Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers

Raquel Bernal[1]
Department of Economics
Universidad de los Andes
Colombia

Michael P. Keane
ARC Federation Fellow
University of Technology Sydney
Research Fellow, Arizona State

First Version: September 15th, 2004
This Version: October 1st, 2007

Abstract

We evaluate effects of home inputs on children's cognitive development using single mothers from the National Longitudinal Survey of Youth (NLSY). Important selection problems arise in trying to assess the impact of maternal time and income on children's development. To deal with this, we exploit the (plausibly) exogenous variation in welfare policy rules facing single mothers. In particular, the 1996 Welfare Reform, and earlier State level policy changes, generated substantial increases in their work and child care use. Thus, we construct a comprehensive set of welfare policy variables, and use them (along with local demand conditions) as instruments to estimate child cognitive ability production functions.

Because welfare rules are very complex, we need many variables to characterize them. Thus, we face a "many instrument problem" (i.e., 2SLS can be severely biased toward OLS). We deal with this in two ways. First, we use LIML, which is more robust in this case. Indeed, 2SLS estimates fall roughly halfway between LIML and OLS. Second, we use factor analysis to condense the instrument set. Using only the most important factors as instruments, 2SLS and LIML results are very similar, and both are very similar to LIML results using the full instrument set.

We find that the effect of child care use is negative, significant and rather sizable. In particular, an additional year of child care is predicted to reduce child test scores by 2.9% (.156 standard deviations). But this general finding masks important differences across types of child care, child age ranges, and maternal education. Indeed, only informal care leads to significant reductions in cognitive outcomes. Formal care (i.e., center based care) does <u>not</u> have any adverse effect. In addition, child care has larger negative effects for older children, and maternal time is more valuable for more educated mothers. Finally, we also provide evidence of a strong link between test scores at ages 4, 5 and 6 and completed education.

---

[1] We acknowledge support from Australian Research Council grants FF0561843 and DP0774247, and Bernal acknowledges financial support from "The Searle Fund". Corresponding author: michael.keane@uts.edu.au.

1. Introduction

The effect of parental time inputs and child care use (and/or quality) on child development has been widely analyzed, especially in the psychology and sociology literatures. Economists have also recognized the importance of this issue. In particular, some recent studies find that the factors determining individuals' labor market success are already largely in place by about age 16.[2] Thus, policies to enhance human capital at later ages (e.g. college tuition subsidies) have, at best, a minor impact. Naturally, such findings focus attention on human capital development at early ages, including the role of child care. But this topic is difficult, due to a range of data and econometric problems. In this paper, we take a small step towards learning more about development of cognitive ability at very young ages (i.e., up until age 6).

Extensive research has shown that children's early achievement is a strong predictor of a variety of later life outcomes, including educational attainment, high earnings, teenage pregnancy, welfare participation and crime. Indeed, we provide new evidence of a strong association between test scores at ages 4, 5 and 6 and completed education, even conditional on a rich set of family background controls. Thus, the issue of what determines cognitive ability at early ages is critical for the design of public policy aimed at improving labor market outcomes. Unfortunately, results from previous literature on determinants of children's cognitive achievement are inconclusive at best.

A major challenge in estimating determinants of achievement is that available data are often deficient. For example, inherited ability cannot be perfectly measured, creating difficult problems of endogeneity and self-selection. In fact, a key reason for the diverse results of previous literature may be failure of many studies to adequately control for biases arising from two factors: (1) Women who work/use child care may be systematically different from those who do not, both in the constraints they face and their tastes; (2) The child's cognitive ability, which is at least partially unobserved by the econometrician, may itself influence the mother's decisions. In general, mothers' work and child care decisions may depend on unobserved characteristics of *both* mothers and children.

To clarify the endogeneity problem, consider two example cases. In case (1), women with higher skills are more likely to have children with high cognitive ability and also more likely to work/use child care. Failure to control for this correlation would cause estimated effects of maternal employment (child care) on child cognitive outcomes to be biased upward. In case (2), mothers of low ability endowment children may compensate by spending more time with them. Then, mothers of high ability children are more likely to work (use child care). Again, the estimated effect of

[2] See, e.g., Keane and Wolpin (1997, 2001) and Cameron and Heckman (1998).

maternal employment (child care) on cognitive outcomes is upward biased. Clearly, such sample selection issues make evaluating the effects of women's decisions on child outcomes very difficult.

In this paper, we estimate child cognitive ability production functions for single mothers in the NLSY. We focus on single mothers because recent major changes in welfare rules led to dramatic and plausibly exogenous variation in work incentives and childcare prices/availability for this group. From 1993 to 1996, 43 States received federal waivers authorizing innovative approaches to welfare reform. Then, in 1996, the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) replaced Aid to Families with Dependent Children (AFDC) with the Temporary Aid to Needy Families (TANF) program, and created the Child Care Development Fund (CCDF). These policy changes greatly increased employment and childcare use among single mothers with children aged 0-5. Indeed, the percent working increased from 59% in 1992 to 78% in 2001.

The main changes in welfare rules under waivers and PRWORA relevant to our exercise can be grouped into four categories: termination and work requirement time limits, earnings disregards, childcare assistance and child support enforcement. Under PRWORA, States operate their own programs, so there is great heterogeneity in the rule changes they have adopted. Thus, we construct an extensive set of State/time-specific welfare rule variables, and use these as instruments in the estimation of the cognitive ability production function. We get leverage for identification by (i) comparing outcomes for children born before 1990 vs. those born later, as welfare waivers and TANF only impacted mothers in the later period, and (ii) comparing outcomes across States with different rules (in pre- or post-reform periods). We also use local demand condition instruments that have good explanatory power for behavior of single mothers over the whole sample period.

An important technical problem arises as the welfare rules are very complex. Thus, we need many variables to characterize them. As a result, we face a "many instrument problem." That is, 2SLS estimates can be severely biased (towards OLS) when the number of over-identifying instruments is large. This issue has received a great deal of attention lately - see, e.g., Stock and Yogo (2004), Anderson, Kunitomo and Matsushita (2005), Hansen, Hausman and Newey (2006), Andrews and Stock (2006). We deal with the many instrument problem in two ways: First, as Hansen et al note, the LIML and Fuller estimators correct the 2SLS bias in the many instrument case, so we use LIML. Our 2SLS estimates fall in between LIML and OLS, suggesting 2SLS does indeed suffer from severe bias. Stock-Yogo provide a test for whether the many instruments problem induces biases (in estimates or test sizes) that are large relative to the OLS bias. The Stock-Yogo test suggests a many instrument problem for 2SLS, but gives no evidence of a problem for LIML.

Second, we also experiment with using factor analysis to reduce the dimension of the instrument set. To our knowledge, this seems to be a novel approach in the current literature on the many instruments problem. Using only the most important factors as instruments, we find that the 2SLS and LIML results are very similar. And both are very similar to LIML results using the full instrument set. This gives us added confidence in the LIML results based on the full instrument set.

The main results indicate that the effect of child care use on children's achievement is negative, significant and rather sizeable. Our estimates imply that one additional year of child care use reduces cognitive ability test scores by approximately 2.9%. This corresponds to 0.156 standard deviations, so it is a substantial effect. This result is quite robust, in that it differs only modestly across a wide range of production function specifications, instrument sets, and samples.

However, this general finding masks important differences across child care types, maternal education, and child ages. Formal child care (i.e., center based care) has no adverse effect on cognitive outcomes. Only informal care (i.e., non-center based care by grandparents, siblings, other relatives, non-relatives) has significant adverse effects. We estimate that an additional year of informal care causes a 3.5% reduction in test scores. Our overall negative estimate of the effect of child care obtains because 75% of single mothers use informal care arrangements. Also, we only find significant negative effects of child care for children who are least 2 years old. And, as expected, the value of mother's time relative to day care is greater for more educated mothers.

Finally, it is interesting to examine how the welfare policy changes of the mid-90s affected test scores of children of single mothers. Reduced form estimates (i.e., substitute the welfare rules for the endogenous variables in the outcome equation) imply test scores were modestly reduced.

This paper is organized as follows. In section 2 we review the relevant literature. In Section 3 we describe in detail the welfare policy and local demand condition variables that we use as instruments to identify effects of child care and other endogenous inputs on child outcomes. Section 4 presents a theoretical framework for interpreting the estimates. In Section 5 we discuss the data and the sample used in this paper. Section 6 presents the estimation results and Section 7 concludes.

2. Literature Review

2.1. The Effect of Maternal Employment and Child Care on Children's Cognitive Outcomes

Many prior studies, mostly in the developmental psychology literature, have used the NLSY to assess effects of maternal employment and child care use on child cognitive development. Recent reviews of this literature include Love et al (1996), Blau (1999a), Lamb (1996), Haveman and Wolf (1994), Ruhm (2000) and Blau and Currie (2004). Less than half of these studies provide results that are interpretable in terms of effects of specific inputs.[3] Most present simple correlations between inputs and child outcomes and do not control for family and/or child characteristics. Furthermore, some of these studies use small samples, often nonrandomly selected. In most cases, no control for self-selection of children into child care arrangements (group of working mothers) is implemented.[4]

Table 1 summarizes recent papers that use the NLSY data to assess effects of maternal employment on child cognitive outcomes.[5] Clearly the evidence is inconclusive. Approximately a third of the studies report positive effects of maternal employment, a third report negative effects and the rest report effects that are either insignificant or that vary by the group studied or the timing of inputs. A similar picture is seen in Table 2, which summarizes recent papers that evaluate the effects of daycare (and/or daycare quality) on child outcomes.[6] Again, effects range from positive to negative, are often insignificant, and vary by group.

