Understanding the Effects of Income and Child Care Subsidies on Children's Academic Achievement*

Jorge Rodríguez†

November 8, 2017

Click here for the most recent version

Abstract

In Milwaukee, Wisconsin (1994-1997) an anti-poverty intervention, "New Hope," randomly assigned an income supplement—similar to the EITC—and a child care subsidy to a group of economically disadvantaged families. The experimental evaluation found positive effects of the program on labor supply, income, and child care use. Notably, the program also boosted various measures of child academic achievement. However, since policies were given in a single package, little is known about the mechanisms by which New Hope affected child outcomes. The goal of this paper is to disentangle the mechanisms that explain the impact of income and child care subsidies on children's academic achievement. To this end, I estimate a dynamic-discrete choice model of the household and child academic achievement using the New Hope data. I use the model to quantify the importance of household inputs in explaining the impact of New Hope and similar policies—such as the EITC and the CCDF—on children's academic achievement. I find that the negative effect on child human capital due to the program-induced increase in labor supply is more than compensated by the positive effects of additional income and a higher probability of using center-based child care. Given the dynamics of the technology of skills production, the effects of income and child care subsidies become economically significant only when children are exposed to these policies for many years.

*I am indebted to Magne Mogstad, Derek Neal, and Alessandra Voena for providing guidance and feedback throughout this project. I also thank Thibaut Lamadon, Petra Todd, Steve Levitt, Stephane Bonhomme, James Heckman, and seminar participants at the SOLE annual meeting, North American Summer Meeting of the Econometric Society, and University of Chicago for valuable comments.

†University of Chicago, jorger@uchicago.edu.

1 Introduction

Over the last 20 years, policymakers have used a variety of strategies to encourage the labor market participation of low-income families. Two of the most important policies have been income and child care subsidies.[1] The empirical evidence indicates that income and child care subsidies have succeeded in meeting their original goals—namely, to promote work and increase family income.[2] However, these policies could have unintended, negative consequences on child outcomes: parents could opt for center-based child care, increase the time spent in the labor market, and reduce the time caring for children.[3]

Some of the most compelling evidence on the effects of income and child care subsidies comes from a randomized controlled trial called New Hope. The program assigned applicants (over 18 years old) to a policy bundle that included an income subsidy—similar to the EITC—and a child care subsidy—which resembled the CCDF—tied to a full-time work requirement. The New Hope literature finds that the program increased annual family income by 7%, the probability of being employed in any given quarter by 8 percentage points (from a baseline of 63%), and the likelihood of using child care for young children by 22 percentage points (from a baseline of 40%). Notably, the program also boosted various measures of child academic achievement.[4]

Given the New Hope experimental design, the program provides reliable evidence on

[1]In the U.S., two prominent examples of such subsidies are the Child Care Development Fund (CCDF)—created by the Personal Responsibility and Work Opportunity Act (PRWORA)—and the Earned Income Tax Credit (EITC). The CCDF is a block grant to states for the provision of child care vouchers to low-income working parents. This program was conceived as a complement to Temporary Assistance for Needy Families (TANF) that would facilitate welfare-to-work transitions. The EITC is a mean-tested cash transfer program for low-income families.

[2]See Grogger and Karoly (2009) and Moffitt (2016) for a review of the evidence. See Hoynes and Rothstein (2016) for a description of the EITC and its impacts on labor supply. See also Chan (2013) and Keane and Wolpin (2010) for evidence of changes in the welfare system within a dynamic framework. See Blundell et al. (2015) for recent evidence on tax credits effects on the labor supply and educational choices of single and married mothers.

[3]See Bernal (2008), Bernal and Keane (2010), Bernal and Keane (2011), and Brilli (2014) for evidence comparing income and time allocation impacts on child development. See Heckman and Mosso (2014) for a review of the evidence on the effects of income on skills accumulation over the child life-cycle.

[4]See Bos et al. (1999), Huston et al. (2003), and Miller et al. (2008) for evidence on the effects of New Hope on child and family outcomes. The numbers I present in this paper are own calculations obtained using the original New Hope data (see Section 4).

the causal effects of New Hope on children's academic achievement—a crucial input for public policy purposes. Nonetheless, the New Hope evidence has at least three limitations. First, because all policies were bundled together, we cannot assess the role played by each individual policy in the induced changes on household choices and child outcomes. Second, for the same reason, the experimental data does not give us enough exogenous sources of variation to assess which household behaviors were more influential in accounting for the impact on child outcomes. Third, by using New Hope data, we are limited to evaluate the effects of policies that only vary within the sample; for example, we cannot predict the effects of the EITC or CCDF on child outcomes.[5]

The goal of this paper is to disentangle the mechanisms that explain the impact of income and child care subsidies on children's academic performance. To address the empirical limitations of the experimental evidence, I posit and estimate a dynamic-discrete choice model of the household and child academic skills. In the model, a single-child unitary household chooses hours of work and child care types (informal home care or formal, center-based child care). Child human capital production follows a dynamic process, where household decisions and the current stock of skills are inputs in this production function. The household's budget set encompasses different mean-tested programs, including the AFDC, the EITC, and New Hope. I estimate the model using non-experimental moments while leaving experimental estimates for model validation. I exploit the estimated model to evaluate how policies—such as New Hope, the EITC, and the CCDF—impact household choices and thereby child skills.

I find that the impact of New Hope on child academic performance is a result of different counteracting channels. The program-induced rise in labor supply has a negative effect on child skills acquisition. However, this negative effect is more than compensated by the positive impact on child skills caused by the increase in income and child care probability. For young children, one year after the program, a higher probability of child care explains

[5] Grogger and Karoly (2009) conclude that experimental studies on welfare reforms and child well-being yield mixed results, where the estimated effects are likely to depend on each program's characteristics. See Heckman (2010) and Keane et al. (2011) for a related discussion on the comparison of structural and reduced-form approaches.

82% of New Hope's effects while the rise in income explains close to 26%. After the first year, the persistence of the production function plays a key role in explaining the evolution of the program's treatment effects on child academic skills. In the first year of the program, the impact on child human capital is relatively small. However, most of the skills acquired in that period are transferred to the next one. Since the program continued to affect income and child care use, more skills were produced on top of the existing stock, augmenting the average effect of the program over time. When the program ended (in the third year), the program's effects on income and child care dissipated and the impact on child human capital started to diminish.

I quantify the importance of each one of the New Hope policies in accounting for the impact of the program on children. My analysis shows that the child care subsidy explains most of the effects of New Hope on child academic achievement. The contribution of this policy to the overall effects on children more than doubles the contribution of the wage subsidy. Furthermore, forcing individuals to work full time to have access to any of these benefits has a negative effect on child human capital; the effects of the program would have been larger (about 0.04 standard deviations) if New Hope had not included a work requirement. The negative effect of the work requirement is explained because this policy, relative to a program without this condition, decreases parental time at home and center-based child care use.

I study the consequences of the EITC and a permanent child care subsidy (similar to Wisconsin's version of the CCDF subsidy) on child academic skills. I find that both policies have an economically significant potential to impact child skills development through its effects on household behavior—especially in the long run. The EITC increases child academic skills by 0.08 standard deviations for children exposed for at least seven years to the policy. A child care subsidy has larger effects in a shorter period of time: over 0.1 standard deviations at the end of the fifth year. Nevertheless, the child care subsidy's effect reaches a plateau as children get older and the child care option is no longer available for them. After seven years, both policies combined boost child academic skills by 0.18 standard deviations. The effects of the EITC are largely explained by the income channel, with year-by-year small additions to the human capital stock that are being accumulated over time due to the human capital production function. Labor supply plays only a negligible role and its negative effect on child skills does not offset the positive effect of having additional income in the household.[6] Similarly, I find that most of the child skills effect of the child care subsidy comes from a higher probability of attending a center-based child care—not from having more disposable income.[7] In both policies, self-productivity sustains an increasing average treatment effect.

This study contributes to the research that explores the intergenerational consequences of mean-tested programs. Specifically, it advances the literature on two fronts. First, this study provides new evidence on the impacts of the EITC and income on child outcomes. Chetty et al. (2011), Maxfield (2013), Hoynes et al. (2015), Manoli and Turner (2015), and Bastian and Michelmore (2017) report relatively large intergenerational effects of the EITC. Along similar lines, Dahl and Lochner (2012) use variation in the EITC schedule across states and time as an instrument for family income. Often, researchers interpret this literature as causal evidence on the effects of money on child outcomes. However, in order to interpret a EITC exposure as a pure income effect, one needs to instrument for other household variables that affect child outcomes and that are being affected by the policy; an exogenous income shock lowers the marginal value of market time while simultaneously raising the marginal value of paying for private child care or preschool. Moreover, in a treatment effects analysis, we do not know for certain if income shifts caused by the EITC can be thought of as temporary or permanent shocks.[8] To understand the reduced-form effects of the EITC, I use the structure of my model to simulate a permanent EITC policy, examine its short- and

[6] Dahl and Lochner (2012) exploits EITC variation to instrument for income and uncover causal effects on child test scores. The authors argue that the IV estimates on the effects of income on child test scores hold once accounting for labor supply effects (Dahl and Lochner, 2016).

[7] Black et al. (2014) find that child care subsidies in Norway impacted child outcomes for the most part due to a larger average disposable income.

[8] Dahl and Lochner (2012) justify their relatively large estimated impact by arguing that the instrument captures permanent income changes.

long-run consequences on household and child outcomes, and decompose its effects on child skills in terms of the different changes in household behavior.

Second, this paper also contributes to the literature on child care subsidies and child skills accumulation. So far, the literature has reached varied results regarding the average effects of child care subsidies on child outcomes (Baker et al., 2008; Herbst and Tekin, 2010a,b; Havnes and Mogstad, 2011, 2015; Black et al., 2014; Cornelissen et al., 2017).[9] An emerging consensus seems to be that children from low-income families benefit the most from child care subsidies and universal child care programs (Havnes and Mogstad, 2011, 2015; Cornelissen et al., 2017), although the channels generating this pattern of heterogeneous effects on children are still unclear.[10] Since mechanisms are not being systematically assessed in these studies, it is difficult to obtain general lessons about the effectiveness of child care subsidies out of the differing results found in the literature. To understand the effects of child care subsidies, I use the structure of my model to reveal the contribution of different sources of household behavior in explaining the impact of child care subsidies. Overall, I find that the effects of child care subsidies are largely consistent with evidence showing sizable impacts of early childhood education on children from economically disadvantaged families (Elango et al., 2016).

On the methodological side, I build upon the literature that combines experimental and quasi-experimental evidence with structural models (Bajari and Hortaçsu, 2005; Todd and Wolpin, 2006; Keane and Wolpin, 2007; Attanasio et al., 2011, 2015; Voena, 2015; Autor et al., 2017). Following this line of research, my framework advances reduced-form studies by delineating the mechanisms that explain the observed policy treatment effects obtained by either experimental or quasi-experimental methods. I also contribute to the structural literature on household behavior and child welfare by exploiting experimental data to validate

[9]See Elango et al. (2016) for a review of the literature.

[10]Elango et al. (2016) suggest that the returns to universal child care programs are higher for children from low-income families because the quality of alternative child care arrangements for them is lower than that of children from high-income families.

the model's capacity to predict the impact of different welfare policies.[11]

The remaining of the paper is structured as follows. Section 2 describes the New Hope program's characteristics. Section 3 provides details on the available data. Section 4 outlines the available reduced-form evidence on New Hope. Section 5 presents the dynamic-discrete choice model and Section 6 discusses its estimation. Section 7 shows the model's estimates and explains its implications for the dynamics of skills acquisition. Finally, Section 8 assesses the consequences of income and child care subsidies on household decisions and child outcomes.

2 The New Hope welfare model and context

Inspired by the welfare debate that dominated the policy agenda in the 90s, New Hope was designed to promote the transition from welfare to work. As a result, the program deepened the incentive to work that families were subjected to at the time the program was implemented.[12]

The participants—recruited in two economically disadvantaged neighborhoods in Milwaukee, Wisconsin—had to meet a few conditions to enter the program and have access to the benefits package.[13,14] To be eligible, individuals had to be at least 18 years old and have a household income equal to or less than 150% of the federal poverty line.[15] Additionally, applicants had to be willing to work at least 30 hours per week (which was considered full-time

[11]See Bernal (2008), Brown and Flinn (2011), Del Boca et al. (2013), Brilli (2014), Del Boca et al. (2014), Mullins (2015), Bruins (2016). This literature uses survey data to estimate structural models and non-experimental moments as model validation.

[12]Later on, many of the policy changes in the U.S. were similar to the New Hope package. See for example Moffitt (2003).

[13]Participants came from the north and south side of the U.S. Highway 94 and the Menomenee River Valley. The New Hope team selected those neighborhoods (which were defined by their postal zip codes) because they had a relatively high poverty rate and ethnically diverse populations. Each area had about 40,000 residents (Bos et al., 1999).

[14]New Hope was heavily promoted during the eligibility period. The New Hope team advertised the program in posters, radio, TV, and newspapers, and they sent personal letters. About 20% of potential participants in the target areas became aware of the program (Brock et al., 1997).

[15]According to Bos et al. (1999), for a household with one adult and two children, the federal poverty threshold was $12,278. For a single-person household, the threshold was $7,929.

employment for the purposes of accessing the program's benefits).[16] Beginning at baseline recruitment and lasting for 36 months, a randomly selected group of applicants had access to various benefits. To receive any of the subsidies of the program in a given month, the participants had to prove that they had worked at least 30 hours a week on average. Each 5th day of the month, New Hope agents enforced this requirement by asking for the last month's wage stubs. After reviewing those wage stubs, New Hope representatives would determine the amount of supplement that the participants were to receive. This process usually lasted 15 days. After this period, the participant would receive the payment by approximately the 20th of the same month.

In this paper, I focus on two elements of the New Hope package: the income supplement and the child care subsidy. Next, I describe these two components and leave the description of the other components of New Hope for Appendix A.

2.1 The income supplement

Figure 1, panel (a), illustrates the income supplement design (the earnings subsidy plus the child allowance) for a family with one earner and one child.[17] To show how the schedule looks across the distribution of ex-ante labor earnings,[18] these figures assume no work requirements. The New Hope income supplement complements the EITC subsidy. In the figure, the income supplement is represented as the difference between the dashed and solid lines—that is, the New Hope subsidy is positive for a worker as long as the New Hope schedule stays above of that of the EITC (and meets the work requirement).[19,20]

[16] Participants entered the program during a period of buoyant economic activity. Between 1992 and 1997, job creation at the Milwaukee Primary Metropolitan Statistical Area (which covers the Milwaukee, Washington, Ozaukee, and Waukesha counties) grew by 8.2%. For the same area, the unemployment rate diminished from 4.8% in 1992 to 3.6% in 1997 (Bos et al., 1999).

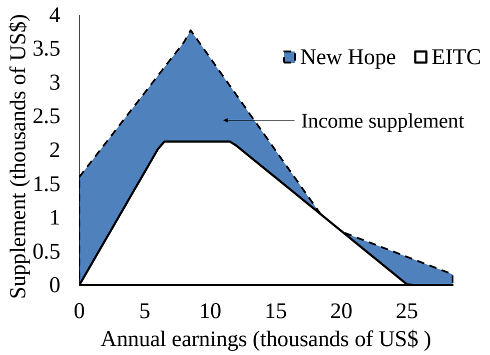
[17] The income supplement corresponds to the sum of two subsidies: an earnings subsidy and a child allowance. Appendix A provides the exact formula of the subsidy.

[18] Even though Figure 1 depicts the income supplement in terms of annual benefits, New Hope beneficiaries received their supplements on a monthly basis. The income supplement was not taxable.

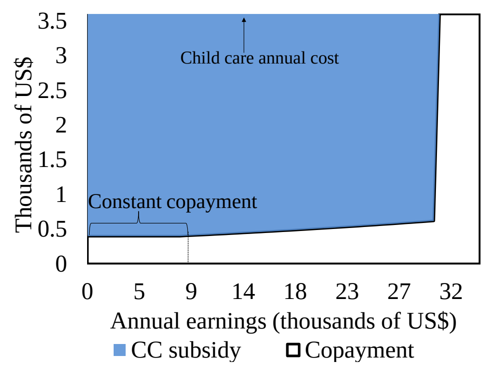
[19] The dashed line shows a discontinuity at $19,000 because the earnings subsidy is zero at that point while the child allowance continues to phase out.

[20] Because the income supplement schedule stayed fixed whereas the EITC schedule expanded while the program was running, the treatment "intensity" varied in time. In graphic terms, the dashed line in Figure

Figure 1: New Hope income supplement and child care subsidy



(a) Income supplement



(b) Child care subsidy

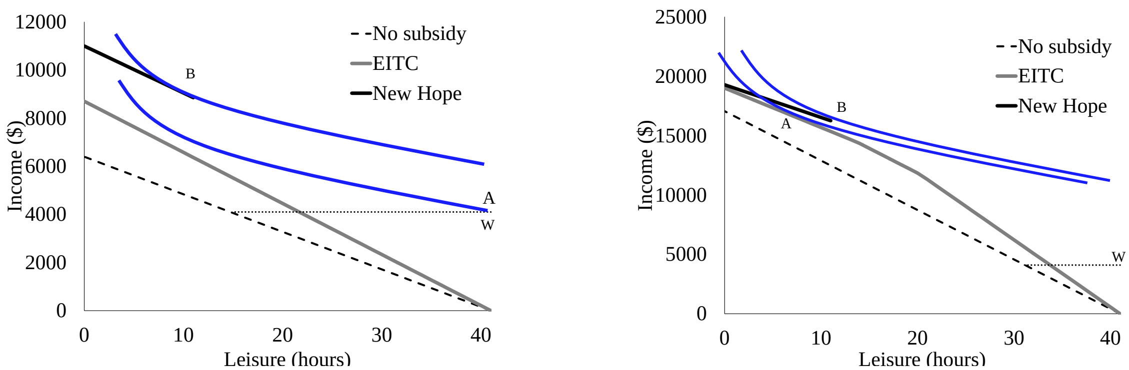
Notes: Panel (a) compares the New Hope and EITC design as a function of annual earnings. The solid line shows the EITC income supplement. The difference between the dashed and the solid line represents the New Hope supplement, for each level of earnings, assuming no work requirement. Panel (b) illustrates the child care subsidy design. For this figure, the child care cost equals 3,600 dollars a year, and it is indicated at the top of the y-axis. The solid line represents the copayment. The subsidy corresponds to the solid area above the copayment line.

To evaluate the economic incentives introduced by the program, consider two individuals, X and Y, choosing between home and labor market time (for simplicity, suppose that they do not have children). Compared to those in the control group, individuals in the treatment group had different incentives to work depending on their wage offer. The choices made by X and Y are illustrated in Figure 2. In these graphs, the horizontal and vertical axis show income and time outside the labor market, respectively. Both figures present the individual's budget set under three different cases: without EITC or New Hope ("No subsidy"), the control group ("EITC"), and treatment group ("New Hope"). Since New Hope requires working 30 hours or more, the New Hope budget set ends at the point of 10 hours of leisure. Additionally, the figure shows what both individuals would earn if they do not work (at point W).[21] X and Y have the same preference towards income and leisure, and so both have an equal set of indifference curves in the income-leisure plane. The only difference between the budget sets of X and Y is that the wage offer of X is lower than that of Y.

1, panel (a), remained constant, whereas the solid line shifted upwards alongside the changing EITC regulations. These modifications in the EITC meant that the treatment group received lower levels of the income supplement in time.

[21]I assume this value to be 4,100 dollars, which equals the sum of the average values the control group received one year after baseline from Aid to Families with Dependent Children (AFDC) and Food Stamps (Bos et al., 1999)

Figure 2: Intensive- and extensive-margin responses to New Hope



(a) Individual X (low wage offer) (b) Individual Y (high wage offer)

Notes: The figure illustrates extensive- and intensive-margin responses to New Hope. For individuals with two different wage rates and same structure of preferences, it presents individuals choices under different budget sets in the income-leisure plane.

The figure indicates that, *ceteris paribus*, the impact of the program on labor supply depends on the wage offer. Without New Hope (if X and Y were in the control group), individuals would allocate at point A. At this point, X would not meet the 30 hours condition whereas Y would work over this threshold. In fact, the wage offer of X is low enough so that she is better off receiving welfare and not working at all. If X and Y were in the treatment group, they would choose to allocate at point B. Compared to point A, X would work more hours and receive more income. Y would earn more as well. However, Y works more or less compared to the scenario of not having the program, depending on the relative magnitudes of substitution and income effects. Figures 2a and 2b illustrate one of many situations in which the income supplement impacts labor supply and income. Overall, the New Hope income subsidy should have a non-negative effect on income and an ambiguous effect on hours worked.[22]

Because the income supplement affects behavior, it can also produce effects on child outcomes. Following the literature, suppose that labor supply causes a negative effect on child human capital, while income causes a positive effect (Bernal, 2008; Dahl and Lochner, 2012). The effect of the program on individual X's child is ambiguous; X has more income

[22]Given a different utility function, individual X may choose to stay at point W even in the New Hope situation. Therefore, for some individuals, the program has no impact whatsoever.

but works more. The impact on individual Y's child is also ambiguous. If Y works more, then she would be in the same situation as X: more income but less hours at home. If Y works less then we can guarantee a positive effect on children, as the individual spends more time at home and has more income. Overall, the program's impact on children depends on the relative strength of intensive- and extensive-margin labor supply responses and the relative productivity of income and time with the child in the production function of child skills.

2.2 The child care subsidy

Figure 1 (panel b) depicts the child care subsidy and copayment for the case of a single-child household paying $3,600 a year for child care. The solid line represents what the household pays after the subsidy (shown in the solid area) across levels of family earnings. Thanks to the subsidy, families paid a relatively small copayment. Up to $8,500 of annual family earnings, this family starts paying $400. After the $8,500 point, the copayment increases by 1% for every extra dollar of family earnings. The household stops receiving a subsidy at the point where family earnings reaches 200% of the federal poverty line or $30,000, whichever comes first.[23]

According to Brock et al. (1997), economically disadvantaged families had access to a number of child care programs offered by the Milwaukee's welfare department—with reimbursement rates and subsidy limits that were similar to the New Hope design.[24] However, families in the New Hope program had some clear advantages over those families using the public system. First, participation in New Hope increased their chances of finding low-cost child care services. Parents in the public system under AFDC, for example, usually faced

[23]The average total expenditures in child care reported is 180 dollars a month, 65% of the average monthly labor earnings of the control group at baseline.

[24]Starting 1997, the CCDF enhanced the low-cost child care supply. Furthermore, the State of Wisconsin supplemented the federal funds from the CCDF to make the child care subsidies available to all eligible families. As a result, the public system began to offer a very similar service to that of New Hope, making the relative gain of the latter system much smaller after 1997. (Wisconsin's TANF program created a child care assistance program named "Wisconsin Shares." This program was implemented in September of 1997, and it provided child care vouchers that could be used at any of the licensed child care providers in the county.)

long waiting lists to apply for public child care subsidies. For families who were not in the welfare system, finding a low-cost child care provider was even harder (for example, obtaining a Head Start slot was almost impossible for these parents).[25] In contrast, New Hope beneficiaries had the possibility of enrolling their children in any of the county- or state-licensed child care centers available in the city. Second, qualitative evidence indicates that families who were eligible for these expanded public child care programs struggled to comprehend and navigate Wisconsin's complex system.[26] The qualitivative evidence suggests that families under New Hope benefited from a simpler and more easily understood system, since New Hope gathered all subsidies into one single program.[27]

Figure 1 (panel b) shows that the child care subsidy produces heterogeneous effects on household choices. Take the case of two individuals ("A" and "B") who would work 30 hours or more with and without the program. Without New Hope, "A" would pay for a child care service while "B" would not. For individual "A", the program only raises her disposable income while for "B" there is an incentive to take up the subsidy to use the child care option. Now consider another individual ("C") who, without New Hope, would work less than 30 hours and would not use center-based child care. If she would like to use child care under New Hope, she would have to work more than 30 hours. For this individual, the economic incentive provided by the income supplement may induce her to do so, and she would be able to take up the child care subsidy if she wishes so.

The child care subsidy may affect child human capital for these three individuals through different mechanisms. Suppose that, relative to home care, child care has a positive impact on child human capital. Even though there is no effect on child care take-up for individual A,

[25] 54% of the New Hope full sample (control and treatment groups) were not under AFDC (Bos et al., 1999).

[26] Individuals had to be aware of the different child care assistance programs for which they would be eligible as their situation changed. For example, if a family had left AFDC, then they would have had to apply to a child care for working parents. If they had become unemployed and fell under AFDC again, they would have had to redo all paperwork to receive the child care assistance from the public system (Bos et al., 1999; Blau, 2003).

[27] Moreover, families could reach out to New Hope representatives whenever they had questions regarding their benefits, or if they could not find suitable child care facilities in the city.

her child would benefit from the child care subsidy because A has more income. In contrast, individual B's child would benefit from the center-based child care if B chooses this option. One could find a negative effect of the child care subsidy in the case of individual C. Because she has to work full time in order to take up the child care subsidy, the impact on child human capital depends on the productivity of child care relative to that of labor supply in the human capital technology. Therefore, as with the income supplement, the theory does not give a clear prediction on the sign of the effect of the child care subsidy on child outcomes.

