Abstract

This paper investigates how government antipoverty programs in the United States can support the development of disadvantaged children. I accomplish this by formalizing and estimating a model in which maternal investments of time and money determine child outcomes. Mothers choose these investments, their labor supply, and their program participation subject to constraints on their time and budget. I use this structure to answer important policy questions regarding the design and delivery of income supports to poor families. My estimates suggest that boosting child outcomes through sustained income transfers is expensive. However, there are significant returns to providing additional income supports to the poorest families in the sample. I propose a policy reform, based on a minimum guaranteed income standard, that is revenue neutral and significantly improves the cognitive and behavioral skills of the most economically disadvantaged children. When associated with high school graduation, the induced change in skills lifts the probability of graduation by 2.9 percentage points for children in the bottom 10 percent of the income distribution.

*Many thanks to Chris Flinn, Gianluca Violante, Pamela Morris, Tobias Salz, Matthew Wiswall, Kevin Thom, David Cesarini, Timothy Christensen, Isaac Baley and Alexander Coutts for many helpful comments and suggestions. All remaining errors are my own.
1 Introduction

The development of capabilities and skills in childhood is a process that crucially shapes human experience. If children face poverty or economic disadvantage, then this may limit their developmental potential, leaving a socioeconomic scar on their adult lives and exacerbating the transmission of inequality across generations. An impetus follows this logic: governments should protect children from disadvantage and deprivation through income supports to families. As a point of departure for this paper, I emphasise that this insight, though it provides a strong ethical grounding for antipoverty initiatives, offers no guidelines with respect to the optimal size or design of such a policy.

This paper offers progress on this question, by studying how the design of government antipoverty programs in The United States affects child development outcomes. Importantly, I perform this analysis using a model in which such outcomes are determined by constrained, economic choices, that are fully articulated in a rich array of welfare policy environments. I estimate the model using longitudinal data on maternal labor supply, time use, earnings, welfare participation, and child outcomes. I then use these estimates to simulate decisions and child outcomes under counterfactual policy arrangements. My estimates, combined with these experiments, produce an important set of findings. Heterogeneity across parents explains the largest amount of variance in child outcomes, suggesting that eliminating economic inequality will not make a large impact on overall inequality in skill development. However, both time and money are found to be critical inputs in skill production. Thus, when current policies are reconfigured to target the most needy families in the sample, this results in appreciable human capital gains. This study highlights the fact that there are potentially important social returns to fine-tuning and strengthening the social safety net.

The problem of how to design income supports can be defined over two key dimensions. First, how fundamentally generous should these programs be? Second, conditional on size, how forcefully should these policies encourage labor supply and self-sufficiency? A very broad array of antipoverty initiatives can be fully characterised across these two dimensions. Some programs, such as the Earned Income Tax Credit (EITC), reward labor supply by phasing in the transfer as earnings increase. The EITC, in this fashion, acts as a negative income tax. Other programs may discourage labor supply by phasing out a baseline transfer as earnings increase. This is true of both incarnations of welfare in The United States: Aid to Families with Dependant Children (AFDC), and Temporary Assistance for Needy Families (TANF), as well as the latter part of the EITC transfer schedule. The phase-out rate of these transfers determines the severity at which additional earnings are effectively taxed. Programs that rapidly phase out transfers with earnings impose high rates of marginal income taxation, which strongly discourages labor supply.

Historically, efforts to alleviate poverty in the U.S. have varied chiefly in these two dimensions. Recent decades have seen a concerted policy shift towards programs that encourage self-sufficiency.
and discourage dependance on government support. First, the 1990s saw a significant expansion in the Earned Income Tax Credit (EITC), particularly in the period between 1993 and 1995. Second, in 1996, the Personal Responsibility and Work Opportunity Reconciliation Act (PROWRA) was signed into law. This act affected a transition in 1997 from the former welfare program Aid to Families with Dependant Children (AFDC) to a new program: Temporary Assistance for Needy Families (TANF). Three aspects of the reform combined to encourage labor force participation among welfare recipients. First, entitlements to TANF are limited to a 60 month time limit for participants\(^{1}\). Second, recipients are required to spend 20 hours a week in work or work-related activities. Failure to meet this requirement is punishable by sanctions. Third, many states restructured their benefit calculation formulae to reduce the effective marginal tax rate on earnings. This series of institutional changes was successful in shifting a sizeable number of participants off welfare rolls and into the labor force (Hoynes, 1996; Grogger, 2002, 2003; Chetty et al., 2013). I fit the model to data taken from a period that encompasses the timing of these policy changes. This provides a valuable opportunity to study the effects of variation in policy arrangements on key variables of interest that interact causally in the model.

In the model, child outcomes are determined by maternal investments of time and money, which enter into a skill production function. Skill formation in the model is both dynamic, in that human capital at maturity is determined over several periods in which current skills combine with investments to beget future skills, and multidimensional, in that it incorporates an array of both cognitive and socioemotional abilities. This model of skill formation joins work at the frontier of research on human capital development, which suggests that both these features are essential (Cunha et al., 2010; Almlund et al., 2011; Heckman et al., 2006).

If child development depends on the availability of both time and money, as it does in this model, then the design choices of governments with respect to antipoverty programs has important implications for this process. I study the problem of program design through this lens, evaluating alternative policies of equivalent size to the suite of initiatives in the U.S. with different degrees of labor supply incentivization. These questions are of critical importance to society and to broader discussions of policy. Prior research suggests that skills and capabilities (which I sometimes refer to in this paper as human capital) are most malleable in childhood (Cunha et al., 2010), and that these skills wield considerable influence over important socioeconomic outcomes \(^{2}\). The extent to which human capital can be shaped by policies, and the extent to which this malleability is concentrated in childhood, has strong implications for the optimal timing and design of social programs. In addition, these insights provide a sense of possibility when faced with the problem of socioeconomic inequality: the nature of skill formation in childhood frees policymakers from the

---

1 These time limits apply only to the block grants awarded to the States by Federal Government. States can, if they choose, independently fund support beyond the time limit.

2 Some primary examples include: high school and college completion, earnings, health, and prosocial behaviors.
constraints of traditional equity-efficiency tradeoffs (Cunha and Heckman, 2007).

To date, many empirical studies of public investment in children have validated this theoretical perspective. Analyses of experimental Early Childhood Education (ECE) initiatives have shown very high returns on initial program costs (Heckman et al., 2010), and that much of these gains can be explained by the change in skills induced by the program (Heckman et al., 2013). The results on the impacts of early childhood investment suggest a pragmatic case both for improving and expanding current programs. At the very least, they encourage a particular prior that improving such programs can bring potentially large gains in outcomes.

The structure of the model I develop allows, in principle, for many different potential answers to the questions of best policy design. My theory proposes that welfare policies affect development because they shape two important economic constraints: the availability of time and money. Thus, my results will hinge on the absolute and relative importance of these factors in producing child outcomes. This amounts to estimating the human capital production function of children, which is the first major input into subsequent policy analysis. On its own, this object is not sufficient to analyze counterfactuals. We need, in addition, the ability to model and predict parental responses to policy changes. This is enabled by specifying parental preferences over relevant outcomes and adapting revealed preference arguments to identify them from data. Both these components - the production technology, and parental preferences - are specified so that they can be statistically identified in a transparent way from the available data. The model then, in addition to providing theoretical insights into the policy design problem, provides a mapping from historical data to alternative simulations, which lies at the heart of this paper’s analysis.

The causal relationship between economic resources and child outcomes can be difficult to infer from the data, since it is likely that unobservable family characteristics determine both family income and child outcomes. To negotiate this issue, several studies have used natural experimental variation in family budgets, either through welfare programs (Duncan et al., 2011) or the EITC (Dahl and Lochner, 2012; Chetty et al., 2011), to estimate a fairly consistent effect size: $1,000 in additional annual income can increase test scores by between 6 and 9 percent of a standard deviation. Additional work suggests that these income effects are not limited to test scores, producing long-term effects in outcomes also (Loken et al., 2012; Hoynes et al., 2014). ³ While some researchers have expressed cynicism about the use of transfers to boost child outcomes (Heckman and Mosso, 2014), these results suggest that economic resources have some role to play in the development process. Although useful and critical, the empirical design of these studies does not permit fully realized policy counterfactuals.

We can use the model as a lens through which to interpret these estimates. One insight is that the same combination of production and preference parameters can easily produce different impacts, depending on the policy variation and research design that is used to estimate them. In ³This finding is not unanimous, however. See Cesarini et al. (2015) for an important example.
particular, a $1,000 increase in annual income will have variable effects depending on the changes in labor supply that each policy induces. Furthermore, while each policy might produce identical effects on cognitive test scores, they may affect other skills differently. Since the formation of skills is dynamic and complementary, there could be stark differences between contemporaneous and long-term impacts.

The structure of the model can be described in more detail as follows. Mothers have preferences defined over the outcomes of their children, their private consumption, and private leisure. They can improve the outcomes of their children by investing either time or money, but are limited in terms of two key resource constraints: their annual budget (the sum of their earnings, non-labor income, and government transfers) and the number of hours available in the period. Thus, investments in children must be made at the expense of private consumption and leisure. Labor supply acts as a means for agents to expand their budgets at the expense of time. Mothers choose the amount of labor they supply to the market, as well as their participation in government programs. These decision variables, along with the mother’s wage, and the relevant government policy in each time period determines the household budget. The decision problem faced by a mother is dynamic, since she must weigh the future return to skills (which develop over time) against her current consumption and leisure. These features of the model build on work by Del Boca, Flinn, and Wiswall (2014), which itself follows in the tradition of Becker and Lewis (1973) and Becker and Tomes (1976).

I extend the modelling framework of Del Boca et al. (2014) in several meaningful ways. First, I adopt a multidimensional skill technology, which broadens the policy implications and importance of the study, since behavioral traits are influential in shaping many life-cycle outcomes (Heckman et al., 2006). Furthermore, theory and evidence suggest that such skills are crucial in the development of future skills (Cunha and Heckman, 2007; Cunha et al., 2010). I also pursue a different approach to identification of the production technology which, in combination with the fact that I estimate the model on a different population of children, allows for new empirical insights on the nature of human capital formation in childhood. The empirical methodology I adopt here, which mediates the developmental effects of programs through their effect on family income and total maternal hours, resembles the approach taken by Bernal (2008) and Bernal and Keane (2010), yet differs markedly in terms of the theoretical foundations for this approach and its analytical focus.

To properly analyze the problems and populations of this paper, my model takes extra precision in reflecting the economic and policy environment that corresponds to each available year in the data. Mothers must elect to participate in welfare, and are subject to eligibility requirements and benefit calculations that approximate the salient policies from the year in which the data is taken. In addition, I incorporate the introduction of work requirements and time limits that were brought on by the reform act. This aspect of the problem bears similarity to the empirical and modeling approaches taken by Hoynes (1996), Keane and Wolpin (2002), and Swann (2005). In addition to
accurately capturing the welfare environment for each year and family in the sample, I ensure that the EITC for each family’s earnings is calculated exactly according to its historical parameters. To date, Chan (2013) is the only paper that takes an equally thorough environmental approach to modeling the labor supply and welfare participation decisions of single mothers. An additional advantage of explicitly modeling the policy changes that take places through the sample period is that I can exploit exogenous program changes to aid in identification of the model.