Reasons for this diversity of results may include the wide range of specifications that are estimated, and that many studies fail to control for endogeneity of employment and child care. To make our exposition of the literature more clear, it is useful to have a specific framework in mind. Consider the following equation, interpretable as a cognitive ability production function:

$$\ln S_{ijt} = \alpha_1 T_{ijt} + \alpha_2 C_{ijt} + \alpha_3 G_{ijt} + \alpha_4 X_{ijt} + \mu_j + \delta_{ij} + \varepsilon_{ijt} \tag{1}$$

Where S_{ijt} is a cognitive outcome (i.e., test score) for child i of mother j at age t. The log is typically taken as test scores are positive. T_{ijt} is a measure of the maternal time input up through age t. This might be a scalar, as in a cumulative specification, or a specification where only average or current inputs matter, or a vector, if inputs at different ages are allowed to have different effects. Similarly, C_{ijt} is a measure of nonmaternal time inputs (i.e., child care), and G_{ijt} represents goods and services used in the production of child's ability. Next, X_{ijt} is a set of controls for the child's initial skill endowment. This may include variables such as mother's age, education, AFQT score, etc. (meant to capture the inherited ability endowment), and initial characteristics of the child such as gender, race and birthweight. Turning to the error components, μ_j and δ_{ij} are family and child effects, which capture parts of the *unobserved* skill endowment of the child. Finally, ε_{ijt} is a transitory error term that may be interpreted as measurement error inherent in the test plus error in recording test results.

---

[3] Some studies show associations between clusters of child care arrangements/attributes and child development instead of assessing the impact of each input (Howes and Rubenstein (1985), Peterson and Peterson (1986), Studer (1992)). And in some cases, coefficient estimates or signs are not provided by authors (e.g., Howes and Rubenstein (1981)).

[4] See for example, Burchinal et al. (1995) and Parcel and Menaghan (1990).

[5] Todd and Wolpin (2003), Rosenzweig and Wolpin (1994) and Rosenzweig and Schultz (1983) discuss estimation of cognitive ability production functions *in general*. We summarize only studies of parental time and child care inputs.

[6] Since the literature contains fewer studies of day care, Table 2 is not restricted to studies that use NLSY data only.

While this general setup seems to underlie, at least implicitly, most papers in the literature, none actually estimate equation (1), and many estimate equations that seem quite far from it. One fundamental problem is that the maternal time input T and the goods inputs G are not directly observed. Most papers ignore the problem that T is unobserved, simply using maternal employment or child care use in place of maternal time.[7]

Similarly, most papers simply ignore G, while a few proxy for it using household income or the NLSY's "HOME" environment index. The latter is problematic because it is based not just on goods inputs (e.g., books in the home) but also on maternal time inputs (e.g., time spent reading to the child). Baydar and Brooks-Gunn (1991) estimate effects of both maternal employment and child care arrangements, but do not include goods/services. Desai et al. (1989) use maternal employment to proxy for T, average number of child care arrangements during the first three years after childbirth to proxy for C and household income to proxy for G. But, as noted by Todd and Wolpin (2003), it is difficult to interpret production function estimates when proxies are used for key inputs. To our knowledge, only James-Burdumy (2005) discusses the relationship between her estimating equation and a child ability production function by pointing out the difficulty of interpreting estimates when proxies are used for maternal time and goods inputs. We discuss this issue in detail in Section 4.

Secondly, most papers in the literature have estimated specifications that include only *current* inputs. This is a strong assumption, especially in light of the tradition in the human capital literature of letting cumulative inputs matter. One could think of the effect of inputs cumulating over time or having a more general specification according to which the whole history of inputs since childbirth matters for the child's outcome at time t. Most papers do not discuss the implications of their assumptions regarding timing of inputs.[8] We also discuss this issue in Section 4.

Finally, most papers estimate equation (1) by OLS, ignoring potential endogeneity of inputs – i.e., potential correlation of maternal work and day care use decisions, and goods inputs, with the unobserved ability endowments μ_j and δ_{ij}. A few recent studies try to overcome this problem by either: (1) using a very extensive set of variables to proxy for unmeasured endowments, (2) using child or family fixed effects, or "value added" models,[9] and/or (3) using instrumental variables.

[7] Also, most papers use one or the other of these variables, and do not examine both. For example, Vandell and Ramanan (1992) estimate the effect of maternal employment on child's cognitive outcomes but do not include child care arrangements as an additional input, while Caughy et al (1994) do the reverse.
[8] Notable exceptions are Blau (1999a) and Duncan (2003). Some papers use maternal employment (or child care use) at different years after childbirth but do not discuss implications of their choice in terms of properties of the underlying production function (e.g., Waldfogel et al. (2002), Vandell and Ramanan (1992), and Baydar and Brooks-Gunn (1991)).
[9] In the value-added approach, the test score in period t (S_{ijt}) is a function of the outcome in period $t-1$ and the inputs in period t, the idea being that the lagged test score proxies for the child's ability at the start of a period.

Consider first the studies that can be classified as using extensive controls for the child's skill endowment. Among others, Han et al (2001), Baydar and Brooks-Gunn (1991), Parcel and Menaghan (1994), Vandell and Ramanan (1992) and Ruhm (2002), use an extensive set of observable characteristics of the child and the mother, including mother's AFQT score. In spite of this, the results of these papers are inconclusive. For example, Ruhm (2002) finds significant *negative* effects of maternal employment on math scores while Parcel and Menaghan (1994) report small *positive* effects of maternal employment on child's cognitive outcomes. Baydar and Brooks-Gunn (1991) find that maternal employment in the child's first year *negatively* affects cognitive outcomes, while Vandell and Ramanan (1992) find *positive* effects of early maternal employment on math achievement, and of current maternal employment on reading achievement.

Next, consider the studies that use fixed effects. Chase-Lansdale et al. (2003) use child fixed effects models to assess the effect of maternal employment on child outcomes. They analyzed 2,402 low-income families during the recent era of welfare reform. Their results suggest that mothers' transitions off welfare and into employment did not cause negative outcomes for preschoolers. They note, however, that this approach does not account for endogeneity of these transitions, and they do not attempt to use welfare rules as instruments for maternal employment as we do here.

James-Burdumy (2005) estimated household FE models using 498 sibling children in the NLSY. Her results suggest that effects of maternal employment vary depending on the particular cognitive ability assessment used and the timing of employment.[10] The use of sibling differences eliminates the mother (or household) fixed effect μ_j from (1) but does not eliminate the child fixed effect δ_{ij}. It is plausible that mothers make time compensations for children depending on their ability type. Using household fixed effects does not solve this problem, as maternal employment is then correlated with the sibling specific part of the cognitive ability endowment. In addition, the FE estimator requires that input choices are unresponsive to prior sibling outcomes. If inputs to child *i'* are responsive to outcomes for child *i*, then ε_{ijt} will be correlated with those inputs.

Blau (1999a) and Duncan and NICHD (2003) both study the effects of child care use and quality on child outcomes. They use similar methodologies, including a wide range of proxies for unmeasured child ability (e.g., mother's AFQT and education), controls for many aspects of the home environment, and use of fixed effects and value added specifications. The main difference is

[10] According to James-Burdumy (2005)'s fixed effects (FE) estimates in her Table 5, an increase in maternal work from 0 to 2000 hours in year 1 of the child's life reduces the PIAT math score (measured at ages 3 to 5) by (-.00117) ×2000 = -2.34 points. This is similar to the effect we estimate for one year of full-time work (-2.9%). But she finds no significant effect of maternal employment after the first year. In contrast, we find maternal time is *more* important in years 2+.

that Blau (1999a) uses the NLSY while Duncan uses the NICHD Study of Early Child Care. Blau (1999a) concludes "child care inputs … during the first three years of life have little impact on … child outcomes …" while Duncan finds modest positive effects of improved child care quality.

From our perspective, a key difficulty in interpreting the Blau and Duncan results is that their specifications don't allow one to estimate the effect of maternal time *per se*. Both studies include the HOME environment index, which contains both goods inputs, like books in the home, and also time inputs, like how often the child is read to, eats meals with parents, or talks with the mother while she does housework. Thus, the coefficients on maternal work or day care measure effects of those variables holding HOME fixed, which means holding some maternal time inputs fixed. In contrast, we are interested in the total impact of maternal time on child outcomes, including how a decline in the time input (from increased work or day care use) affects time reading to the child and so on.

Finally, Currie and Thomas (1995) use the NLSY to look specifically at how pre-school affects outcomes. Using sibling differences and extensive controls for ability endowments, they estimate a year of Head Start increases PPVT scores by roughly 7%, while other types of pre-school have no effect. The Head Start effect persists for whites, but is wiped out by age 10 for blacks.

The Blau, Duncan-NICHD and Currie-Thomas papers all contain useful discussions of the limitations of fixed effects and value added specifications. As they point out, neither approach provides a panacea for dealing with unobserved child ability, as each relies on assumptions that can be stronger than OLS. For example, the household FE estimator requires that input choices be unresponsive to the child specific part of the ability endowment. The value added model faces the problem that estimates of lagged dependent variable models are inconsistent in the presence of fixed effects like μ_j and δ_{ij}.[11] Neither approach, nor child fixed effects, deals with endogeneity arising because current inputs may respond to lagged test score realizations. An IV approach is necessary to deal with these problems.

To our knowledge, only two prior papers have attempted to use IV in this context. These are Blau and Grossberg (1992) and James-Burdumy (2005).[12] Both look at effects of maternal work on child outcomes, and do not examine effects of day care use *per se*. More importantly, both papers

[11] Estimation of a first-differenced version of the value-added specification would eliminate the fixed effects, but Blau (1999a) points out this is difficult or impossible due to limitations of existing data. This would require three outcome observations and two lagged input observations. Even if feasible, this approach would entail a severe efficiency loss.