3 The data

From August 1994 until December 1995, the MDRC (the agency in charge of the experimental evaluation) selected the original New Hope sample. This sample consisted of 1,357 individuals (678 individuals to the treatment group and 679 to the control group). To evaluate the intervention's impact, the MDRC collected data on labor market outcomes and participants' families.

Table 1 presents a description of the available databases. The New Hope data consists of three databases: the parents' and children's survey (the "New Hope surveys"), the teachers' survey, and administrative information from the state of Wisconsin. The first database corresponds to the New Hope surveys. These surveys gathered household information on work and child outcomes. The second source of information is the teacher surveys. In these surveys, teachers gave their assessment on several child academic and behavioral indicators. Finally, I use administrative records from the state of Wisconsin. The administrative records contain information on earnings from the Wisconsin UI system and directly from the firms that hired participants for New Hope's community service jobs.

In my empirical analysis, I use the sub-sample of individuals with at least one child. This sample is referred as the Child and Family Study (CFS). The CFS has information only for participants with at least one child between 1 and 10 years of age (745 adults of the original

Table 1: Available databases on the New Hope intervention

Data	Variables	Periods
New Hope Survey	Hours worked and child care use and expenditures.	Two, five, and eight years after random assignment.
Teachers' survey	Child outcomes: SSRS Academic Subscale	Two, five, and eight years after random assignment.
New Hope's and Wisconsin's administrative records	New Hope: benefits take-up (income supplement, child care payments). Wisconsin: Labor earnings (UI system), AFDC, and Food Stamps payments.	1994-2003.

Notes: This table shows the available databases (first column), the associated variables (second column), and the years in which they were collected (third column).

1,357).[28] Table 2 shows the number of individuals from which we have information across different data sets. The first panel compares the number of adults in the whole sample and in the CFS. The last three rows of the same panel show the number of CFS participants who answered the New Hope surveys. The second panel presents the number of CFS children. Depending on the year, the New Hope surveys have information for 75 to 80% of the adults and 78 to 86% of the children in the CFS sample.[29] The third panel of Table 2 presents the number of children with data from the teachers' survey. Not all teachers filled out the questionnaires; relative to the number of children with information on the surveys from years two, five, and eight, teacher's underreport reduces the sample by 56, 37, and 42%, respectively.

To evaluate if attrition compromised the randomization outcomes in the two experimental groups, Appendix B (Tables B.1-B.5) compares participant baseline characteristics of the original CFS sample with those who answered New Hope and teachers' surveys from years two, five, and eight. Table B.2 shows baseline characteristics of the original CFS sample. The majority of participants in the CFS are women (90%), a little more than half are African-

[28]Up to two children per family were selected to be part of the CFS. According to Miller et al. (2008), if more than two children were potentially eligible to participate in the CFS survey, only two of them were randomly chosen (with more preference given to opposite-sex siblings).

[29]In addition, 50 adults in the CFS database do not match in the youth database, and two children in the youth database do not match in the adult CFS data. I excluded these observations from the analysis.

Table 2: Sample size of the Children and Family Study (CFS) across surveys

Sample	Treatment	Control	Total
Adults			
All	678	679	1,357
CFS	344	351	695
Year two	288	302	590
Year five	282	280	562
Year eight	297	300	597
Children			
All	544	561	1,105
CFS	544	561	1,105
Year two	456	489	945
Year five	432	429	861
Year eight	468	479	947
Children with teachers' report			
Year two	203	217	420
Year five	272	274	546
Year eight	272	277	549

Notes: This table shows the sample size for different surveys. The first panel (Adults) presents the number of adults from the original sample ("All"), the CFS study (adults with at least one 13-year-old or younger), and from the New Hope surveys of years two, five, and eight. The second panel (Children) shows the number of children under the same databases as the previous panel. Finally, the third panel depicts the number of children with teachers' reports in the surveys from years two, five, and eight.

American (58% and 53% in the treatment and control group) and 88% do not cohabitate with a spouse or partner. Moreover, only half of the sample has a high school diploma or GED certification and over 50% earned less than $10,000 in the last 12 months (1994 dollars). For all the variables in Table B.2, there are no statistically significant differences between treatment and control groups. Even though baseline characteristics change when comparing the original CFS group to those with survey data (from Table B.2 to Tables B.3-B.5), there is no evidence to reject the null hypothesis that baseline characteristics differ by treatment status across the different samples.

4 Treatment effects

The goal of New Hope was to test whether a new group of welfare policies could help families to overcome poverty by encouraging work. Because the program was expected to change multiple household behaviors, effects on child skills development were expected as well.

Bos et al. (1999) and Huston et al. (2001) document experimental evidence on the effects of New Hope on household behavior and child outcomes.[30] In this paper, I focus on three household variables—income, labor supply, and child care— and on a group of measures of child academic skills. To obtain the income variable, I consider labor earnings from the Unemployment Insurance (UI) system, simulated EITC payments, New Hope supplements, community service job (CSJ) earnings, and welfare payments (Food Stamps and AFDC or TANF cash transfers).[31] I define quarterly employment as having a positive UI or CSJ record in a given quarter. Finally, I set the child care dummy as1 if the child was enrolled in a center-based child care (Head Start, preschool, nursery school, or another child care other than someone's home) and as 0 if the child received home care (that is, she stayed at home with another family member, or attended an informal child care in someone's home). To measure children's academic outcomes, I use the Social Skills Rating System (SSRS), Academic Subscale, from the teachers' survey. In the SSRS, teachers rank students on a five-point scale—1 being "lowest 10%" and 5 "highest 10%"—according to how well the student performs in a particular subject (Math, Reading, etc.).[32]

Using the public New Hope database and the variables as defined above, I estimate the effect of the program on household choices. I compute the effects on household variables while the program was in effect, that is, three years in total including the baseline year. Even though I am not able to exactly replicate the numbers from previous studies estimates, I find estimates that are around the same order of magnitude.[33] The program increased household

[30] See also Huston et al. (2003), Huston et al. (2005), Epps and Huston (2007), Miller et al. (2008), Grogger and Karoly (2009), and Huston et al. (2011).

[31] The income variable does not include other sources of welfare (such as the WIC or the child tax credit) and income from relatives or from other household members. Thus, in my data, it is possible to have participants with zero income.

[32] See Appendix C for a more detailed definition of the variables and alternative estimation approaches.

[33] The differences in the estimates are explained by the different treatment of the data. First, the MDRC report shows effects only until the second year of the program. Second, I simulate EITC annual payments and assume full compliance, whereas the MDRC elaborates a statistical procedure to approximate the EITC payments. Third, I consider effects on child care use only on young children (six years old or less), while the MDRC report does not differentiate by age (which explains the seemingly large difference in the child care estimates). Notwithstanding these differences, both sets of estimates are similar in revealing moderate-to-

annual income by an average of $1,000 (a 7% increase), quarterly employment probability by 8 percentage points (from a baseline of 63%), and child care use by 22 percentage points for children under six years of age (from a baseline of 40%). In line with past research—and the theoretical prediction from Section 2—I find that the program had a larger impact on employment probability for the group that was unemployed at baseline.[34]

The reported changes in household behavior suggest that the program may have influenced child skills development as well. Appendix C.4 presents the estimated impacts for each measure available in the data.[35] Two years after baseline, depending on the measure, the program increased the probability of being in the top 30% of the class academic ranking by 4 to 10 percentage points from a baseline of 34 to 48%. The effects five years after baseline are lower, and mostly non-significant. By eight years after baseline, effects are close to zero and statistically insignificant. It is hard to compare what the New Hope studies find with my estimates, as previous studies do not take into account the ordinal nature of the data. Nevertheless, both sets of estimates coincide in finding no effects after year five (Huston et al., 2003; Miller et al., 2008).[36]

I perform two additional exercises that confirm the existence of positive effects on child academic achievement.[37] First, following the step-down procedure of Romano and Wolf (2005, 2016), I adjust p-values of the null hypothesis of no effects for a fixed familywise error rate. Confirming the positive results from the raw analysis of academic measures, I

large impacts on household behavior.

[34]See Appendix C.3. Also, the appendix compares the effects on quarterly employment and employment using survey instead of administrative data. Both groups of estimates are consistent in finding positive effects on employment. Finally, the appendix shows quantile treatment effects on hours worked. The analysis suggests that individuals increased their working hours along both the extensive and intensive of labor supply.

[35]Because these measures are based on rankings, I estimate ordered probit regressions, where the observed dependent variable corresponds to an SSRS measure and the independent variables the random assignment dummy and a set of controls measured at baseline (right before random assignment). Using the probit estimates, I estimate the impacts on scoring “4” or above—which stands for being in the top 30% of the class—on different academic performance measures .

[36]The literature finds this fade-out trend in other rank-free measures of academic skills. See Appendix C.4 for an analysis on other groups of child outcomes. Five years after baseline, I find statistically significant positive effects in two reading skills measures. By eight years after baseline, there are no statistically significant effects. Finally, there are no statistically significant effects on classroom behavior measures.

[37]See Appendix C.4 for details.

find that some the estimated effects two years after baseline do survive the familywise error rate control. Second, I perform a Principal Component Analysis (PCA) on the measures of academic performance to summarize the impacts coming from these different measures. The effect of New Hope on the first PCA score is positive and statistically significant two years after baseline. In this year, the size of the effect is about 0.2 standard deviations, and decreasing thereafter.

Overall, the literature suggests that New Hope had a meaningful impact on household behavior and child outcomes. However, the treatment effects analysis cannot quantify the importance of household variables—as intermediate inputs—in explaining the effects of the program on child academic achievement. Next, I provide an explicit economic structure that allows to assess the role of the household in explaining the impacts on child outcomes.

5 A dynamic-discrete choice model of labor supply, child care, and child's skills

This section presents a dynamic model of the household and child human capital. Since the model incorporates all of the economic constraints defines by New Hope, the EITC, and welfare programs as part of the individual's budget set, it is capable to reproduce the effects of income and child care subsidies on child human capital. Thus, the model is able to shed lights into the household mechanisms by which policies such as New Hope, the EITC, and the CCDF impact child academic skills.

The basic timing and features of the model are as follows. At the beginning of time, a participant receives the New Hope "shock" and draws an initial value of child human capital and child age. Each period, the agent observes her household composition, a wage offer, and the current level of child skills. The individual makes labor supply (not working, part-time, or full-time work) and child care choices (center-based child care or home care) up until the child turns 18 years old. These choices are shaped by various shocks to the agent's budget set—New Hope and the welfare system— and by a dynamic production function of child skills. Next, I present the model formally and explain its components in detail.

Utility function. The individual's current-period utility function corresponds to

$$U(c_t, h_t, \theta_t) = \ln c_t + \alpha^p \mathbf{1}\{h_t = 15\} + \alpha^f \mathbf{1}\{h_t = 40\} + \eta \ln \theta_t, \tag{1}$$

where h_t are weekly working hours. It can only take three values: 0, 15 (part-time work), and 40 (full-time work). c_t is per capita consumption; it represents the average consumption family members enjoy after paying for child care services. α^p, α^f capture the psychic costs or benefits of part-time work, full-time work. η is the preference for the current stock of child human capital. θ_t denotes child human capital. Here, parents observe a "true" value of child human capital, that is, an underlying factor that drives academic achievement. The presence of θ_t in equation (1) implies that the individual makes her choices based on a weighted average of the stock of human capital across time—not just on its long-run value.[38]

Single and married individuals have the same utility function. For single agents, equation (1) represents the utility function of the parent that cares for her child's human capital.[39] For married individuals, for all t, the spouse receives no income.[40] All choices are made by the caregiver of the child. Having a spouse affects choices by adjusting consumption per-capita (more mouths to feed) and the budget set (welfare rules differ by marriage status).

Human capital production function. The technology of child human capital follows

$$\theta_{t+1} = \exp\left(\gamma_0 + \gamma_1 cc_t \mathbf{1}\{a_t \leq 6\}\right) \theta_t^{\gamma_2} c_t^{\gamma_3} \tau_t^{\gamma_4} \tag{2}$$

where cc_t is a child care dummy—equal to 1 for center-based child care and 0 for home care or any informal care at someone's home—and τ_t are weekly hours the individual spends

[38]Hence, the fading-out effects of New Hope can emerge as a rational outcome of the model.

[39]Nearly 90% of participants are single or living alone with the child.

[40]Data on spouse's earnings is available only in New Hope surveys from year two. Among those married or cohabiting (10% of the sample), 44 and 36% of the total income of married respondents in the second-year survey comes from the own participant's and her spouse's earnings, respectively.

with the child. The indicator function next to the child care dummy implies that only an individual with a young child (age $a_t \leq 6$) can use the child care option. The coefficients γ_k, for $k = 1, \ldots, 4$, represent the effect of current-period inputs on next-period child human capital. The constant in the production function (γ_0) is normalized so that $E[\ln \theta_t] = 0$ for $t > 0$. γ_1 is a total-factor-productivity parameter. It captures the human capital gain from center-based child care relative to home care.

Equations (1) and (2) imply that per-capita consumption enters individual's utility both directly and indirectly through the production function. We can interpret the indirect effect in two ways. First, part of what the agent purchases can also affect child human capital (e.g. books, food, etc). Second, having more money at home can relieve stress in the household, which can potentially enhance the parent-child relationship.

An additional child in the family does not directly impact utility (equation 1) or the process of human capital formation (equation 2). Having more children influences choices only through per-capita consumption and the budget set; an additional child in the household, all else equal, lowers c_t and changes the eligibility for welfare programs. For a family with more than one child, the adult makes her choices taking into account how they impact the human capital of a representative child.[41]

Wages. Each period, the individual receives an hourly wage offer (w_t). Following Bernal (2008), Chan (2013), and Del Boca et al. (2013), the offer depends on a vector of observable individual characteristics X_t^w. Furthermore, the wage offer also depends on an individual

[41] Household choices would differ from a multiple-children model—as Todd and Wolpin (2006) and Tartari (2015)—only in the case where there are young and old children at the same period in a given household (which occurs in 28% of the cases). Compared to such framework, average choices should not deviate as much (even though, at the individual level, choices would be different). Another option would be to disregard families with more than two children (Bernal, 2008; Del Boca et al., 2013), which would mean losing more than 50% of the sample.

productivity that follows an AR(1) process. Formally, the wage offer process is given by

$$\begin{aligned}\ln w_t &= X_t^{w'}\beta^w + \nu_t^w,\\ \nu_t^w &= \rho\nu_{t-1}^w + \epsilon_t^w\\ \epsilon_t^w &\sim N(0,\sigma_w^2)\end{aligned} \tag{3}$$

where X_t^w includes age, age squared, a dummy variable for high school diploma, a constant, and the log of t. This last term captures a non-linear trend of the wage offer process. The agent knows the coefficient associated with $\ln(t)$ beforehand, and so she anticipates an average trend on her wage offer.

Parental education and child human capital are related via individual choices. The level of parental education is an input in the wage process, which in turn affects labor supply and child care choices. Because human capital is affected by income, time, and child care, parental education has an indirect effect on child skills.

Budget set. The budget set incorporates various features of the welfare system. Income is a function of labor supply, earned income, and various mean-tested programs. Conditional on eligibility, the agent receives payments from New Hope, the EITC, the AFDC or Food Stamps.[42] Eligibility to these programs and payment amounts depend on working hours, earned income, and family composition.

Income can be represented as follows. Let k_t and m_t be the number of children and a marriage indicator (1 if the household has two adults married to each other or living together and 0 otherwise). Income (I_t) is given by

$$\begin{aligned}I_t = {} & w_t h_t \times 52 + EITC_t(w_t h_t \times 52, k_t, m_t) + NH_t(w_t h_t \times 52, k_t, m_t).\\ & + B_t + S_t\end{aligned} \tag{4}$$

In the equation above, $EITC_t(.)$ corresponds to EITC payments. If the individual is

[42]New Hope administrative data does not include other forms of welfare payments.

eligible to receive these payments, she always comply.[43] The same happens with the New Hope payments, $NH_t(.)$. B_t and $SNAP_t$ are cash transfers from AFDC (or TANF) and Food Stamps (now known as SNAP). As with New Hope and the EITC, the individual always chooses to take up the benefits of AFDC and Food Stamps payments (Blundell et al., 2016).[44] However, she faces a random i.i.d. take-up shock, which captures the possibility of being misinformed about the welfare system. Specifically, at the beginning of each period, the individual draws two values, $\rho_t^B, \rho_t^S \in \{0, 1\}$, from a pair of known, time-invariant binomial distributions, indicating whether the individual takes up the corresponding payment or not. These information shocks affect only AFDC and Food Stamps. Hence, the available money from AFDC and Food Stamps follows:

$$B_t \equiv \rho_t^B B_t^*(w_t h_t \times 52, k_t, m_t),$$
$$S_t \equiv \rho_t^S S_t^*(w_t h_t \times 52, k_t, m_t),$$

where $B_t^*(.)$ and $S_t^*(.)$ are the potential AFDC and Food Stamp payments.

Each of the payment functions $EITC_t(.)$, $NH_t(.)$, $B_t^*(.)$, and $S_t^*(.)$ are given by precise formulas determining eligibility and payment levels. They are a function of the level of earnings, labor supply, and family composition (k_t and m_t). These rules may change from year to year.[45] Nonetheless, at $t = 0$, families have perfect information regarding the evolution of rules of the welfare system (the uncertainty faced in this context is in future misinformation shocks about AFDC and Food Stamps).[46] Moreover, eligibility rules are always enforced,

[43]The EITC national take-up rate is estimated at over 80% (Scholz, 1994; Plueger, 2009; Hoynes and Rothstein, 2016). Furthermore, New Hope representatives took care to advise participants about how to take advantage of the EITC (Bos et al., 1999). As for New Hope, assuming eligibility on an annual basis and using the definitions of hours worked and gross income that is consistent with the data for estimating the structural model, I estimate a take-up rate of 92%.

[44]An alternative structural framework would allow individuals to choose whether they want to take up the benefits and include taste parameters ("welfare stigma" coefficients) associated with each program. However, I do not have enough sources of exogenous variation to identify stigma coefficients (Keane and Moffitt, 1998; Chan, 2013). Appendix D (figures D.1 and D.2) shows that take-up rates for the AFDC and Food Stamps do not follow an obvious pattern across income quantiles.

[45]See Appendix D for details.

[46]In the case of New Hope, the program's representatives explained the details of the benefits package to all participants. Furthermore, representatives were available throughout the eligibility period to answer any

including the New Hope work requirement.[47]

During the time frame of the model, AFDC is replaced by TANF. For this sample, the State of Wisconsin implemented the Wisconsin Works program (W-2), which eliminated AFDC's unconditional cash transfers and established time limits for welfare utilization. Specifically, W-2 offered paid community service jobs at a flat rate. So from 1997 onward ($t = 2$), the wage of the state-provided CSJ is part of the pool of potential log-wage offers (equation 3). Participants do not face time limits for W-2.[48]

Per-capita consumption. Consumption is defined as money net from child care expenditures. Agents have access to child care services at a fixed, known price p for all t. Nonetheless, she can get a cheaper price in two ways. First, at the beginning of every period, the individual draws a value $q_t \in \{0, 1\}$ indicating whether she has a free child care offer (in which case $q_t = 0$) from a known binomial process. If $q_t = 0$ and chooses to use a child care, she always takes up the free slot. Both free and paid child care have the same associated TFP coefficient in the production function. Second, if the agent is in the New Hope treatment group and works full time, she gets a lower copayment, $\underline{p} \leq p_t$ which depends on the level of earnings (see Section 2). Formally, the cost function equals:

$$\delta(q_t, p, D, h_t) \equiv \begin{cases} q_t\underline{p}\mathbf{1}\{h_t = 40\} + q_t p\mathbf{1}\{h_t < 40\} & \text{if } D = 1, \\ q_t p & \text{otherwise.} \end{cases} \tag{5}$$

The individual cannot save or borrow.[49] Per-capita consumption is thus given by:

$$c_t = \frac{I_t - cc_t \times \delta(q_t, p, D, h_t)}{1 + m_t + k_t}. \tag{6}$$

questions participants might have had (Brock et al., 1997).

[47]New Hope agents implemented various procedures to ensure that requirements were met. See Section 2.

[48]The model's version of W-2 does not include time limits because it would require having data on labor supply beyond 2003.

[49]There is little evidence suggesting that individuals are able to save for future consumption. In the control group, from the year-five interview, 58% manifested some concern about not having enough money to buy food. Additionally, a large share does not access to banking services, 42% of individuals do not have a checking account, and 52% do not have a savings account.

Parental time. Time with the child (τ_t) depends on labor supply choices (h_t) and child care ($cc_t \in \{0, 1\}$). Let $\overline{T}$ be the total available time the adult has in a week. Mathematically, τ_t is defined as

$$\tau_t \equiv cc_t(\overline{T} - 40) + (1 - cc_t)\left(\overline{T} - h_t\right) \tag{7}$$

The logic behind equation (7) is as follows. If the child spends all week in home care ($cc_t = 0$), then the labor supply choice determines how much time the adult spends with the child. In this case, there are three possible scenarios. If the individual does not work, then she spends all the available time with the child ($\tau_t = \overline{T}$). If she works part time, then she must spend 15 hours a week away from home, so $\tau_t = \overline{T} - 15$. Analogously, $\tau_t = \overline{T} - 40$ if she works full time. If the child spends her time in child care ($cc_t = 1$), then she spends 40 hours a week outside the house being cared in a child care center. Hence, if $cc_t = 1$, then $\tau_t = \bar{T} - 40$ no matter how many hours the adult spends working.

Equations (1), (2), and (7) make evident the benefits and costs the agent faces when choosing labor supply and child care. Intuitively, child care allows the individual working more and thus having more income without reducing child human capital (if $\gamma_4 > 0$ in equation 2). Thus, she can consume more (equation 1) and have more income to produce further child human capital. At the same time, child care has a direct, positive effect on child human capital (if $\gamma_1 > 0$ in equation 2). Hence, the benefits of child care are twofold: it produces child human capital and it lowers the cost of labor supply.

The discussion above draws no distinction between leisure and passive or active time spent with a child. Therefore, following equations (2) and (7), any time allocated outside the labor market (if the child is at home) has a constant effect on human capital.[50] Moreover, given the utility function (equation 1), the adult enjoys her leisure hours (dislikes her work hours) to the same degree regardless of whether or not the child is at a child care center or

[50] Because New Hope data does not have time diaries, I cannot distinguish between passive or active time with the child (Del Boca et al., 2013; Brilli, 2014). Nonetheless, the literature consistently shows that non-working mothers do spend more time with their children than working mothers (Guryan et al., 2008).

remains at home.

Family composition. Marriage formation and childbearing are exogenous processes. Each period, the individual draws a marital status (1 if married and 0 if single) and childbearing values (1 if there is a new child in the family and 0 otherwise) from known binomial distributions with probability parameters m_t^* and k_t^*. These probabilities depend on observed participant characteristics and past family composition, as follows:

$$m_{t+1}^* = f_m(X_t^m, m_t), \tag{8}$$

$$k_{t+1}^* = f_k(X_t^k, k_t, m_t), \tag{9}$$

where m_t equals 1 if the participant is married or living with her partner and 0 otherwise, and k_t indicates the number of children in the household. X_t^m includes a constant, age of the adult, and a dummy variable for high school diploma. X_t^k includes a constant, age, and age squared of the participant.

The dynamic problem. In each period, given a set of state variables, the individual solves a dynamic problem of labor supply and child care choices. The state variables of the problem are collected in the vector $\boldsymbol{s}_t = (D, m_t, k_t, a_t, \theta_t, \boldsymbol{X}, \nu_t^w, \boldsymbol{\rho}_t, q_t, p_t)$, where $\boldsymbol{X}$ contains the wage offer, marriage, and childbearing processes observed control variables, $\boldsymbol{X} \equiv \left(X^w, X^m, X^k\right)$, and $\boldsymbol{\rho}_t$ the misinformation shocks to welfare take-up, $\boldsymbol{\rho}_t \equiv \left(\rho_t^B, \rho_t^S\right)$. For a given $\boldsymbol{s}_t$, each period the agent maximizes the present discounted value of the utility stream by choosing labor supply and child care type. Let $\mathcal{C} = \{0, 1\}$ and $\mathcal{H} = \{0, 15, 40\}$ be the choice sets of child care and labor supply. We can represent the entire choice set, for any period, as $\mathcal{J}(a_t) = \mathcal{C} \times \mathcal{H}$ if the child is young ($a_t \leq 6$) and $\mathcal{J}(a_t) = \mathcal{H}$ otherwise ($a_t > 6$).

Let $u(\boldsymbol{s}_t, j)$ be the current-period utility for a given state $\boldsymbol{s}_t$ and choice $j \in \mathcal{J}(a_t)$. For any t, the problem of the individual is represented in the usual recursive formula

$$V_t(\boldsymbol{s}_t) = \max_{j \in \mathcal{J}(a_t)} \{V_t^j(\boldsymbol{s}_t)\} \quad \text{subject to (1)-(9)},$$
$$V_t^j(\boldsymbol{s}_t) = u(\boldsymbol{s}_t, j) + \beta E[V_{t+1}(\boldsymbol{s}_{t+1}) \mid \boldsymbol{s}_t, j] \qquad t < T(a_0)$$

The model is closed with initial and terminal conditions. At baseline ($t = 0$), individuals draw a a pair of values defining family composition (m_0, k_0) and pair of values of initial age and child human capital (a_0, θ_0). The initial θ_0 is related to the parent's unobserved characteristics. In particular, the initial shocks to unobserved productivity and child human, ε_0^θ and ε_0^w , capital follow a joint normal distribution with correlation coefficient ρ_θ.

