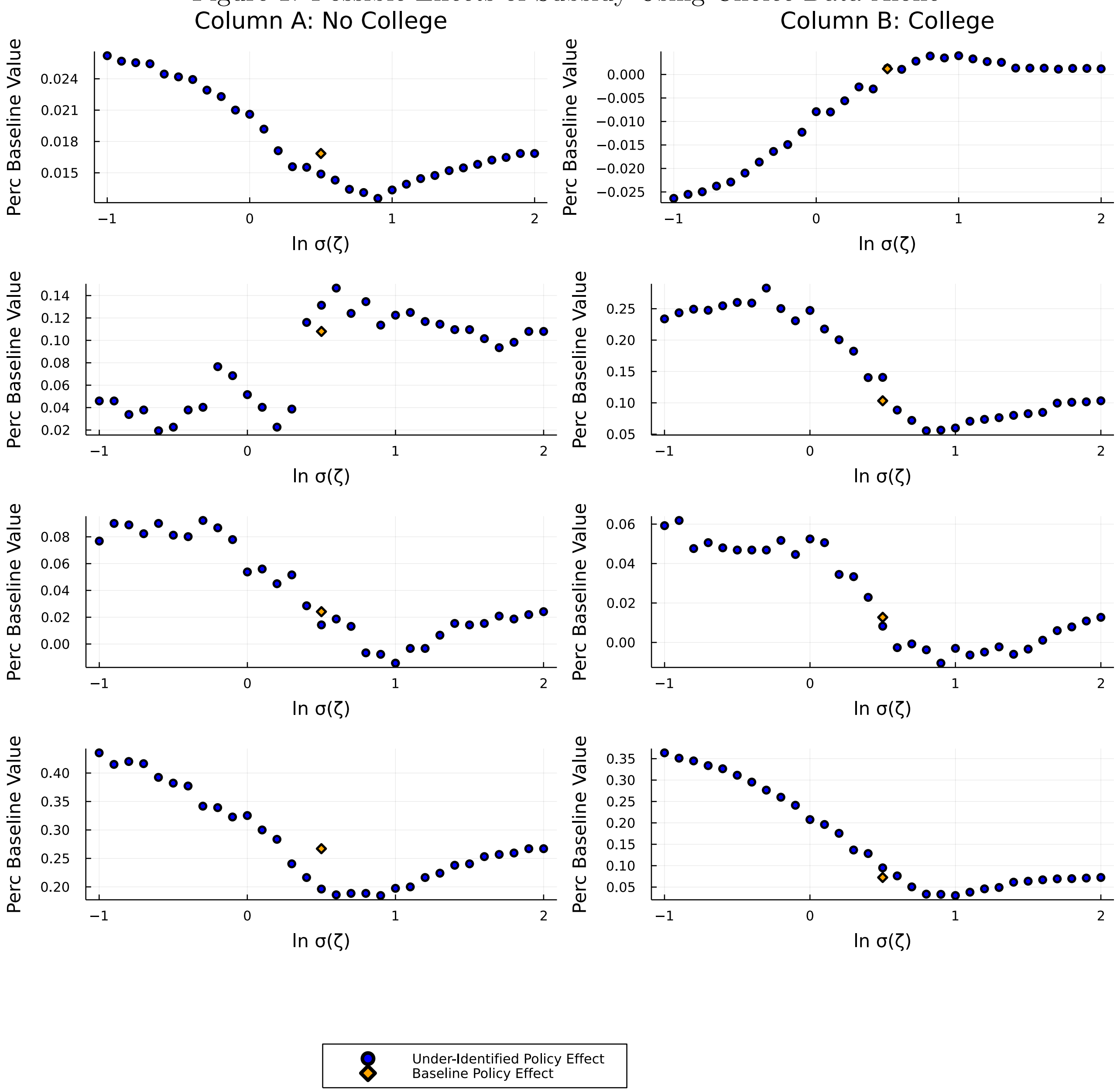
Table 15: Average Effect of $1700 Direct Transfer