Although this paper focuses on questions of general policy design, the historical setting and data employed in estimation permit additional insights into the effect of welfare reforms. In particular, while the post-reform era has seen modest increases in family income, this has arrived in hand with increases in labor force participation and stricter eligibility requirements. The likely impact of this structural change depends on the estimated parameters of the production function, which determine the quantitative importance of time and money. By fitting the model to data, I derive predictions for both the average and distributional impacts of welfare reform. Work by Bernal and Keane (2010; 2011) suggests that there are potential negative impacts on child test scores when mothers join the labor force and use childcare. This of course may be countered by the positive impacts of additional income in the family budget, or by future growth in human capital induced by labor force participation (Blundell et al., 2015). Fang and Silverman (2009) argue that this may be one key benefit of imposing time limits on welfare, since it provides a credible commitment for time-inconsistent agents to join the labor force.

I estimate the model using data on single mothers from the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS). The CDS provides, in three waves (1997, 2002, 2007), important measures of development outcomes for selected children of PSID families. I study the formation of two measures of cognitive skill (verbal and math) and two measures of behavioral problems (externalizing and internalizing) reflected by scales taken from this supplement. Many studies, including the ones mentioned here, rely on evaluating test score impacts in terms of their relative scale. I exploit the longitudinal component of the PSID to tie each score to a human capital outcome of significant interest: high school graduation. This approach, known as “anchoring”, gives a more interpretable scale to estimates of the skill production function (Cunha and Heckman, 2008; Cunha et al., 2010). In addition, when considering the effect of policy counterfactuals, I simulate counterfactual graduation probabilities by estimating a model of high school graduation that depends on all four skills.

This paper offers progress on a number of high level questions. First, what is the causal importance of economic disadvantage in childhood in determining life-cycle outcomes? Second, how do current anti-poverty measures remediate environments for disadvantaged children? Finally, what is the scope (if any) for potential improvements that can be made by adjustments to these programs? In my analysis, I find evidence of important channels for both monetary and maternal time investment to shape the outcomes of children through each stage of their development. However,
the impacts of direct transfers to all sample families are modest: an annual lump sum payment of $1,000 to every family with a child younger than 17 produces positive but small test score gains. When skill outcomes are linked to the probability of graduating from high school, this initiative produces an increase in the graduation rate of 0.44% points. Most importantly however, there are large skill returns to expanding program coverage to poor, working mothers. When I impose a minimum income standard on all families in the sample, I find that the rate of high school graduation is boosted by 2.9% points for children in the poorest 10% of families.

This paper is the first to conduct policy analysis of government income supports and child development using a quantitative, behavioral model. In addition, it contributes a range of methodological and theoretical insights into how government policies can affect child outcomes and how such effects can be identified from data. Finally, I provide some important implications for current federal policies, showing that there are potentially large gains to be made by targeting programs towards the poorest families in the sample. The rest of the paper proceeds as follows. In Section 2 I describe the data used to estimate my model of parental investment, labor supply, and program participation. In addition, I identify important temporal trends in the data that have implications for future changes in policy. In Section 3, I introduce the model, presenting details and properties of its solution. Section 4 provides details regarding the estimation procedure I employ, as well as a discussion of identification. In Section 5 I use these estimated parameters to conduct counterfactual policy analyses, leading to conclusions in Section 6.

2 Data

To answer the empirical and policy questions outlined in Section 1, I use data from the Panel Survey of Income Dynamics (PSID) and its Child Development Supplement. The PSID is a dynastic, longitudinal survey taken annually from 1968 to 1997, and biennially since 1997. It collects important information on a range of economic and demographic indicators. The CDS consists of three waves, collected in 1997, 2002 and 2007. Any child in a PSID family between the ages of 0 and 12 at the time of the 1997 survey was considered eligible. These surveys contain a broad array of developmental scores in cognitive and socioemotional outcomes as well as information on the home environment of the child. One important feature of the survey is the availability of time use data, which is collected by the participants’ completion of time diaries. I provide further details below.

2.1 Sample Selection

Since much of the focus of the antipoverty programs considered here is on single mothers, sample restrictions are made to look more closely at this subpopulation. Since family structure in the data
is quite dynamic and tends to fluctuate, a stand must be taken on how to restrict the sample. In particular, I restrict attention to Mothers of CDS children who report themselves as household head in the 1997 and 2003 waves of the PSID survey. Very few mothers report themselves as household head in every year in which they appear in the survey. This subsample encompasses women with different marital histories and living arrangements, but it provides a reasonable subpopulation on which to focus. In particular, these women are household heads throughout the period of welfare reform. The great majority of them are household heads for the entire period between 1997 and 2003, which is the survey window on which I focus. I exclude any mother who appears in a household unit that is different from any of her CDS children in any year of the survey. Making this restriction leaves 364 mothers and 523 children.

From the PSID survey I use data (when available) on mothers’ labor supply, labor income, total family income, welfare receipt and some demographic variables. The CDS is comprised of several questionnaires. I use two in particular: the child interview and the primary caregiver (PCG) interview. From the child interviews, I use measures of cognitive ability as reflected by a battery of test scores. From the primary caregiver interview I use a number of behavioral scales constructed by the caregiver’s answers to a series of questions regarding the child’s behavior. In addition to this, the PCG completes a passage comprehension test, and a number of items from which scales on the caregiver’s self-esteem (Rosenberg, 1986) and self-efficacy (Pearlin, Lieberman, Menaghan & Mullan, 1981) can be constructed. When the PCG is identified as the mother of the child, I collect these scores also.

To measure child attributes, I utilize two measures of cognitive ability and two measures of behavioral traits. For cognitive ability, I use the Letter-Word (LW) and Applied Problems (AP) modules of the Woodcock-Johnson Aptitude test. For socioemotional or “non-cognitive” abilities, I use constructed scales that measure *externalizing* and *internalizing* behavioral problems. This gives, in total, four measures of child attributes that we use to track human capital outcomes. In the next section, I explore how each measure is related to adult outcomes, as proxied by high school graduation.

Finally, the CDS asks participant children to fill out a “time diary”. This portion of the survey requires participants to record a detailed, minute by minute timeline of their activities for two days of the week: one random weekday and one random day of the weekend. Activities were subsequently coded at a fine level of detail. When necessary, children are assisted in completion of the time diary by the PCG. These diaries provide a unique snapshot into the daily life of the child. From this data I construct a measure of maternal time investment by taking a weighted sum\(^4\) of the total hours of time use in which the mother is recorded as actively participating in each diary activity.

\(^4\)\(\frac{5}{7}\) for the weekday, and \(\frac{2}{7}\) for the weekend
Figure 2.1: Gradient of Skill Outcomes against 10-year mean of Household Income

### 2.2 Skill and Income Gradients

In the cross-section, how is income correlated with human capital outcomes of interest? Additionally, are the measurements of skills used in this project credibly related to important life-cycle outcomes? To answer these questions I examine two sets of gradients in the data: (1) The relationship between family income and skills; and (2) The relationship between skills and High School graduation. To measure income, I take the mean of annual household income for each child in their first 10 years of life. By this variable, I separate children into quintiles and take the mean score for each skill measurement. I plot these against the mean income level for each quintile in figure 2.1. From the graph it is clear that as income increases, cognitive achievement increases and the prevalence of problem behaviors decreases. Most striking is that the steepest portion of each curve is concentrated at the lowest level of income. A principal task of this project is to unpack this relationship into its causal and non-causal parts; however the visible shape of the gradient we find is suggestive of the fact that the marginal impact of income is greatest for families most in need.

In order to give some weight to the scale of cognitive and behavioral scores, I next inspect how each score is associated with the probability of high school graduation. For the cognitive scores, I divide children by the quartile of their Letter-Word and Applied Problems test score, plotting the mean of each bin against the probability of graduating from high school. For the behavioral scales (which are integer-valued), I group children into bins of size two, plotting the mean of each bin against the rate of graduation. Figure 2.2 shows the results. We can see that each skill is strongly associated with the probability of high school completion. The relationship is particularly pronounced for behavioral problems. Those with little indication of externalizing or internalizing behavioral issues have a very good chance of completing high school (about 95%) while those with the most severe behavioral issues in either dimension will only graduate with 65% probability.

---

5In doing so, I exclude those who obtain a GED qualification.
relationship with cognitive test scores is similar, albeit less dramatic: moving from the lowest to the highest quartile of either test score is associated with an increase in the likelihood of graduation between 10 and 15 percentage points.

In accordance with the patterns shown here, I estimate a series of linear probability models to give each skill outcome an interpretable scale. That is, for each skill \( k \), if \( \tilde{\theta}_{i,k} \) is the raw score of skill \( k \) for child \( i \), I estimate the model:

\[
H_{SG,i} = 1\{\text{const}_k + \tilde{\theta}_{i,k}\beta_{HSG,k} - U_i \geq 0\}, \quad U_i \sim \text{unif}[0,1] \tag{2.1}
\]

Thus, I re-weight each raw score to get \( \theta_{i,k} = \beta_{HSG,k}\tilde{\theta}_{i,k} \), so each skill has been “anchored” in the language of Cunha and Heckman (2008), to the probability of high school graduation. Estimates are presented in table A.1.

### 2.3 Trends in Labor Supply, Welfare, and Family Income

In this section I document some of the temporal and cross-sectional patterns of the dataset I employ in this project. I am particularly interested in documenting the effect of changes in the institutional and economic environment faced by mothers through the 1990s on their income, labor supply, and program participation. In addition to these patterns, it is important to examine the role played by welfare in the poorest households. This will provide insight into potential areas for improvement in policy design.

As discussed in Section 1, the 1990s saw significant and relevant changes to the economic environment faced by families with children. Figure 2.3 shows trends in labor supply during this time period. I plot the rate of labor force participation (LFP) and, conditional on labor force participation, the mean weekly hours spent at work by mothers in the sample. This plot demonstrates that, through this time, more mothers joined the labor force and, conditional on doing so, supplied more hours of labor to the market. In technical terms, we see an increase in
labor supply at both the extensive and intensive margins. This trend is likely due to a combination of the following effects: (1) the mean age of children in the sample is increasing, reducing the marginal importance of time at home; (2) an increase in the return to labor through rising wages; and (3) an increase in the returns to labor from expansion of the EITC and the switch from AFDC to TANF.

Mirroring this trend, Figure 2.4 shows a commensurate decline in the reliance on and generosity of welfare. There are two likely sources of this change. First, if mothers are earning more by expanding their labor supply, this decreases the rate of eligibility for welfare and, conditional on eligibility, also decreases receipt in welfare payment formulae. Second, the institutional changes from AFDC to TANF were built around the ambition to decrease reliance on government support and increase reliance on earned income. This is reflected in the visible decline in welfare participation between 1996 and 1997; the period in which the prescriptions of PROWRA were adopted by States. To assess further the likelihood of this effect, I plot the labor force participation rate and mean weekly work hours for welfare participants only. Figure 2.4 shows that, through this time period, welfare participants significantly increased their labor supply on both the extensive and intensive margins. This is an important temporal pattern that reflects the substantive changes to the welfare system in 1996. Notice in particular the significant increase in labor force participation from the year of reform (1996) onward. This is an important feature of the data that the model must capture.