Because agents draw different initial values for the child's age, each individual solves a problem of a different time horizon. Let $T(a_0) \equiv 18 - a_0$ be the terminal period for an individual with a a_0-year-old child. Thus, for a one-year-old child arriving in period $t = 0$, the parent solves the dynamic problem for 17 years after baseline, stopping when the child turns 18. In this final period, the individual can no longer invest in child human capital. The associated terminal value function is such that

$$V_{T(a_0)}^j(s_{T(a_0)}) = \max_{j \in \mathcal{J}_{T(a_0)}} \left\{\tilde{u}(\boldsymbol{s}_{T(a_0)}, j)\right\} + \eta \ln \theta_{T(a_0)}.$$

6 Identification and estimation

6.1 Identification

To better understand the sources of the data that identify the model, I divide the discussion into two related concepts: global (or point) identification and local identification.

Global identification. Identification of the structural parameters builds on distributional and functional form assumptions.[51] Nevertheless, to reduce the dependence on such

[51] See Rust (1994) and Magnac and Thesmar (2002).

structural assumptions, the policy environment within the model offers various sources of exogenous variation. Within the time frame of the model, these shocks are the TANF implementation, EITC expansions, and the New Hope random assignment. Moreover, the money received from New Hope and welfare programs depends on family structure (number of children and marriage status) and there are several discontinuity points in the rules of the different programs. All of these policy changes within the model affect labor supply and child care decisions without directly changing preferences or the wage offer equation. Therefore, global identification hinges on the comparison of family choices and outcomes across periods (before and after policies are implemented), treatment groups, family composition, and at different points of the wage offer distribution.

To identify the production function (equation 2), I use observed data on inputs and the SSRS measure of overall academic achievement. The econometrician observes a noisy measure (M_t) of child human capital. It is a ordinal measure that is given by

$$M_t = \begin{cases} 1 & \text{if } \ln\theta_t + \epsilon^z_{t,m} \leq \kappa_{1,t} \\ 2 & \text{if } \kappa_{1,t} < \ln\theta_t + \epsilon^z_t \leq \kappa_{2,t} \\ \vdots & \qquad\vdots \\ 5 & \text{if } \ln\theta_t + \epsilon^z_t > \kappa_{4,t}, \end{cases} \tag{10}$$

where θ_t is observed by the families but not by the econometrician and ϵ^z_t is measurement error that follows a known distribution.

A key identifying assumption is that the production function is constant in time and it does not vary by age (except in the child care component). Observing human capital measures for different samples at different points in time helps to identify the production function and the structure of the measurement system. Appendix E provides the formal identification argument. Intuitively, identification follows in two steps. First, one can show identification of the technology of young children at home care with data on inputs (consumption and hours at home) and one measure of human capital observed for two periods. This result follows because the production function does not vary by period. Once the technology for a young child at home is identified, we can use the fact that cutoffs ($\kappa_{j,t}$, for $j = 1, 2, 3, 4$) are constant across child age and child care types to identify the TFP parameter.[52]

All cutoffs $\kappa_{j,t}$, can vary in time. The assumption of time-varying cutoffs restricts the family of productions functions that can be estimated. In particular, this assumption implies that the constant term in the production function (equation 2) cannot vary freely.[53]

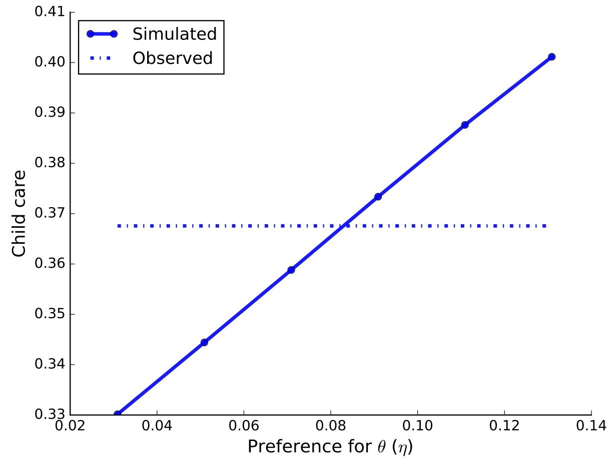
Local identification. Some non-experimental moments contain identifying information at the local level; around a vicinity of the true value of a structural parameter, and holding the rest of parameters constant, there is only one value of the parameter that generates the sample moment. To illustrate the sources in the data that contribute to the identification of the structural parameters, Figure 3 plots the relationship between a simulated moment and a structural parameter, holding the rest of the parameters fixed.[54] Here, the simulated moment has a monotonic relationship with the structural parameter, crossing the observed moment only once. In this way, the chosen moment contains sufficient identification power at the local level—around a vicinity of the optimal parameter value, holding other parameters constant—to pin down the structural parameter. In this case, I plot the preference for human capital (η in equation 1) and the proportion in the sample who use child care. A higher preference for θ means that the agent is willing to sacrifice more consumption for higher levels of child human capital. Thus, for a given child care cost and holding other parameters fixed at their estimated values, a bigger η implies a larger probability of choosing center-based child care. The simulated probability of child care monotonically increases, crossing

[52]Factor loadings—the coefficient associated with $\ln \theta_t$—equal 1 for every period. This assumption is necessary for identification, given that there are no baseline measures of child human capital. By assuming factor loadings are known, the requirement of having multiple measures at baseline is no longer needed. A byproduct of this assumption is that I only need one measure per period to identify the production function. Furthermore, I do not need to impose the requirement that the coefficients of the production function sum up to one. See Agostinelli and Wiswall (2016) for the necessary conditions to identify a production function with an unknown scale.

[53]Only with constant cutoffs across time (or that follow a predictable pattern), one would be able to identify γ_0 as a free parameter and thus let the production function have a time-varying TFP (Agostinelli and Wiswall, 2016).

[54]Voena (2015) and Autor et al. (2017) follow the same approach to show local identification.

Figure 3: Local identification of the preference for human capital



Notes: The figure plots the proportion of children in child care from the model against the structural parameter of preference for human capital (solid line). This parameter is locally identified at the crossing point with the corresponding observed moment (depicted in the dotted line).

the observed value in the data only once. Hence, at this crossing point, η is locally identified. As I explain next, to exploit these sources of identification, this and other moments that meet the single-crossing property are used directly in the estimation procedure.

6.2 Estimation

For estimation purposes, I proceed in two steps.[55] In the first step, I estimate the parameters of some of the exogenous processes straight from the data. In the second step, I estimate the rest of the structural models using the simulated method of moments.

External estimation and calibration. Table 3 summarizes the sources for external estimation an calibration. To obtain the parameters governing the probability of being married (m^*_{t+1}) and of childbearing (k^*_{t+1}), consider the following linear probability models:

$$m_{t+1} = X^m_t \beta^m + m_t \gamma^m + \epsilon^m_{t+1}, \tag{11}$$

[55]This method is a standard practice in the literature. For instance, see Gourinchas and Parker (2002), De Nardi et al. (2010), Voena (2015), and Blundell et al. (2016). The goal with the two-step procedure is to keep estimation computationally feasible.

$$k_{t+1} - k_t = X_t^k \beta^k + k_t \gamma^k + m_t \gamma^{k,m} + \epsilon_{t+t}^k. \tag{12}$$

Since I do not have marriage data for two years in a row, I cannot directly estimate the parameters of equations (11) and (12). To circumvent this problem, I estimate a linear probability model of m_{t+1} on m_{t-1} and X_{t-1}^m, and use the resulting reduced-form parameters to identify β^m and γ^m.[56] I implement a similar method to identify β^k, γ^k, and $\gamma^{k,m}$.[57] Given the estimated parameters of equations (11) and (12), the parameters of the the binomial distribution determining the probabilities of marriage and childbearing (equations 8 and 9) are given by $m_{t+1}^* = X_t^m \widehat{\beta}^m + m_t \widehat{\gamma}^m$ and $k_{t+1}^* = X_t^k \widehat{\beta}^k + k_t \widehat{\gamma}^k + m_t \widehat{\gamma}^{k,m}$.

I determine the rest of the parameters of the exogenous processes to match different observed statistics. I set the probability of obtaining a free child care to 0.57, corresponding to the the share of individuals in the control group, conditional on $cc_t = 1$, who declare not paying for child care. To calibrate the child care price, I take the value reported in Bos et al. (1999) corresponding to the average sum of individual copayments and child care subsidies paid by the program ($750 a month). I define the probability of receiving AFDC and Food Stamps—conditional on being eligible—as the average take-up observed in the data (60% and 70%, respectively). Finally, I follow Chan (2013) and set the discount factor to $\beta = 0.86$, which is a middle point between the equivalent parameters of Swann (2005) and Keane and Wolpin (2010).

[56]To estimate the regression I use data for the second-year survey and baseline information.

[57]Precisely because I do not have data for two periods in a row, I was unable to implement a logistic form to estimate the marriage and childbearing processes.

Table 3: Calibrated and externally estimated parameters

Parameter/equation	**Source for estimation/calibration**	**Values**
Probability of being married	OLS: m_{t+1} on m_{t-1} and X^m	$m^*_{t+1} = 0.21 - 0.002age + 0.8m_t$.
Probability of childbearing	OLS: $(k_{t+1} - k_t)$ on m_{t-1}, k_{t-1} and X^k	$k^*_{t+1} = -0.11 + 0.05age -$ $+0.0003age^2 - 0.006k_t - 0.1m_t$
Probability of free child care	Share of participants reporting not paying for child care.	0.57
Child care price	Bos et al. (1999)	\$750 monthly
Take-up probabilities of AFDC and SNAP	Average AFDC and SNAP take-up conditional on eligibility	0.6 and 0.7

Notes: The table describes the sources for estimation or calibration of the structural parameters determined outside the estimation procedure.

Internal estimation. In a second step, I use the Simulated Method of Moments to estimate the rest of the parameters. The procedure compares the estimated moments of an auxiliary model using observed data on choices and exogenous variables with equivalent estimates from the model-simulated data (Gourieroux et al., 1993).

I use backward induction to solve the model and obtain paths of simulated choices. Because $\boldsymbol{s}_t$ contains continuous variables, obtaining an exact solution for $V_t(\boldsymbol{s}_t)$ at every point of the state space is computationally unfeasible. Thus, following Keane and Wolpin (1994) and Keane et al. (2011), I compute $V_t(\boldsymbol{s}_t)$ for a grid of the state space and then use linear interpolation (in which I include polynomial terms of the state variables) to approximate $V_t(\boldsymbol{s}_t)$ for values outside the grid. The size of the grid equals 4,536. Finally, I use Monte Carlo integration (with 50 draws) to estimate the multivariate integral.[58]

The estimation problem can be stated as follows. Let $\hat{g}$ be the vector of moments extracted from the data. I solve the model $M = 1,000$ times for a sample size of $n = 691$ children and compute the required moments from simulated data. Let $\{\boldsymbol{\epsilon}^m_t\}^M_{m=1}$ denote the structural random shocks (fixed across the estimation procedure). Let ψ be the vector of structural parameters of the model. Define $\{y^m_{it}(\psi)\}^M_{m=1}$ as the simulated choices associated with the M draws of structural random shocks. Let $\hat{g}^m(\psi)$ be the equivalent moment

[58] I set the grid size and the number of draws for the Monte Carlo integration in order to balance precision and computational time.

associated with the m draw. I estimate the structural parameters ψ by solving

$$\hat{\psi} = \arg\min_{\psi \in \Psi} [\hat{g} - \hat{g}(\psi)]' W [\hat{g} - \hat{g}(\psi)],$$

where $\hat{g}(\psi) = \frac{1}{M} \sum_{m=1}^{M} \hat{g}^m(\psi)$ and $\widehat{g}$ is the vector of auxiliary estimates from the data.

Following Del Boca et al. (2013) and Blundell et al. (2016), I define W as the inverse of the diagonal of the estimated variance-covariance matrix of $\hat{g}$. I do not use the efficient weighting matrix because of its poor small-sample properties (Altonji and Segal, 1996). I use the bootstrap method (1,000 samples) to estimate W.

Targeted moments. Table 4 lists the set of auxiliary estimates used in estimation. Estimation exploits a set of unconditional and conditional moments. The matched moments are non-experimental, while experimental moments are left for validation. Based on the argument given in Section 6.1 (see Figure 3), estimation targets moments that provide identification at the local level.[59] To estimate preferences for hours of work and child human capital, I use the labor supply and child care choices of children's parents. To estimate the production function process, I use the correlation of consumption and parental time with the child (as defined in equation 7) with the raw SSRS rankings. To estimate the measurement system, I include various moments capturing the distribution of children in the SSRS rankings from two and five years after random assignment.[60] Finally, to estimate the wage offer process, the auxiliary model includes the OLS coefficients of a regression of log wages onto the variables discussed in the context of equation (3). Table 4 shows that model successfully replicates most of the target moments.

[59]Appendix F.2 shows that each of the chosen moments locally identify a structural parameter.

[60]I use the "overall" SSRS measure. This variable shows the reported rankings based on the overall academic performance of the child in a classroom. In the PCA analysis from Section 4, this measure has the highest correlation with the first PCA component in years two and five.

7 Model estimates

7.1 Estimates

Table 5 presents the estimated structural parameters. Panel A shows the parameters of the utility function (equation 1). Since this function is expressed in log-consumption units, taste parameters represent what the agent is willing to pay (in terms of a percentage change in

Table 4: Target moments

Moments	**Simulated**	**Data**	**S.E. data**
A. Labor supply and child care decisions			
$Pr(\text{child care}_t \mid RA = 0)$, $t = 1$ age ≤ 6	0.366	0.368	0.033
$Pr(\text{part-time}_t \mid RA = 0)$, $t = 0$	0.288	0.296	0.014
$Pr(\text{full-time}_t \mid RA = 0)$, $t = 0$	0.442	0.446	0.016
B. $log(wage_t) = X_t'\beta + \epsilon_t$			
Coefficient on age	-0.023	-0.020	0.006
Coefficient on age squared	0.000	0.000	0.000
Coefficient on high school dummy	0.243	0.247	0.043
Coefficient on $log(t)$	0.374	0.382	0.038
Constant	1.539	1.549	0.063
σ^2	0.476	0.466	0.053
AR(1) shock (ρ)	0.382	0.368	0.056
C. SSRS and household choices			
$Corr[SSRS_2, SSRS_5]$	0.379	0.377	0.058
$Corr[\text{consumption}_1, SSRS_2]$	0.037	0.027	0.037
$Corr[\text{time}_1, SSRS_2]$	-0.061	-0.043	0.043
$E(SSRS_2 \mid cc = 1) - E(SSRS_2 \mid cc = 0)$	0.214	0.253	0.137
$Corr(SSRS_2, \ln w_0)$	-0.022	-0.012	0.061
D. SSRS ($t = 2$)			
$Pr(SSRS_2 = 2)$	0.153	0.152	0.021
$Pr(SSRS_2 = 3)$	0.360	0.360	0.025
$Pr(SSRS_2 = 4)$	0.226	0.226	0.023
$Pr(SSRS_2 = 5)$	0.151	0.149	0.020
E. SSRS ($t = 5$)			
$Pr(SSRS_2 = 2)$	0.217	0.216	0.020
$Pr(SSRS_2 = 3)$	0.292	0.292	0.023
$Pr(SSRS_2 = 4)$	0.173	0.174	0.019
$Pr(SSRS_2 = 5)$	0.188	0.188	0.019

Notes: This table compares the simulated and observed estimated moments that are targeted in estimation. $SSRS_t$ corresponds to overall SSRS measure of academic achievement in period t. In this measure, teachers rank children in a five-point scale based on the overall academic performance in the classroom. $time_t$ corresponds to time with the child (τ_t). cc_t is child care in period t for children who are less than six years old. The rest of the variables are constructed following Appendix F.1.

current-period consumption) to compensate a marginal increase in one input while keeping others inputs fixed. The coefficient on full-time work equal -0.21, which means that the individual is willing to accept working full time if she increases its consumption today by 21%. This is lower than that of Blundell et al. (2016); for single mothers, the equivalent parameter lies in the 33-38% interval (it varies across education levels). This difference may be explained by the nature of New Hope sample. Individuals were explicitly willing to work full time, and so their utility cost of working may be lower than the analogous cost from a broader sample. Consistent with this hypothesis, the coefficient on part-time work is also lower than the Blundell et al. estimates. Child human capital is positively valued by the agent, with $\eta = 0.081$. This value implies that the individual is willing to sacrifice 9% ($= 0.081/0.88$, where 0.88 is the estimated standard deviation of $\ln \theta_t$) of her consumption for a one-standard-deviation increase in child human capital.[61]

Panel B shows the estimated parameters of the wage offer process (equation 3). For a given period, the estimates imply a negative and almost linear effect of age on the wage offer. Nonetheless, the wage offer increases for everyone, capturing a growing labor demand.[62] A high school diploma increases the wage offer by 23%. This estimate is higher than the return to high school graduation for men estimated Heckman et al. (2016a) and Heckman et al. (2016b)(which is approximately 10%). The variance of the wage process (0.31) is higher and the autocorrelation coefficient of the unobserved component of the wage offer (0.65) is lower compared to what Blundell et al. (2016) find for women without a high school diploma. Hence, individuals in my sample face a larger degree of uncertainty regarding future wage shocks.

I show the estimated parameters of the production function and measurement system in Panels C-E. Consumption has a positive effect on human capital, although somewhat lower

[61] Here, the literature provides a wide range of estimates: Bernal (2008) estimates an almost 0 coefficient, while Del Boca et al. (2013) documents that a one-percent increase in child human capital is more valued than the same increase in consumption.

[62] Employment probability for both treatment and control groups grows throughout the covered period 1994-2003. See Miller et al. (2008) and Section 2.

Table 5: Estimated structural parameters

Parameter	Estimate	S.E.
A. Utility function		
Preference for part-time work (α^p)	-0.010	0.054
Preference for full-time work (α^f)	-0.209	0.063
Preference for human capital (η)	0.081	0.060
B. Wage offer		
Age	-0.022	0.008
Age^2	0.000	0.000
High school	0.227	0.042
log(t)	0.384	0.035
Constant	1.449	0.103
AR(1) error term	0.647	0.047
Variance of error term	0.310	0.111
C. Production function		
Child care TFP (γ_1)	0.288	0.236
Lagged human capital (γ_2)	0.948	0.037
Consumption (γ_3)	0.037	0.040
Time at home (γ_4)	0.121	0.694
$Corr(\varepsilon_0^\theta, \varepsilon_0^w)$	-0.076	0.219
D. SSRS ($t = 2$)		
κ_1	-1.518	0.157
κ_2	-0.714	0.197
κ_3	0.574	0.137
κ_4	1.552	0.175
E. SSRS ($t = 5$)		
κ_1	-1.200	0.188
κ_2	-0.248	0.201
κ_3	0.733	0.283
κ_4	1.424	0.277

Notes: This table shows the estimated parameters of the model presented in Section 5. The utility function follows $U(c_t, h_t^p, h_t^f, \theta_t) = \log c_t + \alpha^p h_t^p + \alpha^p h_t^f + \eta\theta_t$. The wage offer obeys $\ln w_t = X_t^{w\prime}\beta^w + \epsilon_t^w$, where X_t^w includes a constant, age, age squared, a dummy for high school diploma and $\epsilon_t^w \sim N(0, \sigma_w^2)$. The production function is given by $\theta_{t+1} = \exp\left(\gamma_0 + \gamma_1 cc_t \mathbf{1}\{a_t \leq 6\}\right)\theta_t^{\gamma_2} c_t^{\gamma_3} \tau_t^{\gamma_4}$.

than what previous studies have shown. The standard practice is to compute the effect of a 1,000-dollar increase on human capital. In my model, this marginal effect depends on baseline consumption, labor supply, the initial human capital stock, and family composition. Assuming no behavioral responses and holding constant labor supply, a 1,000-dollar boost rises child skills by 0.6% of a standard deviation, which is smaller than what is reported in Dahl and Lochner (2012) and Dahl and Lochner (2016).[63] Nonetheless, as Dahl and Lochner (2012) suggests, their estimate may capture a long-term impact. As I show below, I can reproduce a bigger impact on child human capital when policies giving cash to families are in place for a long time.

Time at home has a positive effect on children's human capital. Conducting a similar experiment as the paragraph above, I compute the effect of going from $h_t = 40$ to $h_t = 0$ (*ceteris paribus*), assuming no one uses child care. This labor supply change rises child human capital by only 4% of a standard deviation. This small effect of time with the child is at odds with the literature that places a relative large value of time inputs in the production function of child skills (Cunha et al., 2010; Del Boca et al., 2013; Attanasio et al., 2015). Nonetheless, newer evidence suggests that the productivity of time inputs may be even negative when within-home interactions are of low quality (Elango et al., 2016).

I find that child care has a sizable effect on child human capital. My estimates imply that choosing child care instead of home care in period t increases child skills by 0.32 standard deviations in period $t + 1$. This large effect coincides with findings showing sizable effects of child care relative to home care. Using an index of cognitive skills measures, Kline and Walters (2016) find that the impact of attending Head Start on those who are drawn from home care equal 0.37 standard deviations. For a different cognitive skills measure, Feller et al. (2016) estimates an effect of 0.23 standard deviations of Head Start versus home care.

I find that human capital production function contains substantial persistence. The estimated AR(1) coefficient in the production function —the so called "self-productivity" coefficient in the context of the human capital technology (Cunha and Heckman, 2006)—equals 0.95. The relatively high persistence in the production function is a consistent finding in the literature (Cunha and Heckman, 2006; Cunha et al., 2010; Attanasio et al., 2015) In particular, for a similar linear production function, Cunha and Heckman (2008) find an autoregressive coefficient equal to 0.97. Coupled with dynamic complementarities, a strong

[63] Bernal (2008) and Del Boca et al. (2013) also find that money plays a modest role in explaining child cognitive outcomes.

self-productivity component implies that any shock to the human capital process at early ages has almost permanent consequences for skills production in the future. As I show below, this feature of the human capital technology has important implications for predicting the effects of policies such as the EITC on human capital acquisition in the long run.

7.2 Validation

Before analyzing the counterfactual experiments, I evaluate the model's capacity to predict non-targeted moments. The moments that were not used in estimation come from experimental moments. This form of validation is rarely used in the structural literature on child outcomes and household behavior (Bernal, 2008; Del Boca et al., 2013). Table 6 compares the model-generated impact of New Hope on child care, hours worked, and consumption with the estimated effects from the experimental data.[64] The model predicts a higher impact on hours worked (eight hours a week in the model versus 4 to 5 hours in the data), a similar impact on child care (17 percentage points in the model versus almost 20 percentage points in the data), and a higher impact on log per capita consumption (0.6 log-consumption units versus 0.3 in the data).

The lack of predictive power in some of the experimental moments should not affect overall conclusions. The impact of New Hope on child human capital (as I show next) is mostly explained by the child care component, with income and time playing a minor role. This result is explained because the productivity parameters of income and time are relatively small (see the discussion in the previous section). Hence, the upward bias in predicting the effects of New Hope on income and labor supply are heavily discounted in the production function. Taking into account the overshooting in the treatment effects on income and hours worked would only reinforce the general qualitative conclusions about the importance of child care in explaining the effects of New Hope on child outcomes.

[64] Appendix F presents the auxiliary model used for estimation and compares the simulated with observed auxiliary estimates. The model is able to match with high precision the observed auxiliary estimates.

Table 6: Simulated and observed treatment effects on household variables

Treatment effect	Simulated	Observed
Hours worked ($t = 0$)	7.921	4.621
		(1.051)
Hours worked ($t = 1$)	8.124	4.381
		(1.173)
Child care ($t = 1$)	0.167	0.196
		(0.052)
Log consumption ($t = 1$)	0.601	0.313
		(0.096)

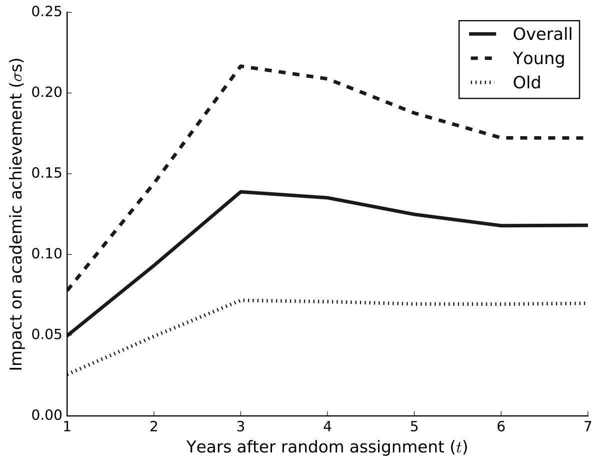
Notes: This table compares the simulated treatment effects on household variables against that of the actual data.

8 Explaining the impact of income and child care subsidies

8.1 Understanding the effects on New Hope

Figure 4 depicts the impact of the program on child human capital (in standard deviations of $\ln \theta_{t+1}$). For the overall sample, the simulated impact is lower (about 0.1 standard deviations) than the PCA-based estimate of 0.2 standard deviations for $t = 2$ (see Section 4). However, the effect on children attains a maximum by $t = 3$, where it reaches 0.15 standard deviations. The figure also shows the effects on child human capital for two groups: children who were six years old or less after the first two years of the program ("young") and children who were more than six years old at that time ("old"). Since the younger group had the child care option while the program was running ($t = 0, 1, 2$), we can expect a bigger effect of the program on this group's child human capital. The figure shows substantial differences in the predicted effects. By period $t = 3$, the program boosts the young group's average human capital by 0.22 standard deviations, almost four times the estimated effect in the "old"

Figure 4: The impact of New Hope on child human capital (θ_t)



Notes: The figure shows the impact of New Hope on θ_t in standard deviation units. "Young" corresponds to children who are six years of age or less by $t = 2$, while "Old" children over six years old by $t = 2$.

group. Finally, the model predicts fading-out treatment effects. Nonetheless, the fading-out is not as pronounced as the one estimated from the actual data. The next section offers a framework to understand how these effects are generated through changes in household behavior.