[12] James-Burdumy's preferred specification uses sibling differences to control for household FE, and does not use IV. However, she notes that maternal employment may be endogenous in the differenced equation, due to correlation of the time-varying parts of the errors in the child outcome and maternal employment equations. Of course, another source of endogeneity is correlation between unobserved child ability and the mother's decisions about work and day care.

suffer from the problem that the instruments are extremely weak. As a result, the standard errors on the maternal work variables in their 2SLS regression are so large that no plausibly sized effect could possibly be significant (i.e., in each case, to attain 5% significance, maternal work over a three year period would have to change a child's tests scores by roughly 50 points or 3 standard deviations).[13] Thus, we would argue that their attempts to implement IV were not successful. Similarly, Currie and Thomas (1995) report they attempted to use IV but could not find sufficiently powerful instruments.

The main advantage of our approach is that the welfare policy and local demand instruments that we employ are much stronger. Indeed, the first stage marginal R^2 values we obtain using these instruments (i.e., about .09) are fairly large relative to what one often sees in the IV literature, and, in the second stage, the standard error on day care does not "explode" when these instruments are used.

Bernal (2006) takes a different approach by estimating a structural model of work and child care choices of *married* women. She estimates a child cognitive ability production function – which includes mother's work and child care use as inputs – jointly with the mother's work and child care decision rules, thus implementing a selection correction. Her results suggest rather sizable effects of maternal employment and child care use on child cognitive ability. In particular, one full year of maternal work and child care use causes a 1.8% reduction in test scores of children ages 3 - 7.[14]

It is interesting to extend this work to *single* mothers for several reasons. The first is to see if results generalize. Second, single mothers are of special policy relevance, as welfare reform led to large increases in their work/day care use. Third, welfare rules have large effects on work/day care use by single mothers, so as instruments they provide a strong basis for identification. It is difficult to find plausibly exogenous variables that impact behavior of married women so strongly.

2.2 Relationship between Test Scores and Subsequent Outcomes (Wages, Education, etc.)

Several studies have examined the relationship between test scores during childhood and subsequent outcomes like educational attainment and wages. While causality is difficult to ascertain, this research has shown that children's cognitive achievements are strong *predictors* of a variety of outcomes later in life. This highlights the importance of understanding what determines ability of individuals at early stages of life, particularly for the design of public policy aimed at improving labor market outcomes. We summarize some of these studies in this section.

---

[13] For Blau and Grossberg (1992), who use work experience prior to childbirth to instrument for maternal employment, compare columns 1 and 2 of their Table 2. For James-Burdumy (2005), who uses the percentage of the county labor force employed in services to instrument for maternal employment, compare columns FE and IV-FE from her Table 3.

[14] Liu et al. (2003) also adopt a structural approach to estimate effects of maternal employment and school inputs on test score outcomes for 5 to 15 year olds in the NLSY. They also find a negative effect of maternal employment on child outcomes. Obviously, the focus in Bernal (2006) and here is rather different, as we are interested in pre-school inputs.

First, consider studies using U.S. data. In the NLSY, Neal and Johnson (1996) find that scores at ages 14 to 21 on the Armed Forces Qualifying Test (AFQT), an IQ-type measure, are highly significant predictors of wages at ages 26 to 29. Murnane, Willett and Levy (1995) use two longitudinal surveys of high school seniors to document a strong relationship between their math test scores and wages at age 24. Zax and Rees (1998) use the Wisconsin Longitudinal Study (WLS) to document that age 17 IQ is a strong predictor of wages at ages 35 and 53.

Additional studies use the British National Child Development Study (NCDS). Hutchinson, Prosser and Wedge (1979) use the NCDS to link test scores at age 7 with scores at age 16. Similarly, Connolly, Micklewright and Nickell (1992) find a significant positive relationship between test scores at age 7 and earnings at age 23 (in a sample of young men who left school at age 16). More recently, Robertson and Symons (1996) and Harmon and Walker (1998) find a positive association between age 7 test scores and earnings at age 33. And Currie and Thomas (2001) show that a one standard deviation increase in age 16 math scores is associated with a 14% higher wage rate and a 7% higher employment rate at age 33 (for low or medium-SES individuals). In addition, they provide evidence that age 7 (math) test scores are strong predictors of age-16 math test scores.

From our perspective, a limitation of these studies is they all look at test scores at age 7 or older (14 or older in the U.S. case). Do tests scores measured at earlier ages also predict subsequent achievement? In Appendix 1 we present evidence from the NLSY that the PPVT at age 4 and PIAT reading and math scores at ages 5 and 6 are significantly correlated with educational attainment of youth who are at least 18 years old. For example, consider a one-point increase in the math score at age 6 (i.e., roughly a 1% increase, as the mean score is 99.7). Holding parental background variables like mother's education fixed, this is associated with increased educational attainment (measured at age 18 or later) of approximately .019 years. Similarly, a one-point (or roughly 1%) increase in the reading score at age 6 is associated with an increase in highest grade completed of approximately .025 years. These estimated impacts are fairly substantial, since our estimates imply that a year of full-time maternal work combined with informal day care use reduces test scores by roughly 3.5%. This translates into an effect on completed schooling of roughly .067 to .088 years, a large effect.[15]

A striking aspect of the Appendix 1 results is that mother's AFQT score is <u>not</u> a significant predictor of completed education. Thus child test scores, even at ages 4-6, are better predictors of later outcomes than mother's scores.

[15] The following back-of-the-envelop calculation helps put these figures in perspective: Say people are of two types, those who finish high school (12 years) and those who finish college (16 years), and that 20% finish college. To increase average completed schooling by .08 years, the percentage finishing college must increase to 22%, a 10% increase.

3. Construction of Instruments using Welfare Rules and other Policy Variables

To deal with endogeneity of maternal work/child care (see Section 2.1), we propose using welfare policy rules as instruments to estimate cognitive ability production functions for children of single mothers. Welfare rules are known to have a large impact on labor supply of single mothers (see Moffitt (1992)). To construct our instruments, we collect detailed information on State welfare policies from many sources. Bernal and Keane (2007) describe these sources, and construction of the instruments, in detail. Here, in the interest of space, we only briefly highlight the key aspects of Section 1115 welfare waivers and the 1996 Welfare Reform that are relevant to this work. Table 3 presents the complete instrument list, including all the policy variables. Each instrument has up to three subscripts: *i* for individual, *s* for State and *t* for period (quarter in our case).

3.1. Benefit Termination Time Limits

Under AFDC, single mothers with children under 18 were *entitled* to receive benefits if they met income and asset eligibility requirements. But under Section 1115 Waivers and TANF, States can set time limits on benefits. Indeed, PWRORA forbids States from using federal funds to provide benefits to adults beyond a 60-month lifetime time limit, and it allows states to set shorter limits. For instance, California sets a 5-year limit, and Texas and Florida set time limits in the 2-3 year range.

Time limits have direct and indirect effects. The direct effect is clear (i.e., when a woman hits the time limit she becomes ineligible). The indirect effect arises if women are forward-looking and "bank" months of eligibility for later use. We use <u>eight</u> variables to capture both effects of time limits (see Table 3). They measure time limits created under both TANF and AFDC waivers. We include, for example, a dummy for whether a single mother's State of residence had imposed a time limit (TLI_{st}) in time *t*, a dummy indicating if the time limit could be binding for a particular woman (TL_HIT_{ist}), and her maximum potential remaining time of eligibility ($REMAIN\_TL\_ELIG_{\mathrm{ist}}$).

It is worth emphasizing that when possible we construct instruments that are person specific. For example, consider TL_HIT_{ist}. Say a woman resides in a State that had imposed a 5-year time limit 6 years earlier. Then it is possible that she could have hit the limit, provided her oldest child is at least 5. If not, she could not have participated in AFDC/TANF for 5 years, and therefore could not have hit the limit. Crucially, however, we do not use a woman's actual welfare participation history to determine if she had hit a time limit, because the actual history is endogenous. Our individual specific instruments are functions of policy rules and demographics alone. Regarding demographics, we assume ages of a woman and her children are exogenous (conditional on age controls in the main equation). However, as we discuss later, we treat fertility (number of children) as endogenous.

3.2. Work Requirement Time Limits and Work Requirement Exemptions

Work requirements increase time and utility costs of receiving welfare. Under PRWORA, recipients must participate in "work activities" within two years in order to continue receiving TANF benefits. But many States adopted shorter work requirement time limits. Due to variation in when States implemented TANF, and in the length of their work requirement clocks, there is substantial cross-State variation in how early single mothers could have been subject to binding work requirements. Also, States may exempt single parents with children up to 1 year old from work requirements, and may provide exemptions to other families. Thus, within a State, there is variation across women in whether work requirements can be binding, based on age of the youngest child.

We constructed a total of <u>nine</u> variables, listed in Table 3, meant to capture these various effects. For example, WR_HIT_{ist}, is an indicator for whether the woman could have been subject to work requirements (based on the length of the work requirement, time elapsed since the requirement had been implemented, age of her oldest and youngest child, etc.), and $CHILD\_EXEM_{st}$ is a dummy variable that indicates whether state s has an age of youngest child exemption in place at t.

3.3. AFDC/TANF Benefit Levels, Earnings Disregards and Benefit Reduction Rates

AFDC/TANF benefits are, roughly speaking, determined by a formula where a State specific grant level, which is increasing in number of children under 18, is reduced by some percentage of the recipient's income. Two variables we use to characterize the system are the maximum *potential* real monthly AFDC/TANF benefit for families with one and two children ($BEN(1)_{st}$ and $BEN(2)_{st}$), assuming zero earnings, in the mother's State of residence. We do not condition on a mother's <u>actual</u> family size, because we treat fertility as endogenous.[16]

Under AFDC, benefits were reduced as income increased according to a "benefit reduction rate" (BRR) that changed several times over the history of the program. Under waivers and the TANF program, the BRR was made State specific, and it now varies considerably across States.