What does the distribution of income in the sample look like, and how does this change over time? To answer this question, I compute the 10th, 20th, and 30th percentile of total annual income over the period between 1990 and 2002. Figure 2.6 plots the results. We can see that there is modest growth at each point of the income distribution over this time. Although there are issues with calculating total household income, and particularly so in poorer households, this graph makes it clear that there is a significant portion of CDS sample families who are surviving...
Figure 2.4: Trends in the rate of Welfare Program Participation and Welfare Receipt of CDS Mothers

Figure 2.5: Labor Force Participation and Hours of Labor Supplied, Conditional on Welfare Program Participation
on very few economic resources. For example, in 1997, the year of the CDS supplement, nearly 10% of the sample report living on less than $10,000 for the year. A further 10% report living on less than nearly $18,000 a year. Although this calculation does not take into account potential receipt of tax credits, unreported transfers, or SNAP benefits, it still gives a useful sense of how many members of the sample face legitimate economic constraints. To add further perspective, I note that the federal poverty guideline for a four-person household in 2001 is $17,650.

What role do labor supply and income supports play in the financial lives of the poorest families in this sample? To examine this, I compute several statistics of interest over three brackets for total annual income. I examine individuals in the sample from three separate brackets: defined by the 10th, 20th, and 30th percentile of total annual income. Figure 2.7 shows the mean labor force participation, weekly work hours, and wages for each income bracket. Although there is little difference in the labor supply behavior of the upper two brackets, we can observe that mothers in the lowest decile work less on both the extensive and intensive margins than their higher-income counterparts. The large jump in the participation rate of these women mirrors our previous observations on the effect of welfare reform during this time. Predictably, the ranking of mean wages corresponds to the position of each bracket in the income distribution. Thus, we observe that mothers in the lowest end of the income distribution work less and earn less for every hour supplied to the market than those in higher income households.

Next, I look for cross-sectional and temporal patterns in how households rely on federal income supports to supplement their budget. Figure 2.8 shows the rate of reported participation in
AFDC/TANF, the average annual receipt of payments for participating households, and the predicted level of EITC payment for which each household is eligible. To compute the latter statistic, I use the formula for federal EITC on sample mothers’ labor income. Since EITC parameters vary by year and by the number of children in the household, I use this information also. Since this excludes the labor income of any co-filing partner, it may not be a reliable estimate of the actual tax credit received by families. However, it should give a useful indication of the size and relevance of this income transfer for lower income households.

Turning to figure 2.8, several important features of the data appear. First, all three income deciles rely on welfare support to some extent, although the ranking of participation rates corresponds to their ranking in the income distribution in the way that we would expect. Second, families in the lowest income decile, who depend most on welfare, experienced the greatest decline in participation during this time. Third, conditional on participating in AFDC/TANF, there is
no appreciable difference in the amount of income received. This, despite the fact that payment formulae should, in theory, adjust for other sources of income. Fourth, although the EITC saw a significant expansion in generosity during this time period, the size of the credit is still less than the amount in welfare received by participating families. Finally, the lowest income households receive significantly less in tax credits than their upper two deciles. This is, naturally, driven by the fact that the EITC is increasing in labor earnings. This is, in turn, at least partly related to the fact that mothers in the lowest decile face lower wages in the labor market. To this extent, we may say that the EITC is regressive in its design. If one goal of antipoverty programs is to allocate resources to the most needy, this figure suggests that the EITC may not satisfy this goal.

I conclude this section by examining how total family income is composed in the cross-section and over time. Figure 2.9 plots the mean share (as reflected by its fraction of total reported income) of earnings, welfare, and the EITC in the household budget for each income bracket. Several important patterns are revealed. Households in the lowest decile rely much more on welfare and less on labor income. Second, these households experienced a significant decline (from 40% to less than 10%) in the fraction of their budget that is composed by welfare receipt. This decline is particularly sharp in the years of welfare reform. In this same time period, labor income begins to comprise a much larger share of total income. These patterns align with my other findings on the stark temporal effects of welfare reform on labor supply and welfare participation.

3 A Model of Child Development

As discussed in Section 2, the focus of this paper is on female-headed households. The model formulated in this section reflects this by positioning mothers as sole decision-makers in a family unit.
Although the model includes several stylized features, there are some clear advantages to the structure presented here. First, this approach can include an arbitrary number of child attributes at virtually no additional computational cost. This allows for a rich structure of skill formation. Second, the model admits a closed form, which establishes a clean link between potential instruments and identification of the production parameters of the model.

I abstract away from fertility and divorce decisions. In this setting, family structure is irrelevant to the production problem, while fertility is taken as exogenous. This is a direction for future extension. In addition, we abstract away from any concept of child autonomy in the family environment. Finally, to simplify exposition, I develop the case for a mother with one child. In section B, I show how this can be extended to multiple children with relative ease.

3.1 Environment, Production and Preferences

Time in the model is discrete and indexed by $t$ at an annual frequency. Let $\theta_t$ be a $N_\theta$-dimensional vector\(^6\) that describes the behavioral traits and capabilities of a child at time $t$. A period after birth, the child reaches maturity. Maturity is defined as the stage beyond which an individual’s skills ($\theta$) no longer develop. I let $t = 0$ be the year of birth, such that in period $A^7$, the production problem terminates. In every period $t = 0, ..., A - 1$, mothers choose their consumption ($c$), leisure ($l$), labor supply ($h$), and program participation ($I_w$). Importantly, they can also choose to make investments in their child in the form of money ($x$) and time ($\tau$). Period $A$ signifies the end of the mother’s decision problem, and at this point they receive a terminal payoff from final child abilities, which can be written as $V_A(\theta)$. In each period $t < A$, utility is derived from consumption, leisure and child abilities, which has cardinal value determined by the function, $u$. Finally, future payoffs are discounted exponentially by a factor, $\beta$. I assume that mothers make their decisions to maximize this discounted stream of payoffs.

There are two key ingredients that map parental decisions to the outcomes we wish to analyze: (1) resource constraints; and (2) the technology of child skill formation.

The economic substance of the problem is introduced by resource constraints. First, mothers cannot consume and invest beyond their budget constraint, which is in turn determined by the hours of labor, $h$, she supplies to the market, her non-labor income and her participation in the Government’s transfer or welfare program. Second, the sum of mothers’ leisure, time investment, and labor supply must equal the number of hours in one period, which we can normalize to 1. Thus, mothers face a trade-off between earned income through the supply of labor and available hours that may be spent in child-relevant activities or private leisure. I let $\mathcal{H}$ indicate the set of hours from which mothers can choose. In the empirical application that follows, I make labor supply a discrete choice with $\mathcal{H} = \{0, 10, 20, 30, 40\}$. Finally, let $I_w \in \{0, 1\}$ indicate the mother’s decision to

\(^6\)In my empirical application, I set $N_\theta = 4$
\(^7\)I set $A = 17$ in this paper.
participate in the Government’s welfare program, the structure of which I formalize below. In what follows, let $S_t$ be a vector that tracks all variables that are relevant for determining the household budget in period $t$. With this variable, we can write the budget function as $B(h, I_w, S_t)$. $S_t$ will contain variables that are relevant to the determination of wages, non-labor income, as well as the parameters that summarize federal transfer policies.

The evolution of each child’s skills $\theta_{i,t}$ over time is determined by a production function. I specify that this technology takes the familiar Cobb-Douglas form:

$$\theta_{k,t+1} = \psi_t x_{t,k,t} \delta_{x,k,t} \tau^{\delta_{x,k,t}} \prod_{j=1}^{N_k} \theta_{j,t}^{\delta_{g,k,j}}$$

(3.1)

Here, $\delta_{x,k,t}$ ($\delta_{r,k,t}$) is the Cobb-Douglas share of monetary (time) investment in the production of skill $k$ at age $t$. Similarly, $\delta_{g,k,j}$ indicates the share of the current stock of skill $j$ in the production of next period’s future stock of skill $k$. The coefficient $\psi_t$ represents total factor productivity in production, which I assume is determined as:

$$\log(\psi_t) = X \beta_t + \eta_t,$$

(3.2)

where $X$ are observable measures of the mother’s human capital and $\eta_t$ is a time-varying component that is known to the mother, but unobservable in our data.

This specification introduces particular dynamics to skill production. For example, the returns to monetary investment at time $t$ are not solely determined by the shares $\delta_{x,t}$, but also by how skills in period $t + 1$ shape the formation of skills in period $t + 2$, and so on.

One should also note the dependance of the development process on $t$. This is a natural way to specify production, since it is well known that developmental sensitivities change as children age. Later, I will take a stand on how the age of the child maps to the relevant developmental stage.

### 3.2 Family Budget and The Policy Environment

To implement the solution described above it is necessary to unpack the components of the budget function, $B(h, I_w, S_t)$, and specify the members of the state vector, $S_t$. We can specify this as:

$$B_{i,t} = w_{i,t} h_{i,t} + N_{i,t} + T^E(w_{i,t} h_{i,t}; \pi^E_t) + I_w T^W(N_{i,t}, w_{i,t} h_{i,t}, h_{i,t}; \pi^W_t)$$

(3.3)

In this expression, $w_{i,t} h_{i,t}$ is earned income from labor hours supplied to the market, $N_{i,t}$ is non-labor income, and $(T^E, T^W)$ are functions that map these variables to a tax credit and welfare payment, respectively. Each function $T$ is indexed by policy parameters $\pi_t$ that vary over time.

The tax credit is defined by a phase-in rate, a maximum credit, a phase-out income level, and a phase-out rate: $\{\pi_{0,i,t}, \pi_{1,i,t}, \pi_{2,i,t}, \pi_{3,i,t}\}$. It can be written as:

$$T^E_{i,t} = \begin{cases} 
\min\{\pi_{0,i,t} h_{i,t} w_{i,t}, \pi_{1,i,t}\} & \text{if } B_{i,t} \leq \pi_{2,i,t} \\
\max\{\pi_{1,i,t} - \pi_{3,i,t} B_{i,t}, 0\} & \text{if } B_{i,t} > \pi_{2,i,t} 
\end{cases}$$

(3.4)
In my empirical implementation of the model, these parameters are allowed to modify according to the number of children in the family. Similarly, welfare policy is defined by two variables: eligibility and receipt. Although it should be noted that there are precise, institutional functions that map the circumstances of each mother to welfare eligibility and payments, these are highly complicated, require more information than is available in our dataset, and are often subject to the discretion of case workers. In addition to this, such formulae may be accurate at a monthly frequency, but it is not clear if they hold any advantage over an approximation when aggregated to an annual frequency.