8.1.1 Mediation analysis

In this section, I assess the role of labor supply, income, and child care in New Hope's impact on child skills.[65] To this end, I implement a mediation analysis based on the structural dynamic model.[66] Because the objective is to analyze the mediating roles of labor supply, money, and child care, all of the subsequent analyses is based on the sample who were six years of age by period $t = 2$—and so they were exposed to the child care subsidy policy throughout the New Hope period.

[65]With the exception of Epps and Huston (2007), previous literature does not have a formal analysis of the mediating factors that lead to the observed impacts on child outcomes. See Huston et al. (2001, 2005, 2011).

[66]This is similar to the approach taken in Heckman et al. (2013) and Attanasio et al. (2015). To understand the factors mediating the effects on child skills from early childhood interventions, the authors posit a framework that relies less on structural assumptions to estimate the production function. However, since these papers do not explicitly model parents behavior, they cannot generate the types of policy counterfactuals that I present in this paper.

Consider the following representation of the academic achievement production function:

$$\theta_{t+1}^d = f(\theta_t^d, \tau_t^d, cc_t^d, c_t^d)$$

where $d \in \{0, 1\}$ indicates assignment to the treatment group. The individual-level treatment effect of the program corresponds to $\theta_{t+1}^1 - \theta_{t+1}^0 = f(\theta_t^1, \tau_t^1, cc_t^1, c_t^1) - f(\theta_t^0, \tau_t^0, cc_t^0, c_t^0)$. This term can be decomposed as

$$\begin{aligned}
\theta_{t+1}^1 - \theta_{t+1}^0 &= \underbrace{[f(\theta_t^1, \tau_t^1, cc_t^1, c_t^1) - f(\theta_t^1, \tau_t^1, cc_t^1, c_t^0)]}_{\text{explained by consumption}} \\
&+ \underbrace{[f(\theta_t^1, \tau_t^1, cc_t^1, c_t^0) - f(\theta_t^1, \tau_t^1, cc_t^0, c_t^0)]}_{\text{explained by child care}} \\
&+ \underbrace{[f(\theta_t^1, \tau_t^1, cc_t^0, c_t^0) - f(\theta_t^1, \tau_t^0, cc_t^0, c_t^0)]}_{\text{explained by time}} \\
&+ \underbrace{[f(\theta_t^1, \tau_t^0, cc_t^0, c_t^0) - f(\theta_t^0, \tau_t^0, cc_t^0, c_t^0)]}_{\text{explained by self-productivity}} \qquad (13)
\end{aligned}$$

where each term on the right-hand side identifies the contribution of the corresponding input in explaining the effect of the program.[67]

Figure 5 presents the results of decomposing the impact of New Hope as in equation (13). In this figure, the contribution of a change in one input in period t to the program's impact on child human capital in period $t+1$ is shown in each area. The total effect of the program on human capital equals the sum of all areas. For period $t = 1$, child care explains most of the effect of the program (82%), while income and labor supply explain a smaller proportion (26% and -8%). Because of the low estimated productivity of time at home, the negative effect produced by the decrease in time spent at home is more than compensated by the positive effects from child care and income. This interaction explains the positive impact of New Hope on child human capital in period $t = 1$.

In the first year after baseline, the total impact on human capital is relatively small: the

[67] Given the linearity of $f(.)$ (see equation 2), the order of the terms in equation (13) does not affect the estimate of the contribution of each input.

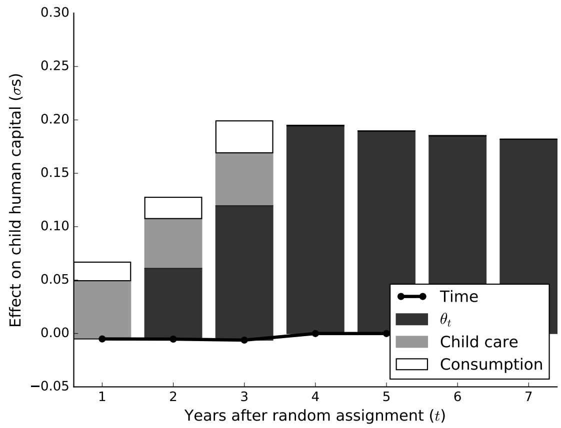
sum of the contributions adds up to 0.07 standard deviations. However, the effect of New Hope on child human capital becomes larger once we consider the dynamic of human capital accumulation in the long run. We can see this process by looking at the impact at two years after baseline. In this period, the contribution of income and child care to the effect of New Hope on human capital is similar to that of period $t = 1$: an additional 2% and 5% of a standard deviation of human capital, respectively. Nevertheless, given the autoregressive coefficient in the production function, a large share of the human capital acquired in period $t = 1$ remains for period $t = 2$; human capital accumulation in $t = 2$ is mostly explained by the human capital gain from the previous period (52%). In addition, the program increased income and child care use at $t = 1$ as well. This additional "investment" implies that the difference between the treatment and control groups is larger in $t = 2$ than in $t = 1$. Therefore, as the program induces changes in behavior leading to further increases in child human capital, the program's impact on child human capital grows with time. As I show in Section 8.2, this process has important implications for interpreting the effect of permanent policies, such as the EITC and the CCDF.

After the program ends, the model predicts a decline in the treatment effects of the program on child skills. The fade-out is explained by the convergence in the levels of income, working hours, and child care use after individuals exit the program. In Figure 5, we can see this process starting $t = 4$, where the contributions to the human capital treatment effects of income, time, and child care are nearly zero. After this period, human capital slowly depreciates (following the AR(1) coefficient of the technology). Even though the model predicts a decreasing treatment effect after the program ends, it does so at a slower rate than the one that can be estimated using the New Hope data.[68]

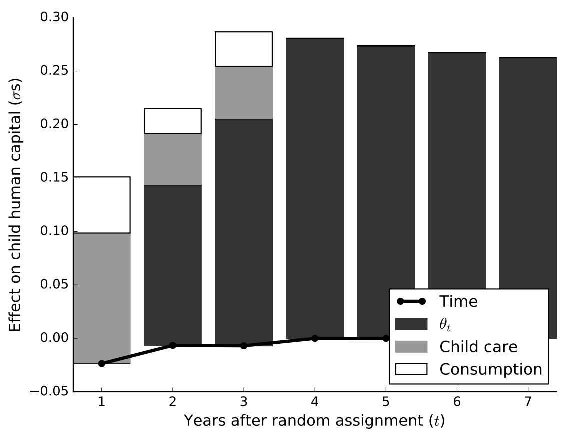
Across all periods, the induced impact on working hours is not large enough to cause an overall negative effect on child human capital. However, this average effect might mask

[68]One potential explanation for a more rapid fade-out is that teachers may have diverted educational resources toward less-advantaged students. The teacher's responses are not modeled here, and so a more pronounced fade-out cannot be exactly reproduced.

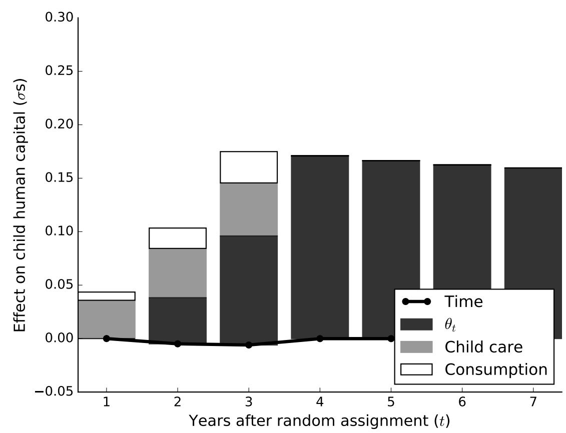
Figure 5: Decomposition of the impact of New Hope on child human capital (θ_t)



(a) Overall



(b) Extensive margin



(c) Intensive margin

Notes: The figure plots the share of each input in explaining the effect of the program on the child human capital (θ_t). Panel (a) plots the decomposition for the whole sample. Panel (b) shows the decomposition for the group which were induced to work full time with the program (that is, they would have work for less than 40 hours a week without it). Panel (c) depicts the decomposition for the group that is not induced to work full-time (would work 40 hours a week without the program or would not work 40 hours with the program).

heterogeneity of responses across the wage offer distribution (see Section 4). To investigate heterogeneous impacts, panels (b) and (c) of Figure 5 plot the structural decomposition of equation (13) for two groups: those who would work full time with the program but not without it (the extensive-margin group) and those who would not work full time with the program or would work full time in both scenarios (the intensive-margin group). Relative to the overall average, children of participants who are induced to work full time have a larger negative effect originated from the additional parental labor supply. Nevertheless, for this group, the positive contributions of income and child care more than offset that negative labor supply effect. Summing up the individual contribution, by $t = 3$, the program's impact on the extensive-margin group is larger than the average, reaching 0.28 standard deviations. In contrast, for the intensive-margin group, in period $t = 3$, the program increased child human capital by 0.17 standard deviations. The reason for the worse results for the intensive-margin group lies in the lower effects of the program on income and child care use in period $t = 1$ relative to that of the extensive-margin group.

One can use the result of this exercise to quantify the effect of having more income on child skills, holding constant labor supply and child care effects. The mediation analysis provides the necessary ingredients to compute this marginal effect. From equation (5), I compute $f(\theta_t^1, \tau_t^1, cc_t^1, c_t^1) - f(\theta_t^1, \tau_t^1, cc_t^1, c_t^0)$ and compare it to the actual rise in consumption $c_t^1 - c_t^0$. The model predicts that New Hope raises consumption by \$900 and the contribution of this change in the total effect on skills $(f(\theta_t^1, \tau_t^1, cc_t^1, c_t^1) - f(\theta_t^1, \tau_t^1, cc_t^1, c_t^0))$ equals 0.02 standard deviations. To put this number in perspective, Dahl and Lochner (2012) find that a 1,000 increase in income leads to a raise in 0.06 boost in math and reading test scores. If we assume a three-person household, the Dahl and Lochner estimate implies that a \$900 increase in income per capita rises test scores by 0.16 standard deviations.

Why are there such large differences in the estimated effects of income on child human capital? My framework suggests two potential reasons. First, the extent to which we can interpret reduced-form studies as pure income effects depends on controlling for other factors. Second, as the authors themselves recognize, the Dahl and Lochner instruments may capture the effects of permanent policies (or temporary policies that affect permanent income). To provide a more direct comparison with the literature, in Section 8.2, I predict the effects of the EITC on child human capital. Here, I show that the effect of the EITC does resemble the Dahl and Lochner results.

8.1.2 The effects of New Hope policy bundle

Grogger and Karoly (2009) review the impacts of a series of welfare experiments on children human capital. They suggest that the varied results coming from these experiments may be explained by the different types of policies that each program included. In this section, I study which components in New Hope were more influential in changing behavior within the family and thus children outcomes. I simulate different versions of New Hope and analyze its effects on labor supply, child care use, income, and child academic skills.

Table 7 presents the average treatment effect of various combinations of the New Hope policies. In each row, the table depicts the effect of a particular policy on average labor supply, child care use, and consumption per capita from $t = 0$ to $t = 2$, and on child human capital for all periods. Column (6) shows the effects of New Hope as it was originally conceived. To study the full potential effect of these policies, I only analyze the sample of children under six years old in period $t = 2$ (older children are not eligible for the child care subsidy).

Table 7: The effects of the New Hope policy bundle

ATE	(1)	(2)	(3)	(4)	(5)	(6)
Consumption (US$)	877	952	928	63	147	1127
Part-time	0.115	0.124	-0.086	-0.001	-0.028	-0.106
Full-time	0.026	0.017	0.189	-0.006	0.055	0.231
Child care	0.031	0.211	0.035	0.203	0.132	0.172
θ_t (σs)	0.072	0.202	0.061	0.143	0.101	0.168
Wage subsidy	✓	✓	✓			✓
Child care subsidy		✓		✓	✓	✓
Work requirement			✓		✓	✓

Notes: The table shows the impact of New Hope on consumption, part-time work, full-time work, child care, and child human capital. The sample corresponds to children who are six years of age or less by $t = 2$. To estimate impacts, I take averages of annual effects across $t = 0$ to $t = 2$. Each policy (indicated with "✓") is compared to a counterfactual scenario where no policy is implemented.

The exercises from Table 7 reveal that New Hope's effect on child outcomes is mostly explained by the child care subsidy component. The child care subsidy alone increases human capital by 0.14 standard deviations (column 4). The wage subsidy, as a unique policy, increases human capital by 0.07 standard deviations (column 1).

The New Hope work requirement causes a negative effect on children's human capital. To see this result, consider columns (4) and (5). In these experiments, I compute the impacts of the child care subsidy with and without the work requirement. Combined with the work requirement, the child care subsidy increases child skills by 0.10 standard deviations (column 5), 0.04 less than the child care subsidy as a unique policy (column 4). The smaller impact of the combined policy package is explained by a larger effect on employment and a lower child care take-up. The child care subsidy alone has almost no effects on full-time employment, but adding the work requirement implies a six-percentage point increase. In addition, the effect of the policy bundle on the probability of child care use is seven percentage points lower than that of the child care subsidy regime alone. The resulting higher labor supply and lower child care take-up of the bundled policy suggest that, for a group of participants, the child care option is valuable only if they are not forced to work full time.

Even though the work requirement may be detrimental to child human capital, the policy could be used nevertheless as a tool to promote work. Yet, as the counterfactual experiments suggest, the work requirement may not be as useful as it may seem to promote employment and increase income. The work requirement makes the wage subsidy less attractive for a sample of individuals who would prefer staying out of the labor market but would work part-time under a program without the work requirement. The wage subsidy policy alone (column 1) boosts employment by 14 percentage points, whereas the wage subsidy tied to the work requirement (column 3) increases employment by 10 percentage points. Moreover, the effects on income of both policies are similar.

Table 7 shed lights on the potential complementarities arising from putting wage and child care subsidies together. Consider a wage subsidy tied to a work requirement (column 3). This policy raises full-time employment by 19 percentage points. If we add to this policy bundle the child care subsidy (column 6), we see that the effect on full-time work increases an additional four percentage points. Hence, a child care subsidy causes a positive effect on full-time employment, only when it is bundled together with other work incentives. Notably, the child care policy in isolation does not increase full-time employment by these 4 percentage points—in fact, there is a slightly negative effect. This additional impact on full-time employment is explained because some individuals would work only if they have access to a low-cost child care—otherwise, working full time would mean sacrificing too much child human capital. Since parents are spending less time at home, this mechanism causes a (probably small) negative impact on child skills. Furthermore, depending on the context, one could find that this channel may dominate the positive effects of spending more time at a formal child care center and end up with an overall negative effect of child care.[69] These results suggest that evaluating policies in isolation may ignore other channels—that could potentially reverse the sign of the total impact—derived from the complementarities of policies that are implemented one on top of the other.[70]

8.2 The EITC and child care subsidy

For the 2013 fiscal year, 68 billion dollars were spent in the EITC, making it one of the largest mean-tested cash transfers programs in the U.S. (Hoynes and Rothstein, 2016; Moffitt, 2016). Additionally, since the establishment of the CCDF in 1996, child care subsidies have been an important component in the U.S. policy directed at low-income families. In this section, I estimate the impacts of the EITC and a child care subsidy as permanent policies. I evaluate effects on child human capital and examine its causes in terms of income, time, and child care.

In this experiment, the "treatment" consists in having the EITC or a child care subsidy (or both).[71] At $t = 0$, the individual receives an unexpected policy change. The agent

[69] This situation may occur if the productivity of child care in the production function is sufficiently low (see discussion from Section 2).

[70] The bulk of the literature evaluating welfare reforms does not consider that combined policies may not be equal to the reported effects of different policies from different studies. See for example Moffitt (2003) and Moffitt (2016).

[71] Instead of calibrating the actual parameters of the child care subsidy implemented in Wisconsin ("Wisconsin Shares"), I expand the policy implemented in New Hope for all years (without a work requirement). Wisconsin Shares followed a similar structure to that of New Hope (Bos et al., 1999).

Table 8: The effect of the EITC and a child care subsidy

ATE	(1)	(2)	(3)
Consumption (US$ 1,000)	0.13	0.02	0.16
Part-time	0.00	0.01	0.01
Full-time	0.02	0.01	0.02
Child care	0.03	0.11	0.14
EITC (1995-2003)	✓		✓
Child care subsidy		✓	✓

Notes: The table shows the average effect of the EITC and a child care subsidy on consumption, part-time work, full-time work, child care, and child human capital. The sample corresponds to children who are six years of age or less by $t = 2$. To estimate impacts, I take averages of annual effects across all periods for consumption, part-time work, and full-time work, and over $t = 0$ to $t = 2$ for child care.

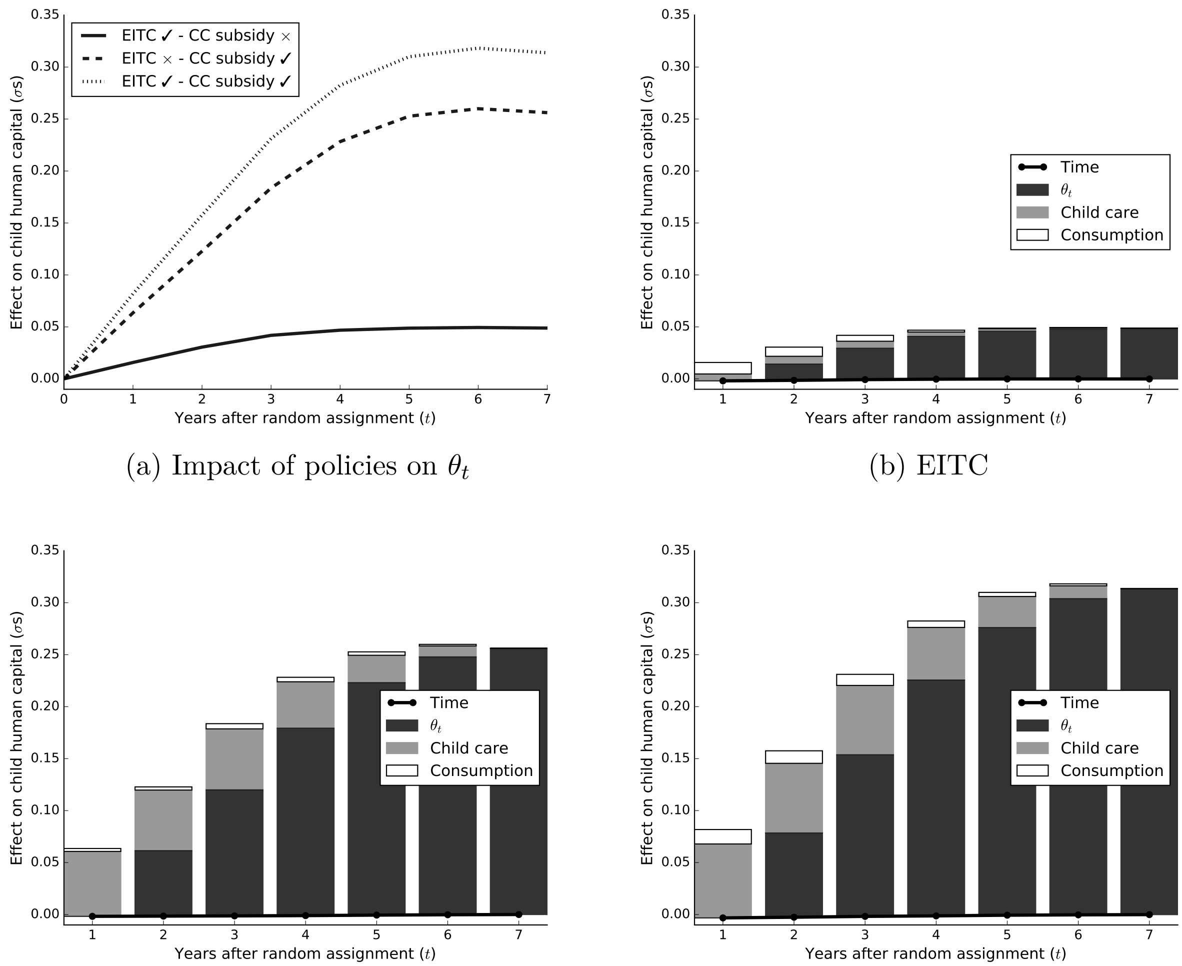
knows exactly how the policy parameters evolve for the rest of the periods. New Hope and work requirements are shut down in this simulation. In contrast to New Hope, the EITC and the child care subsidy are permanent policies.[72] As with the previous simulations, this experiment uses the sample of children under six years of age or less by $t = 2$.

Table 8 and Figure 6 present the results of these experiments. Table 8 presents the average effects for all periods considered (from $t = 0$ to $t = 8$) on household variables. The EITC increases annual per-capita consumption by 130 dollars, the probability of being employed by two percentage points, and the probability of child care by three percentage points. The child care subsidy has similar effects on employment but raises child care probability by eleven percentage points. Both policies combined have larger effects on the household (column 3).

Figure 6, panel (a), depicts the impact on human capital (in standard deviation units) across years after random assignment. One year of being exposed to any policy does not produce economically large effects on human capital; even for both policies combined, the impact is close to 0.05 standard deviations in period $t = 1$. However, these policies have larger impacts once children are being exposed to them for several years. The EITC treatment effect

[72]In this model, I am silent about the labor supply effects for married individuals coming from intra-household responses to spouse's income. In this regard, Eissa and Hoynes (2004) find that the EITC has a negative effect on married women's labor force participation, but the impact is relatively small. Moreover, approximately 90% of the women in my sample are single at baseline.

Figure 6: The impact of the EITC and child care subsidy on child human capital (θ_t)



(a) Impact of policies on θ_t

(b) EITC

(c) Child care subsidy

(d) Both policies

Notes: The figure presents the impact of the EITC and child care subsidy on child human capital. Panel (a) collects the impacts of the policies in isolation and of both combined on θ_t, for $t = 1, ..., 8$. Panels (b)-(d) plot the share of each input in explaining the effect of a policy on child human capital.

on human capital depicts a concave pattern, reaching 0.05 standard deviations by seven years after random assignment (almost exactly the Dahl and Lochner estimate). The effects of the child care subsidy also show a concave profile—as children become older and the child care option is not on the individual's choice set—attaining a maximum of 0.26 standard deviations by six years after baseline. The two policies together produce an increasing, positive effect on child human capital. By seven years after baseline, the impact of the EITC and child care subsidy bundle equals 0.31 standard deviations.

Panels (b)-(d) illustrate the results from the mediation analysis (equation 13) applied to the EITC and the child care subsidy experiments. As with Figure 5, each area represents the contribution of the change in one household input in explaining the total effect of the policy on child human capital. Panel (b) explores the mechanisms behind the impact of the EITC. In the first year after baseline, the small impact of the policy on human capital is almost entirely explained by the income component, followed by a smaller child care effect. Even though the EITC produces a positive effect on labor supply, the negative contribution to the overall impact on human capital is not big enough to offset the positive effects from income and child care. A large part of the added human capital in period $t = 1$ passes over to the next period, where additional boosts on income and child care probability increase the child human capital stock even further.

Panel (c) shows the decomposition analysis on the child care subsidy policy. The policy's impact is nearly entirely explained by the child care contribution. The additional disposable income (from individuals who were already using child care) and higher labor supply (given a reduced marginal cost of working full-time) are not economically relevant to explain the effects on human capital.

Finally, Panel (d) depicts the decomposition analysis on the combination of the EITC and child care subsidy. Similar to New Hope, the bulk of the human capital effect is accounted by child care. Once again, self-productivity sustains human capital growth from one period to the next, and human capital investment—in the form of higher income and child care use—increases the stock as time passes by. This process implies large effects of both policies taken together in the long run.[73]

[73] Johnson and Jackson (2017), find similar evidence on the complementarity of policies (Head Start and public school spending) in explaining the effects on child outcomes.

8.3 Explaining the reduced-form effects of income and child care subsidies through the lens of the structural model

A ubiquitous threat when it comes to interpreting the findings of the literature on the intergenerational effects of welfare policies is that other household behavioral responses to the policy changes are usually not controlled for. Because some of these household responses directly affect child outcomes, it is impossible to disentangle the effect of one particular household input on child outcomes. As a result, reduced-form studies do not always reveal which mechanisms account for the observed effects.

The problem of understanding the mediating effects of income and child care subsidies on child outcomes can lead to misleading conclusions about the effectiveness of a particular policy. Suppose one wishes to evaluate the efficacy of giving (unconditionally) cash to families in terms of its effects on children and parental employment. In my model, one must take into account the increased incentives to spend less time in the labor market and a higher likelihood of choosing child care. In a reduced-form strategy, in order to isolate each possible confounding channel, one must have at least one instrument for every possible input that can affect child outcomes—a demand that is often not met in the literature. The structural framework I follow helps to understand the sources of the reduced-form effects. In particular, I find that most of the effect of the EITC actually comes from an income effect,[74] and most of the child care subsidy effect comes from a higher use of child care.