AFDC also incorporated "earnings disregards" to encourage work among participants. That is, if a recipient started work, then for a period of time, a fraction of her earnings was not subject to the BRR. Generally, the disregard consisted of a "flat" component (e.g., the first $30 of monthly earnings) and a "percentage" part (e.g., one-third of earnings beyond the flat part). Both would be eliminated after a certain number of months of work. Starting in late 1992, many states obtained waivers to increase income disregards. Under PRWORA, States are not required to adopt any

[16] Benefits are put in real terms using a region-specific CPI. Since 1980, the BLS computed the CPI for 24 metropolitan areas. The potential benefits of individuals in other areas were deflated using a region-specific (western, south, midwest and northeast) CPI.

particular disregards, so a great deal of State heterogeneity has emerged. A few States expanded disregards and allowed them to apply indefinitely. We code the BRR and the percentage disregard together in the variable $PERC\_DISREGARD_{st}$. Flat disregards are coded in $FLAT\_DISREGARD_{st}$.

3.4. Child Support Enforcement

Child support is an important source of income for single mothers, despite widespread non-payment by non-custodial fathers.[17] The Child Support Enforcement (CSE) program, enacted in 1975, has implemented programs to locate absent parents and establish paternity. CSE expenditures increased significantly from $2.9 billion in 1996 to $5.1 billion in 2002. These expenditures are an important indication of how likely a single woman is to collect child support. We include a measure of State level CSE activity by taking the State CSE expenditure and dividing it by the State population of single mothers ($ENFORCE_{st}$).

3.5. Child Care Subsidies and the Child Care Development Fund (CCDF)

The CCDF is a block grant to states to provide subsidized child care programs for low-income families, including those who are not current or former cash assistance recipients. Under the CCDF, states have autonomy to design child care assistance programs for low-income families and a great deal of heterogeneity has emerged across States in the design of their programs. As an additional policy instrument, we use the State CCDF expenditure per single mother ($CHILDCARE_{st}$). This variable measures the availability and generosity of child care subsidies in a State.[18]

3.6. Other Instruments: Earned Income Tax Credit (EITC) and Local Demand Conditions

The EITC, enacted in 1975, is a refundable Federal income tax credit that supplements wages for low-income working families. It was originally a minor program, but a major expansion occurred in 1994-96, after which EITC became a sizable wage subsidy to low and moderate-income families. Thus, EITC may provide an important work incentive.[19] The EITC subsidy rate varies by family size. Thus, we use as instruments the EITC subsidy rates for families with one and 2+ children ($EITC(1)_{st}$ and $EITC(2)_{st}$ respectively), using Federal and State EITC rules. As with benefit levels (see Section 3.3) we do not condition on actual family size, which we view as endogenous.

Finally, we use two variables that measure local demand conditions as instruments: the State unemployment rate and the 20th percentile wage rate in the woman's State of residence at time *t*.

---

[17] In 2002, child support accounted for approximately 6.5% of single mother's real incomes (March CPS).

[18] An alternative way to measure generosity of State child care programs would be to use detailed program parameters, such as monthly income eligibility criteria, reimbursement rate ceilings or the co-payment rates, which are State specific and have also varied over time. We opt not to use these measures due to problems associated with rationing.

[19] For example, in 2003, the phase-in and phase-out rates for a family with *one child* were 34% and 15.98%, respectively. As of 2003, 17 States had enacted State earned income tax credits that supplement the federal credit.

4. The Child's Cognitive Ability Production Function

In the human capital production framework (see Ben-Porath (1967)) current and past inputs interact with an individual's genetic ability endowment to generate human capital. Leibowitz (1974) first used this framework to examine how investments in children add to preschool stocks of human capital. The acquisition of preschool human capital is analogous to the acquisition of human capital through schooling or on-the-job training, except that, at preschool ages, inputs are generated by *joint* parental/child decisions (e.g., child tastes presumably affect parent input choices), not by choices of the child alone. Here, we focus on the cognitive ability component of human capital.

Let A_{it} be child i's cognitive ability t periods after birth. We write a production function:

(2) $$\ln A_{it} = A(\tilde{T}_{it}, \tilde{G}_{it}, \tilde{C}_{it}, \omega_i)$$

where $\tilde{T}_{it}$, $\tilde{G}_{it}$ and $\tilde{C}_{it}$ are vectors of period-by-period inputs of maternal time, goods and child care time, respectively, up through period t, and ω_i is the child's ability endowment. Goods inputs may include nutrition, books and toys that enhance cognitive development, etc.. Day-care inputs capture contributions of alternative care providers' time to child cognitive development. These may be more or less effective than mother's own time. Also, care in a group setting may contribute to child development by stimulating interaction with other children, learning activities at pre-school, etc..

Several difficult issues arise in estimation of (2). First, estimation of a completely general specification, where inputs may have a different effect at each age t, and where the endowment ω_i may differentially affect ability at each age, is infeasible due to proliferation of parameters.[20] Thus, we obviously need to restrict how inputs enter (2).

One simplification, familiar from the human capital literature, is to assume that: (i) only cumulative inputs matter, rather than their timing, and (ii) the effect of the permanent unobservable is constant over time (e.g., in a Mincer earnings function, only cumulative education and experience affect human capital, and the unobserved skill endowment has a constant effect). We first consider a specialization of (2) that adopts these assumptions, and consider some feasible relaxations later. Letting $\widehat{X}_{it} = \sum_{\tau=1,t} X_{it}$ be the cumulative amount of input X up through time t, and assuming that cumulative inputs affect $\ln A_{it}$ linearly, we obtain a special case of (2) that takes the form:

(3) $$\ln A_{it} = \alpha_0 + \alpha_1 \widehat{T}_{it} + \alpha_2 \widehat{C}_{it} + \alpha_3 \ln \widehat{G}_{it} + \omega_i$$

[20] For instance, if the effect of just one input (say, maternal time) is allowed to differ between every pair of input and output periods t and t', and we examine outcomes for 20 quarters after birth, we obtain 20·21/2=210 parameters describing effects of that input alone on ability.

We now consider problems of estimating the production function in the special case of (3).[21]

The second issue we face is the selection (or endogeneity) problem that arises because inputs may be correlated with the child ability endowment ω_i. Simply to clarify this problem (but not to develop our specific estimating equation), assume the ability endowment is given by the equation:

$$(4) \qquad \omega_i = \beta_0 + \beta_1 E_i + \hat{\omega}_i ,$$

where E_i is a vector of mother characteristics, like education, that are correlated with child ability, and $\hat{\omega}_i$ is the part of the ability endowment that is mean independent of mother characteristics. Next, as an illustration, assume a mother's decision rule for child care time at time t, C_{it}, can be written:

$$(5) \qquad C_{it} = \pi_0 + \pi_1 E_i + \pi_2 \hat{\omega}_i + \pi_3 cc + \pi_4 R_{it} + \varepsilon_{it}^c ,$$

where cc is the price of day care,[22] R_{it} is a set of welfare rules facing the mother at time t, and ε_{it}^c is a stochastic term subsuming tastes for child care use (both permanent and transitory taste shocks), and shocks to child care availability and the mother's offered wage rate. The presence of $\hat{\omega}_i$ in the decision rule means $\widehat{C}_{it}$ is endogenous in (3), and we will require instruments that affect C_{it} yet are uncorrelated with $\hat{\omega}_i$ and ε_{it}^c. Below we argue the welfare rules R_{it} can plausibly play this role.

The third key issue in estimating (3) is measurement of maternal time and goods inputs. One can imagine a model where mothers decide how much "quality" time to devote to the child while at home (e.g., children's time is divided between day-care, "quality" time with the mother, and time spent watching TV while she does housework). But, we don't observe actual contact time between mothers and children (let alone how much is "quality" time), so we simply side-step the issue by assuming that $T_{it} = T\text{-}C_{it}$, where T is total time in a period. Thus, we distinguish between only two types of time (i.e., time with the mother and time in child-care). Then, we can rewrite (3) as:

$$(6) \qquad \ln A_{it} = \alpha_0 + (\alpha_1 T)\cdot t + (\alpha_2 - \alpha_1)\widehat{C}_{it} + \alpha_3 \ln \widehat{G}_{it} + \omega_i$$

Thus, we can only estimate $\alpha_2 - \alpha_1$, the effect of time in child-care *relative* to that of mother's time.

An issue we abstract from here is that maternal work time may influence how much of $T\text{-}C_{it}$ is "quality time." For example, a mother who uses child care but does not work might devote more of $T\text{-}C_{it}$ to "quality time." Thus, maternal work time might enter the production function directly,

[21] Letting cumulative goods enter in log form is analytically convenient, for reasons that will become apparent later.
[22] That the price of child-care cc is assumed constant over mothers/time is not an accident. A key problem confronting the literature on child-care is that the geographic variation in cc seems too modest to use it as an IV for child-care usage.

independently of how it affects the goods input (via the budget constraint) or how it affects C_{it}. However, for single mothers it is very difficult to address this issue, because child care and maternal work time are extremely highly correlated (ρ=.94).[23] Thus, attempts to include both in the model fail due to severe colinearity. But we make some attempt to address this issue in section 6.6.