Accordingly, I approximate eligibility and welfare receipt as follows. For all years $t$ prior to 1996, I use an approximation to AFDC that follows:

$$T_{i,t}^W = E_{i,t} \times \max\{\pi_{0,0} + X_{i,t} \beta_{WF} + \epsilon_{WF,i,t} - \pi_{DR,0} h_t w_t, 0\}$$

(3.5)

$$E_{i,t} = 1\{X_{i,t} \beta_E + \epsilon_{E,i,t} > 0\}$$

(3.6)

For 1996 and all later years, I adjust the payment and eligibility standards as follows:

$$T_{i,t}^W = E_{i,t} \times \max\{\pi_{0,1} + X_{i,t} \beta_{WF} + \epsilon_{WF,i,t} - \pi_{DR,1} h_t w_t, 0\}$$

(3.7)

$$E_{i,t} = 1\{TL_{i,t} \leq 5\} \times 1\{X_{i,t} \beta_E + \pi_{WR} 1\{h_t \leq 20\} + \epsilon_{E,i,t} > 0\}$$

(3.8)

$$TL_{i,t+1} = TL_{i,t} + I_{i,t}$$

(3.9)

$$TL_{i,96} = 0$$

(3.10)

In both cases, I allow for a vector $X_{i,t}$ of observables to affect both payment and eligibility in the pre and post-reform era. Some of the members of $X_{i,t}$ are natural candidates for institutional reasons (for example, the number of dependant children in the household and the total income earned by the household), others are included with the shocks $(\epsilon_{WF,i,t}, \epsilon_{E,i,t})$ to capture the fact that, despite being technically eligible, many mothers do not participate in welfare. The welfare schemes differ in four important ways. First, they differ in the extent to which earned income is disregarded in the calculation of the payment: $(\pi_{DR,0}, \pi_{DR,1})$. Second, they differ in their base level of payment generosity: $(\pi_{0,0}, \pi_{0,1})$. Third, the post-reform scheme imposes a punishment, in terms of eligibility, on those that work less than 20 hours a week: $\pi_{WR}$. This detail is taken directly from the stipulations of the reform act, however there is considerable variation in terms of how exceptions to this rule are circumstantially applied. Lastly, the latter scheme imposes a time limit of 5 years on welfare receipt, in line with the requirements of TANF.

The policy parameters discussed here are all estimated to fit observed patterns on welfare participation and receipt in the data.

### 3.3 The Family Production Problem

Having described the environment and family preferences, we can move on to stating the dynamic programming problem faced by the family unit. I proceed by stating the problem, then describing
the relevant notation. Recall that the terminal period of the problem occurs at $A$, the time period at which the child matures. I state below the problem for every time period $t < A$:

$$
V_t(\theta_t, S_t, \eta_t) = \max_{c,l,x,h} \left\{ u(c, l, \theta_t) + \beta \mathbb{E}[V_{t+1}(\theta_{t+1}, S_{t+1}, \eta_{t+1}) | S_t] \right\}
$$

(3.11)

Subject to the constraints:

$$
c + x \leq B(s_t, h, I_w)
$$

(3.12)

$$
l + \tau + h = 1
$$

(3.13)

$$
\theta_{k,t+1} = \psi_t x^\Delta r_{k,t} \tau^\Delta r_{k,t} \prod_{j=1}^{4} \theta_{j,t}^{\Delta \theta_{k,j}} \text{ for } k = 1, 2, \ldots, N_{\theta}
$$

(3.14)

In this setup, equation (3.12) is the budget constraint. The function $B$ defines the budget set for the family as a function of parental labor supply, $h$, program participation, $I_w$, and relevant economic state variables, $S_t$. In this setup, $S_t$ is permitted to evolve dynamically with parental decision making. Most importantly, $S_t$ includes all variables that summarize the policy environment faced by mothers, including the welfare and tax policy applicable at time $t$. One particularly important dynamic component of the decision making problem will involve the introduction of time limits, which I detailed in section 3.2. Finally, (3.13) describes the time constraint faced by the mother.

To close the model, we specify the form of utility, $u$, and the terminal value function:

$$
u(c, \tau, \theta) = \alpha_c \log(c) + \alpha_l \log(l) + \alpha_\theta \cdot \log(\theta)
$$

(3.15)

$$
V_A(\theta) = (1 - \beta)^{-1} \alpha_\theta \log(\theta)
$$

(3.16)

Finally, for notational convenience, we can write the production function in the following vector notation:

$$
\log(\theta_{t+1}) = \Delta x_{t} \log(x) + \Delta r_{t} \log(\tau) + \Delta \theta_{t} \log(\theta_t) + \eta_t
$$

(3.17)

For reasons made clear in Appendix B, this assumption is pivotal in reducing the computational complexity of the problem. Since $\theta$ is a vector, $\Delta x_{t}$ and $\Delta r_{t}$ are $N_\theta$ dimensional vectors of Cobb-Douglas shares. Assuming log utility, as in (3.15), facilitates the derivation of closed-form expressions for investment policies. This is also demonstrated in the model solution found in Appendix B. Finally, I specify that the terminal value $V_A$ is equal to the discounted present value of utility derived from each child’s final abilities over an infinite horizon 8.

Assumptions (3.15) and (3.14) lead to the following simplification of the dynamic program. In Appendix B, I formally derive the following expressions in the general case that allows for multiple children. Let me relegate technical details to that section, presenting here the substantive model implications and some informal discussion. The value function can be written as additively separable in the skill vector $\theta$:

---

8In many cases, this specification corresponds exactly to the right value for a fully-specified infinite horizon problem.
\[ V_t(\theta, s_t, \eta_t) = \alpha_{V,t} \log(\theta_t) + \alpha_{V,t+1} \log(\eta_t) + \nu(S_t) \]  
(3.18)

\[ \nu(s_t) = \max_{h \in H_i} \left\{ \alpha_{c,t} \log(B(h, I_w, S_t)) + \alpha_{l,t} \log(1 - h) + \beta \mathbb{E}[\nu(S_{t+1}) \mid S_t] \right\} \]  
(3.19)

In this formulation, the parameters \( \alpha_{c,t} \) and \( \alpha_{l,t} \) are aggregates that represent the total value of income and leisure, respectively. They can be written as:

\[ \alpha_{c,t} = \alpha_c + \beta \alpha_{V,t+1} \Delta_{x,t} \]  
(3.20)

\[ \alpha_{l,t} = \alpha_l + \beta \alpha_{V,t+1} \Delta_{r,t} \]  
(3.21)

We see that the utility derived from income is composed of two terms: \( \alpha_c \) is of course the value derived from private consumption, while \( \beta \alpha_{V,t+1} \Delta_{x,t} \) is the marginal return of investment to each skill, scaled by the value derived from each skill next period, \( \alpha_{V,t+1} \), and discounted by \( \beta \). The value derived from skills in each period, \( \alpha_{V,t} \) can itself be expressed recursively:

\[ \alpha_{V,t} = \alpha_\theta + \beta \alpha_{V,t+1} \Delta_{\theta,t} \]  
(3.22)

The value from skills is the sum of utility derived from each skill today (\( \alpha_\theta \)) and the return of each skill to the production of future skills, scaled by the value of these in the next period (\( \beta \alpha_{V,t+1} \Delta_{\theta,t} \)). Since terminal utility \( V_A \) also takes this log-linear form, we can see how this recursion holds in each period. Finally, preservation of the log-additivity of the value function leads to the following proportional investment rules:

\[ x_t = \frac{\beta \alpha_{V,t+1} \Delta_{x,t}}{\alpha_{c,t}} B(h, I_w, S_t) \]  
(3.23)

\[ \tau_t = \frac{\beta \alpha_{V,t+1} \Delta_{r,t}}{\alpha_{l,t}} (1 - h) \]  
(3.24)

This formulation of the problem greatly simplifies computation, since the value function in terms of the skill vector \( \theta \) can be solved in closed form. The key to this result is that, subject to a log transformation, the current realization of \( \theta \) does not affect the productivity of investments. Thus, the value of current \( \theta \) can be expressed as a linear combination of period utility and the discounted return to future skills, as defined by the share of \( \theta \) in production, \( \Delta_{\theta,t} \). This is formalized in expression (3.22). Since we have log-utility, these coefficients describe in (3.23) the relative share of income spent on the child and in (3.24) the relative share of non-labor hours spent in time investment. The proportional investment rules, when substituted into the dynamic program, allow us to simplify the problem to one of labor supply and program participation, as shown in (3.18) and (3.19). The coefficients \((\alpha_{c,t}, \alpha_{l,t})\) define the labor supply problem, and adjust each period to reflect the relative importance of time and money in the production of child skills.
4 Estimation

4.1 Heterogeneity and Preferences

The model admits a normalization in the scale of utility. Thus, I set $\alpha_c = 1$. In addition, I let mothers value their leisure $\alpha_l$ in the same way, while allowing the extent to which they privilege their child’s outcomes ($\alpha_\theta$) to vary in the population.

In this setup, I allow preferences over child outcomes to be heterogeneous. Although $\alpha_\theta$ is an $N_\theta$-dimensional vector, I let the heterogeneity across mothers be defined by a scalar, $\zeta$, that scales parental utility derived from $\theta$.

$$\alpha_{\theta,i} = \zeta_i \bar{\alpha}_{\theta}, \quad \zeta \sim G_\zeta \tag{4.1}$$

Indexing preference heterogeneity by a scalar greatly reduces the complexity of the estimation problem, however it also provides a convenient simplification of the identification problem, which I discuss in the following section.

4.2 Identification

The results of later policy experiments hinge crucially on the importance of time and money investments on child outcomes (summarised by parameters $\Delta_c$ and $\Delta_r$) and on the dynamic complementarities of skill production (embodied in $\Delta_\theta$). In this section I briefly discuss how these parameters can be identified without leaning too heavily on the structure or assumptions of the behavioral model.

To see this, begin by taking the outcome equation (3.14) and substituting in the investment policies (3.23) and (3.24). Let $B_t$ indicate total family income in period $t$. We obtain the expression:

$$\log(\theta_{k,t+1}) = \Delta_{c,ak} \log(B_t) + \Delta_{r,ak} \log(1 - h_t) + \Delta_{\theta,ak} \log(\theta_{k,t}) + \varepsilon_{k,t} + \eta_{k,t} \tag{4.2}$$

$$\varepsilon_{k,t} = \Delta_{c,ak} \log(\phi_{c,t}) + \Delta_{r,ak} \log(\phi_{r,t}) \tag{4.3}$$

This expression relates child outcomes to two important observables: total income, $B_t$, and total available hours $1 - h_t$. In addition, the expression incorporates unobservables $\varepsilon_{k,t}$ and $\eta_{k,t}$. While the inclusion of $\eta_{k,t}$ is innocuous in this setup, $\varepsilon_{k,t}$ warrants further comment. First, note that it serves as a “fixed effect” in the sense that the error term is a function of the investment shares $(\phi_{c,t}, \phi_{r,t})$ which are themselves functions of time-invariant preferences. However, this mapping is time-varying due to the dynamic structure of skill production. Second, equation (4.3) shows that the $i$th component of $\varepsilon_{k,t}$ is the sum of the Cobb-Douglas share of each investment type for skill $i$, multiplied by the fraction of either dollars or hours that is dedicated to this investment type. Thus, parents that have a greater preference weight on child outcomes will have higher realizations of $\varepsilon_{k,t}$ and hence will appear more productive in this specification. This presents an
identification problem because $\varepsilon_{k,t}$, if indeed unobservable, is correlated with realizations of income and labor supply. There are two reasons for this. First, all three variables are a function of maternal preferences. Second, these preferences may also be correlated with state variables $S_t$ that govern the budget constraint and hence both endogenous variables.