The reduced-form evidence on the EITC effects on children usually finds a positive, large effect (Maxfield, 2013; Hoynes et al., 2015; Bastian and Michelmore, 2017). A general consensus is that a $1,000-boost in income from the EITC increases child test scores by about 0.06 standard deviations. However, the literature has been silent about the timing of this $1,000-boost. If we take this value as a short-term gain, we end up with a substantial effect of the policy (Nichols and Rothstein, 2016): in five years, the impact would be 0.3

[74] Dahl and Lochner (2016) update results from Dahl and Lochner (2012) using further instruments for labor supply. They find that results regarding the impact of income on test scores hold even if controlled for labor supply changes.

standard deviations. To explain such large effects, Dahl and Lochner (2012) suggest that changes in EITC are more correlated to permanent rather than transitory income. My results are consistent with this hypothesis. Given a relatively small productivity of income in the production function of skills, the impact of having the EITC for one year is relatively small. However, self-productivity accumulate these impacts over time. Figure 6, panel (b), which depicts the impacts of the EITC alone on child human capital in time, implies a medium-run effect (five years) that is close to what the reduced-form literature finds.

The current research on the effects of a child care subsidy on the family shows little consensus. One common, emerging result is that the impact of child care subsidies on children from low-income families is larger than on the average family (Havnes and Mogstad, 2015; Cornelissen et al., 2017). My model predicts that similar, large effects are explained by the fact that the productivity gap of child care relative to home care is sizable (see Figure 6, panel c). The large TFP estimate mirrors almost equivalent results from Kline and Walters (2016) and to a lesser extent from Feller et al. (2016). Overall, my results show strong effects from child care policies that are consistent with the literature that documents large short- and long-run effects of early childhood education on children from economically disadvantaged families (Elango et al., 2016).

Finally, there is no evidence about how wage and child care subsidies complement each other in explaining child outcomes. Most of the reduced-form evidence builds on exogenous policy shocks to estimate the effect of either an income or a child care subsidy in isolation. In the context of New Hope, as policies were tied together, one cannot reach definitive conclusions about the effectiveness of income or child care subsidies as unique policy instruments. In Section 8.1.2, I exploit the structure of my model to show that most of the effect of New Hope on children is explained by the child care component. Moreover, as I show in Section 8.1.2, there may be underlying complementarities between policies, adding further complications when evaluating the effects of income and child care policies.

9 Conclusions

In this paper, I present new evidence on the impact of child care and income subsidies on child outcomes. To this end, I use experimental data from New Hope—an anti-poverty program implemented in Milwaukee (1994-1997) which involved both income and child care subsidies that were tied to a minimum full-time work requirement. With these data, I estimate a dynamic-discrete choice model of the household and child academic achievement. I use the model to explain the channels by which New Hope impacted academic achievement and predict the effect of permanent policies such as the EITC and the CCDF on child outcomes.

The structural framework followed in this paper allows for a better understanding of the separate impacts of income and child care subsidies on child human capital. The results suggest that most of the observed New Hope impact on child outcomes is explained by a positive mediating effect of income and child care which more than compensates the negative effect coming from the increase in labor supply. Consistently, I find that policies such as the EITC and the CCDF have positive effects on children's academic achievement.

A common result across the counterfactual experiments is that, after just a few years, these policies only weakly affect child outcomes. This phenomenon occurs because income has a relatively low productivity in the skills production function. However, self-productivity maintains the bulk of the skills stock acquired in a certain period onto the next one. This process implies that small human capital additions caused by income and child care subsidies accumulate in the long run. If policies are shut down, skills fade out in time, as it is reported in the New Hope case.

Two limitations narrow the external validity of my results. First, my findings are only relevant for those who were willing to participate in the New Hope program. Compared to those who were not interested in participating in the program, New Hope's applicants may be better equipped with observed and unobserved characteristics. Second, because of the scale of the New Hope experiment, I cannot analyze general equilibrium effects.[75]

[75] These two issues are also likely to be found in papers using structural models to explain findings from

Notwithstanding the limitations due to the characteristics of the New Hope experiment, the findings from this paper suggest that income and child care subsidies have an economically significant potential to impact children's academic achievement through the mediating effects on household behavior. Future research should quantify the importance of income, labor supply, and child care—and other channels—in explaining the impacts of these policies in more general settings.

randomized controlled trials. See for example Todd and Wolpin (2006), Attanasio et al. (2011), and Attanasio et al. (2015).

References

Agostinelli, F. and M. Wiswall (2016). Estimating the Technology of Children's Skill Formation. NBER Working Paper No. 22442.

Altonji, J. G. and L. M. Segal (1996). Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2015). Estimating the Production Function For Human Capital: Results From a Randomized Control Trail in Colombia.

Attanasio, O. P., C. Meghir, and A. Santiago (2011). Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA. *The Review of Economic Studies 79*(1), 37–66.

Autor, D., A. Kostol, M. Mogstad, and B. Setzler (2017). Disability Benefits, Consumption Insurance, and Household Labor Supply.

Bajari, P. and A. Hortaçsu (2005). Are structural estimates of auction models reasonable? evidence from experimental data. *Journal of Political Economy 113*(4), 703–741.

Baker, M., G. Jonathan, and M. Kevin (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Bastian, J. and K. Michelmore (2017). The Intergenerational Impact of the Earned Income Tax Credit on Education and Employment Outcomes. Forthcoming at the *Journal of Labor Economics*.

Bernal, R. (2008). The Effect of Maternal Employment and Child Care on Children's Cognitive Development. *International Economic Review 49*(4), 1173–1209.

Bernal, R. and M. P. Keane (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics 156*(1), 164–189.

Bernal, R. and M. P. Keane (2011). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics 29*(3), 459–512.

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006). What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. *American Economic Review 96*(4), 988–1012.

Black, S. E., P. J. Devereux, K. V. Løken, and K. G. Salvanes (2014). Care or Cash? The Effect of Child Care Subsidies on Student Performance. *Review of Economics and Statistics*.

Blau, D. M. (2003). Child Care Subsidy Programs. Chicago: University of Chicago Press.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female Labor Supply, Human Capital, and Welfare Reform. *Econometrica 84*(5), 1705–1753.

Blundell, R., M. Dias, C. Meghir, and J. Shaw (2015). Female labour supply, human capital and welfare reform. NBER Working Paper No. 19007.

Bos, J. M., A. C. Huston, G. J. Granger, Robert C. Duncan, T. W. Brock, and V. C. McLoyd (1999). New Hope for People with Low Incomes. Two-Year Results of a Program to Reduce Poverty and Reform Welfare.

Brilli, Y. (2014). Mother's time allocation, child care and child cognitive development.

Brock, T., F. Doolittle, V. Fellerath, and M. Wiseman (1997). Creating New Hope: Implementation of a Program to Reduce Proverty and Reform Welfare. Unpublished manuscript.

Brown, M. and C. J. Flinn (2011). Family law effects on divorce, fertility and child investment.

Bruins, M. (2016). TANF, Child Care, and Child Well-Being in Sole-Parent Families.

Chan, M. K. (2013). A Dynamic Model of Welfare Reform. *Econometrica 81*(3), 941–1001.

Chetty, R., J. Friedman, and J. Rockoff (2011). New Evidence on the long-term impact of tax credits. In *Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, Volume 104, pp. 116–124.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2017). Who benefits from universal childcare? Estimating marginal returns to early childcare attendance. Forthcoming *Journal of Political Economy.*

Cunha, B. F. and J. Heckman (2006). The Technology of Skill Formation. *The American Economic Review 97*(2), 31–47.

Cunha, F., J. Heckman, and S. Schennach (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica 78*(3), 883–931.

Cunha, F. and J. J. Heckman (2008). Formulating , Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. (December 2006).

Dahl, G. and L. Lochner (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *The American Economic Review 102*(5), 1927–1956.

Dahl, G. B. and L. Lochner (2016). Correction and Addendum: The impact of family income on child achievement: Evidence from the earned income tax credit. Unpublished manuscript.

De Nardi, M. C., E. French, and J. B. Jones (2010). Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy 118*(1), 39–75.

Del Boca, D., C. Flinn, and M. Wiswall (2013). Household Choices and Child Development. *The Review of Economic Studies 81*(1), 137–185.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Transfers to Households with Children and Child Development. Forthcoming at the Economic Journal.

Eissa, N. and H. W. Hoynes (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics 88*(9-10), 1931–1958.

Elango, S., J. L. Garcia, J. J. Heckman, and A. Hojman (2016). Early Childhood Education. In R. A. Moffitt (Ed.), *Economics of Means-Tested Transfer Programs in the United States, Volume II*. The University of Chicago Press.

Epps, S. R. and A. C. Huston (2007). Effects of a Poverty Intervention Policy Demonstration on Parenting and Child Behavior: A Test of the Direction of Effects. *Social Science Quarterly 88*(2), 344–365.

Feller, A., T. Grindal, L. Miratrix, and L. C. Page (2016). *The Annals of Applied Statistics 10*(3), 1245–1285.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect Inference. *Journal of Applied Econometrics 8*(S1), S85–S118.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption Over the Life Cycle. *Econometrica 70*(1), 47–89.

Grogger, J. and L. A. Karoly (2009). *Welfare Reform: Effects of a Decade of Change*. Harvard University Press, Cambridge: MA.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental Education and Parental Time with Children. *Journal of Economic Perspectives 22*(3), 23–46.

Havnes, T. and M. Mogstad (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy 3*(2), 97–129.

Havnes, T. and M. Mogstad (2015). Is universal child care leveling the playing field? *Journal of Public Economics 127*, 100–114.

Heckman, J. J. (2010). Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of economic literature 48*(2), 356–398.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016a). Dynamic treatment effects. *Journal of Econometrics 191*(2), 276–292.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016b). Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking. NBER Working Paper No. 22291.

Heckman, J. J. and S. Mosso (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics 6*, 689–733.

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). Understanding the mechanisms through which an influential early Childhood programm boosted Adult outcomes. *American Economic Review 103*(6), 1–35.

Herbst, C. M. and E. Tekin (2010a). Child care subsidies and child development. *Economics of Education Review 29*(4), 618–638.

Herbst, C. M. and E. Tekin (2010b). The Impact of Child Care Subsidies on Child Well-Being: Evidence from Geographic Variation in the Distance to Social Service Agencies. NBER Working Paper No. 16250.

Hoynes, H., D. Miller, and D. Simon (2015). Income, the Earned Income Tax Credit, and Infant Health. *American Economic Journal: Economic Policy 7*(1), 172–211.

Hoynes, H. and J. Rothstein (2016). Tax Policy Toward Low-Income Families. NBER Working Paper 22080.

Huston, A., G. Duncan, R. Granger, J. Bos, V. McLoyd, R. Mistry, D. Crosby, C. Gibson, K. Magnuson, J. Romich, and A. Ventura (2001). Work-Based Antipoverty Program for Parents Can Enhance the School Performance and Social Behavior of Children. *Child Development 72*(1), 318–336.

Huston, A. C., G. J. Duncan, V. C. McLoyd, D. A. Crosby, M. N. Ripke, T. S. Weisner, and C. A. Eldred (2005). Impacts on children of a policy to promote employment and reduce poverty for low-income parents: new hope after 5 years. *Developmental psychology 41*(6), 902–918.

Huston, A. C., A. E. Gupta, J. T. Walker, C. J. Dowsett, S. R. Epps, A. E. Imes, and V. C. McLoyd (2011). The long-term effects on children and adolescents of a policy providing work supports for low-income parents. *Journal of Policy Analysis and Management 30*(4), 729–754.

Huston, A. C., C. Miller, L. Richburg-Hayes, G. J. Duncan, C. A. Eldred, T. S. Weisner, E. Lowe, V. C. McLoyd, D. A. Crosby, M. N. Ripke, and C. Redcross (2003). New Hope for Families and Children. Five-Year Results of a Program to Reduce Poverty and Reform Welfare.

Johnson, R. C. and C. K. Jackson (2017). Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending. NBER Working Paper Series No. 23489.

Keane, M. P. and R. A. Moffitt (1998). A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review 39*(3), 553–589.

Keane, M. P., P. E. Todd, and K. I. Wolpin (2011). The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 331–461.

Keane, M. P. and K. I. Wolpin (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *The Review of Economics and Statistics 76*(4), 648–672.

Keane, M. P. and K. I. Wolpin (2007). Exploring the Usefulness of a Nonrandom Holdout

Sample for Model Validation: Welfare Effects on Female Behavior. *International Economic Review 48*(4), 1351–1378.

Keane, M. P. and K. I. Wolpin (2010). The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women. *International Economic Review 51*(3), 851–892.

Kline, P. and C. Walters (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Magnac, T. and D. Thesmar (2002). Identifying Dynamic Discrete Decision Processes. *Econometrica 70*(2), 801–816.

Manoli, D. and N. Turner (2015). Cash-on-hand and College Enrollment: Evidence from population Tax Data and Policy Nonlinearities. Unpublished Manuscript.

Maxfield, M. (2013). The Effects of the Earned Income Tax Credit on Child Achievement and Long-Term Educational Attainment. Unpublished Manuscript.

Miller, C., A. C. Huston, G. J. Duncan, V. C. McLoyd, and T. S. Weisner (2008). New Hope for the Working Poor. Effects After Eight Years for Families and Children.

Moffitt, R. A. (2003). *Means-Tested Transfer Programs in the United States.* Chicago and London: National Bureau of Economic Research.

Moffitt, R. A. (2016). *Means Tested Programs in the US, Volume II.* Chicago: The University of Chicago Press. Forthcoming.

Mullins, J. (2015). Improving Child Outcomes Through Welfare Reform. Unpublished manuscript.

Nichols, A. and J. Rothstein (2016). The Earned Income Tax Credit. In R. A. Moffitt (Ed.), *Economics of Means-Tested Transfer Programs in the United States, Volume I.* Chicago.

Plueger, D. (2009). Earned Income Tax Credit participation rate for tax year 2005. IRS Research Bulletin.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters 113*, 38–40.

Rust, J. (1994). Structural Estimation of Markov Decision Processes. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume IV, Chapter 51, pp. 3139–3139.

Scholz, J. K. (1994). The earned income tax credit: Participation, compliance, and antipoverty effectiveness. *National Tax Journal 47*(1), 64–87.

Swann, C. A. (2005). Welfare Reform When Recipients Are Forward-Looking. *Journal of Human Resources XL*(1), 31–56.

Tartari, M. (2015). Divorce and the Cognitive Achievement of Children. *International Economic Review 56*(2), 597–645.

Todd, P. E. and K. I. Wolpin (2006). Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. *The American Economic Review 96*(5), 1384–1417.

Voena, A. (2015). Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples? *American Economic Review 105*(8), 2295–2332.

A The benefits of New Hope

Table 9 compares the New Hope benefits to the public system's welfare services. The table illustrates the actual New Hope "treatment:" the benefits given to participants compared to what the control group had access to. New Hope had three main advantages: it gave an income supplement that was larger than the EITC schedule, it increased the affordable child care supply for low-income working families, and it lowered health care costs.

Table 9: New Hope versus Wisconsin's social assistance

Components	**New Hope (treatment group)**	**Wisconsin's public services (control group)**	**New Hope's value-added**
Cash assistance	Income supplement: wage subsidy + child allowance.	Earned Income Tax Credit.	Increase disposable income (earnings plus cash assistance) up to 200% depending on the level of annual earnings.
CSJs	New Hope assigned unemployed participants to temporary CSJs.	CSJs available for welfare recipients.	The New Hope CSJs were paid, and it qualified for hours worked to receive New Hope benefits.
Child care	Child care subsidy with a low copayment.	Child care subsidies to welfare receipients and for families in transition out of welfare. Head start was available as well.	Limited supply of public child care slots. In practice, NH increased supply of affordable child care.
Health insurance	Health plans with low copayment through local HMOs.	Medicaid, employer-funded plans.	New Hope complemented employer plans. Also available for families not in AFDC.

Notes: This table summarizes the main components of New Hope. It compares the New Hope benefits with equivalent services available in Wisconsin.

A.1 Income subsidy

The income subsidy is defined as the sum of an earnings subsidy and a child allowance. The earnings subsidy increases at low levels of earnings and phases out (at a slower rate) until reaching zero benefits. Let E be the annual labor earnings for a given year. The earnings subsidy (ES) is determined by the following formula:

$$ES^* = \begin{cases} 0.25 \times E & \text{if } E \leq 8,500 \\ \max\{0.25 \times 8,500 - 0.2(E - 8,500), 0\} & \text{if } E > 8,500, \end{cases}$$

and so the earnings subsidy equals zero at 19,125 dollars of earnings. These parameters do not depend on family composition or other sources of income.

Unlike the earnings supplement, the child allowance component considers *family* annual labor earnings. Let FE denote family earnings and n the number of children in the family. The per-child child allowance (CA) is given by

$$CA = \begin{cases} x_n^* & \text{if } FE < 8,500 \\ \max\{x^* - r(\bar{e})(FE - 8,500), 0\} & \text{if } FE \geq 8,500 \end{cases}$$

where x_n^* is the subsidy maximum level and $r(\bar{e})$ is the phase-our rate. This rate is implicitly defined by the level of earnings at which the child allowance phases out completely ($\bar{e}$).[76] This last parameter is determined as follows:

$$\bar{e} = \begin{cases} 30,000 & \text{if } n < 4 \\ 30,000 + e* & \text{if } n \geq 4 \end{cases}$$

where e^*varies by year of the program (starts at \$300 and reaches \$2,100 by the third year). The maximum level of child allowance depends on the number of children, as follows:

$$x_n^* = \begin{cases} x_{n-1}^* + (x_{n-1}^* - x_{n-2}^* - 100) & \text{if } n \leq 4 \\ x_{n-1}^* & \text{if } n > 4 \end{cases},$$

where $x_0^* = 0$ (child allowance when the family has no children) and $x_1^* = 1600$. Thus, the maximum level reaches $1,600$ dollars for the first child, an extra $1,500$ for the second, and so on. The maximum subsidy stays fixed at x_4^* for families with more than four children.

The New Hope income supplement ($ES + CA$) complements the EITC. Specifically, let $EITC$ be the amount of EITC for a given level family earnings. The total income supplement (IS) follows:

$$IS = \begin{cases} (ES + CA) - EITC & \text{if } (ES + CA) > EITC \\ 0 & \text{if } (ES + CA) \leq EITC \end{cases}$$

A.2 Child care subsidy

New Hope provided child care vouchers with a relatively low copayment. To have had access to the subsidy, families must have met three basic conditions. First, only individuals with children with children under age 13 were eligible. Second, beneficiaries had to have worked at least 30 hours a week on average in a particular month.[77] For two-parents families,

[76] $r(\bar{e})$ corresponds to the rate r at which $x_n^* - r \times (\bar{e} - 8,500) = 0$.

[77] The New Hope representatives designed a standardize procedure to minimize fraud. Each month, the participant and the provider sign a voucher indicating the hours and the cost of the services. By the end of the month, the child care provider submits these vouchers to New Hope representatives to receive their payments. New Hope pays the subsidy directly to the child care provider. The participant pays the copayment to the provider as well. If the participant does not submit the wage stubs, New Hope would cover only 75% of the child care cost of the month. If the participant does not submit the wage stub for the second month in a row, New Hope reps would suspend the subsidy.

in addition to the full-time requirement of the primary earner, the second earner had to have worked at least 15 hours a week. If the participant had been unemployed, she would have received a subsidy covering a portion of a part-time child care (up to three hours, for a maximum of three weeks). Finally, participants who were eligible to receive the child care benefit were able to enroll their children only in a state- or county-licensed provider. This definition included preschool and daycare centers for younger children and after-school programs for children in school ages.

Let p be the child care cost offerred at a child care facility. The copayment ($\underline{p}$) follows (numbers are in term of monthly dollars):

$$\underline{p} = \begin{cases} 400 & \text{if } p > 400 \text{ and Earnings} \leq 8,500 \\ 315 + 0.01 \times \text{Earnings} & \text{if } p > 400 \text{ and Earnings} > 8,500 \\ p & \text{if } p \leq 400 \end{cases}$$

A.3 Community Service Jobs (CSJ)

New Hope staff advised participants in finding local job openings. If after a period of eight weeks the participant had not find a job, New Hope would assigned her to a paid CSJ for a maximum of six months.[78] The CSJ's paid was minimum wage. Importantly, the hours worked in these CSJs qualified for the income supplement, child care subsidy, and the health insurance subsidy.

According to Brock et al. (1997), other forms of CSJs were available at that time in Milwaukee. However, unlike the New Hope program, these types of CSJs did not qualify for the state's EITC. Indeed, the state CSJs positions were meant for individuals who needed them to receive welfare grants, not as a mean to earn a salary. The New Hope CSJs were given to people regardless of their employment status, while the state CSJs were not usually offered to unemployed individuals.

A.4 Health insurance

New Hope financed part of the health insurance for workers with no employer-granted health insurance or Medicaid. To have access to the health insurance, individuals must have worked at least 30 hours a week every month. If a participant became unemployed or reduce her working hours below 30, New Hope kept their health insurance up to three weeks.[79]

New Hope provided health insurance through a Health Maintenance Organizations (HMO). The program's representatives displayed a number of plans and explained in detail the ups and downs of every plan. Beneficiaries would pick from any of those plans. Most of the participants choose to stay with the HMO that had a contract with Milwaukee County to provide Medicaid services.

To receive health insurance through New Hope, participants had to pay a small share of its cost. The copayment was a function of household income and size. The copay began at

[78]The Milwaukee Private Industry Council acted as the former employer, although funds came from New Hope.

[79]In practice, New Hope representatives would kept the health insurance eligibility up to three months if the participant would have demonstrated active job search efforts.

$72 and $168 a year for a single person and households with three members or more. The maximum copay was $600 and $1,548 for single- and three-person households, respectively. If an individual had an employer health plan, New Hope would cover for the difference between the insurance's premium and the New Hope copayment. Moreover, if the participant did not have a dental coverage under her employer health plan, she had the option of choosing from the New Hope available dental plans.

Many of participants opted out from the New Hope health insurance plan, as some families choose Medicaid instead. To be eligible to Medicaid, families under AFDC had to make less than 185% the federal poverty line.[80] As many New Hope families met these requirements and given that Medicaid had no premiums, the Medicaid option seemed more convenient. Nonetheless, take-up was still considerable: 47.6% of participants were covered by a New Hope health insurance at some point during the 36-months eligibility period.

[80] After PRWORA, individuals that were eligible to Medicaid as of August 1996 maintained their eligibility status.