The fourth key issue in estimation of (3) is that goods inputs G_{it} are largely unobserved. For example, the NLSY contains information on books in the home, but not nutrition, health care, tutors, recreation, etc.. To deal with this, consider a decision rule for the cumulative goods input into the child's ability (conditional on work, income and child-care usage decisions) given by:

$$\ln \widehat{G}_{it} = \gamma_0 + \gamma_1 E_i + \gamma_2 \hat{\omega}_i + \gamma_3 \widehat{C}_{it} + \gamma_4 \ln \widehat{I}_{it}(W, H; R) + \gamma_5 t + \varepsilon_{it}^g \; . \tag{7}$$

This is a conditional decision rule, obtained as the second stage of an optimization process, where, in stage one, a mother chooses childcare time C and hours of market work H. The notation $\widehat{I}_{it}(W, H; R)$ highlights the dependence of income on wages, hours of market work, and the welfare rules R that determine how benefits depend on income. Equation (7) can be thought of as a linear approximation to a more complex decision rule generated by a dynamic model. The key thing captured by (7) is that a mother's decisions about goods inputs into child development may be influenced by (i.e., made jointly with) her decisions about hours of market work and child-care. (7) also captures the notion that per-period inputs depend on the mother's characteristics E (which determine her human capital), and the child's ability endowment $\hat{\omega}_i$.[24] The time trend in (7) captures growth of cumulative inputs over time. The stochastic term, ε_{it}^g, captures the mother's idiosyncratic tastes for investment in the form of goods.[25] Now, substituting (7) into (6) we obtain:

$$\begin{aligned} \ln A_{it} &= \alpha_0 + (\alpha_1 T)\cdot t + (\alpha_2 - \alpha_1)\widehat{C}_{it} \\ &\qquad + \alpha_3[\gamma_0 + \gamma_1 E_i + \gamma_2 \hat{\omega}_i + \gamma_3 \widehat{C}_{it} + \gamma_4 \ln \widehat{I}_{it} + \gamma_5 t + \varepsilon_{it}^g] + \omega_i \\ &= (\alpha_0 + \alpha_3\gamma_0) + (\alpha_1 T + \alpha_3\gamma_5)\cdot t + (\alpha_2 - \alpha_1 + \alpha_3\gamma_3)\widehat{C}_{it} \\ &\qquad + \alpha_3\gamma_4 \ln \widehat{I}_{it} + \alpha_3\gamma_1 E_i + (1+\alpha_3\gamma_2)\hat{\omega}_i + \alpha_3 \varepsilon_{it}^g \\ &= \varphi_0 + \varphi_1 \cdot t + \varphi_2 \widehat{C}_{it} + \varphi_3 \ln \widehat{I}_{it} + \varphi_4 E_i + \hat{\hat{\omega}}_i + \hat{\varepsilon}_{it}^g \end{aligned} \tag{8}$$

[23] Obviously, single mothers must use day care to work, and most cannot afford day care otherwise. In contrast, for married women, use of day care while not working is fairly common (see Bernal (2006)).
[24] Note that the child's ability endowment may matter for two reasons: Either mother's may choose good inputs based on the child's ability (e.g., they may buy educational toys to compensate a child who is having certain learning problems) or because child ability affects the types of inputs a child demands (e.g., a high ability child may request more books).
[25] This would arise due to heterogeneous preferences for child quality. ε_{it}^g may also be influenced by the child's tastes.

Equation (8) is estimable, because all the independent variables are observable. However, we must be careful about the appropriate estimation method and interpretation of the estimates. As we have noted, child care utilization may be correlated with the unobserved part of the child ability endowment $\hat{\omega}_i$. Furthermore, child care use may also be correlated with $\hat{\varepsilon}_{it}^{g}$, the unobserved taste shifter in equation (7), if tastes for child care usage ε_{it}^{c} in (5) are correlated with tastes for goods investment in children, as seems plausible.[26] Thus, estimation of (8) using OLS is not appropriate. To our knowledge, it has not been previously noted that consistent estimation of an equation like (8) requires instruments that are not only uncorrelated with the unobserved part of the child's skill endowment, $\hat{\omega}_i$, but also with the mother's tastes for goods investment in the child, ε_{it}^{c}. In order for the welfare rule parameters R_{it} to be valid instruments for cumulative child care in estimating (8), they must be uncorrelated with these two error components. This seems like a plausible exogeneity assumption.[27] We would make a similar argument for local demand conditions.

The cumulative income variable in (8) is also potentially endogenous, for multiple reasons. First, income depends on the jointly made child care use and work decisions. Hence it is potentially correlated with child ability for the same reasons as were operative for child care usage. Second, income depends on the mother's wage rate, which depends on her ability endowment. To the extent that this ability endowment is not perfectly captured by mother's education, and the residual part is correlated with the child ability endowment, this will also generate correlation between the mother's income and $\hat{\omega}_i$. Thus, we need to instrument for mother's income as well. Again, we will argue that welfare rules R_{it} and local demand conditions provide plausibly valid instruments, as they have important effects on work decisions, yet are plausibly uncorrelated with child ability endowments.

Assuming that instrumental variables provides consistent estimates of (8), it is important to recognize that the child care "effect" that is estimated is $\beta_2 = \alpha_2 - \alpha_1 + \alpha_3 \cdot \gamma_3$. This is the effect of child care time (α_2) *relative* to the effect of mother's time (α_1) plus the effect of any change in goods inputs that the mother may choose as a result of using day care ($\alpha_3 \cdot \gamma_3$). In light of this, it is important to understand the limitations of IV estimates of (8). For instance, such estimates cannot tell us how a

[26] For instance, a mother with a high taste for child quality may both spend more time with the child (i.e., use less day care) and invest more in the child in the form of goods. This would tend to bias estimated effects of day care usage in a negative direction, since not only the maternal time input but also the goods input is lower for children in day care.

[27] In Appendix 5 we report means of child test scores prior to 1990 by State, broken down by whether the State subsequently implemented welfare waivers (i.e., moved towards Welfare reform early), and whether the State implemented strict or lenient welfare rules after 1996. There is no significant difference in average pre-reform test scores between "strict" and "lenient" States. And, in fact, the mean pre-reform test score were higher in States that failed to adopt waivers or set longer time limits. This is opposite to the direction of bias one would worry about in our results.

policy like child care subsidies would affect child cognitive ability outcomes. Such subsidies would not only alter day care use, but also goods inputs, and in a way not captured by $\alpha_3 \cdot \gamma_3$. The problem arises because, while α_1, α_2, and α_3 are structural parameters of the production technology (3), the parameter γ_3 comes from the decision rule for goods inputs (7), which is not policy invariant.

Thus, when interpreting estimated effects of child care use on child cognitive outcomes, one must be careful to view them as applying only to policy experiments that do not alter the decision rule for goods in investment in children (7). As this decision rule is conditional on work, income and child-care usage decisions, it will be invariant to policies that leave the budget constraint conditional on those decisions unchanged. A work requirement that induces a woman to work and use child care, but that leaves her wage rate and the cost of care unaffected, would fall into this category.

Besides mother's education in (4), we use a rich set of additional controls for the child's cognitive ability endowment. Letting Z_i be a vector of mother/child characteristics that may be correlated with the child's skill endowment (e.g., mother's AFQT score, child gender), we obtain:

$$\text{(8')} \qquad \ln A_{it} = \varphi_o + \varphi_1 \cdot t + \varphi_2 \widehat{C}_{it} + \varphi_3 \ln \widehat{I}_{it} + \underline{\varphi}_4 Z_i + \hat{\hat{\omega}}_i + \hat{\varepsilon}_{it}^g$$

A detailed description of the variables included in Z_i can be found in Table 4.

Finally, note that the econometrician does not observe actual cognitive ability of children, but instead has available a set of (age adjusted) cognitive ability test scores that measure ability with error. Let S_{it} be the test score observed in period t and let measurement error be specified as:

$$\text{(9)} \qquad \ln S_{it} = \ln A_{it} + \eta_1 d_{i1t} + \eta_2 d_{i2t} + \varepsilon_{it}$$

where d_{1t} and d_{2t} are cognitive ability test dummies, which capture the mean differences in scores across the three tests we use (PPVT, PIAT-math, reading),[28] and ε_{it} is measurement error.

By substituting (9) into (8') we obtain:

$$\text{(10)} \qquad \ln S_{it} = \varphi_o + \varphi_1 \cdot t + \varphi_2 \widehat{C}_{it} + \varphi_3 \ln \widehat{I}_{it} + \underline{\varphi}_4 Z_i + \eta_1 d_{i1t} + \eta_2 d_{i2t} + \upsilon_{it}$$

where $\upsilon_{it} = \hat{\hat{\omega}}_i + \hat{\varepsilon}_{it}^g + \varepsilon_{it}$. Equation (10) is the baseline specification that we estimate.

While we have considered particular functional forms in order to clarify estimation issues, in our empirical work we consider many generalizations and alternative formulations of (10). For instance, the effect of income differences may cumulate over time, or it may be current income that matters. Similarly, we do not know *a priori* whether it is the cumulative or current day care usage

[28] In particular, d_{1t}=1 if S_t corresponds to the Peabody Picture Vocabulary Test (PPVT) and d_{2t}=1 if S_t corresponds to the Peabody Individual Achievement Test-Math Section (PIAT-math). The PIAT-reading test is the base case.

that matters for current test scores. For this reason, we estimate different specifications and let the data tell us whether cumulative or current inputs matter most for children's development.

In addition, we estimate models that allow for heterogeneous treatment effects in the form of interactions between child care use and observed characteristics of the mother (such as education and AFQT score). This captures the notion that the effect of home inputs on child's cognitive ability might vary depending on the type of mother. We also test for differences in the effect of separation from the mother depending on characteristics of the alternative child care provider (i.e., formal vs. informal). And we allow the effect of mother's time to differ by age of the child.

5. Data

5.1. Individual Work and Child Care Histories and Construction of the Sample

We use data from the National Longitudinal Survey of Youth 1979 youth cohort (NLSY79). The data contain 12,686 individuals, approximately half of them women, who were 14-21 years of age as of Jan. 1, 1979. The NLSY79 consists of a core random sample and oversamples of blacks, Hispanics, poor whites and the military. Interviews have been conducted annually since 1979. In 1986 a separate survey of all children born to NLSY79 female respondents began. The child survey includes a battery of child cognitive, socio-emotional, and psychological well-being questions that are administered biennially, including the tests that we use in our analysis (see Sect. 5.2).