In this paper I propose two methods of identification. First, let the state $S_t$ be decomposed into $[Z_t, W_t]$ where $Z_t$ represents variables that are exogenous to $\zeta$, which indexes preference heterogeneity in this model. We can identify production parameters using the moment condition:

$$
\mathbb{E}[\log(\theta_{k,t+1}) - \Delta_{c,a_k} \log(B_t) - \Delta_{r,a_k} \log(1 - h_t) - \Delta_{\theta,a_k} \log(\theta_{k,t}) \mid Z_t] = 0. \quad (4.4)
$$

Since there are two important and exogenous policy changes that transpire during the relevant sample period, these propose a natural set of instruments from which to identify parameters.

In addition to this strategy, I exploit the availability of unique data on time investment from the CDS sample. In the model solution, equation (B.13) shows that the share of available time devoted to time investment, $\phi_{r,t}$, is a function of preferences, $\alpha_\theta$, and the ages of each child in the family, $\mathbf{a} = \{a_k\}_{k \leq N_k}$. In addition, my assumption that heterogeneity in preferences is indexed by a scalar variable, $\zeta$, implies that there is a monotonic one-to-one mapping between $\zeta$ and $\phi_{r,t}$ for any given $\mathbf{a}_t$. From the time diary supplement, we can use the time investment aggregate, $\tau_{97}$, to construct a measurement of $\phi_{r,97}$:

$$
\phi_{r,97} = \frac{\tau_{97}}{1 - h_{97}} \quad (4.5)
$$

By the argument above, since the mapping from $\zeta$ to $\phi_{r,97}$ is monotonic, we can invert this function to give:

$$
\zeta = \Phi^{-1}(\phi_{r,97}, \mathbf{a}) \quad (4.6)
$$

This permits the derivation of the following condition:

$$
\mathbb{E}[\varepsilon_{k,t} \mid h_t, B_t, \phi_{r,97}, \mathbf{a}] = \mathbb{E}[\varepsilon_{k,t} \mid \phi_{r,97}, \mathbf{a}] = \varphi(\phi_{r,97}, \mathbf{a}) \quad (4.7)
$$

Thus, the addition of a nonparametric control function in the specification (4.2) allows us to add the typical orthogonality restrictions of non-linear regression to the moment restrictions given by the instruments:

$$
\mathbb{E}[\log(\theta_{k,t+1}) - \Delta_{c,a_k} \log(B_t) - \Delta_{r,a_k} \log(1 - h_t) - \Delta_{\theta,a_k} \log(\theta_{k,t}) - \varphi(\phi_{r,97}, \mathbf{a}) \mid h_t, B_t] = 0 \quad (4.8)
$$

In combination, these moment conditions allow us to identify the production parameters of the model, $(\Delta_c, \Delta_r, \Delta_\theta)$, without using assumptions regarding the co-dependance of unobservables on the choice variables, $(B_t, h_t)$.

### 4.3 Estimation Method

I estimate the model using minimum distance on a collection of statistics from the data. Most importantly, I include the vector of covariances that collectively define the Instrumental Variable
and Control Function estimates implied by equations (4.4) and (4.8). I use applicable EITC and welfare parameters in each year, as well as state-level unemployment rates as potential instruments. I use a 2nd order polynomial in the empirical measure of $\phi_{r,g_t}$ described in the previous section. In addition to these statistics, I match a series of interactions between explanatory variables with wage and non-labor income. I also match labor supply, welfare receipt and participation in the pre and post-reform eras. In total, I compute 152 moments from the data. Table A.5 shows a subset of these moments. Let $\Gamma$ be the full vector of moments and statistics computed from the data. In addition to the parametric assumptions described above, to simulate the model I assume that $\epsilon_{TW}, \epsilon_w, \epsilon_N$ are normally distributed with mean zero and standard deviations $(\sigma_{TW}, \sigma_w, \sigma_N)$. I additionally assume that $\zeta$ is log-normally distributed with mean $\mu_{\alpha,\theta}$ and standard deviation $\sigma_{\alpha,\theta}$. Let $\omega$ signify the full vector of parameters of the model, and let $\Gamma(\omega)$ be the vector of simulated moments that corresponds to $\Gamma$. The classical minimum distance estimator, $\hat{\omega}_{MD}$ is produced as:

$$\hat{\omega}_{MD} = \arg\min_{\omega} (\Gamma(\omega) - \Gamma)'W(\Gamma(\omega) - \Gamma).$$

I choose the weighting matrix $W$ to have zeros in the off-diagonal entries, and $1/V(\Gamma_i)$ in the diagonal entry at position $(i,i)$, where $V(\Gamma_i)$ is the variance of the $i$th statistic, calculated by bootstrap resampling.

### 4.4 Results

Table A.2 presents the results for the baseline model. In this specification, we allow non-labor income and wages to vary according to: mother’s education ($ED_i$), the mother’s score on a passage comprehension test ($PC_i$) and a self esteem scale ($SE_i$) administered in the PCG survey, mother’s age ($age_{it}$), the state level unemployment rate ($unemp_{it}$), and the year ($Y_{it}$). The results indicate that each variable is important in explaining some portion of the variation of wages and non-labor income. I estimate the standard deviation in preferences over child skills ($\log(\alpha_a)$) to be 0.5625 which suggests that some of the variation in child outcomes is driven by how mothers invest in their children, even when conditioning on financial and temporal resources. In addition, I estimate significant variation in the determination of wages, non-labor income, and welfare transfers ($\sigma_w, \sigma_N, \sigma_{TW}$). The latter is particularly striking, with $\sigma_{TW} = 3.4627$. The minimum payment standard for families in the Pre-TANF era (represented by the constant contribution to $T_{it}^W$) is estimated at about $5,860 a year (equivalent to about $113 a week) with close to a 100% marginal tax rate on this payment, as reflected by $\pi_{DR,0}$ being estimated so close to 1. By contrast, the approximate welfare system under TANF displays a much lower marginal tax rate ($\pi_{DR,1} = 0.63$) with a less generous income standard ($$65 a week) and a work requirement standard ($\pi_{WR} = 126.3$) which effectively rules out any payment when labor supply is less than 20 hours a week. Taken broadly, these estimates seem to reflect the real policy changes occurring during this time.
In the baseline, production parameters are specified to vary with age according to:

\[ \delta_{S,J,z,a} = \exp(\gamma_{S,J,z,0} + \gamma_{S,J,z,a} \times a) \] for \( S, J \in \{LW, AP, BE, BN\} \), for \( z \in \{x, \tau\} \) \hspace{1cm} (4.10)

By contrast, estimates of \( \Delta_{\theta} \) are not permitted to vary by age, but are permitted to take negative values. Table A.3 presents the estimates of these parameters, in addition to the contribution of mother’s education and passage comprehension to skill production.

However recall that the production technology, when taken in logs, adopts a VAR structure. Thus, there are significant dynamic interactions over time that make these parameter estimates difficult to interpret. For example, the marginal response of skills in period \( t+s \) to monetary investment in period \( t \) can be expressed as:

\[ \frac{\partial \log(\theta_{t+s})}{\partial \log(x_t)} = \left( \prod_{j=1}^{s-1} \Delta_{\theta,t+j} \right) \Delta_{x,t} \Delta_{s-1} \Delta_{x,t}, \] \hspace{1cm} (4.11)

where the second equality above follows under the simplifying assumption that \( \Delta_{\theta,t} = \Delta_{\theta,s} \) for all \( t, s \). Thus, to interpret first the overall importance of each investment type, I calculate the total contribution of each investment type to final skills. This is facilitated by the representation of final period skills as:

\[ \log(\theta_{TM}) = \sum_{t=0}^{T_{M}-1} \Delta_{\theta,t}^{T_{M}-t-1} (\Delta_{x,t} \log(x_t) + \Delta_{\tau,t} \log(\tau_t) + X \beta_{\theta}) \] \hspace{1cm} (4.12)

Thus, in Table A.4 I report the following set of coefficients:

\[ \sum_{t=0}^{T_{M}-1} \Delta_{\theta,t}^{T_{M}-t-1} \Delta_{x,t}, \quad \sum_{t=0}^{T_{M}-1} \Delta_{\theta,t}^{T_{M}-t-1} \Delta_{\tau,t}, \quad \sum_{t=0}^{T_{M}-1} \Delta_{\theta,t}^{T_{M}-t-1} \beta_{\theta} \]

The first two columns of the table show the overall contribution of mother’s education and passage comprehension score to final skill outcomes. Signs are mixed in this case, indicating that there may not be strong patterns between maternal human capital and child outcomes in this sample. The third and fourth columns show the contribution of money and time investments to final skills. Each coefficient in these columns can be interpreted as the percentage change in skills (scaled by its association with high-school graduation) in response to a 1% increase in each investment type. For example, a 1% increase in monetary investment in every period leads to a decrease in externalizing behaviors that is associated with a 10% point increase in the rate of high school graduation. This magnitude of elasticity is striking, however it is important to remember two caveats. First, although my empirical strategy argues for a causal interpretation of the relationship between investment and skills, the anchoring relationship between skills and graduation is not. Second, the elasticities presented here cannot be interpreted as policy parameters. To use the above example, a lump-sum transfer of income, aimed at increasing each household’s budget by 1%, would cause a reduction in labor supply, offsetting the intended effect of the transfer and causing a substitution away from monetary investment into time investment. In the case of externalizing
behaviors, table A.4 indicates that time investment is less productive and thus, a transfer of this kind cannot replicate the numbers drawn from our interpretation of the production parameters.

Table A.4 demonstrates that both investment types have a significant role to play in the formation of child skills. In particular it appears that in aggregate, time is more influential in shaping cognitive outcomes, while money is more influential when it comes to behavioral problems. To get a more precise picture of how investment shares vary with age, I plot the contribution of time and money investments to final skills in figure 4.1. Recall that, according to expression (4.12), the contribution of monetary and time investments in time $t$ to final skills can be expressed as:

$$\Delta T_{M}^{t-1} \Delta x_{t}, \quad \Delta T_{M}^{t-1} \Delta \tau_{t}$$

Figure 4.1 shows interesting dynamic patterns in the importance of different investments by age. In the production of behavioral traits, time is initially more influential, but cedes importance to monetary investments later in the development process. Contrastingly, contributions of either investment to the production of cognitive skills do not become significant until late childhood or early adolescence.

The estimates of production parameters imply that time and money investments both play a crucial role in the formation of cognitive and behavioral traits. This should serve as confirmation that differently designed labor market and income transfer policies have the potential to produce
markedly different human capital outcomes. However, the production parameters discussed above do not provide a complete picture of how to leverage policies to improve child outcomes. In the next section, I consider the potential effect of several policy initiatives that attempt to improve income support for poor families.

5 Policy Analysis

With estimates of the model in hand I next consider how the production parameters, in concert with the labor market supply decisions of mothers, determine the effectiveness of particular policy initiatives. In particular, I consider the effect of policies that are aimed at increasing mothers’ financial resources. To do this, I calculate the mean in total income ($B_t$), hours ($h_t$), welfare receipt ($T_W^t$) and total program cost per mother, per year. To measure the developmental effect of these initiatives, I calculate the mean change in each skill, as anchored by their association with the probability of high school graduation. In addition, I calculate a counterfactual graduation rate by estimating a linear probability model of high school graduation on all skills, jointly. That is, I estimate:

$$HSG_i = 1\{\theta_i \beta_{HSG} - U_i \geq 0\}, \quad U \sim \text{unif}[0, 1],$$

from the data, and report the average change in graduation probabilities implied by the change in skills. Estimates for this model are presented in Table A.1. In the discussion that follows, I suppose that this outcome equation is valid, and therefore discuss counterfactual graduation rates with causal language. However, one can alternatively suppose that this equation is simply a means by which we can weight aggregate skill outcomes: by their respective association with high school graduation in a jointly estimated model.