B Baseline characteristics by sample

Table B.1: Baseline characteristics: whole sample

Variable	(1) Treatment	(2) Control	(3) T-C
Age	31.35	31.06	0.30
	[9.52]	[9.05]	(0.50)
Female (%)	71.39	71.87	-0.48
	[45.23]	[45.00]	(2.45)
African-American, non-Hispanic (%)	51.77	50.96	0.81
	[50.01]	[50.03]	(2.72)
Hispanic (%)	25.81	27.10	-1.29
	[43.79]	[44.48]	(2.40)
White, non-Hispanic (%)	12.83	13.11	-0.28
	[33.47]	[33.77]	(1.83)
Others (%)	9.59	8.84	0.75
	[29.46]	[28.40]	(1.57)
Never married (%)	59.44	60.24	-0.80
	[49.14]	[48.98]	(2.66)
Married living w/ spouse (%)	12.54	11.93	0.61
	[33.14]	[32.44]	(1.78)
Married living apart (%)	9.44	9.72	-0.28
	[29.26]	[29.65]	(1.60)
Separated, divorced or widowed (%)	18.58	18.11	0.47
	[38.93]	[38.54]	(2.10)
Highschool diploma or GED (%)	45.72	45.21	0.51
	[49.85]	[49.81]	(2.71)
Highest grade completed	10.79	10.76	0.03
	[287.20]	[266.64]	(15.05)
$0 (%)	30.24	32.11	-1.87
	[45.96]	[46.72]	(2.52)
$1-999 (%)	17.40	14.14	3.27*
	[37.94]	[34.87]	(1.98)
$1,000-4,999 (%)	24.19	26.22	-2.03
	[42.85]	[44.01]	(2.36)
$5,000-9,999 (%)	16.08	17.38	-1.30
	[36.76]	[37.92]	(2.03)
$10,000-14,999 (%)	8.26	7.36	0.90
	[27.55]	[26.14]	(1.46)
$15,000 or more (%)	3.83	2.80	1.04
	[19.22]	[16.50]	(0.97)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to the original New Hope group of applicants. The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.2: Baseline characteristics: CFS sample

Variable	(1) Treatment	(2) Control	(3) T-C
Age	29.04	28.53	0.51
	[7.14]	[6.64]	(0.52)
Female (%)	90.12	91.74	-1.62
	[29.89]	[27.57]	(2.18)
African-American, non-Hispanic (%)	58.14	53.85	4.29
	[49.40]	[49.92]	(3.77)
Hispanic (%)	27.03	29.06	-2.02
	[44.48]	[45.47]	(3.41)
White, non-Hispanic (%)	10.47	14.53	-4.06
	[30.65]	[35.29]	(2.51)
Others (%)	4.36	2.56	1.80
	[20.45]	[15.83]	(1.39)
Never married (%)	62.21	62.39	-0.18
	[48.56]	[48.51]	(3.68)
Married living w/ spouse (%)	11.05	9.69	1.36
	[31.39]	[29.62]	(2.31)
Married living apart (%)	9.88	11.11	-1.23
	[29.89]	[31.47]	(2.33)
Separated, divorced or widowed (%)	16.86	16.81	0.05
	[37.49]	[37.45]	(2.84)
Highschool diploma or GED (%)	50.00	45.87	4.13
	[50.07]	[49.90]	(3.79)
Highest grade completed	11.24	11.10	0.14
	[2.15]	[2.04]	(0.16)
$0 (%)	36.63	36.75	-0.12
	[48.25]	[48.28]	(3.66)
$1-999 (%)	16.86	14.53	2.33
	[37.49]	[35.29]	(2.76)
$1,000-4,999 (%)	23.26	23.65	-0.39
	[42.31]	[42.55]	(3.22)
$5,000-9,999 (%)	13.66	14.81	-1.15
	[34.40]	[35.58]	(2.66)
$10,000-14,999 (%)	6.98	6.84	0.14
	[25.51]	[25.28]	(1.93)
$15,000 or more (%)	2.62	3.42	-0.80
	[15.99]	[18.20]	(1.30)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample. The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.3: Baseline characteristics: CFS sample in the year-two New Hope survey

Variable	(1) **Treatment**	(2) **Control**	(3) **T-C**
Age	28.91	28.51	0.40
	[7.02]	[6.57]	(0.56)
Female (%)	89.58	92.72	-3.13
	[30.60]	[26.03]	(2.34)
African-American, non-Hispanic (%)	59.38	53.31	6.06
	[49.20]	[49.97]	(4.08)
Hispanic (%)	25.69	29.14	-3.44
	[43.77]	[45.52]	(3.68)
White, non-Hispanic (%)	11.11	15.23	-4.12
	[31.48]	[35.99]	(2.79)
Others (%)	3.82	2.32	1.50
	[19.20]	[15.07]	(1.42)
Never married (%)	62.50	62.25	0.25
	[48.50]	[48.56]	(4.00)
Married living w/ spouse (%)	10.42	9.93	0.48
	[30.60]	[29.96]	(2.49)
Married living apart (%)	10.42	10.26	0.15
	[30.60]	[30.40]	(2.51)
Separated, divorced or widowed (%)	16.67	17.55	-0.88
	[37.33]	[38.10]	(3.11)
Highschool diploma or GED (%)	51.74	46.03	5.71
	[50.06]	[49.92]	(4.12)
Highest grade completed	11.37	11.09	0.28*
	[2.05]	[2.08]	(0.17)
$0 (%)	37.15	37.09	0.07
	[48.41]	[48.38]	(3.99)
$1-999 (%)	15.63	14.57	1.06
	[36.37]	[35.34]	(2.95)
$1,000-4,999 (%)	23.61	25.17	-1.55
	[42.54]	[43.47]	(3.54)
$5,000-9,999 (%)	14.24	13.91	0.33
	[35.00]	[34.66]	(2.87)
$10,000-14,999 (%)	6.60	5.96	0.64
	[24.87]	[23.71]	(2.00)
$15,000 or more (%)	2.78	3.31	-0.53
	[16.46]	[17.92]	(1.42)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample who responded the second-year survey. The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.4: Baseline characteristics: CFS sample in the year-five New Hope survey

Variable	(1) Treatment	(2) Control	(3) T-C
Age	29.20	28.53	0.67
	[7.34]	[6.78]	(0.60)
Female (%)	90.78	92.50	-1.72
	[28.98]	[26.39]	(2.34)
African-American, non-Hispanic (%)	57.80	53.93	3.87
	[49.48]	[49.93]	(4.19)
Hispanic (%)	28.01	28.21	-0.20
	[44.99]	[45.08]	(3.80)
White, non-Hispanic (%)	10.64	15.36	-4.72*
	[30.89]	[36.12]	(2.83)
Others (%)	3.55	2.50	1.05
	[18.53]	[15.64]	(1.45)
Never married (%)	61.35	61.43	-0.08
	[48.78]	[48.76]	(4.11)
Married living w/ spouse (%)	11.70	10.36	1.34
	[32.20]	[30.52]	(2.65)
Married living apart (%)	10.28	10.71	-0.43
	[30.43]	[30.98]	(2.59)
Separated, divorced or widowed (%)	16.67	17.50	-0.83
	[37.33]	[38.06]	(3.18)
Highschool diploma or GED (%)	50.71	46.79	3.92
	[50.08]	[49.99]	(4.22)
Highest grade completed	11.29	11.10	0.18
	[2.20]	[1.96]	(0.18)
$0 (%)	34.75	38.93	-4.18
	[47.70]	[48.85]	(4.07)
$1-999 (%)	17.38	15.36	2.02
	[37.96]	[36.12]	(3.13)
$1,000-4,999 (%)	24.47	21.79	2.68
	[43.07]	[41.35]	(3.56)
$5,000-9,999 (%)	13.48	13.21	0.26
	[34.21]	[33.93]	(2.87)
$10,000-14,999 (%)	6.74	7.50	-0.76
	[25.11]	[26.39]	(2.17)
$15,000 or more (%)	3.19	3.21	-0.02
	[17.61]	[17.67]	(1.49)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample who responded the fifth-year survey. The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.5: Baseline characteristics: CFS sample in the year-eight New Hope survey

Variable	(1) **Treatment**	(2) **Control**	(3) **T-C**
Age	28.82	28.47	0.35
	[6.92]	[6.57]	(0.55)
Female (%)	90.57	92.33	-1.76
	[29.27]	[26.65]	(2.29)
African-American, non-Hispanic (%)	61.28	55.33	5.95
	[48.79]	[49.80]	(4.04)
Hispanic (%)	25.59	27.00	-1.41
	[43.71]	[44.47]	(3.61)
White, non-Hispanic (%)	9.43	15.67	-6.24**
	[29.27]	[36.41]	(2.71)
Others (%)	3.70	2.00	1.70
	[18.92]	[14.02]	(1.36)
Never married (%)	62.63	63.67	-1.04
	[48.46]	[48.18]	(3.96)
Married living w/ spouse (%)	10.77	10.00	0.77
	[31.06]	[30.05]	(2.50)
Married living apart (%)	9.09	10.33	-1.24
	[28.80]	[30.49]	(2.43)
Separated, divorced or widowed (%)	17.51	16.00	1.51
	[38.07]	[36.72]	(3.06)
Highschool diploma or GED (%)	47.81	45.33	2.48
	[50.04]	[49.86]	(4.09)
Highest grade completed	11.28	11.18	0.10
	[2.18]	[1.97]	(0.17)
$0 (%)	36.70	38.33	-1.63
	[48.28]	[48.70]	(3.97)
$1-999 (%)	17.17	14.33	2.84
	[37.78]	[35.10]	(2.98)
$1,000-4,999 (%)	23.23	22.67	0.57
	[42.30]	[41.94]	(3.45)
$5,000-9,999 (%)	13.13	14.33	-1.20
	[33.83]	[35.10]	(2.82)
$10,000-14,999 (%)	6.73	6.67	0.07
	[25.10]	[24.99]	(2.05)
$15,000 or more (%)	3.03	3.67	-0.64
	[17.17]	[18.83]	(1.48)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample who responded the eighth-year survey. The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.6: Baseline characteristics: CFS sample in the year-two Teachers' survey

Variable	(1) **Treatment**	(2) **Control**	(3) **T-C**
Age	30.25	29.39	0.85
	[7.05]	[6.13]	(0.83)
Female (%)	91.20	93.70	-2.50
	[28.44]	[24.39]	(3.34)
African-American, non-Hispanic (%)	60.80	55.12	5.68
	[49.02]	[49.93]	(6.23)
Hispanic (%)	23.20	26.77	-3.57
	[42.38]	[44.45]	(5.47)
White, non-Hispanic (%)	12.80	15.75	-2.95
	[33.54]	[36.57]	(4.42)
Others (%)	3.20	2.36	0.84
	[17.67]	[15.25]	(2.08)
Never married (%)	57.60	61.42	-3.82
	[49.62]	[48.87]	(6.20)
Married living w/ spouse (%)	12.80	13.39	-0.59
	[33.54]	[34.18]	(4.27)
Married living apart (%)	12.80	7.87	4.93
	[33.54]	[27.04]	(3.84)
Separated, divorced or widowed (%)	16.80	17.32	-0.52
	[37.54]	[37.99]	(4.76)
Highschool diploma or GED (%)	50.40	47.24	3.16
	[50.20]	[50.12]	(6.32)
Highest grade completed	11.31	10.84	0.47
	[2.49]	[2.05]	(0.29)
$0 (%)	37.60	30.71	6.89
	[48.63]	[46.31]	(5.98)
$1-999 (%)	14.40	14.96	-0.56
	[35.25]	[35.81]	(4.48)
$1,000-4,999 (%)	24.80	22.05	2.75
	[43.36]	[41.62]	(5.35)
$5,000-9,999 (%)	10.40	18.90	-8.50*
	[30.65]	[39.30]	(4.44)
$10,000-14,999 (%)	8.00	8.66	-0.66
	[27.24]	[28.24]	(3.50)
$15,000 or more (%)	4.80	4.72	0.08
	[21.46]	[21.30]	(2.69)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample with teacher survey data (year two). The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.7: Baseline characteristics: CFS sample in the year-five Teachers' survey

Variable	(1) **Treatment**	(2) **Control**	(3) **T-C**
Age	29.52	28.52	1.00
	[7.98]	[6.46]	(0.78)
Female (%)	90.06	91.95	-1.90
	[30.01]	[27.28]	(3.09)
African-American, non-Hispanic (%)	53.80	51.72	2.08
	[50.00]	[50.11]	(5.39)
Hispanic (%)	29.82	28.74	1.09
	[45.88]	[45.38]	(4.91)
White, non-Hispanic (%)	12.28	18.39	-6.11
	[32.92]	[38.85]	(3.88)
Others (%)	4.09	1.15	2.94*
	[19.87]	[10.69]	(1.71)
Never married (%)	61.40	63.22	-1.81
	[48.83]	[48.36]	(5.23)
Married living w/ spouse (%)	13.45	10.92	2.53
	[34.22]	[31.28]	(3.53)
Married living apart (%)	11.11	9.20	1.92
	[31.52]	[28.98]	(3.26)
Separated, divorced or widowed (%)	14.04	16.67	-2.63
	[34.84]	[37.38]	(3.89)
Highschool diploma or GED (%)	53.22	44.25	8.96*
	[50.04]	[49.81]	(5.38)
Highest grade completed	11.36	11.18	0.18
	[1.98]	[1.92]	(0.21)
$0 (%)	33.33	41.95	-8.62*
	[47.28]	[49.49]	(5.21)
$1-999 (%)	16.37	13.22	3.16
	[37.11]	[33.97]	(3.83)
$1,000-4,999 (%)	24.56	22.41	2.15
	[43.17]	[41.82]	(4.58)
$5,000-9,999 (%)	16.37	12.64	3.73
	[37.11]	[33.33]	(3.80)
$10,000-14,999 (%)	5.26	6.90	-1.63
	[22.40]	[25.41]	(2.58)
$15,000 or more (%)	4.09	2.87	1.22
	[19.87]	[16.75]	(1.98)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample with teacher survey data (year five). The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table B.8: Baseline characteristics: CFS sample in the year-eight Teachers' survey

Variable	(1) **Treatment**	(2) **Control**	(3) **T-C**
Age	28.37	27.96	0.41
	[6.69]	[6.36]	(0.70)
Female (%)	92.57	93.60	-1.03
	[26.30]	[24.54]	(2.73)
African-American, non-Hispanic (%)	60.57	54.07	6.50
	[49.01]	[49.98]	(5.31)
Hispanic (%)	26.29	26.74	-0.46
	[44.14]	[44.39]	(4.75)
White, non-Hispanic (%)	10.29	16.86	-6.57*
	[30.46]	[37.55]	(3.67)
Others (%)	2.86	2.33	0.53
	[16.71]	[15.12]	(1.71)
Never married (%)	64.00	66.28	-2.28
	[48.14]	[47.41]	(5.13)
Married living w/ spouse (%)	9.71	9.30	0.41
	[29.70]	[29.13]	(3.16)
Married living apart (%)	7.43	11.63	-4.20
	[26.30]	[32.15]	(3.15)
Separated, divorced or widowed (%)	18.86	12.79	6.07
	[39.23]	[33.50]	(3.92)
Highschool diploma or GED (%)	50.29	47.67	2.61
	[50.14]	[50.09]	(5.38)
Highest grade completed	11.27	11.26	0.02
	[2.46]	[1.98]	(0.24)
$0 (%)	35.43	38.37	-2.94
	[47.97]	[48.77]	(5.19)
$1-999 (%)	13.71	15.70	-1.98
	[34.50]	[36.48]	(3.81)
$1,000-4,999 (%)	24.57	20.93	3.64
	[43.17]	[40.80]	(4.51)
$5,000-9,999 (%)	14.86	15.70	-0.84
	[35.67]	[36.48]	(3.87)
$10,000-14,999 (%)	6.86	6.40	0.46
	[25.34]	[24.54]	(2.68)
$15,000 or more (%)	4.57	2.91	1.66
	[20.95]	[16.85]	(2.04)

Notes: This table compares baseline characteristics of treatment and control groups. The variables were measured at baseline. The last six rows present the proportion of individuals in different groups defined by labor earnings (past 12 months). I show sample means of each variable for treatment and control group in columns 1 and 2. The third column tests the null hypothesis that the means are equal. The sample corresponds to participants in the CFS sample with teacher survey data (year eight). The first two columns present standard deviations in square brackets. In the third column, robust standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

C Treatment effects of New Hope

C.1 Child care

I construct the child care variable using the New Hope surveys. In the second-year survey, individuals were asked about all regular child care arrangement for the past two years. Possible child care arrangements in the survey are: (i) Head Start; (ii) preschool, nursery school, or a child care center other than Head Start; (iii) school-based extended day program; (iv) another child care other than in someone's home; (v) a person other than a member of the household; (vi) another member of the family of household; and (vii) no arrangements. Participants reported the number of months spent in each case (except for number (vii)). I consider a formal child care arrangement categories (i)-(iv), and an informal child care the rest of them. I define $cc_t = 1$ if the child (as declared by the parent) spent the maximum number of months in categories (i)-(iv), and 0 otherwise. Using this information, I obtain child care choices for period $t = 1$ (even though this question would cover both period $t = 0$ and $t = 1$).

To recover child care choices at the fifth year, the procedure is similar. In this year, the child care options are: (i) by someone 16 years of age or younger; (ii) by an adult at home; (iii) by an adult in someone else's home; (iv) in a child care center, before or after school program, community center, or Head Start; (v) child's own supervision; (vi) by sibling; (vii) others. I define $cc_t = 1$ (for period $t = 4$) if the child spent the higher number of months in (iv).

C.2 Income

I construct a proxy for family income using administrative information on the different sources of income. I define household income for individual i at year t as follows:

$$I_{it} = E_{it} + EITC_{it} + D_i(Sup_{it} + CSJ_{it}) + W_{it},$$

where E_{it} are labor earnings, $EITC_{it}$ is the earned income tax credit, D_i the treatment group dummy, Sup_{it} is the New Hope income supplement, CSJ_{it} are earnings from CSJs, and W_{it} are welfare payments. Sup_{it} and CSJ_{it} can be earned only by the treatment group. The Unemployment Insurance system (UI) of the State of Wisconsin collects quarterly data of E_{it}. I construct yearly measures of the nominal values of E_{it} to simulate the corresponding amount of EITC for every family. Finally, New Hope administrative data has information on Sup_{it} and CSJ_{it} on a quarterly basis.[81] For all period, I express income (after simulating the EITC) as annual 2003 dollars. Finally, W_{it} contains Food Stamps money and AFDC (replaced by "Wisconsin Works" after TANF) cash transfers.

Family income from administrative databases does not include several sources of income. Some of the excluded source of income are the unemployment insurance, child support, and others payments from social programs. Furthermore, it does not consider income from other family members. The New Hope surveys collect these and others sources of income. Unfortunately, the New Hope surveys do not track income for every year. Additionally,

[81] The income of the New Hope CSJs does not show up in the UI records. The CSJs that New Hope offered were limited in time (no longer than 6 months), and so they were not eligible to UI.

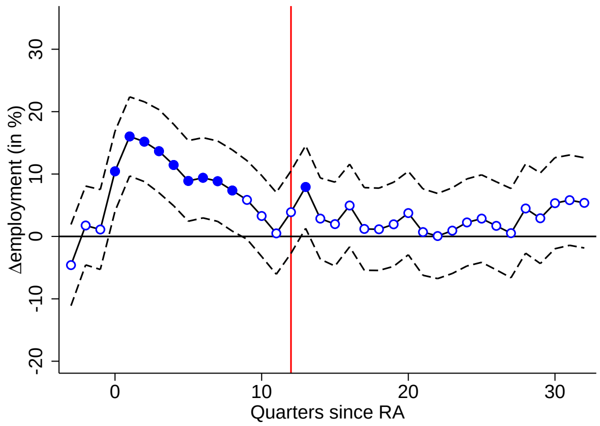
the year-two survey only asks about "last month's income," so income from administrative sources and surveys cannot be directly compared.

C.3 Labor supply

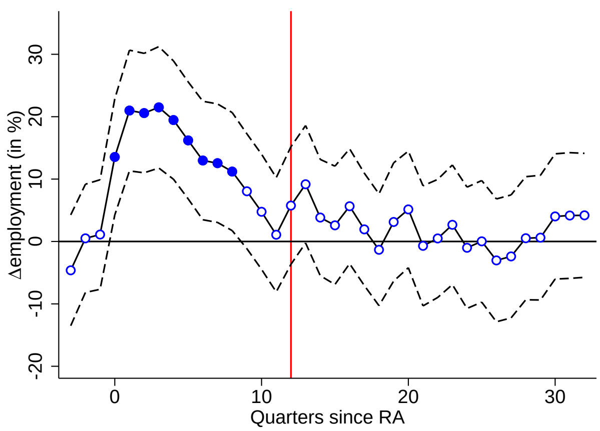
I define the employment measure using the Wisconsin UI records and the New Hope client database containing earnings in New Hope Community Service Jobs (CSJ). The employment dummy equals 1 if there is a positive wage in the UI or client database in a given period, and 0 otherwise.

In the body of the paper, I show average effects using a yearly measure. Figure C.1 shows the effect of New Hope on quarterly employment probability. In the figure, an individual is employed if she had a positive UI or CSJ income in a particular quarter. The figure suggests an overall positive employment effect. During the eligibility period, New Hope raised participants' employment probability by 9 percentage points on average (from a baseline average of 68%). After the the program ended, the impacts are statistically insignificant. Effects are explained by the unemployed-at-baseline group.

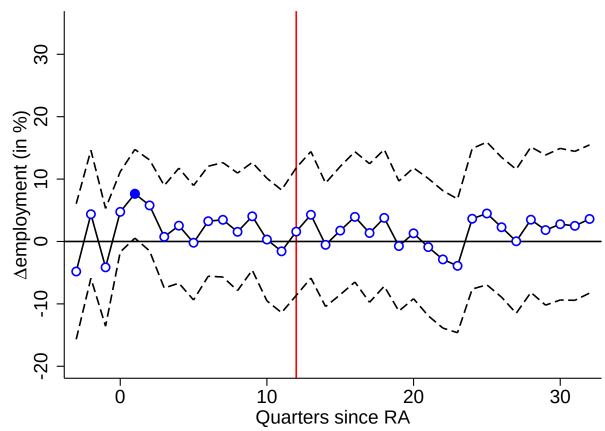
Figure C.1: The impact of New Hope on employment probability (administrative records)



(a) Overall



(b) Unemployed at baseline



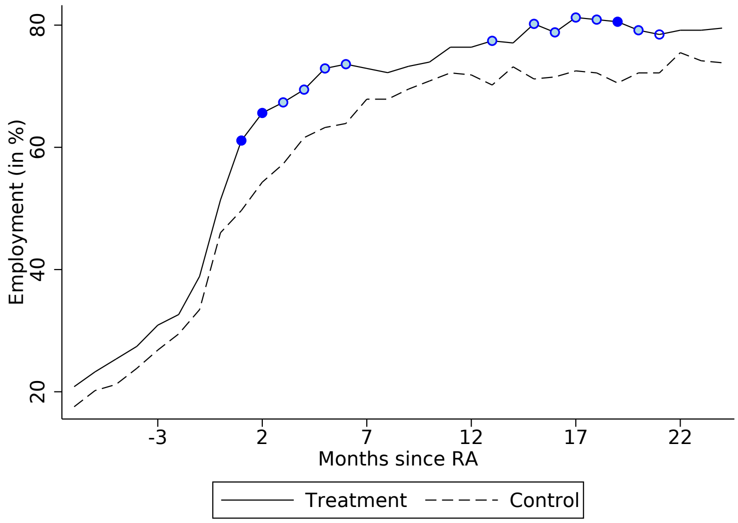
(c) Employed at baseline

Notes: The figure shows the effect of New Hope on employment probability by quarters since random assignment. The outcome variable is a dummy variable indicating a positive UI earnings record or CSJ wage. The solid vertical lines indicate the quarter where New Hope ends. I estimate $Employment_{it} = \alpha_t + \beta_t RA_i + X_i'\phi_t + +\varepsilon_{it}$, where RA_i equals 1 if family i is in the treatment group and 0 otherwise. For each quarter t, I show the estimated β_t with a 95%-level confidence interval ("•" indicates a statistically significant estimate).

An additional employment variable can be constructed using survey data. Figure C.2 plots the proportion of individuals employed across months since random assignment but using survey data instead. According to these estimates, the participant's employment probability increased 5 percentage points on average during the first two years of the program. Table C.1 shows the (full- and full- or part-time) employment probability for the five- and eight-year surveys. As the evidence from administrative sources, this table indicates that the program's effect five and eight years after random assignment is not statistically significant.[82]

[82] The surveys from years five and eight do not have the individual's working history, thus I cannot extend Figure C.2 up until eight years after baseline.

Figure C.2: The impact of New Hope on employment probability (New Hope surveys)



Notes: The figure shows the effect of New Hope on employment probability by quarters since random assignment. I use New Hope survey from the year-two follow-up to construct family income. For this figure, I use the sample of participants with at least one seven-year-old boy or younger. The solid and dashed lines show the probability that the individual reports at least one job spell in each month, for treatment and control groups. The solid vertical line indicates the quarter where New Hope ended (three years after random assignment).

Consider the following regression: $Employment_{it} = \alpha_t + \beta_t RA_i + X_i' \phi_t + \varepsilon_{it}$, where RA_i equals 1 if family i is in the treatment group and 0 otherwise. For each month t, the figure shows if the estimate of β_t is significant at the 10 (○), 5 (○), or 1% (●) level.

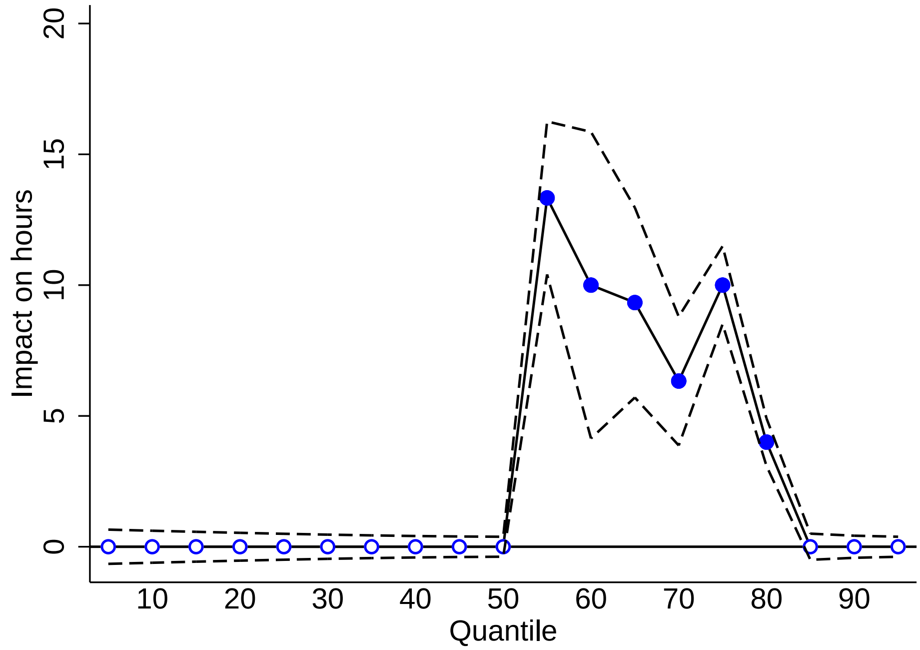
Table C.1: The effect of New Hope on employment probability (years five and eight)

	Treatment	**Control**	**Difference**
Year five			
Full-time employment	68.94	62.90	0.06
	[46.45]	[48.50]	(5.94)
Full- and part-time employment	76.52	73.39	0.03
	[42.55]	[44.37]	(5.44)
Year eight			
Full-time employment	55.64	55.04	0.01
	[49.87]	[49.94]	[6.17]
Full- and part-time employment	66.92	68.22	-0.01
	[47.23]	[46.74]	(5.81)

Notes: This table shows the impact of New Hope on employment probability (in percentage points). For this table, I use the CFS samples of the surveys from years five and eight. I consider only individuals with at least one seven-year-old boy or younger. Let $employment_{i,t}$ equals 1 if individual i declares to be "currently employed" at the year-t survey. The first and second column show the average of $employment_{i,t} \times 100$ for the treatment and control group. In square brackets, I show the corresponding standard deviation. Consider the following regression $employment_{i,t} = \alpha_t + \gamma_t RA_i + \varepsilon_{i,t}$. The third column depicts estimates of γ by year of the survey and whether $employment_{i,t}$ consider only full-time or both full- and part-time jobs. Standard errors are in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

To approximate effects on the intensive- and extensive-margin of labor supply, I follow Bitler et al. (2006) and compute Quantile Treatment Effects (QTE) on hours worked. Figure 4 illustrate the results from this exercise. In line with the reported estimated effects on employment, the figure shows statistically significant effect starting from the 55th quantile (the proportion of individual who do not work is 56%). Similar to Bitler et al. (2006), QTEs are positive beyond the 55th quantile, reaching a zero impact starting the 85th quantile. This pattern suggests that the program increased hours worked on both the extensive and intensive margin.