Using the NLSY79 work history file, we construct a detailed employment history for each mother in the sample for the period surrounding the birth of each child, up to four quarters before birth and 20 quarters after birth (for a period of five years). We use the geocode data to identify the State of residence of each mother in order to construct State specific welfare rule parameters.

For child care, retrospective data were gathered during 1986, 1988, 1992, and 1994-2000 that allow us to construct complete quarterly child care histories for the first three years of a child's life. In addition, data on whether the mother used child care or not during the 4 weeks prior to the interview date are available for the 1982-86, 1988, 1992 and 1994-2000 survey years. This allows us to construct partial histories of child care for the fourth and fifth years after birth.

We use the sample of single mothers in the NLSY to estimate the child's ability production function. We focus on single mothers because their labor supply behavior and child care utilization decisions were greatly affected by the welfare and other policy changes that occurred during our sample period. For instance, the percent of single mothers with children aged 0-5 who work increased from 59% in 1992 to 78% in 2001 (see Fang and Keane (2004)).

Thus, we require that women in our sample be single (or not cohabitating with a male) during five years following the birth of the child, and that we observe at least one test score for the child. There are 1,464 mother/child pairs in the NLSY79 who satisfy these criteria, and they had a total of 3,787 test score observations (an average of 2.59 per child).

In our sample, 251 women had children between 1990 and 2000, so waivers and TANF impacted their labor supply behavior before the children reached school age. Much of our leverage for identification comes from comparing outcomes for these children to those of the 1,213 children born too early for their mothers to be impacted by welfare reform before they reached age 5. However, it is important to note that even in the pre-reform period some of our instruments, like AFDC grant levels and local demand conditions, varied greatly across States and over time, also providing an important source of identification. And, in the post-reform period, we also get leverage for identification by comparing children in States with "strict" vs. "lenient" welfare rules.

Table 5 compares the single mothers in our sample with all single mothers, as well as all mothers, in the NLSY79. Does using only women who remain single for 5 years after childbirth lead to a very select sample? As we see in Table 5, the single mothers in our sample are very similar to the set of all single mothers in the NLSY79. Of course, the single mothers in our sample differ in significant ways from typical mothers (married or single). They are younger by 1.7 years, less educated by 0.8 years, and have a low wage rate ($5.08 in 1983$). They are more likely to be Hispanic or black, and less likely to work during the first year after childbirth (39% vs. 47%).

Figure 1 displays employment and child care choices for 5 years after birth for women in our sample. During the first quarter after birth, about 73% of single mothers stay at home and do not use child care. The remainder use child care, with 10% working full-time, 5% part-time and 12% staying home. By the end of 16 quarters, only 38% continue to stay at home and not use child care. 29% work full-time, 17% part-time and 26% stay home and use child care.

5.2. Measuring Maternal Time and Other Inputs, and Measuring Child Cognitive Ability

Unfortunately for our purposes, the NLSY does not report the actual number of hours a child spent in child care (rather than maternal care). The child care variable is simply an indicator for whether the mother used child care for at least 10 hours per week during the last month.[29] This information is inadequate to determine whether a child was in child care full or only part-time. However, by combining the child care variable with work history information, we can make a reasonable determination about full vs. part-time care. Specifically, we use the following procedure:

[29] In 1982, 1983 and 1984, mothers were asked how many hours the youngest child was in daycare. But there is a serious missing data problem (e.g., only 115 of the 1,464 mother-child pairs in our sample have non-missing data in 1982).

If a woman reports using 10+ hours per week of child care, we assume she used child care during the quarter. If she worked full-time (i.e., 375+ hours in the quarter), we assume the child care must have been full-time, which seems straightforward. However, if the mother did not work (i.e., <75 hours in the quarter) but still reported using child care, it seems highly likely the child care was only part-time. More difficult is making a reasonable assignment if the mother worked part-time (75-375 hours in the quarter). We decided to assume the child care was part-time in this case. We admit this assignment is not so obvious. However, we experimented with assigning full-time day care in this case, and found it had almost no effect on the results. Thus, we define the function:

$$I_t^c = \begin{cases} 1 & \text{if mother works full} - \text{time and used child care} \\ 0.5 & \text{if mother works part} - \text{time and used child care} \\ 0.5 & \text{if mother did not work and used child care} \\ 0 & \text{otherwise} \end{cases}$$

and form cumulative child care, $\hat{C}_t$, average child care, $\overline{C}_t$, and current child care, C_t ,as follows:

$$\hat{C}_t = \sum_{\tau=1}^{t} I_\tau^c, \quad \overline{C}_t = \frac{1}{t}\sum_{\tau=1}^{t} I_\tau^c \quad \text{and} \quad C_t = I_t^c$$

where *t* is the age of the child.

As we noted earlier, complete child care histories are only available for the first three years after childbirth. Thus, we impute child care choices in years 4 and 5 after childbirth based on current work and work/child care histories. First, we set I_t^c=1 or 0.5 for mothers who work full or part-time, respectively, in a given period *t* after the third year. Second, for mothers who do not work in a given period *t*, we impute the child care choice based on the predicted probability of using child care from a probit model that we estimate using observed work and child care histories. As the probit coefficients in Appendix 2 indicate, day care use by non-working mothers is very well predicted by (i) having used day care a lot in the past and (ii) having not worked a lot since child birth. The pseudo R-squared is very large, suggesting these are very good predictors.

Another input into the child production function (10) is real household income. We measure it by summing income from all sources including wages, public assistance, unemployment benefits, interest or dividends, pension, rentals, alimony, child support and/or transfers from family or relatives. Household income is deflated using a region-specific CPI, just as we did for welfare benefits (see Section 3.3). We will experiment with different specifications of (10) using cumulative income since childbirth and/or current income.

An issue we did not discuss in deriving (10) is that mothers may have multiple children, which may affect resources allocated to any one child. Thus, we include the number of children in Z_i in (10), and also interact the child care time input with number of children (an income interaction was insignificant). This allows day care use to affect child outcomes differently depending on the number of children. Also, the number of children may be endogenous in (10), e.g., if there is a quality/quantity tradeoff, so we instrument for these variables using the instruments in Table 3.

Finally we turn to the child cognitive ability measures in the NLSY79. The measures we use are scores on the Peabody Picture Vocabulary Test (PPVT) at age 3, 4 and 5, and the Peabody Individual Achievement Test (PIAT) at ages 5 and 6. The later consists of reading and math subtests, PIAT-R and PIAT-M, respectively. The PPVT and PIAT are among the most widely used assessments for preschool and early school-aged children. The PPVT is a vocabulary test for standard American English and provides a quick estimate of verbal ability and scholastic aptitude. The PIAT-M consists of eighty-four multiple-choice items of increasing difficulty. It begins with such early skills as numeral recognition and progresses to measuring advanced concepts in geometry and trigonometry. Finally the PIAT-R measures word recognition and pronunciation ability.

Appendix 3 contains descriptive statistics for test scores in our sample. Note that there is no clear age pattern in the mean scores, as they are age adjusted. Mean scores on the PPVT, PIAT-M and PIAT-R are roughly 80, 95 and 101. Standard deviations seem to vary more by age than by test. For instance, at 5, the one age where we see all three tests, the standard deviations are quite similar: 17.5, 14.3 and 15.3, respectively. Thus, we decided to merge information from the three tests, allowing for mean differences. Interestingly, score differentials between children who are white/non-white and who have high-school graduate vs. high-school drop out parents are already apparent in the PPVT at age 3, and there is no discernable pattern of these differentials growing over time.

5.3. Descriptive Statistics

In Table 6 we present means and standard errors of the variables used in the analysis. For example, the average log test score in the sample is 4.50 with a standard deviation of 0.186 (after adjusting for mean differences across tests). 64% of women in the sample worked prior to giving birth at an average hourly rate of $4.39 (1983 dollars). Average work experience was 4.7 years prior to childbirth, and 72% of women had never been married. Average annual real household income is $10.9 thousand (1983) dollars. During the 20 quarters after childbirth mothers use .355 units of child care per quarter, for a total of 7.1 quarters, on average. However, if we compare the ’79-’93 and post-’93 periods, the child care usage rate increases 10 points (from 59% to 69%).

6. Estimation Results

6.1. The Reduced Form Regressions for the Endogenous Variables

The first stage of 2SLS, and the reduced form equations in LIML, use the instruments listed in Table 3, along with all the exogenous variables that appear in (10) – see Table 4 - to predict the three endogenous variables in the model (e.g., cumulative child care since birth, cumulative income since birth, number of children). The procedure is complicated by the fact that the instruments are time varying, and cumulative child care and income from birth of the child up through age *t* are presumably functions of the values of the instruments for all periods from birth up through time *t*. Thus, the set of instruments grows with *t*. We describe this structure in equation (1) of Appendix 4.

Table 7 examines the correlation of the instruments with the endogenous variables. The first column shows the partial correlation squared, while the second shows Shea's partial correlation[30] squared. For cumulative child care these are .1749 and .1495, respectively. The third column reports R^2s of regressions including only the exogenous variables, while the fourth shows incremental R^2s from adding the instruments. For cumulative child care, R^2 increases by .0917 when the instruments are added, and the F-test of their joint significance is 15.56, compared to a 1% critical value of 1.47. These results suggest that our welfare policy and local demand instruments are reasonably powerful – marginal R^2s reported in applied microeconomics are often smaller than what we see here, especially in earlier attempts to use IV to study effects of maternal employment (see Section 2.1).