I first consider two simple interventions that are designed to give a sense of the potency of additional income in the general population. These are (1) a lump sum transfer of $1,000 per year, to each family in the sample; and (2) a wage subsidy, valued at $1/hour, offered to each mother in the sample. Note that these policies are not designed to be progressive and are not targeted at the bottom end of the income distribution. In this sense they are clearly not optimal, but they will assist in giving insight to how particular policy initiatives will transfer into impacts. Table 5.1 shows the results from these simulations.

As was expected, the lump sum policy does not succeed in increasing family income by the same magnitude as the transfer: the mean impact is about $650 a year. This is because mothers adjust their labor supply, to the order of just over one hour per week. The decrease in labor earnings that results offsets the income effect of the transfer. This, in turn, increases welfare payments for participating families, which explains why the additional cost of the initiative is in fact a little more than the original $1,000 per family. There is a strong contrast between these impacts and the**

---

9Cost is computed in thousands of dollars, per family, per year.
Table 5.1: Overall Impacts of Policy Experiments

<table>
<thead>
<tr>
<th>Case</th>
<th>$B_t$ ($000s)</th>
<th>$h_t$</th>
<th>$T^W_t$ ($000s$)</th>
<th>Cost$^b$</th>
<th>$\Delta LW$</th>
<th>$\Delta AP$</th>
<th>$\Delta BE$</th>
<th>$\Delta BN$</th>
<th>$\Delta HSG$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>24.8936</td>
<td>23.7907</td>
<td>0.7769</td>
<td>1.9367</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Lump Sum</td>
<td>25.5149</td>
<td>22.7117</td>
<td>0.8399</td>
<td>2.9819</td>
<td>0.0014</td>
<td>0.0023</td>
<td>0.0047</td>
<td>0.0057</td>
<td>0.0044</td>
</tr>
<tr>
<td>Wage Subsidy</td>
<td>26.4180</td>
<td>24.5511</td>
<td>0.7327</td>
<td>3.1862</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0068</td>
<td>0.0075</td>
<td>0.0049</td>
</tr>
<tr>
<td>Welfare Changes</td>
<td>25.1425</td>
<td>21.5692</td>
<td>1.2602</td>
<td>2.3613</td>
<td>0.0005</td>
<td>-0.0035</td>
<td>0.0081</td>
<td>0.0144</td>
<td>0.0025</td>
</tr>
<tr>
<td>Minimum Income</td>
<td>24.1996</td>
<td>17.3335</td>
<td>1.9528</td>
<td>1.9528</td>
<td>0.0023</td>
<td>0.0021</td>
<td>0.0021</td>
<td>0.0066</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

As we can see, the wage subsidy is more successful in boosting income, since the transfer is attached to increases in labor supply. On average, mothers work an extra three quarters of an hour every week, and enjoy an extra $1,500 dollars a year in their budget set. An additional consequence of their increased labor supply is a decrease in their reliance on welfare support. Thus, my estimates predict that additional work incentives provided to mothers would be successful in boosting incomes, increasing labor force attachment, and decreasing reliance on welfare. In this sense, policymakers might be inclined to prefer this method of income support over the unconditional transfer, which does not increase income at a dollar-for-dollar rate, discourages labor supply, and increases reliance on transfers. However, this perspective misses an important part of the story, which is made clear when we look at the developmental impacts of each policy change.

Returning to Table 5.1, we see that both policies have roughly the same total effect on high school graduation. However, the direct transfer achieves this at a cost of $237 per percentage point increase in graduation rates, while the wage subsidy costs $255 per percentage point. This magnitude is much larger when aggregated across the number of students to which such a policy might apply. This result is intuitively reasonable and unsurprising: a lump sum transfer is unattached income, which means that resources increase without any adjustment in labor supply. Since time inputs are also critical in this model, unconditional transfers are by definition more effective than transfers of equivalent size that are attached to increased labor supply (which reduces time investment).

Of further note is the difference in how the developmental impacts of these policies are distributed across skills. The lump sum transfer has greater impacts on cognitive abilities, while the wage subsidy is more impactful in preventing behavioral problems. This difference is explained by the relative importance of time and money in the production function. Increasing the wage induces substitution away from total leisure hours towards income. Thus, according to the relative contributions of money and time (summarised in Table A.4), behavioral traits flourish at the relative
expense of cognitive abilities. The converse logic applies to the lump sum, which allows mothers to reduce their labor supply and increase their time investment.

I next consider two policy experiments that aim for greater efficacy by specifically targeting the poorest mothers in the sample, since this group will enjoy the greatest gains from expansion in income support programs. First, as a quantitative experiment, I simulate an adjusted version of the 1996 reforms. In this experiment, I introduce in 1996 an alternative version of TANF. Under the new reform, I simulate the same changes in payment standards and earnings disregards, but I do not impose the 5 year time limit, nor do I impose the component of payment standards designed to simulate work requirements. As indicated by Table 5.1, these adjustments to welfare reduce the labor supply, and increase the welfare reliance of mothers. We see modest changes in cognitive skills, and large gains in behavioral traits. These effects combine to produce a program cost of $177 per percentage point gain in high school graduation rates. Of particular note is that this initiative is more efficient than the previous two quantitative experiments. This is because the policy change directs resources towards poorer families, for whom the marginal return to extra income and hours is higher.

This result implores us to pause and reassess, from section 2.3, my findings on recent trends in the budget set of families. It is particularly important that we consider these patterns in light of the model estimates and the results of these policy experiments. Recall that we saw, in figures 2.7, 2.8 and 2.9, that the poorest mothers in our sample have dramatically reduced their welfare participation and replaced this income with increased labor earnings. By 1998, less than 15% of mothers in the lowest income decile have welfare coverage, while nearly 40% of them work. In the same year, these mothers report annual budgets of less than $12,500. This poses an important question: what are the returns to extending coverage of federal income supports to mothers who occupy this section of the income distribution? We have seen already that providing income assistance to the poorest families yields the highest returns.

Additionally, as many income assistance programs have become increasingly tied to labor force participation, we must ask how relatively effective this means of income support is when compared to unconditional transfers.

To evaluate these questions, I next propose a policy reform that radically alters the welfare landscape for poor mothers. Under this alternative system, every mother is guaranteed a minimum income of $6,250 a year. This minimum payment is phased out at a rate of 50% as mothers earn labor income or claim other sources of income. Of course, such a payment system must address questions of implementation in practice. For now, I focus on the potential benefits of the program, subject to implementability.

The aggregate results of this reform can be found in the bottom line of Table 5.1. First, note that the program is close to being revenue neutral. Thus, we can think of the benefits of the program as being attainable at nearly no extra cost. Note that the lump sum, combined with
an effective marginal tax rate of 50%, provides a strong disincentive to work. This is reflected in the significant reduction of mean labor hours. This effect is, in fact, strong enough to offset any potential additions to income that expanding program coverage might have achieved. However, the developmental results of the reform speak for themselves: children enjoy aggregate increases in both cognitive and behavioral skills, with an effective increase in the high school graduation rate of 0.2 percentage points.

These results are more compelling when we look at how impacts are distributed across income levels. To do this, I divide children into income brackets, defined by the 10th,20th and 30th percentile of total income\textsuperscript{10} in the Baseline simulation. I then calculate the mean changes in skills in each income bracket. Table 5.2 shows the results. As a revenue neutral experiment, the Minimum Income system induces an almost 3% point increase in the graduation rate for the poorest children in the sample. These gains are spread relatively evenly across skill types.

As a point of comparison, Table 5.2 presents the same distributional impacts from the other policy experiments. Of particular interest is the fairly modest progressivity of the modified TANF

\textsuperscript{10}To define this variable over multiple years, I take the mean family income for each child over the simulated panel.
program, which I have named the “Welfare Changes” experiment. There are two features that largely drive this result. First, the proposed reforms in this experiment do not negotiate the issues of coverage and eligibility that I have discussed. By contrast, the minimum income standard is highly effective in providing support to the neediest families, by virtue of the fact that all households are eligible by default. Second, a large portion of anti-poverty spending in this system is allocated to the EITC. The model predicts that the EITC has been quite effective in encouraging labor supply and boosting income, but the effective transfer is by no means concentrated in the bottom end of the distribution. In fact, since the EITC is phased in as a percentage of total earnings, the transfer is least generous for mothers with the lowest wages. On this portion of the credit schedule, the EITC acts as a proportional wage subsidy: mothers with higher rents on the labor market face a greater incentive to work and are rewarded more generously for doing so. In this sense, the EITC is regressive. The minimum income standard does not possess the same properties of regressivity: all mothers who are eligible for the payment face the same marginal tax rate.

As would be expected, the lump sum transfer and the wage subsidy are less successful still in distributing improvements to the neediest families, since they have little to no design progressivity. However, it is still the case that, proportionally, poor mothers receive the most assistance from these initiatives, and hence, their children benefit the most in terms of skill outcomes. One lesson from these simulation experiments is clear: policies that seek to maximize the impact of income supports should be designed to ease the budget constraints of the poorest families. Programs that incentivize labor force participation are effective in boosting incomes, but they should be designed in a way that maximizes their impact on mothers with the least human capital and labor supply opportunities.

6 Conclusion

In this paper I investigated how household economic resources and government transfer policies influence child outcomes. To do this, I developed a dynamic, empirical model of maternal labor supply and investment in children. A key advantage of this approach is that the parental response to a rich portfolio of welfare and labor market policies is fully articulated. I combine this structure with a transparent approach to identification of the technology of skill formation. I find that both household income and time investment play an important causal role in shaping the behavioral and cognitive traits of children: increasing either investment good by 1% can, in select years, boost final cognitive and behavioral skills by up to 7.5% points, when anchored to the probability of high school graduation.

Next, I used estimates of the model to see how effectively government transfer programs can be used to improve child skill outcomes. I find that attempting to boost child skills through direct income supports is expensive, since it necessitates a persistent increase in resources over
every period, and because the returns to additional family income and maternal time quickly diminish. Consequently, programs are most effective when they assist the poorest mothers in the sample, for whom the marginal returns to additional resources are greatest. In this spirit, I design a counterfactual income support based on a minimum income standard that is revenue neutral. That is, it costs the same in simulation as AFDC/TANF and the EITC combined. Evidence from simulations suggested that such an initiative is highly promising: skills in the bottom decile of the income distribution - families who face extreme economic disadvantages and only partial welfare coverage - are lifted substantially, inducing a 3% point increase in high school graduation rates.