	Baseline Avg		Avg Effect		Avg Effect \ |Baseline|	
	No College	College	No College	College	No College	College
All Families						
Skills t = T	−0.116	0.079	0.000	0.000	0.000	0.000
Nonparent Quality	−0.022	0.029	−0.000	−0.000	−0.002	−0.000
Paid expenditures	0.121	0.301	0.000	0.000	0.001	0.001
MLFP	0.255	0.415	−0.018	−0.010	−0.069	−0.024
Time with mom	0.908	0.880	−0.001	0.000	−0.001	0.000
Time in paid care	0.055	0.085	0.002	−0.000	0.033	−0.001
No Relative Available						
Skills t = T	−0.109	0.082	−0.000	0.000	−0.000	0.001
Nonparent Quality	−0.004	0.013	−0.001	0.000	−0.142	0.008
Paid expenditures	0.131	0.331	−0.002	0.000	−0.016	0.000
MLFP	0.240	0.398	−0.020	−0.009	−0.084	−0.022
Time with mom	0.940	0.908	−0.001	0.000	−0.001	0.000
Time in paid care	0.060	0.092	0.001	−0.000	0.010	−0.003
Relative Available						
Skills t = T	−0.134	0.069	0.000	−0.000	0.002	−0.000
Nonparent Quality	−0.065	0.081	0.001	−0.000	0.017	−0.005
Paid expenditures	0.099	0.204	0.005	0.001	0.055	0.006
MLFP	0.290	0.469	−0.011	−0.013	−0.039	−0.028
Time with mom	0.831	0.791	−0.000	0.000	−0.001	0.001
Time in paid care	0.045	0.061	0.005	0.000	0.106	0.005

Table 16: Average Effect of Free High-Quality Paid Care

	Baseline Avg		Avg Effect		Avg Effect \ |Baseline|	
	No College	College	No College	College	No College	College
All Families						
Skills t = T	−0.116	0.079	0.008	0.000	0.065	0.006
Nonparent Quality	−0.022	0.029	0.018	0.011	0.822	0.386
Paid expenditures	0.121	0.301	−0.121	−0.301	−1.000	−1.000
MLFP	0.255	0.415	−0.001	−0.001	−0.005	−0.003
Time with mom	0.908	0.880	−0.035	−0.017	−0.038	−0.020
Time in paid care	0.055	0.085	0.037	0.019	0.671	0.226
No Relative Available						
Skills t = T	−0.109	0.082	0.008	0.001	0.075	0.008
Nonparent Quality	−0.004	0.013	0.019	0.013	5.047	1.052
Paid expenditures	0.131	0.331	−0.131	−0.331	−1.000	−1.000
MLFP	0.240	0.398	−0.004	0.001	−0.015	0.003
Time with mom	0.940	0.908	−0.039	−0.020	−0.042	−0.022
Time in paid care	0.060	0.092	0.039	0.020	0.659	0.215
Relative Available						
Skills t = T	−0.134	0.069	0.006	−0.000	0.045	−0.003
Nonparent Quality	−0.065	0.081	0.016	0.004	0.250	0.049
Paid expenditures	0.099	0.204	−0.099	−0.204	−1.000	−1.000
MLFP	0.290	0.469	0.004	−0.008	0.013	−0.018
Time with mom	0.831	0.791	−0.024	−0.009	−0.029	−0.012
Time in paid care	0.045	0.061	0.032	0.017	0.709	0.280

Figure 1: Possible Effects of Subsidy Using Choice Data Alone



A Parametric Specification

In my empirical application, utility shifters take the form of

$$\Psi_\ell(X, \zeta) = \exp(\psi_{\ell,0} + \psi_{\ell,1} \times (\text{both parents} = 1)) \tag{38}$$

$$\Psi_m(X, \zeta) = \exp(\psi_{m,0}) \tag{39}$$

$$\Psi_q(X, \zeta) = \exp(\psi_{q,0} + \psi_{q,1} \times (\text{mom college} = 1)) \tag{40}$$

$$\Psi_p(X, \zeta) = \exp(\zeta_p) \tag{41}$$

$$V_T(d_T) = \exp(AT_0 + AT_1 \times (\text{mom college} = 1)) \times d_T \tag{42}$$

B Identification

Here I consider a simple model that illustrates the identification argument above. Key ingredients of the model are that mothers may only choose relative care if it is available, and that the analyst does not observe availability.