We do not report the reduced form regressions, to conserve on space.[31] But it is worth noting that the 78 policy instruments have reasonable coefficients. The strongest predictors of cumulative child care use are: (i) if a State had implemented a work requirement, which has a strong positive effect on work/day care use (as expected), and a t-stat of about 3, (ii) the number of work requirement exemptions the State allows, which has a strong negative effect (as expected), and a t-stat of –6, (iii) the remaining time a woman is categorically eligible for welfare, which has a negative effect (t-stat of –4), and (iv) child support enforcement, which has a positive effect (t=2.7). Notably, interactions of education and AFQT with the welfare policy variables are always opposite in sign to the main effects, indicating behavior of more educated mothers is less influenced by welfare rules. Also, the 20th percentile wage interacted with AFQT has a positive effect (t-stat 2.6), and an analogous pattern holds for the unemployment-AFQT interaction. Thus, a stronger labor market increases the probability of employment, and this effect is stronger for more skilled women.

[30] This partials out the correlation of an endogenous variables with fitted values of the other endogenous variables.

[31] These contain 97 variables, of which 19 are exogenous variables also appearing in the main equation, and 78 are excluded instruments. As the main equation contains 3 endogenous variables, there are 75 over-identifying restrictions.

6.2. Baseline Specification of the Child Cognitive Ability Production Function

First, we seek to assess whether it is cumulative or current child care that matters most for children's achievement. In Table 8 we present estimates of alternative specifications of equation (10), including either cumulative or current child care (along with cumulative income). The table reports only a few key coefficients of interest (Recall the equation includes all variables in Table 4).

OLS estimates of the effect of child care on children's achievement appear to be upwardly biased, regardless of the child care measure. The OLS estimate of the effect of cumulative child care is essentially zero and insignificant, while that for current child care is *positive* and significant. In contrast, the LIML estimates are *negative* both for current and cumulative child care. However, only cumulative child care is statistically significant. Thus, we conclude cumulative child care is the more important determinant of cognitive ability, and adopt Table 8 column 1 as our baseline model.[32]

The baseline model implies an additional quarter of full-time day care reduces a child's test scores by roughly 0.73%. Thus, a year of full-time child care reduces scores by about 2.9%.[33] This corresponds to roughly .0291/.1861=0.156 standard deviations of the score distribution. Viewed another way, given our estimates in Appendix 1, a 2.9% test score reduction at age 6 translates into a .055 to .073 year reduction in completed schooling.[34] [Recall that this estimate should be interpreted as the effect of child care time (α_2) *relative* to the effect of mother's time (α_1) plus effects of any change in goods inputs the mother may choose as a result of using day care ($\alpha_3 \cdot \gamma_3$)].

In the bottom panel of Table 8, we see that the estimated effect of cumulative income since childbirth is quantitatively small and statistically insignificant, given controls for mother's education and AFQT score. For instance, the point estimate of .0069 in Table 8 column 1 implies that, at the mean of the data, doubling cumulative income (which means increasing its log by .69 – see Table 6) would increase test scores by only .0069·.69 = 0.5% In contrast, mother's education and AFQT are highly significant and quantitatively important. This is consistent with a view that lifetime income is much more important than transitory income in determining parental investment in children, and hence child achievement.[35] But we do not attempt to disentangle the extent to which the education

[32] In results not reported, we also found that the average day care variable defined in section 5.2 was not significant. Its coefficient was -.0638, with a standard error of .0643. We also experimented with using current instead of cumulative income, and found this change had essentially no impact on the child care coefficients.
[33] As a point of comparison, for married women, Bernal (2006) estimates that one full year of maternal work and day care use reduces scores by about 1.8%. And, James-Burdumy (2005) estimates that a full year of maternal work in year 1 of the child's life reduces the PIAT math score (measured at ages 3 to 5) by about 2.3%.
[34] Repeating the calculation at the end of Section 2, this gives about a 2 point increase in the percent who attend college.
[35] This is reminiscent of findings by Keane and Wolpin (2001) and Cameron and Heckman (1998) to the effect that transitory fluctuations in parental income have little effect on college attendance decisions by youth. In addition, it is

and AFQT coefficients reflect genetic transmission of maternal ability vs. the impact of maternal permanent income on investment in children.

6.3. Heterogeneity in the Effect of Maternal Time Inputs

To assess how effects of maternal time inputs on children's ability vary with characteristics of the mother/household, we next include interactions between cumulative child care and mother's education, AFQT score and number of children. Results are presented in Table 9. The variables are de-meaned before being interacted with cumulative child care. Thus, the child care coefficient can be interpreted as the mean affect of cumulative child care for a typical mother/household.

In Table 9 note first that the interaction between mother's education and child care use is *negative*, as expected (maternal time is a less valuable input into child cognitive ability production for less educated mothers). Its t-statistic is –1.74, so it attains significance at the 8% level. The point estimate is fairly substantial. It implies, e.g., that if a mother's education is 4 years above the sample average, then the negative day care effect goes from –.62% to –1.31%. The later estimate has a standard error of .41, and hence a t-stat of –3.19.[36] Thus, we have strong evidence that day care has more negative (*positive*) effects for children whose mothers have higher (lower) levels of education. The same pattern holds for AFQT (i.e., maternal time is more valuable for high AFQT mothers). Finally, the interaction between cumulative child care and number of children in the household is very small and insignificant, implying the child care effect does not depend on number of siblings.

6.4. Robustness of the Results to Alternative Estimation Methods

As we noted in the introduction, we rely on the limited information maximum likelihood (LIML) estimator of Anderson and Rubin (1949) because both theory and Monte Carlo evidence suggest it is less subject to the bias induced in 2SLS by the use of a large number of instruments. In Table 10 we compare estimates of our baseline model obtained using five alternative estimators: OLS, GMM, 2SLS, Fuller and LIML. As noted earlier, the OLS estimate of the effect of a quarter of maternal work/day care use is essentially zero, while the LIML estimate is -0.73%. The 2SLS estimate is -0.48%, and is significant at almost the 1% level. Thus, relative to LIML, the 2SLS estimate is shifted about 40% of the way towards OLS. This is not surprising, as with 78 instruments, we would expect 2SLS to retain some part of the OLS endogeneity bias. Note that the GMM estimate (-0.73%, t=-3.0) is very similar to the 2SLS estimate, while the Fuller estimate is essentially identical to the LIML estimate.

---

consistent with findings by Blau (1999b) and Carneiro and Heckman (2002) according to which permanent household income is significant in determining investments in children while transitory income is not.

[36] The reason this is so highly significant is due to the negative covariance between the main effect and interaction term.

An alternative approach to the "many instrument" problem is to condense the size of the instrument set. Simply dropping instruments would lose efficiency. So instead, we attempt to summarize the information contained in the instruments using a smaller number of variables. We do this using factor analysis. Specifically, we used the principal factor method with varimax rotation. The factor scoring coefficients are calculated using the regression method.

A typical rule of thumb in factor analysis is to retain factors with eigenvalues greater than one, of which there are 13. However, in the present context we are not interested in obtaining a set of factors that best summarize the correlations of the 78 instruments *per se*. Rather, we are interested in finding the factors that best explain the endogenous variables. To do this, we regressed each of the endogenous variables on the full set of factors, and retained those that were most highly significant.

For cumulative child care, the most important factors (t>3), ordered by eigenvalue, where 6, 8, 12, 19, 21, 24, 26 and 47. Based on which variables load heavily on each factor, we see they capture the following: Factor 6 captures benefit levels and child support enforcement. Factor 8 captures remaining eligibility, the local wage rate and EITC. Factor 12 captures the local unemployment rate. Factor 19 captures work requirements, sanctions and the strictness of welfare rules. Factor 21 captures prior work experience interacted with child age.[37] Factor 24 captures CCDF spending, child support enforcement and EITC. Factor 26 captures benefit levels, remaining eligibility, EITC and local wages. Factor 47 captures time limits, work requirements and disregards.

For cumulative income, the important factors were 1, 3, 6, 7, 10, 24, 26, 37 and 47. For number of children, the important factors were 1, 2, 3, 6, 8, 9, 12, 21, 23, 26 and 47. In the interest of space we will not provide interpretations of all these factors. We simply note there is some overlap among the factors that are important in explaining each of the three endogenous variables. Taking the union of the sets of factors that matter for each variable, we obtain a set of 16 variables to use as instruments. Thus, we have reduced the instrument set from 78 to 16.

Table 11B examines the correlation of these 16 factor instruments with the endogenous variables. For instance, Shea's partial correlation squared is .1014 (compared to .1495 for the full set of 78 instruments), and the F-test of their joint significance is 15.03 (compared to 15.56 for the full instrument set). The factors have sensible coefficients in the reduced form equation for cumulative

[37] Note: While prior work experience and child age are both included in the main equation, their interaction is not. Recall that prior work experience is included as a proxy for the mother's skill endowment (which is presumably correlated with the child's skill endowment). Thus, excluding the age interaction from the main equation follows logically from the assumption that the child skill endowment has an age invariant effect in the log test score equation (This assumption is common in the human capital literature – see discussion prior to equation (3)). The interaction is useful for predicting cumulative child care, because, obviously, mothers who work more prior to childbirth also tend to work more afterward.

child care use (not reported). For instance, Factor 6 has a substantial negative coefficient (t=-15), suggesting that higher benefit levels reduce work and child care use as expected. Factor 12 also has a substantial negative coefficient (t=-9.8), implying higher unemployment reduces work and child care use. And factor 21 has a positive coefficient (t=8.6), suggesting that work experience/cumulative child care grows more quickly with child age for mothers with more prior experience.