Despite the strength of these results, there are some necessary caveats. My results suggest that mothers are highly responsive to labor supply incentives. As such, the unconditional provision of a minimum income standard has strong negative impacts on labor supply attachment. Since, in this paper, I have not assigned any policy importance to labor supply, I do not see this as a weakness in the program. However, the political substance of the PROWRA reforms was to discourage welfare dependance and encourage labor supply. To this extent, the significant boost in child skills that my proposal induces can be weighed against its impact on maternal labor force attachment.

Furthermore, it is worth nothing that the parameters of the minimum income standard can be adjusted to provide stronger labor supply incentives. The tradeoff, as has been established by my work here, is that these programs are not maximally progressive. A revenue neutral shift from unconditional income to earnings supplements shifts transfers away from mothers with the lowest wages and the greatest preference for leisure. For programs of equivalent cost, work incentives are less helpful for the most disadvantaged children. This study has shown that the magnitude of such effects cannot be ignored in any policy discussion on this topic.

Finally, this paper represents an important first step in a broader research program aimed at modelling the developmental role of income supports in poor families. Directions in which to extend this analysis include, but are not limited to: endogenous household formation, endogenous fertility, child care, involuntary unemployment, and including a richer developmental production technology. For example, the technology and preferences assumptions in this model make the number of time use and expenditure categories irrelevant. However, it seems intuitively true that different categories of expenditure and time use play different developmental roles that may vary according to the child’s current stock of cognitive and behavioral traits. For example, food expenditure is likely to be an investment category which is essential, up to a level of sustenance, with little productive return beyond sufficiency. An additional example: a child that suffers from behavioral problems may benefit from more remedial time use activities (for example, those that allow for active supervision) compared to a less troubled child, who may benefit more from self-directed study. This extra richness is intricately tied to the economic realities of poverty, in which frequent income shocks may affect parental investment strategies and time use habits. Such shocks may also have long-lasting impacts, which the structure of this model does not allow.
There is sufficient evidence in this work to justify a more detailed look at the best ways for government antipoverty programs to buffer against the severe levels of resource deprivation observed in the data. As a first pass, this paper provides compelling evidence that there are sizeable returns to providing economic support for society’s most financially vulnerable children.
## A Tables

### Table A.1: Estimates of anchoring equations

<table>
<thead>
<tr>
<th>Specification</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WJ-LW</td>
<td>0.0657</td>
<td></td>
<td></td>
<td></td>
<td>0.0229</td>
</tr>
<tr>
<td></td>
<td>[0.0506, 0.0841]</td>
<td></td>
<td></td>
<td>[0.0046, 0.0413]</td>
<td></td>
</tr>
<tr>
<td>WJ-AP</td>
<td>0.0925</td>
<td></td>
<td>[0.0752, 0.1099]</td>
<td></td>
<td>0.0596</td>
</tr>
<tr>
<td>BPI-E</td>
<td>-0.0464</td>
<td></td>
<td></td>
<td>[-0.0545, -0.0383]</td>
<td>[-0.0528, -0.0324]</td>
</tr>
<tr>
<td>BPI-N</td>
<td>-0.0368</td>
<td></td>
<td></td>
<td></td>
<td>0.0099</td>
</tr>
<tr>
<td></td>
<td>[0.0482, -0.0254]</td>
<td></td>
<td></td>
<td>[-0.0040, 0.0238]</td>
<td></td>
</tr>
</tbody>
</table>

Parameters are estimated using a linear probability model, with 95% confidence intervals shown. These estimates are calculated the full sample of PSID-CDS children with available data on scores and high school completion.

### Table A.2: Model Estimates I

<table>
<thead>
<tr>
<th>Preferences</th>
<th>Variances</th>
<th>Welfare</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_l$</td>
<td>$\mu_{\alpha,0}$</td>
<td>$\sigma_{\alpha,0}$</td>
<td>$\sigma_w$</td>
</tr>
<tr>
<td>2.3022</td>
<td>-2.1271</td>
<td>0.5625</td>
<td>0.6937</td>
</tr>
</tbody>
</table>

Coefficients:

- $\text{Const}$
- $\text{ED}_i$
- $\text{PC}_i$
- $\text{SE}_i$
- $\text{age}_i$
- $\text{unemp}_i$
- $\text{Y}_t$
- $\text{1}\{N_K > 1\}$
- $\text{1}\{Y_t \geq 1997\}$

<table>
<thead>
<tr>
<th>$\log(w_{it})$</th>
<th>$\log(N_{it})$</th>
<th>$\text{WR}_{it}$</th>
<th>$T^{w}_{it}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0689</td>
<td>1.0648</td>
<td>0.4076</td>
<td>5.8640</td>
</tr>
<tr>
<td>0.0985</td>
<td>0.0535</td>
<td>0.0956</td>
<td>0.0985</td>
</tr>
<tr>
<td>0.0053</td>
<td>0.0859</td>
<td>0.0440</td>
<td>0.0644</td>
</tr>
<tr>
<td>0.0081</td>
<td>0.0859</td>
<td>0.3959</td>
<td>0.3959</td>
</tr>
<tr>
<td>-0.0024</td>
<td>-0.0081</td>
<td>0.0675</td>
<td>0.0675</td>
</tr>
<tr>
<td>0.0224</td>
<td>-0.0555</td>
<td>-1.3377</td>
<td>-1.3377</td>
</tr>
<tr>
<td>-1.2471</td>
<td>-2.4624</td>
<td>-1.2471</td>
<td>-2.4624</td>
</tr>
</tbody>
</table>
### Table A.3: Model Estimates II: Fundamental Production Parameters

<table>
<thead>
<tr>
<th></th>
<th>ED&lt;sub&gt;i&lt;/sub&gt;</th>
<th>PC&lt;sub&gt;i&lt;/sub&gt;</th>
<th>LW</th>
<th>AP</th>
<th>BE</th>
<th>BN</th>
<th>γ&lt;sub&gt;x,0&lt;/sub&gt;</th>
<th>γ&lt;sub&gt;x,a&lt;/sub&gt;</th>
<th>γ&lt;sub&gt;γ,0&lt;/sub&gt;</th>
<th>γ&lt;sub&gt;γ,a&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>LW</td>
<td>-0.0586</td>
<td>0.0439</td>
<td>0.5489</td>
<td>0.3433</td>
<td>-0.0295</td>
<td>0.0511</td>
<td>-33.5584</td>
<td>-26.3890</td>
<td>-79.8984</td>
<td>0.1032</td>
</tr>
<tr>
<td>AP</td>
<td>0.0347</td>
<td>0.0246</td>
<td>-0.0377</td>
<td>0.6853</td>
<td>-0.0285</td>
<td>0.0452</td>
<td>-2.2210</td>
<td>-0.0362</td>
<td>-3.1389</td>
<td>0.2009</td>
</tr>
<tr>
<td>BE</td>
<td>-0.0089</td>
<td>-0.0133</td>
<td>0.1594</td>
<td>-0.0703</td>
<td>0.9579</td>
<td>0.0062</td>
<td>-4.6291</td>
<td>0.2475</td>
<td>-23.0786</td>
<td>-0.2103</td>
</tr>
<tr>
<td>BN</td>
<td>0.0536</td>
<td>-0.0510</td>
<td>0.0116</td>
<td>-0.0077</td>
<td>0.1795</td>
<td>0.6270</td>
<td>-7.6537</td>
<td>0.5261</td>
<td>3.0456</td>
<td>-1.6073</td>
</tr>
</tbody>
</table>