Figure C.3: Quantile Treatment Effects on hours worked



Notes: The figure shows quantile treatment effects on hours worked. The dependent variable corresponds to the average weekly hours worked from baseline until 15 months after baseline. "•" indicates a statistically significant estimate at the 5% level.

C.4 Child outcomes

Measures. The Teachers' reports contain have information on teachers' perceptions about the child academic outcomes. It has data on three measures: the academic subscale of the Social Skills Rating System (SSRS), the Classroom Behavior Scale and the Mock reports cards.

In the SSRS academic subscale (Gresham and Elliot, 1990), the teacher ranks the child in several subjects. These are reading skills, math, intellectual functioning, motivation, oral communication, classroom behavior, and parental encouragement. Each variable takes the following values: 1 (bottom 10%), 2 (next lowest 20%), 3 (middle 40%), 4 (next highest 20%), and 5 (highest 10%).

The Classroom Behavior Scale is a shorter version of School Adjustment Scale (Wright and Huston, 1995) and focuses on how the child behaves in class. The teacher must answer several questions on three topics: (i) behavior skills (for example, "complies with teacher requests, behaves so as not to disturb peers?"), (ii) independent skills (for example, "remains on-task with minimal supervision, manages free time constructively?"), and (iii) o transition skills (for example, "recognizes transition cues and stops ongoing behavior, moves quickly to next activity?"). All variables take discrete values from 1 (almost never) to 5 (almost always).

Finally, the mock report card provide an alternative measure of academic performance. In these cards, the teacher rates the child's academic performance regarding the following subjects: reading, oral language, written language, math, social studies, and sciences. The scale of this measure goes from 1 (below average) to 5 (excellent).

The New Hope survey reports gathers information on child skills directly with the child or using parents' reports. In the first case, the Woodcock-Johnson Achievement Battery (Woodock and Johnson, 1990) was administered. It measures four types of skills: letter-word identification, passage comprehension, applied problems, and calculation. The first two items cover reading skills while the last two measure mathematics skills. This test score is available in the surveys from year 5 and 8. However, the calculation item is not available in year 8. For the same survey, I obtain the the comprehension item with the overall measure and assumed that this last variable is a simple average of the three sub-items.

The second source of child outcomes from the New Hope surveys are the parent reports on academic performance. In these questionnaires, parents rate their children on the subjects of reading, math, and written work. The scale takes 5 values: 1 ("not well at all"), 2 ("below average"), 3 ("average"), 4 ("well"), and 5 ("very well").

Impacts on child outcomes. Table C.2 shows the estimated effects of the program on child outcomes two years after random assignment. In the table, I present the results from regressions with and without control variables (model 1 and 2, respectively). Of the 13 outcomes measured, the estimates indicate that the program had positive, statistically significant impact at the 5% level in three measures, and at the 10% level in three additional outcomes. If we control for baseline characteristics, the estimates show five cases where the impact is significant at the 5% level and two additional at the 10% level. All of the statistically significant effects are found for the SSRS Academic Subscale measure. Depending on the measure, the evidence suggests that the program increased the probability of being in the top 30% of the academic ranking of the class by 8 to 10 percentage points, from a baseline of 35 to 48%.

Table C.3 and C.4 present estimates for five and eight years after individuals entered the program. The first table presents p-values for 27 different outcomes. For the regression without control variables, I find only one statistically significant impact at the 5% level, and only one additional at the 10% level. When I add control variables, five impacts become statistically significant at the 5% level. The size of the estimated effects at year five are in general lower than that of year two. Eight years after random assignment (Table C.4), I find no statistically significant impacts at the 5% level and only one at the 10% level. Furthermore, the SSRS estimates are mostly lower than in the five-year estimates.

Table C.2: The effect of New Hope on measures of child development (two years after RA)

Measure	Baseline	Model 1 Estimate	Model 1 p-value	Model 2 Estimate	Model 2 p-value
Panel A. SSRS Academic Subscale					
Overall ($n = 411$)	0.343	0.073	0.073	**0.081**	0.043
Reading ($n = 403$)	0.349	**0.077**	0.038	**0.076**	0.047
Math ($n = 402$)	0.353	0.053	0.181	0.070	0.070
Reading grade expectations ($n = 405$)	0.351	0.043	0.281	0.041	0.311
Math grade expectations ($n = 402$)	0.348	0.034	0.376	0.045	0.250
Motivation ($n = 415$)	0.421	0.044	0.298	0.056	0.196
Parental encouragement ($n = 381$)	0.458	**0.087**	0.045	**0.090**	0.049
Intellectual functioning ($n = 410$)	0.415	**0.088**	0.035	**0.100**	0.021
Classroom behavior ($n = 414$)	0.440	0.073	0.074	0.083	0.055
Communication skills ($n = 413$)	0.477	**0.077**	0.043	**0.089**	0.018
Panel B. Classroom Behavior Scale					
Behavior skills ($n = 418$)	0.694	0.010	0.807	0.024	0.549
Independent skills ($n = 418$)	0.593	0.009	0.824	0.032	0.440
Transitional skills ($n = 417$)	0.581	0.018	0.660	0.032	0.430

Notes: The table shows the impact of the program on the probability that the teacher reports child i in the top 30% of the class academic performance distribution. The estimates are based on an ordered probit estimation. Model 1 does not include any covariate. The independent variables are a random assignment indicator (RA_i), age of the parent, marital status, ethnicity, a dummy variable indicating whether the parent has a high school diploma, the highest grade completed, and earnings in the past year. I show the estimates in percentage points. To construct standard errors, I draw 1,000 bootstrap samples, calculate the marginal effect of RA_i on being in the top 30% for each dataset, and compute its standard deviation across samples. I present standard errors in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table C.3: The effect of New Hope on measures of child development (five years after RA)

Measure	Baseline	Model 1		Model 2	
		Estimate	p-value	Estimate	p-value
Panel A. Woodcock-Johnson					
Letter-Word ($n = 773$)	-0.330	0.162	0.084	**0.231**	0.019
Comprehension ($n = 765$)	-0.159	0.058	0.453	0.126	0.102
Calculation ($n = 749$)	-0.739	0.068	0.467	0.085	0.314
Applied problems ($n = 769$)	-0.200	0.062	0.451	0.090	0.296
Panel B. SSRS Academic Subscale					
Overall ($n = 531$)	0.377	0.024	0.470	0.043	0.198
Reading ($n = 492$)	0.347	0.055	0.130	**0.077**	0.040
Math ($n = 479$)	0.349	0.016	0.673	0.026	0.501
Reading grade expectations ($n = 491$)	0.350	0.052	0.149	0.074	0.053
Math grade expectations ($n = 481$)	0.326	0.024	0.478	0.051	0.179
Motivation ($n = 536$)	0.406	0.001	0.976	0.010	0.789
Parental encouragement ($n = 494$)	0.454	0.015	0.706	0.051	0.197
Intellectual functioning ($n = 531$)	0.399	-0.003	0.928	0.010	0.787
Class behavior ($n = 536$)	0.426	0.040	0.262	0.042	0.273
Communication skills ($n = 532$)	0.384	0.033	0.383	0.039	0.311
Panel C. Teachers' Mock Reports Cards					
Reading ($n = 490$)	0.329	0.025	0.477	0.046	0.226
Oral language ($n = 497$)	0.293	0.023	0.521	0.041	0.282
Written language ($n = 497$)	0.221	0.036	0.238	0.059	0.060
Math ($n = 480$)	0.287	0.000	0.989	0.017	0.567
Social studies ($n = 424$)	0.248	0.005	0.872	0.014	0.706
Science ($n = 427$)	0.255	-0.001	0.985	0.015	0.643
Panel D. Classroom Behavior Scale					
Behavior skills ($n = 537$)	-0.011	0.018	0.837	0.044	0.644
Transitional skills ($n = 535$)	-0.039	0.074	0.394	0.109	0.243
Independent skills ($n = 535$)	0.009	-0.022	0.799	0.013	0.885
Panel D. Parents' Reports					
Reading ($n = 849$)	0.504	**0.080**	0.005	**0.073**	0.018
Math ($n = 849$)	0.551	0.023	0.427	0.033	0.255
Written work ($n = 845$)	0.479	0.017	0.557	0.012	0.696
Overall ($n = 848$)	0.603	0.003	0.904	0.012	0.690

Notes: The table shows the impact of the program on the probability that the teacher reports child i in the top 30% of the class academic performance distribution. The estimates are based on an ordered probit estimation. Model 1 does not include any covariate. The independent variables are a random assignment indicator (RA_i), age of the parent, marital status, ethnicity, a dummy variable indicating whether the parent has a high school diploma, the highest grade completed, and earnings in the past year. I show the estimates in percentage points. To construct standard errors, I draw 1,000 bootstrap samples, calculate the marginal effect of RA_i on being in the top 30% for each dataset, and compute its standard deviation across samples. I present standard errors in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Table C.4: The effect of New Hope on measures of child development (eight years after RA)

Measure	Baseline	Model 1 Estimate	Model 1 p-value	Model 2 Estimate	Model 2 p-value
Panel A. Woodcock-Johnson					
Letter-Word ($n = 767$)	-0.430	0.092	0.312	0.137	0.152
Comprehension ($n = 766$)	-0.378	-0.026	0.685	0.024	0.699
Panel B. SSRS Academic Subscale					
Overall ($n = 544$)	0.348	0.004	0.914	0.019	0.595
Reading ($n = 517$)	0.362	0.024	0.500	0.038	0.321
Math ($n = 421$)	0.305	-0.008	0.832	0.006	0.880
Reading grade expectations ($n = 514$)	0.357	0.000	0.993	0.016	0.680
Math grade expectations ($n = 423$)	0.291	-0.015	0.671	-0.004	0.907
Motivation ($n = 539$)	0.385	0.010	0.771	0.030	0.429
Parental encouragement ($n = 485$)	0.496	-0.070	0.055	-0.043	0.302
Intellectual functioning ($n = 545$)	0.435	0.023	0.510	0.031	0.412
Class behavior ($n = 546$)	0.464	-0.005	0.894	0.001	0.988
Communication skills ($n = 546$)	0.433	0.044	0.225	0.054	0.161
Panel C. Teachers' Mock Reports Cards					
Reading ($n = 507$)	0.315	-0.011	0.737	-0.009	0.825
Oral language ($n = 513$)	0.350	0.027	0.442	0.022	0.568
Written language ($n = 511$)	0.235	0.009	0.762	0.010	0.749
Math ($n = 413$)	0.265	-0.016	0.610	-0.018	0.597
Social studies ($n = 400$)	0.239	-0.003	0.931	-0.005	0.888
Science ($n = 386$)	0.219	-0.012	0.706	-0.006	0.876
Panel D. Classroom Behavior Scale					
Behavior skills ($n = 542$)	-0.006	0.011	0.894	0.031	0.737
Transitional skills ($n = 541$)	-0.008	0.015	0.858	0.042	0.650
Independent skills ($n = 543$)	-0.019	0.039	0.647	0.045	0.636
Panel D. Parents' Reports					
Reading ($n = 909$)	0.534	0.033	0.258	0.014	0.664
Math ($n = 909$)	0.514	0.029	0.282	0.025	0.406
Written work ($n = 909$)	0.486	0.044	0.104	0.048	0.115
Overall ($n = 909$)	0.562	0.027	0.318	0.027	0.357

Notes: The table shows the impact of the program on the probability that the teacher reports child i in the top 30% of the class academic performance distribution. The estimates are based on an ordered probit estimation. Model 1 does not include any covariate. The independent variables are a random assignment indicator (RA_i), age of the parent, marital status, ethnicity, a dummy variable indicating whether the parent has a high school diploma, the highest grade completed, and earnings in the past year. I show the estimates in percentage points. To construct standard errors, I draw 1,000 bootstrap samples, calculate the marginal effect of RA_i on being in the top 30% for each dataset, and compute its standard deviation across samples. I present standard errors in parenthesis. *,**,*** indicates significance at the 10, 5, and 1% level.

Tables C.5 and C.6 present the results from the stepdown procedure that controls for a fixed family-wise error rate. I apply the procedure on the estimates from two and five years after random assignment and for each panel representing a set of variables from a common measure (the two continuous and the three ordinal measures). Within each panel, I order the tests from higher to the lower t-statistic. The first column shows the adjusted p-values from models that do not include any control variable, while the second controls for the same set of covariates from the previous results (Tables C.2-C.4). Table C.5 presents the adjusted p-values for the second-year child outcomes. In the SSRS measures (Panel A), three of the ten outcomes survive the procedure in regressions with no control variables and seven out of ten in a model with control variables. The impacts on the classroom behavior scale (Panel B) are nos significant. Table C.6 shows results from five years after baseline. The regressions with no control show no significant results, while those with control variables reveal statistically significant effects in reading measures.

Table C.5: Step-down procedure (two years after RA)

Measures	**No controls**	**W/ controls**
Panel A. SSRS Academic Subscale		
Intellectual functioning	0.068	0.046
Communication skills	0.096	0.081
Parental encouragement	0.098	0.086
Overall	0.113	0.086
Classroom behavior	0.113	0.086
Reading	0.113	0.086
Math	0.190	0.086
Motivation	0.252	0.173
Math grade expectations	0.252	0.177
Reading grade expectations	0.252	0.177
Panel B. Classroom Behavior Scale		
Transitional skills	0.497	0.340
Independent skills	0.507	0.340
Behavior skills	0.507	0.340

Notes: The table shows p-values for the null hypothesis of no difference between treatment and control groups adjusted to control for family-wise error rate. I show p-values from regresions with and without control variables.

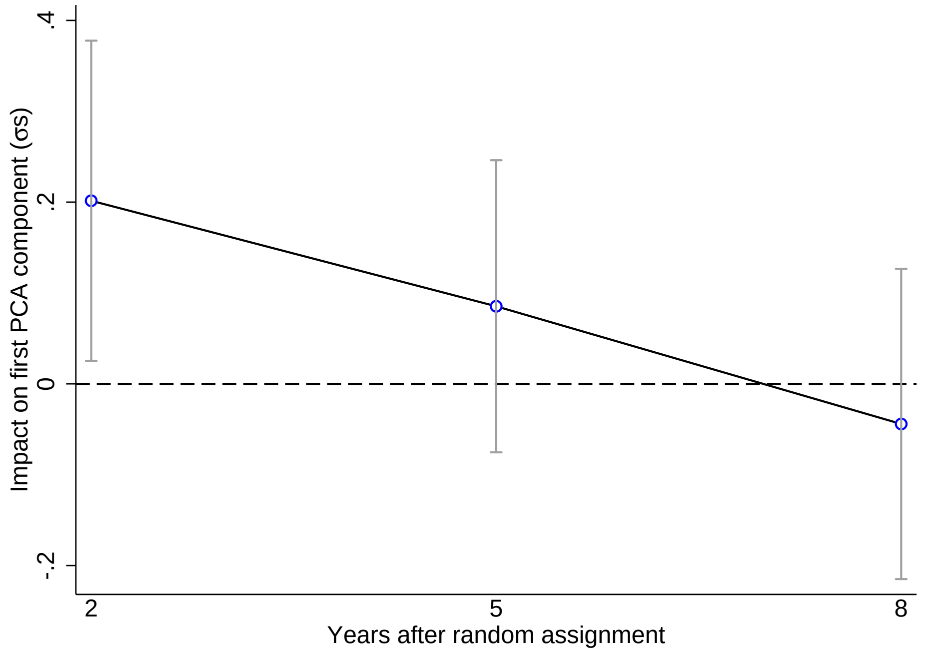
Table C.6: Step-down procedure (five years after RA)

Measures	No controls	W/ controls
Panel A. Woodcock-Johnson		
Letter-Word	0.118	0.025
Comprehension	0.419	0.114
Applied problems	0.419	0.251
Calculation	0.419	0.251
Panel B. SSRS Academic Subscale		
Reading	0.206	0.100
Reading grade expectations	0.231	0.113
Math grade expectations	0.365	0.291
Parental encouragement	0.408	0.309
Overall	0.445	0.309
Classroom behavior	0.445	0.331
Communication skills	0.540	0.331
Math	0.540	0.415
Intellectual functioning	0.589	0.503
Motivation	0.589	0.503
Panel C. Teachers Mock' Reports Cards		
Written language	0.237	0.074
Reading	0.441	0.246
Oral language	0.441	0.259
Math	0.600	0.466
Science	0.627	0.466
Social studies	0.627	0.466
Panel D. Classroom Behavior Scale		
Independent skills	0.271	0.185
Behavior skills	0.496	0.412
Transitional skills	0.579	0.436
Panel E. Parents' Reports		
Reading	0.015	0.035
Math	0.360	0.261
Overall	0.375	0.475
Written work	0.459	0.475

Notes: The table shows p-values for the null hypothesis of no difference between treatment and control groups adjusted to control for family-wise error rate. I show p-values from regresions with and without control variables.

To summarize the effects of New Hope on the different SSRS scores, I perform a PCA analysis. The PCA components are a weighted average of all the available information about academic performance skills. To implement the PCA, I transform the SSRS to a dummy variable that equals 1 if the child ranks 3 or above (top 70% of the class distribution) and 0 otherwise. The first principal component accounts for almost 50% of the variance of the academic measures from years two, five, and eight. For the same years, the second principal component explaines less than 15% of the overall variance. Hence, I use the first component and interpret it as the underlying "child human capital" or "academic skills." Figure C.4 depicts average treatment effects on the principal component factor, expressed as standard deviation units. The estimated effect equals 0.2 standard deviations by year two, close to 0.1 standard deviations by year five, and almost zero by year eight.

Figure C.4: Impact of New Hope on the first PCA component of academic measures



Notes: The figure presents the impact on the principal component factor of the SSRS measures (across years) with a 10%-level confidence interval of the estimated effect.

D Welfare parameters

In this appendix, I show the welfare functions that determine disposable income (equation 4). I consider three mean-tested programs: the EITC, AFDC, and Food Stamps payments.

D.1 The EITC

The EITC parameters vary by state, year, and the number of children (k_t).[83] Denote annual gross earnings as $E_t = w_t h_t \times 52$. Following Chan (2013), there are four key parameters for the federal EITC: the phase-in and phase-out rates ($r^k_{1,t}$ and $r^k_{2,t}$), and the bracket thresholds ($b^k_{1,t}$ and $b^k_{2,t}$), where the index k denote the number of children. k goes from 1 to 3, since the parameters of the EITC schedule do not vary for families with more than three children. In year t and for a family with $k_t = k$ number of children, the federal EITC payment ($EITC^f_t$) follows:

$$EITC_t = \begin{cases} r^k_{1,t} E_t & \text{if } E_t < b^k_{1,t} \\ r^k_{1,t} b^k_{1,t} & \text{if } b^k_{1,t} \leq E_t < b^k_{2,t} \\ \max\left\{ r^k_{1,t} b^k_{1,t} - r^k_{2,t}(E_t - b^k_{2,t}), 0 \right\} & \text{if } E_t \geq b^k_{2,t} \end{cases}$$

In the case of Wisconsin, the state EITC payment ($EITC^s_t$) is determined as a fraction of the federal payment: $r^k_{s,t} EITC^f_t$, where $0 < r^k_{s,t} < 1$ varies by number of children and year. The total EITC payment equals $EITC_t = EITC^f_t + EITC^s_t$.

D.2 The AFDC and TANF

The AFDC parameters vary by family composition and by year. Starting 1997, the state of Wisconsin implemented "Wisconsin Works" (W-2), under the TANF umbrella. Instead of giving cash transfers like most states did, W-2 offered paid CSJs for up to 5 years. In terms of the model then, a W-2 salary becomes part of the potential wage offer (equation 3).

Until 1996 (that is, periods $t = 0$ and $t = 1$), the standard AFDC program was in place. Let B^*_t be the cash transfer an individual under welfare could get, given by:

$$B^*_t = \max\left\{ \min\left\{ \overline{B}, \overline{B} - (E_t - 30) \times .67 \right\}, 0 \right\},$$

where $\overline{B}$ is the so-called "benefit standard," the maximum amount of welfare an individual is entitled to. Individuals enter the program if $E_t \leq c$. The parameters c and $\overline{B}$ vary by family size and state.[84] This formula captures the \$30-and-a-third policy implemented in 1967: The recipient may keep the first 30 dollars she makes. Above that value, for each dollars she earns, she must "pay a tax" of 0.67 (the marginal tax rate is 67%). In practice, the formula is designed for monthly figures, so I adapted parameters to accommodate for annual income.

[83] The federal parameters can be found at http://www.taxpolicycenter.org/sites/default/files/legacy/taxfacts/content/PDF The parameters for the state of Wisconsin are obtained from http://users.nber.org/~taxsim/state-eitc.html.

[84] The exact values can be found at the Welfare Rules Database for Wisconsin, Area 1.

D.3 SNAP

The Supplemental Nutrition Assistance Program (SNAP)—formerly known as Food Stamps—is the largest nutrition program in the U.S. The program provides money vouchers to eligible individuals to spend food in grocery stores.

Unlike the AFDC, SNAP eligibility and voucher parameters have not changed much in time. It does not vary by state either. Let E_n be net income, E gross earned income, B welfare payments (including AFDC and New Hope cash transfers), SD a standard deduction, and e the poverty guideline. To receive SNAP, a household must meet the gross and net income tests:[85]

$$E < 1.3e,$$

$$E_n < e,$$

where net income follows[86]

$$E_n = 0.8E + B - SD$$

The SNAP benefits are determined by the following formula:

$$S^* = \max\left\{MaxB - 0.3E_n, 0\right\},$$

where $MaxB$ is the Maximum allotment. All income thresholds and other parameters are adjusted following Social Security's Cost-of-Living Adjustments.[87]

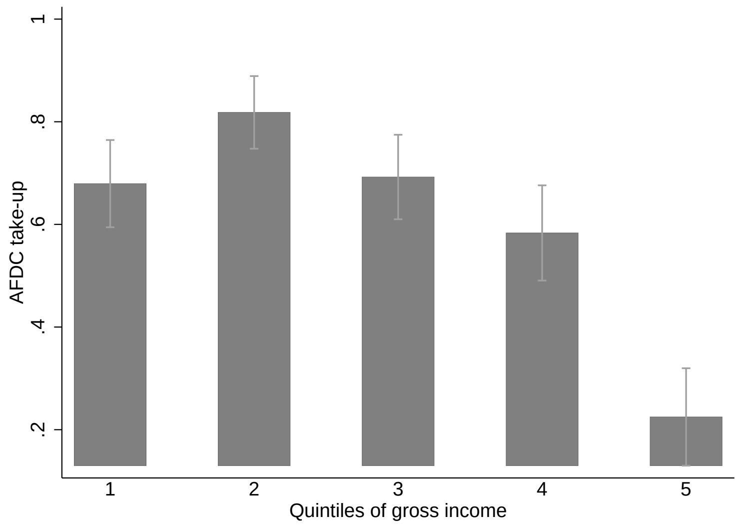
[85] Also, if a family is living only with AFDC payments, then it is automatically eligible. For the purpose of this paper, I assume that if a participant is not working, then she is eligible for SNAP payments.

[86] The actual formula includes a standard shelter deduction, which I assume to be zero for all families.

[87] Maximum allotments ($MaxB$) are taken from http://www.fns.usda.gov/snap/fy-2004-allotments-and-deduction-information. See https://www.ssa.gov/oact/cola/colaseries.html for cost-of-living adjustments. For poverty guidelines, see https://www.ssa.gov/policy/docs/statcomps/supplement/2014/3e.html#table3.e8.

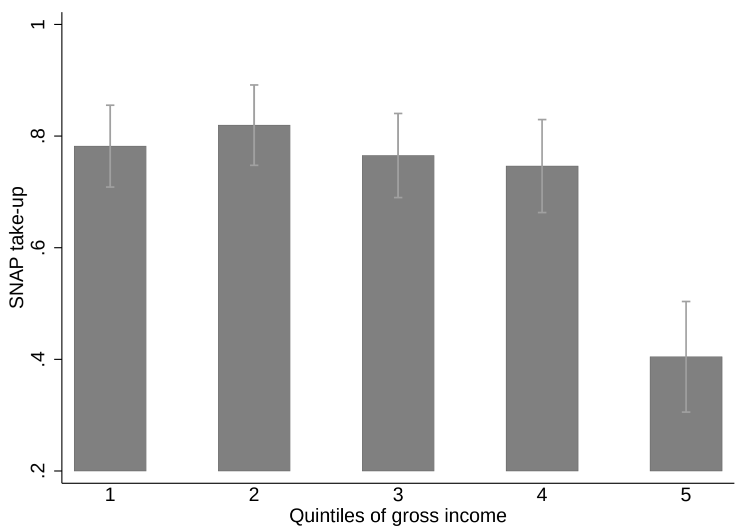
D.4 Take-up rates of AFDC and SNAP

Figure D.1: Take-up rate of AFDC



Notes: The figure shows the estimated take-up rate across gross income quintiles during the year of random assignment (1994-1995). The take-up rate is defined as the proportion of eligible families who received a AFDC payment during the year. Gross income corresponds to the sum of quarterly earnings in the year of random assignment. The bars show the point estimate of the take-up rate with a 95% boostrap confidence interval (1,000 draws).