Assume mothers face a single choice of whether to use relative care $D = 1$ or an alternative care arrangement $D = 0$. Unobserved relative availability is denoted $R \in \{0, 1\}$ and $R = 0 \implies D = 0$. Let X_i be a scalar covariate and $Z_i \in \{0, 1\}$ denote a binary instrument (grandmother death pre-birth). I assume the data consist of $\{D_i, X_i, Z_i\}$ and that it is generated by a simple latent utility framework in which the utility of the alternative arrangement is normalized to 0:

$$U_{1i} = X_i\beta + \nu_r \tag{43}$$

$$U_{0i} = 0 \tag{44}$$

where $\nu_r \sim N(\mu, \sigma)$ is an unobserved taste shock. I assume that relative availability is determined by the following latent variable

$$R_i^* = X_i\delta + Z_i\gamma + \nu_r \tag{45}$$

with $R_i^* \geq 0 \implies R_i = 1$. The fact that ν_r enters both 43 and 45 induces a selection problem. In this simple model, the only information we can exploit is choice probabilities of the form:

$$\pi_1(x, z) = P(D_i = 1 | X_i = x, Z_i = z) \tag{46}$$

We have that $D_i = 1$ if $R_i^* \geq 0$ and $U_{1i} \geq 0$. Therefore (dropping the conditioning for convenience), we can write the complementary probability as the union of two disjoint events: $R_i^* < 0$ or ($R_i^* \geq 0$ and $U_{1i} < 0$). Then using the formula for a Normal random variable truncated from below, we have

$$1 - \pi_1 = P(X_i\delta + Z_i\gamma + \nu_r < 0) + P(X_i\delta + Z_i\gamma + \nu_r \geq 0, X_i\beta + \nu_r < 0) \tag{47}$$
$$= P(\nu_r < -X_i\delta - Z_i\gamma) + P(\nu_r < -X_i\beta | \nu_r \geq -X_i\delta - Z_i\gamma)P(\nu_r \geq -X_i\delta - Z_i\gamma) \tag{48}$$
$$= \Phi(\xi_r) + \frac{\Phi(\xi_u) - \Phi(\xi_r)}{1 - \Phi(\xi_r)} \tag{49}$$

where $\xi_r = \frac{-X_i\delta - Z_i\gamma - \mu}{\sigma}$ and $\xi_u = \frac{-X_i\beta - \mu}{\sigma}$. Therefore we have

$$\pi_1 = 1 - \left(\Phi(\xi_r) + \frac{\Phi(\xi_u) - \Phi(\xi_r)}{1 - \Phi(\xi_r)} \right) \tag{50}$$

The left-hand side of 50 is identified directly from the data and the right-hand side is a function of five parameters: $\delta, \gamma, \beta, \mu, \sigma$. Normalizing μ, σ, the remaining parameters δ, γ, β are identified from at least four choice probabilities ($X_i \in \{x_1, x_2\}$ crossed with $Z_i \in \{0, 1\}$.

To see why the instrument Z_i is necessary for identification, consider an alternative model with

$$R_i^* = X_i\delta + \nu_r \tag{51}$$

The choice probabilities would then take the same form as 50 but with $\tilde{\xi}_r = \frac{-X_i\delta - \mu}{\sigma}$:

$$\pi_1 = 1 - \left(\Phi(\tilde{\xi}_r) + \frac{\Phi(\xi_u) - \Phi(\tilde{\xi}_r)}{1 - \Phi(\tilde{\xi}_r)} \right) \tag{52}$$

Again fixing μ, σ, we have a single equation with two unknowns: δ, β. To see that these parameters are not identified from choice probabilities alone, pick any (δ, β) that satisfy 52 for two different values of X_i. Now conisder $\delta' = \delta - \epsilon$ with $\epsilon > 0$. Since both $\Phi(\tilde{\xi}_r)$ and $\frac{\Phi(\xi_u) - \Phi(\tilde{\xi}_r)}{1 - \Phi(\tilde{\xi}_r)}$

are increasing in $\tilde{\xi}_r$, we have that

$$\left(\Phi(\tilde{\xi}_r) + \frac{\Phi(\xi_u) - \Phi(\tilde{\xi}_r)}{1 - \Phi(\tilde{\xi}_r)}\right) < \left(\Phi(\tilde{\xi}_r') + \frac{\Phi(\xi_u) - \Phi(\tilde{\xi}_r')}{1 - \Phi(\tilde{\xi}_r')}\right) \tag{53}$$

because

$$\tilde{\xi}_r = \frac{-X_i\delta - \mu}{\sigma} < \frac{-X_i\delta' - \mu}{\sigma} = \tilde{\xi}_r' \tag{54}$$

Notice that the only place in 52 that β enters is in $\Phi(\xi_u)$. $\Phi(\xi_u)$ is increasing in ξ_u, which is itself decreasing in β. Therefore, if we can choose $\beta' > \beta$ to reduce $\Phi(\xi_u)$ so that equality obtains in 53, we will have shown the observational equivalence of (δ, β) and (δ', β'). Continuity of both $\Phi(\tilde{\xi}_r)$ and $\frac{\Phi(\xi_u)-\Phi(\tilde{\xi}_r)}{1-\Phi(\tilde{\xi}_r)}$ implies that we can choose ϵ small enough so that there is a $\beta' > \beta$ that yields equality in 53. Therefore, the model without an instrument shifting R_i but not U_i is not identified.