### Table A.4: Model Estimates III: Contribution of Inputs to Final Skills

<table>
<thead>
<tr>
<th></th>
<th>ED&lt;sub&gt;i&lt;/sub&gt;</th>
<th>PC&lt;sub&gt;i&lt;/sub&gt;</th>
<th>Money</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>LW</td>
<td>-0.0003</td>
<td>0.0077</td>
<td>0.0184</td>
<td>0.0797</td>
</tr>
<tr>
<td>AP</td>
<td>0.0128</td>
<td>0.0042</td>
<td>0.0260</td>
<td>0.2177</td>
</tr>
<tr>
<td>BE</td>
<td>-0.0118</td>
<td>-0.0001</td>
<td>0.1006</td>
<td>0.0665</td>
</tr>
<tr>
<td>BN</td>
<td>0.0011</td>
<td>-0.0051</td>
<td>0.1475</td>
<td>0.0269</td>
</tr>
</tbody>
</table>
Table A.5: Model Fit for Select Sample Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>E[h_{it}</td>
<td>Y_i ≤ 1996]</td>
<td>18.5490</td>
</tr>
<tr>
<td>E[h_{it}</td>
<td>Y_i &gt; 1996]</td>
<td>28.0637</td>
</tr>
<tr>
<td>E[h_{it} w_{it}</td>
<td>Y_i ≤ 1996]</td>
<td>8.4955</td>
</tr>
<tr>
<td>E[h_{it} w_{it}</td>
<td>Y_i &gt; 1996]</td>
<td>14.7487</td>
</tr>
<tr>
<td>E[L F_i]</td>
<td>0.7505</td>
<td>0.7572</td>
</tr>
<tr>
<td>E[log(w_{it})]</td>
<td>2.0304</td>
<td>2.0526</td>
</tr>
<tr>
<td>E[log(N_{it})]</td>
<td>1.2579</td>
<td>1.0453</td>
</tr>
<tr>
<td>E[log(w_{it}) · ED_i]</td>
<td>0.3188</td>
<td>0.3742</td>
</tr>
<tr>
<td>E[log(N_{it}) · ED_i]</td>
<td>0.5390</td>
<td>0.2367</td>
</tr>
<tr>
<td>E[log(w_{it}) · PC_i]</td>
<td>0.7455</td>
<td>0.6149</td>
</tr>
<tr>
<td>E[log(N_{it}) · PC_i]</td>
<td>1.6822</td>
<td>1.1814</td>
</tr>
<tr>
<td>E[log(w_{it}) · SE_i]</td>
<td>0.0422</td>
<td>0.0339</td>
</tr>
<tr>
<td>E[log(N_{it}) · SE_i]</td>
<td>0.0484</td>
<td>0.0631</td>
</tr>
<tr>
<td>E[log(w_{it}) · Age_{it}]</td>
<td>0.9014</td>
<td>0.9900</td>
</tr>
<tr>
<td>E[log(N_{it}) · Age_{it}]</td>
<td>3.0124</td>
<td>3.6324</td>
</tr>
<tr>
<td>E[log(w_{it}) · unemp_{it}]</td>
<td>-0.0181</td>
<td>-0.0113</td>
</tr>
<tr>
<td>E[log(N_{it}) · unemp_{it}]</td>
<td>0.0191</td>
<td>0.0555</td>
</tr>
<tr>
<td>E[log(w_{it}) · Y_t]</td>
<td>0.7825</td>
<td>0.8085</td>
</tr>
<tr>
<td>E[log(N_{it}) · Y_t]</td>
<td>0.1299</td>
<td>-0.5339</td>
</tr>
<tr>
<td>V[log(w_{it})]</td>
<td>0.6071</td>
<td>0.5141</td>
</tr>
<tr>
<td>V[log(N_{it})]</td>
<td>3.3203</td>
<td>2.6162</td>
</tr>
<tr>
<td>E[T_{it}^W · 1]</td>
<td>4.1234</td>
<td>4.3341</td>
</tr>
<tr>
<td>E[T_{it}^W · 1{N_K &gt; 1}]</td>
<td>2.1111</td>
<td>1.9146</td>
</tr>
<tr>
<td>E[T_{it}^W · 1{Y_i ≥ 1997}]</td>
<td>0.7034</td>
<td>0.7693</td>
</tr>
<tr>
<td>E[T_{it}^W · h_{it} w_{it}]</td>
<td>5.3973</td>
<td>4.7839</td>
</tr>
<tr>
<td>V[T_{it}^W]</td>
<td>9.3838</td>
<td>8.7076</td>
</tr>
<tr>
<td>E[I_{w,it} · 1]</td>
<td>0.2102</td>
<td>0.1827</td>
</tr>
<tr>
<td>E[I_{w,it} · 1{Y_i ≥ 1997}]</td>
<td>0.0404</td>
<td>0.0495</td>
</tr>
<tr>
<td>E[I_{w,it} · log(B_{it})]</td>
<td>0.4635</td>
<td>0.4300</td>
</tr>
<tr>
<td>E[I_{w,it}</td>
<td>Y_i ≤ 1996, h_{it} = 0]</td>
<td>0.5633</td>
</tr>
<tr>
<td>E[I_{w,it}</td>
<td>Y_i &gt; 1996, h_{it} = 0]</td>
<td>0.4346</td>
</tr>
<tr>
<td>E[I_{w,it}</td>
<td>Y_i ≤ 1996, h_{it} ≤ 20]</td>
<td>0.5419</td>
</tr>
<tr>
<td>E[I_{w,it}</td>
<td>Y_i &gt; 1996, h_{it} ≤ 20]</td>
<td>0.3547</td>
</tr>
<tr>
<td>E[h_{it}</td>
<td>Y_i &gt; 1996, I_{w,it} = 1]</td>
<td>6.7960</td>
</tr>
<tr>
<td>E[h_{it}</td>
<td>Y_i &gt; 1996, I_{w,it} = 1]</td>
<td>8.8958</td>
</tr>
<tr>
<td>E[\phi_{r,97,i}]</td>
<td>22.7921</td>
<td>19.8172</td>
</tr>
<tr>
<td>E[\phi_{r,97,i} · a_{97,k}]</td>
<td>-25.1536</td>
<td>-34.7381</td>
</tr>
<tr>
<td>E[\phi_{r,97,i}]</td>
<td>0.1897</td>
<td>0.2180</td>
</tr>
<tr>
<td>Q_{25}(\phi_{r,97,i})</td>
<td>0.1059</td>
<td>0.0772</td>
</tr>
<tr>
<td>Q_{75}(\phi_{r,97,i})</td>
<td>0.2654</td>
<td>0.2656</td>
</tr>
</tbody>
</table>
In this section, I derive the model solution for an arbitrary number of skills, \( N_\theta \), and an arbitrary number of children \( N_K \). Let \( k = 1, 2, 3, ..., N_K \) index each child in ascending order of birth. In this section I assume that birth years are exogenously given and known to the mother. However, as we move through the solution it should be clear that extending the model to relax this assumption...
need not forbid the key simplification of additive separability in \( \log(\theta) \), the skill vector. In fact, this is shown in greater detail in Brown, Flinn, and Mullins (2015), which includes endogenous fertility and marriage decisions. Let \( B = \{b_1, b_2, \ldots, b_{N_K}\} \) indicate the year in which each child is born. As before, every child matures at age \( A \). I let the problem begin, as before, with the birth of the first child. Therefore we can set \( b_1 = 0 \), and set the terminal period of the problem at \( T_M = b_{N_K} + A \). Now, \( \theta \) refers to the full vector of skills for each child, and is therefore of dimension \( N_\theta \times N_K \). We let \( \theta_{k,j} \) refer to the \( j \)th skill of the \( k \)th child in the family and, similarly, \( \theta_k \) is the full skill vector for child \( k \). Finally, let \( a_k(t) = t - b_k \) indicate the age of child \( k \). Noting that age is collinear in time, I suppress the dependance of \( a_k \) on \( t \) for notational simplicity.

**Preferences and Technology**

To solve the problem, we must extend preferences to include multiple children, which we do simply as:

\[
u(c, l, \theta) = \alpha_c \log(c) + \alpha_l \log(l) + \sum_{k: a_k \geq 0} \alpha_\theta \log(\theta_k)
\]

\[
V_{TM}(\theta) = (1 - \beta)^{-1} \sum_k \alpha_\theta \log(\theta_k)
\]

Next, I have to take a stand on the rivalrous nature of time and monetary investment. Ideally, one could assume the existence of public investment categories as well as categories that each child would benefit from privately. In the case of time investment, this would be empirically plausible since categories of time use are observable\(^{11}\). Further, this setup can handle an arbitrary number of investment categories: we will see that each would be determined by a proportional investment rule. Yet this exact property removes this assumption of any empirical content in our context, since child outcomes are driven through changes in the log of total income (\( B_t \)) and log of total leisure hours (\( 1 - h_t \)). When investment rules are proportional to total income and total leisure time, we can derive labor supply, program participation, and child outcomes in terms of aggregates that depend only on the total cobb-douglas shares of each category. This logic extends to the case of multiple children. If time use is rivalrous, it is true that a child with siblings receives a lesser share of spare leisure hours than an only child, *ceterus paribus*, however this proportion bears no impact once we take logs (in effect, looking at percentage changes in investment). Thus, I assume in this paper that there is only one monetary investment and one time investment category, and that each is non-rivalrous across siblings.

\(^{11}\)Less so, for expenditure categories.
Model Solution

Given this set of assumptions, we can write the dynamic program in the following fashion:

\[
V_t(\theta_t, S_t, \eta_t) = \max_{c, l, x, h} \left\{ u(c, l, \theta_t) + \beta \mathbb{E}[V_{t+1}(\theta_{t+1}, S_{t+1}, \eta_{t+1}) | S_t, h, I_w] \right\}
\]  

(B.3)

Subject to the constraints:

\[
c + x \leq B(s_t, h, I_w)
\]  

(B.4)

\[
\tau + l + h = 1
\]  

(B.5)

\[
\theta_{k,t+1} = \Delta x, a_k \log(x) + \Delta \tau, a_k \log(\tau) + \Delta \theta, a_k \log(\theta) + \eta_{k,t}, \quad \forall k : a_k \geq 0
\]  

(B.6)

Note that in this general formulation, the vector \( S_{t+1} \) is permitted to evolve according to the current state \( S_t \) in addition to the labor supply \( (h) \) and program participation \( (I_w) \) decisions of the mother. Some of this model’s convenient representation could quite conceivably break if the evolution of \( S_t \) was further allowed to depend on investments \( (x, \tau) \). I first propose the following simplification of the model and show that it holds. As before, the key is that the value function is additively separable in \( \log(\theta) \):

\[
V_t(\theta, S_t, \eta_t) = \sum_{k=1}^{N_h} \alpha_{\theta, a_k} \log(\theta_{k,t}) \]  

(B.7)

\[
\nu(S_t) = \max_{h \in H, I_w} \left\{ \alpha_{c,t} \log(B(S_t, h, I_w)) + \alpha_{\tau,t} \log(1 - h) + \beta \mathbb{E}[\nu(s_{t+1}) | S_t, h, I_w] \right\}
\]  

(B.8)

\[
\alpha_{c,t} = \alpha_c + \beta \sum_{k: 0 \leq a_k < A} \alpha_{\theta, a_k + 1} \Delta x, a_k
\]  

(B.9)

\[
\alpha_{\tau,t} = \alpha_c + \beta \sum_{k: 0 \leq a_k < A} \alpha_{\theta, a_k + 1} \Delta \tau, a_k
\]  

(B.10)

\[
\alpha_{V,a} = \alpha_{\theta, a} + 1 \{ 0 \leq a_k < A \} \cdot \beta \alpha_{V,a + 1} \Delta \theta, a
\]  

(B.11)

\[
x_t = \frac{\beta \sum_{k: 0 \leq a_k < A} \alpha_{V,a_k + 1} \Delta x, a_k B(S_t, h, I_w)}{\alpha_{c,t}}
\]  

(B.12)

\[
\tau_t = \frac{\beta \sum_{k: 0 \leq a_k < A} \alpha_{V,a_k + 1} \Delta \tau, a_k}{\alpha_{\tau,t}} (1 - h)
\]  

(B.13)

We prove this by first showing that the recursion holds. That is, assume this form holds for \( V_{t+1} \) and show that it is preserved at time \( t \). The problem can be stated as:

\[
V_t(\theta, S_t, \eta_t) = \max_{c, l, x, h, I_w} \left\{ \alpha_c \log(c) + \alpha_l \log(l) + \sum_{k: a_k \geq 0} \alpha_{\theta} \log(\theta) \right. \\
left. + \beta \mathbb{E} \left[ \nu(S_{t+1}) + \sum_k \alpha_{\theta, a_k + 1} \log(\theta_{k,t+1}) + \sum_{0 \leq a_k + 1 < A} \alpha_{V,a_k + 2} \log(\eta_{k,t+1}) | S_t, h, I_w \right] \right\}
\]  

(B.14)
subject to the constraints given above. First, we can substitute in the production function to get:

\[
V_t(\theta, S_t, \eta_t) = \max_{c,l,x,\tau,h,I_w} \left\{ \alpha_c \log(c) + \alpha_l \log(l) + \sum_{k: a_k \geq 0} \alpha_{\theta} \log(\theta_{k,t}) + \sum_{k} (\alpha_{\theta} + 1 \{0 \leq a_k < A\} \cdot \alpha_{V,a_k+1} \Delta_{\theta,a_k}) \log(\theta_{k,t}) \right. \\
+ \beta \sum_{k: 0 \leq a_k < A} \alpha_{V,a_k+1} \left( \Delta_{x,a_k} \log(x) + \Delta_{\tau,a_k} \log(\tau) + \log(\eta_{k,t}) \right) \\
\left. + \beta \mathbb{E} \left[ \nu(S_{t+1}) + \sum_{0 \leq a_k + 1 < A} \alpha_{V,a_k+2} \log(\eta_{k,t+1}) \mid S_t, h, I_w \right] \right\}
\]

(B.15)

As is indicated in the second line of this equation, this step is sufficient to define the recursion for \(\alpha_{V,a}\), the value derived from \(\log(\theta)\) for a child at age \(a\). Next, we inspect the first order conditions for \(\tau\) and \(x\), subject to choices of \(h\) and \(I_w\). These yield:

\[
\frac{\alpha_c}{c} = \frac{\beta \sum_{k: 0 \leq a_k < A} \alpha_{V,a_k} \Delta_{x,a_k}}{c} \quad \text{(B.16)}
\]

\[
\frac{\alpha_l}{l} = \frac{\beta \sum_{k: 0 \leq a_k < A} \alpha_{V,a_k} \Delta_{\tau,a_k}}{\tau} \quad \text{(B.17)}
\]

Rearranging these equations gives the proportional investment rules shown in (B.12) and (B.13). Substituting those rules into the value function (B.15) and collecting terms gives the final expression of the value function in (B.7) and (B.8). We can complete this exposition for the solution by noting that the terminal period value function \(V_{T_M}(\theta)\) also keeps this additive form, and hence we have the necessary conditions to initiate the recursion.

**References**