Figure D.2: Take-up rate of SNAP



Notes: The figure shows the estimated take-up rate across gross income quintiles during the year of random assignment (1994-1995). The take-up rate is defined as the proportion of eligible families who received at least one SNAP check over the course of the year. Gross income corresponds to the sum of quarterly earnings in the year of random assignment. The bars show the point estimate of the take-up rate with a 95% boostrap confidence interval (1,000 draws).

E Identification of the production function

In this appendix, I show how we can identify the human capital production function (equation 2) using ordinal measures of academic achievement. The proof borrows insights from Cunha et al. (2010) and Agostinelli and Wiswall (2016) in showing the identification of a production function under a particular measurement error structure.

I focus on the following production function:

$$\ln \theta_{t+1} = f(\theta_t, c_t, \tau_t) + \mu c c_t \mathbb{I}\{a_t \leq 6\} \tag{14}$$

where $f(.)$ is such that, for some point $(\theta_t^{'}, c_t^{'}, l_t^{'})$, $f(\theta_t^{'}, c_t^{'}, l_t^{'})$ does not depend on an unknown coefficient.[88] The above function describes a technology that varies by child care choice (and thus by age), but it is otherwise constant in time—a key feature of the identification result.

The measures of academic achievement are ordinal variables. These measures rank the child in the classroom academic achievement distribution. For a measure M_t in period t, we have

$$M_t = \begin{cases} 1 & \text{if } \ln \theta_t + \epsilon_t^z \leq \kappa_{1,t} \\ 2 & \text{if } \kappa_1 < \ln \theta_t + \epsilon_{t,m}^z \leq \kappa_{2,t} \\ \vdots & \\ 5 & \text{if } \ln \theta_t + \epsilon_t^z > \kappa_{4,t}. \end{cases} \tag{15}$$

where

$$\epsilon_t^z \sim N(0,1) \qquad \forall t. \tag{16}$$

Underlying equation (16) there are two essential assumptions. First, the coefficient associated to $\ln \theta_t$ equals 1 for all t—a classical measurement-error model. Second, the cutoffs are constant across child age and child care types.

The problem is to identify (14) and the parameters of the measurement system (15), given (16). I follow Agostinelli and Wiswall (2016) to prove the following Lemma.

Lemma 1. *Suppose the production function and the measurement system follow (14), (15), and (16). Then* $\Phi^{-1}\left[Pr\left(M_{t+1} = 5 \mid \ln \theta_t = \bar{\theta}, \ln c_t = \bar{c}, \ln l_t = \bar{l}\right)\right]$ *(where* $\Phi(.)$ *denotes a standard normal cdf) is identified with two measures* M_t *and* M_{t+1} *and it is equal to* $f\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right) - \kappa_{4,t+1}$.

Proof. To simplify notation, let us abstract for a moment of any age-induced difference in the technology of human capital, and so $\ln \theta_{t+1} = f(\theta_t, c_t, l_t)$. First, note that the cutoffs from period t $(\kappa_{j,t})$ are identified by the normality assumption on the measurement error.[89] Second, given equations (14) and (15),

[88] Agostinelli and Wiswall (2016) introduces the concept of "known location and scale" (KLS). A production function is KLS if, for two non-zero vectors $(\theta_t^{'}, c_t^{'}, l_t^{'})$ and $(\theta_t^{''}, c_t^{''}, l_t^{''})$ such that $\theta_t^{'} \neq \theta_t^{''}$, $c_t^{'} \neq c_t^{''}$, and $l_t^{'} \neq l_t^{''}$, $f(\theta_t^{'}, c_t^{'}, l_t^{'})$ and $f(\theta_t^{''}, c_t^{''}, l_t^{''})$ do not depend on unknown parameters. For the production function from equation (2), the KLS property holds only if the sum of the coefficients sum up to 1, that is, a constant-return-to-scale production function. Because I am assuming a classical measurement error structure, I do not need the production function to be KLS.

[89] The proof follows the standard discrete-choice analysis. In general, a constant term and the variance are not identified in an discrete-ordered framework.

$$\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right] = f\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right) - \kappa_{4,t+1}.$$

Similarly for period t, we can express $\ln\theta_t = \bar{\theta}$ as $\bar{\theta} = \Phi^{-1}\left[Pr\left(M_t = 5 \mid \theta_t = \bar{\theta}\right)\right] + \kappa_{4,t}$, where $\kappa_{4,t}$ is known and $\Phi^{-1}\left[Pr\left(M_t = 5 \mid \theta_t = \bar{\theta}\right)\right]$ can be constructed using observed data.

Hence,

$$\begin{aligned}\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right] = \Phi^{-1}\{Pr(M_{t+1}=5 \mid \bar{\theta} = \Phi^{-1}\left[Pr\left(M_t=5 \mid \theta_t=\bar{\theta}\right)\right] \\ + \kappa_{4,t}, c_t=\bar{c}, l_t=\bar{l})\}\end{aligned}$$

Because the expression $\Phi^{-1}\left[Pr\left(M_t = 5 \mid \theta_t = \bar{\theta}\right)\right] + \kappa_{4,5}$ is known for any $\bar{\theta}$, the equation above shows that we can identify $\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right]$ with data on M_t, c_t, and l_t. □

Now we can use $\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right]$ to identify the production function. This result is summarized in the following proposition.

Proposition 1. *Suppose that (i) the conditions from Lemma 1 hold and (ii) for some point* $(\theta_t^{'}, c_t^{'}, l_t^{'})$, *f(.) is known (do not depend on unknown coefficients). Then the technology of academic skills formation (equation 14) is identified.*

Proof. Let us start with the production function of a young child ($a_t \leq 6$) in home care ($cc_t = 0$). Using Lemma (1), we can identify $\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right]$ for any $\bar{\theta}, \bar{c}$, and $\bar{l} \in \mathbb{R}$. Furthermore, we can exploit the the fact that we know the value of $f(.)$ for some point to eliminate the unknown parameter $\kappa_{5,t+1}$. Choose a point $(\hat{\theta}, \hat{c}, \hat{l})$ such that $f\left(e^{\hat{\theta}}, e^{\hat{c}}, e^{\hat{l}}\right) = \alpha$ is known, and note that

$$\begin{aligned}\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right] - \Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\hat{\theta}, c_t=\hat{c}, l_t=\hat{l})\right] = \\ f_y\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right) - \alpha.\end{aligned}$$

Therefore, given that the left-hand side is identified because of Lemma 1, we can identify the production function $f(.)$ in home care by varying $\bar{\theta}, \bar{c}$, and $\bar{l}$ over their support. Since, by assumption, the production of old and young children in home care is the same, we have also identified the production function of old children.

To identify the production function of a child in child care, we need to identify the TFP parameter (μ). To this end, we can exploit the fact that the cutoffs do not vary by child care choices.[90] Given that $f(.)$ is identified for any $\bar{\theta}, \bar{c}$, and $\bar{l}$ and by Lemma 1, we can obtain the TFP term as follows:

$$\Phi^{-1}\left[Pr(M_{t+1}=5 \mid \theta_t=\bar{\theta}, c_t=\bar{c}, l_t=\bar{l})\right] - f\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right) - \kappa_{4,t+1} = \mu,$$

[90]The KLS property precludes from estimating production functions with TFP dynamics (for example, $f(.)+\mu$). By using that the cutoffs are constant within the group of young children, we can recover the TFP parameter in the young child production function. Agostinelli and Wiswall (2016) states a similar argument to identify a general class of production functions.

and so μ is identified.

Finally, $\kappa_{1,t+1}$, $\kappa_{2,t+1}$ and $\kappa_{3,t+1}$ can be identified using any production function. For example, using a similar reasoning to that of Lemma 1, we can identify $\Phi^{-1}[1 - Pr(M_{t+1} \geq 4 \mid \theta_t = \bar{\theta}, c_t = \bar{c}, l_t = \bar{l})]$. Given that $f\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right)$ is identified, then

$$\Phi^{-1}[1 - Pr(M_{t+1} \geq 4 \mid \theta_t = \bar{\theta}, c_t = \bar{c}, l_t = \bar{l})] - f\left(e^{\bar{\theta}}, e^{\bar{c}}, e^{\bar{l}}\right) = -\kappa_{3,t+1}.$$

Following an analogous argument, we can identify $\kappa_{1,t+1}$, $\kappa_{2,t+1}$ and $\kappa_{3,t+1}$. $\square$

F Auxiliary model

F.1 Data for estimation

To obtain the trajectories of the key variables of the model, I combine administrative and survey data. Administrative data is available throughout the period (from baseline until eight years after), while surveys were collected only at specific years (two, five, and eight years after baseline). This section describes how I combine data from different sources to construct the main variables predicted by the model.

Weekly hours worked. Using the second-year survey, I compute the average hours worked in a week for the baseline and one year after random assignment ($t = 0$ and 1). In this survey, individuals reported the usual hours worked in every job they had in the last two years (thus, covering baseline and year $t = 1$). For every job they had, respondents reported weekly hours worked at the beginning and at the end of the job. Using the reported dates for each job spell, I compute monthly weekly hours worked. If more than one job was reported in a particular month, I assume that there are no overlapping in spells and take the average of all jobs. If the individual did not report having a job in a particular month, I set hours worked to zero. Then, for each calendar year, I compute the annual average of weekly hours worked—including the zeros corresponding to the months that the individual did not work. From the fifth- and eighth-year surveys, I recover the hours worked from periods $t = 4$ and 7. In these surveys, individuals reported the average hours worked at the current or most recent job in the last 12 months. I weight the reported average hours worked with a variable capturing the proportion of quarters employed in a year. I compute this variable using administrative data from the UI database and calculating the proportion that individuals stayed employed in year ($4^{-1} \sum \mathbf{1}\{wage_q > 0\}$, where $wage_q$ is quarterly labor earnings). Finally, I discretize hours worked variable in three categories: 0 if hours worked equals 0, 15 if hours is greater than 0 but less than 30, and 40 if hours worked is above 30.

Hourly wages. To construct this variable, I combine administrative with survey data to compute weekly average gross earnings (in the numerator) and weakly average hours worked (in the denominator). I obtain weekly average gross earnings by averaging quarterly earnings in a particular year (from the UI data) with any salary earned in a CSJ (for those in the treatment group), adjusted to 2003 dollars. I divide weekly earnings by average hours worked in a week from survey data (see paragraph above). Because hours worked are available for period $t = 0, 1, 4$ and 7, so is hourly wages. For $t = 4$ and $t = 7$, the state CSJs from TANF enter the pool of possible wage offers. Thus, I incorporate the CSJs payments in the hourly wage calculation of $t = 4$ and $t = 7$.

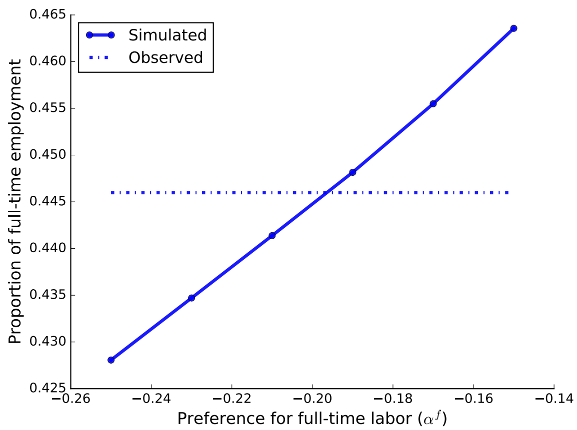
Child care use. See Appendix (C.1) for details on the construction of the child care variable.

Family consumption. To construct annual family consumption, I use information on (i) total income, (ii) child care payments, and (iii) family composition. First, I obtain total annual income as the sum of UI earnings, AFDC or TANF payments (depending on the year), and potential EITC payments. The first three sources of income are observed from administrative data, while for the last source I compute the potential EITC payment following the EITC schedule and assuming full take-up. Second, I compute child care payments as the previous paragraph describes. Then, total family consumption in a year equals total income minus child care payments. These monetary values are expressed in 2003 dollars. Finally, to compute per-capita consumption I divide total family consumption by the household size (parents and number of children).

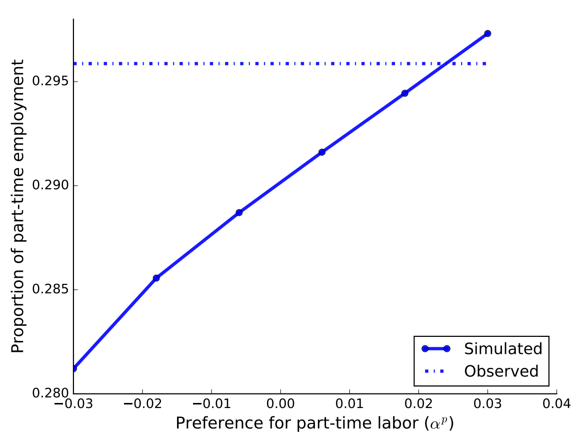
Child human capital. I use the set of SSRS variables to measure child academic performance (see Appendix C.4). I construct dummy variables associated to each measure indicating whether the child is in the top 70% of the class. I take the first PCA score as a composite measure of child academic achievement.

F.2 Local identification from targeted moments

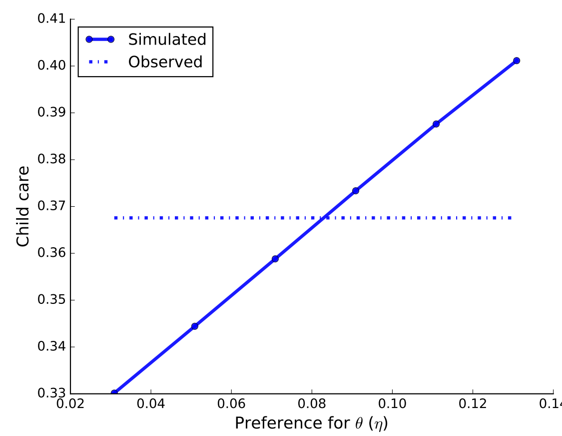
Figure E.1: Target moments locally identify structural parameters: utility function



(a) α^f



(b) α^p

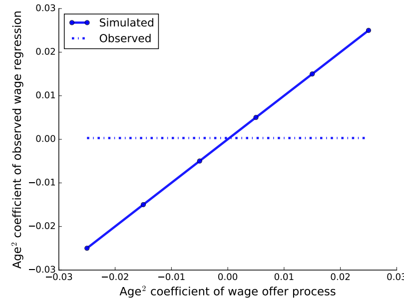


(c) η

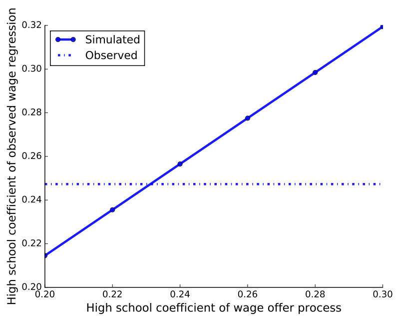
Notes: The figure depicts how a target moment is able to locally identify a structural parameter. The solid line illustrates the simulated moment as a function of a structural parameter. The dashed line shows the observed moment.

Figure E.2: Target moments locally identify structural parameters: wage offer

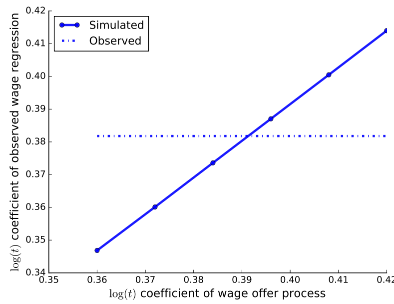
(a) Age



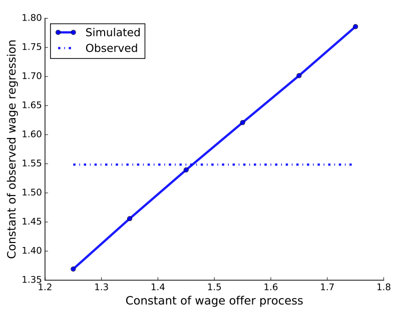
(b) Age^2



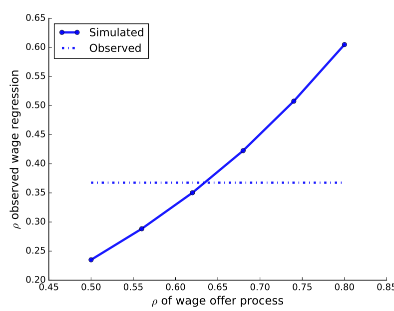
(c) High school



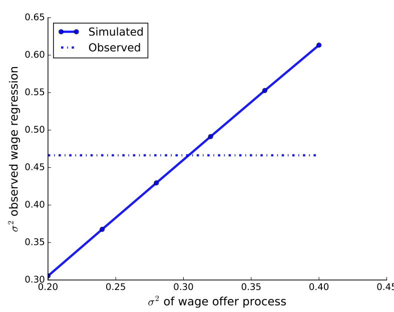
(d) $\ln t$



(e) Constant



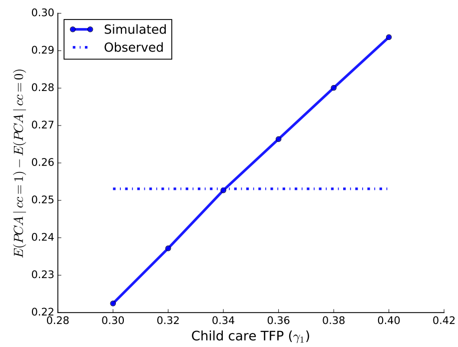
(f) ρ



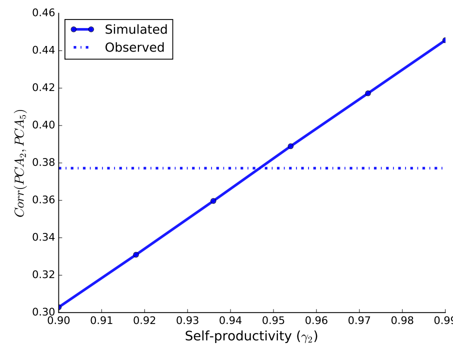
(g) Variance

Notes: The figure depicts how a target moment is able to locally identify a structural parameter. The solid line illustrates the simulated moment as a function of a structural parameter. The dashed line shows the observed moment.

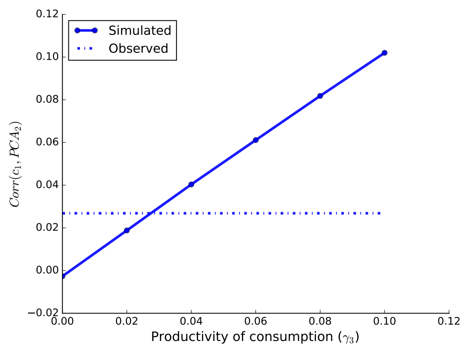
Figure E.3: Target moments locally identify structural parameters: production function



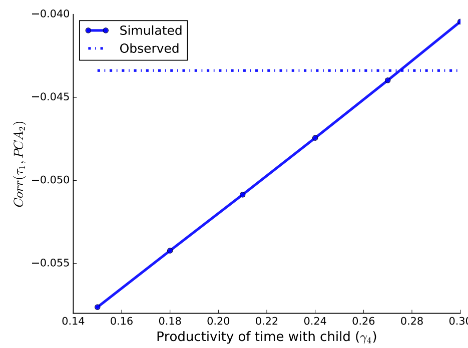
(a) $\gamma 1$



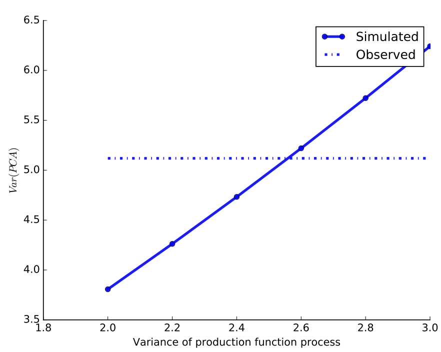
(b) $\gamma 2$



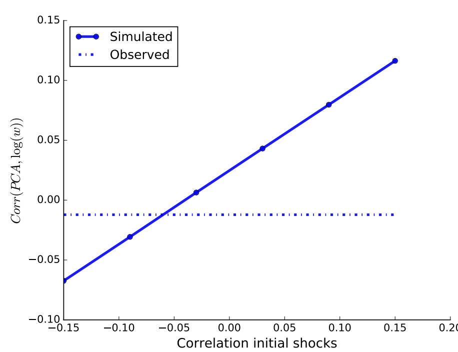
(c) γ_3



(d) γ_4



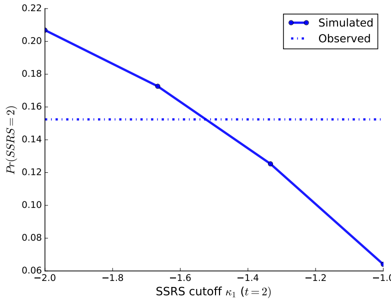
(e) Variance



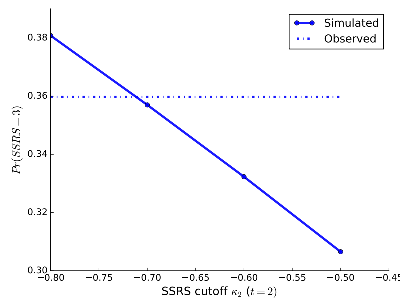
(f) $Corr(\varepsilon_0^{\theta}, \varepsilon_0^{w})$

Notes: The figure depicts how a target moment is able to locally identify a structural parameter. The solid line illustrates the simulated moment as a function of a structural parameter. The dashed line shows the observed moment.

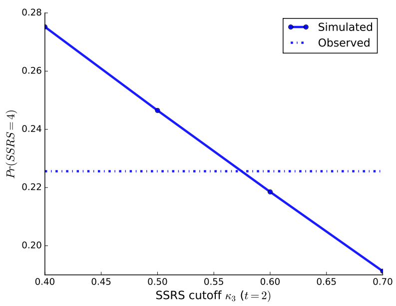
Figure E.4: Target moments locally identify structural parameters: measurement system



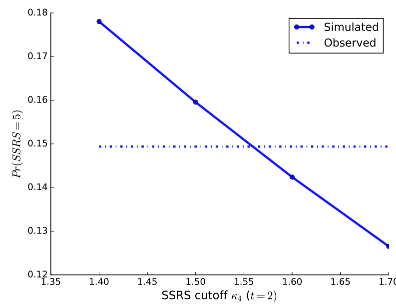
(a) κ_1 $(t = 2)$



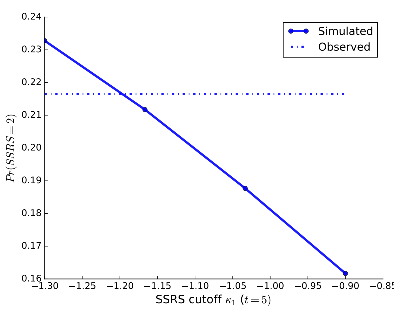
(b) κ_2 $(t = 2)$



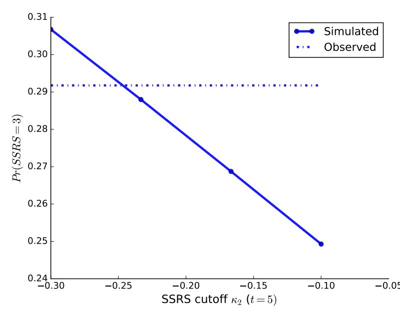
(c) κ_3 $(t = 2)$



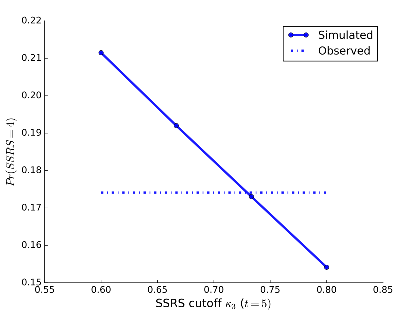
(d) κ_4 $(t = 2)$



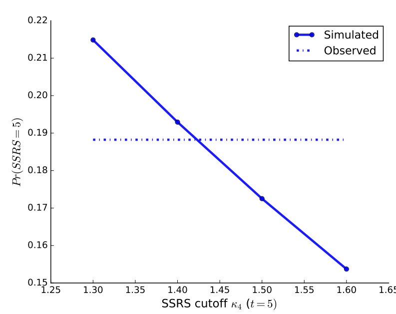
(e) κ_1 $(t = 5)$



(f) κ_2 $(t = 5)$



(g) κ_3 $(t = 5)$



(h) κ_4 $(t = 5)$

Notes: The figure depicts how a target moment is able to locally identify a structural parameter. The solid line illustrates the simulated moment as a function of a structural parameter. The dashed line shows the observed moment.