The intuition for this underidentification result is simple. The conditional probability of choosing the outside option is generated by two groups: one group in which a relative is not available, and one group in which the relative is available but not preferred. These groups are not separately observed: we only see their union. With no exclusion restriction, observed choice probabilities conditional on X_i can be rationalized by multiple combinations of these two groups. Without some variable to shift one group and not the other, we cannot disentangle the two.

C Additional Tables and Figures

Figure 2: Education Gradients in Care Hours

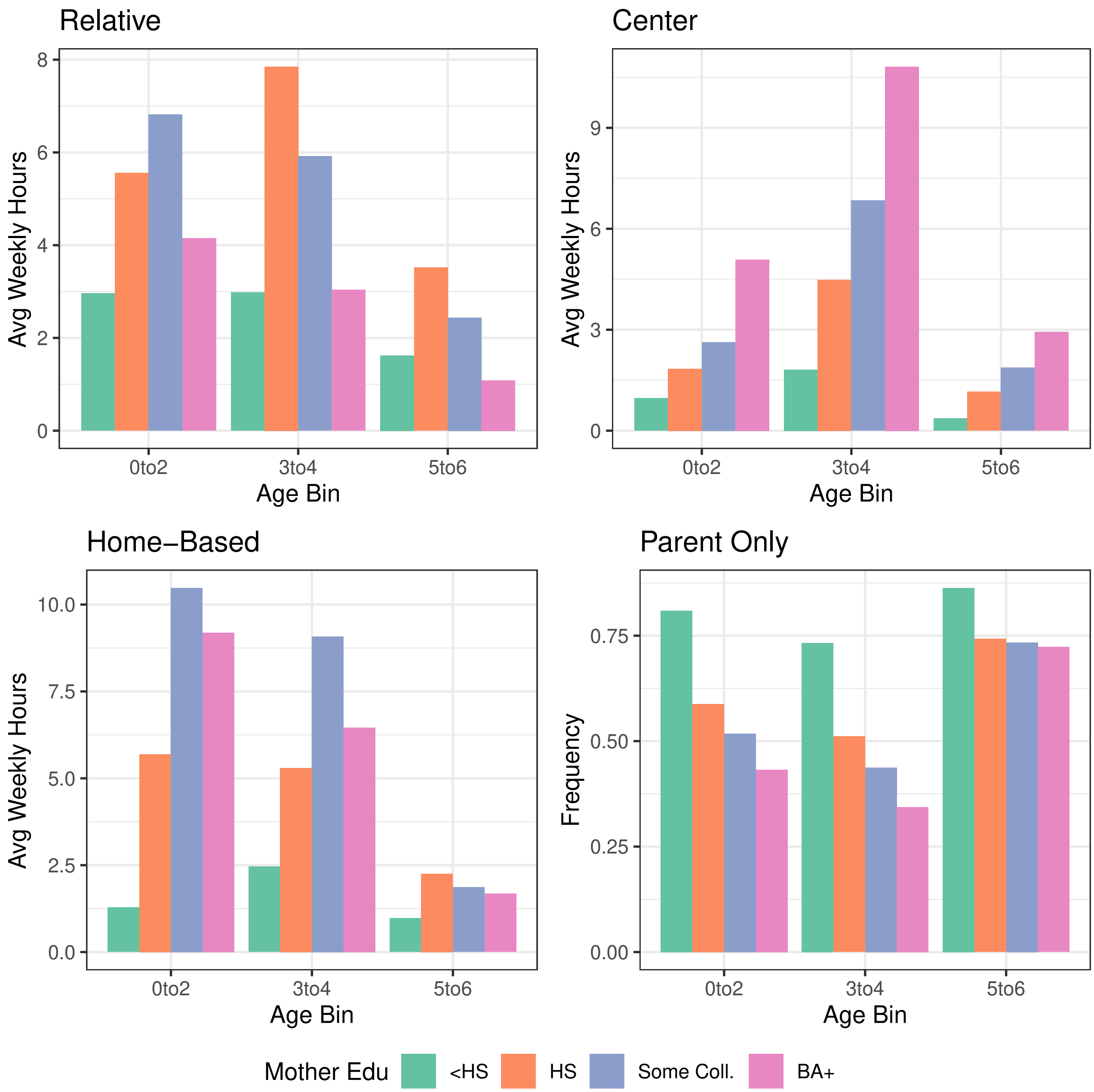


Figure 3: Education Gradients in Care Hours - Mother Not in Labor Market

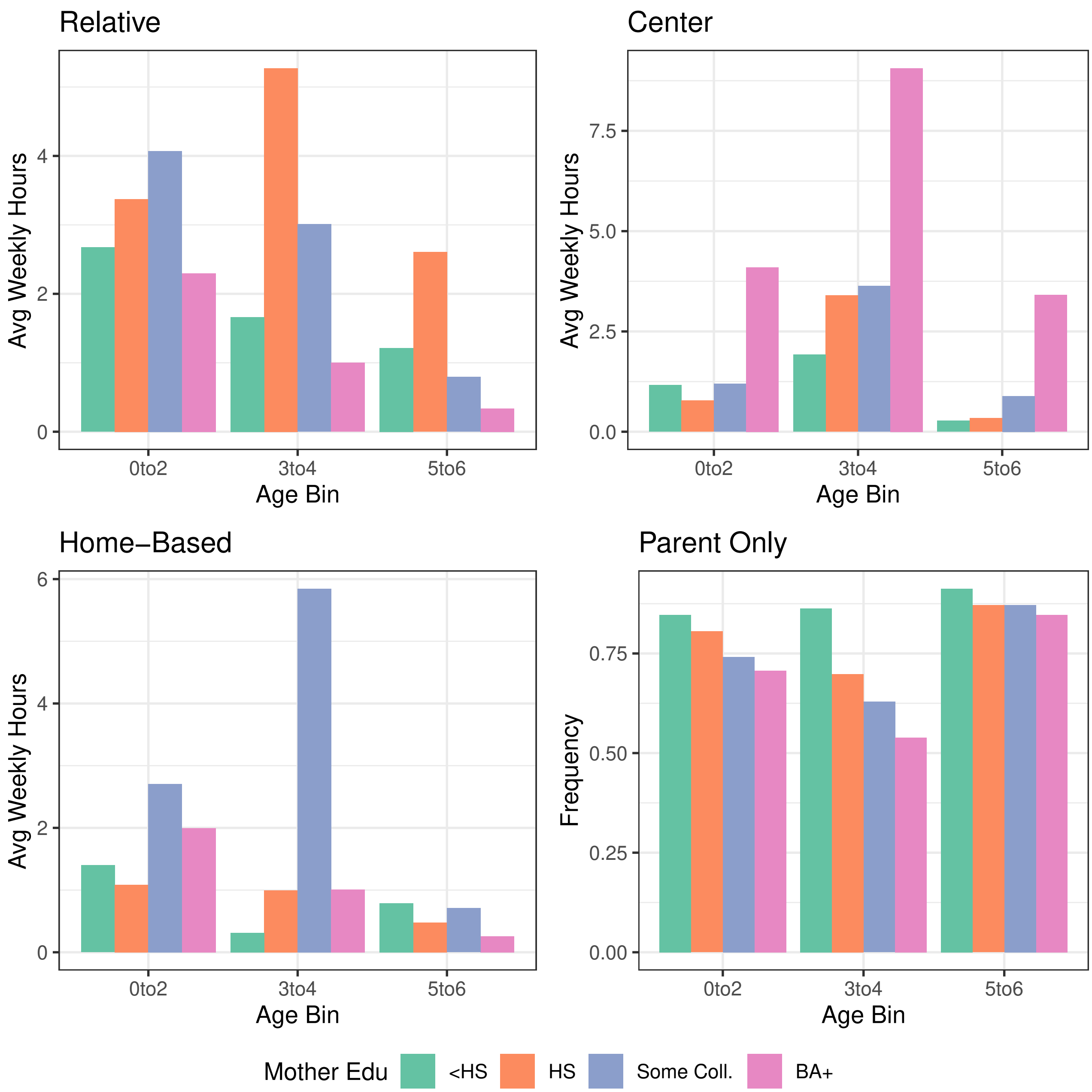


Figure 4: Education Gradients in Care Hours - Mother in Labor Market