Table 11A reports results of the LIML, Fuller, 2SLS and GMM estimators using the 16 factors as instruments, and compares these to the LIML results using the full set of 78 instruments. Three aspects of the results are notable: First, regardless of the instrument set we use, the LIML estimates of the effect of cumulative child care are identical to 5 decimal places. Second, using the factors as instruments actually leads to an increase in efficiency, as the t-statistic on the LIML child care coefficient increases from -2.4 to -2.8. Third, the LIML and 2SLS estimates are now quite similar. The 2SLS estimate is now -.00658 with a t-statistic of -2.7. Thus, the 2SLS estimate is now shifted only 10% of the way towards OLS (compared to 40% when we use all 78 instruments).

The bottom row of Table 11A reports the Cragg-Donald (1993) weak instrument test statistic, which is 5.81 when the full set of 78 instrument is used but increases to 16.45 when we use the reduced set of 16 instruments. Stock and Yogo (2004) develop critical values of this statistic, for testing the null hypothesis that the asymptotic maximal bias of the 2SLS estimator may exceed some percentage of the OLS bias (under many instrument asymptotics). For the case of 3 endogenous variables and 16 excluded instruments, the critical values for the null that the 2SLS bias may exceed 20%, 10% or 5% of the OLS bias are 5.94, 10.41 and 18.94, respectively. Thus, using the 16 factors as instruments, we clearly reject the null that the 2SLS asymptotic bias may exceed 10% of the OLS bias, but we do not quite reject the null that it may exceed 5%.[38] These results suggest that using factor analysis to reduce the size of the instrument set is an effective way to reduce the 2SLS bias.

Finally, while Stock and Yogo (2004) do not report exact values for all cases, we can extrapolate from their figures that the critical value for the null that bias of the Fuller estimator is no greater than 5% of the OLS bias is roughly 3.7 in the 16 instrument case and about 1.8 in the 78 instrument case. Thus, with Cragg-Donald statistics of 16.45 and 5.81, we easily reject the null in either case. And the critical value for the null that bias in size for the LIML test statistics is no greater than 10% of the OLS bias is roughly 3.8 in the 16 instrument case and about 5.4 in the 78 instrument case. Thus, size distortions do not appear to be a problem for LIML in either case.

---

[38] For the case of 3 endogenous variables and 78 excluded instruments, the critical values for the null that the 2SLS bias may exceed 20%, 10% or 5% of the OLS bias are 5.65, 10.76 and 20.82, respectively. Thus, using all 78 instruments, we only barely reject the null that the 2SLS asymptotic bias may exceed 20% of the OLS bias (i.e., 5.81 vs. 5.65).

6.5. Robustness of the Results with Respect to Age of the Mother at Childbirth

A potential concern with our results is that our welfare policy instruments are correlated with mothers' age at childbirth, due to the timing of waivers/welfare reform. Welfare waivers were first implemented in some States in '92-'93. But few women were substantially impacted until a few years later. Fang and Keane (2004, p. 32) note that binding work requirements first hit significant numbers of women in 1995-6. Suppose a woman had a child in 1990, reaching age 5 during 1995. Her labor supply/child care decisions could have been impacted by waivers when the child was 5 years old, possibly affecting test scores at ages 5 and 6. But, for children born prior to 1990, it is unlikely that waivers could have influenced the mother's labor supply before the child was 6. In the NLSY79, women who had children prior to 1990 tend to be younger at the time of childbirth than women who had children later. Indeed, from 1990 onward, all births are to mothers in their 20s and 30s, while, prior to 1990, a significant fraction were to teenage mothers.[39]

To understand the potential bias created by this correlation, consider the following scenario: Since work requirements positively affect maternal work/day care use, and work requirements are positively correlated with age at childbirth, our reduced form predicted values for maternal work/day care use will be *positively* correlated with maternal age at childbirth. Then, if (i) mother's age at birth has a positive effect on child cognitive ability, and (ii) we fail to adequately control for mother's age in the main equation, this will generate a spurious *positive* effect of maternal work/day care use on child cognitive test scores. Thus, since we actually find a negative effect of maternal work/day care use, the concern is that we *understate* this negative effect.

To deal with this concern, Table 12 presents several specifications of the main equation. The results in columns (1)-(5) show the estimated effect of cumulative child care use is very robust to the inclusion/exclusion of different controls for mother's age at childbirth – including age, age squared, and dummies for if she was under 20 or over 33. In fact, the estimated effect ranges from -0.68% per quarter with no controls for age at all, to -0.76% with all four controls. The estimate is -0.73% in the baseline model in column (4), which includes only the teenager and over 33 indicators.

A striking aspect of the results is that, conditional on measures of the mother's human capital (i.e., education, AFQT), there is absolutely no evidence of a *positive* association between maternal age at birth and child cognitive ability outcomes. Indeed, column (2) shows a significant *negative* relationship; and the estimates in column (4) imply that, *ceteris paribus*, children of teenage mothers

[39] Indeed, the youngest women in the NLSY – i.e., those who were 14 in Jan. 1979 – would be 24 by Jan. 1990. Thus, the large majority of mothers who would have been affected by welfare reform would have been at least 24 at childbirth.

have 2.6% *higher* cognitive ability. In hindsight, this is unsurprising. For almost all of human history teenage childbearing was the norm rather than the exception, so it would be strange to find a negative effect of teenage childbearing *per se* (controlling for economic resources).

To further address this issue, we also estimated the baseline specification on subsamples of women based on the age of the mother at birth. The results are in Table 13. In column (1) we restrict the sample to women who were 24 to 34 years old at childbirth. In column (2) we restrict the sample to women who were at least 24 years old. Using these sub-samples actually increases the estimated effect of child care somewhat (to -3.5 or -4.0% per year). Again, there is no indication that the age/welfare policy correlation leads us to exaggerate the size of the day care effect.

6.6. Robustness of the Results with Respect to Specification of the Main Equation

We now return to Table 12 and consider sensitivity of our results to four other changes in the specification of the main equation. First, in column (6) we drop the mother's AFQT score. This has essentially no effect on the estimated day care coefficient. But it does produce a large increase in the coefficient on cumulative income, which becomes highly significant. This seems consistent with the view that lifetime income is more important than transitory income in determining parental investment in children, and hence children's achievement. With AFQT omitted, income appears significant, as it proxies for the mother's permanent income/skill endowment. But AFQT is a better proxy, so when it is included the income variable drops out.

Even without AFQT, the implied effect of income remains modest. The point estimate implies that, at the mean of the data, a doubling of cumulative income would increase test scores by about (.062)·(.69)=4.3%. Recall that column (6) still includes such variables as mother's education and her pre-childbirth wage, which also proxy for her permanent income/skill endowment.

Next, in column (7), we attempt to determine if maternal work has a separate affect from day care. As noted in Section 4, married women often use day care while not working, so it is possible to estimate the effect of maternal work time holding day care time fixed. But single mothers rarely use day care while not working. Thus, the correlation between cumulative work and daycare time is very high (ρ=.94), making it nearly impossible to separate their effects. Instead, we construct a measure of maternal employment equal to 1 if the mother works continuously after childbirth, 0.5 if she works part of the time, and 0 if she did not work at all. This variable (which we treat as endogenous) is positive but insignificant. With its inclusion, the day care coefficient increases to an implausibly large value of -.0147 that is imprecisely estimated (and only significant at the 10% level). Clearly, collinearity makes it infeasible to sort out effects of work vs. day care time for single mothers.

Third, in column (8), we consider sensitivity of our results to controls for the ages of other children in the household. Our baseline model (4) controls for the number of siblings, not their ages. If young children have a different effect on the mother's time constraint than older children, our main equation is simply misspecified, which may bias our estimated child care effect.[40] Thus, column (8) includes in the main equation separate regressors for the number of children aged 0-5, and the number aged 6-17 (both treated as endogenous). Interestingly, these variables have very similar coefficients. Their inclusion has essentially no impact on the estimated child care effect.

Finally, we consider the issue of aggregate time effects. It is possible that, during our sample period, some omitted time varying factor both influenced child test scores and was correlated with the increasing stringency of welfare rules. Column (9) includes an aggregate time trend to address this concern. Interestingly, the time trend is negative and significant, implying an aggregate factor not included in our model drove down average test scores by 0.3% per year during the sample period. Including the time trend reduces the estimated day care effect slightly, from -2.9% per year under the baseline to -2.7%. Thus, any bias from ignoring aggregate time effects appears to be minor. Results were essentially identical using a quadratic time trend or unrestricted time dummies.

6.7. Robustness of the Results with Respect to State Fixed Effects

Next, we examine robustness of our results to inclusion of State fixed effects. The argument for including State effects is to deal with potential cross-State correlation between the instruments and *unobserved* child skill endowments; e.g., States where children had relatively low unobserved skill endowments may have adopted stricter welfare reform. This would bias the day care effect negatively. However, we are skeptical of fixed effects for two reasons: First, in the child production function context, we are skeptical of the strict exogeneity assumption required for consistency.[41] Second, Keane and Wolpin (2001) show fixed effects can lead to very misleading results if expected future values of policy variables matter for current decisions.[42] Hence, we do not adopt fixed effects as the baseline specification, but we report fixed effects estimates in Table 14.

Adding State fixed effects to the main equation increases the estimate of the cumulative day care effect from -.73% to a rather implausibly large value of -1.43%. It also reduces the precision of

[40] A related concern is that, as we discussed in Section 3, some of our instruments are functions of ages of the mother's children. If child ages are endogenous it is a threat to the validity of those instruments. We deal with this issue in Section 6.8 where we experiment with different instrument sets, including ones that drop instruments that depend on child age.
[41] The strict exogeneity assumption will fail if children's test score realizations at age *t* affect future inputs into child production, and/or how the welfare policy rules evolve. [See also the discussion of this point in Section 2].
[42] A State fixed effect controls for a State's average level of welfare generosity. Thus, using State fixed effects, we estimate the impact of a transitory change in work/day care use induced by a transitory change in welfare rules, holding a State's expected future rules fixed. This effect may be radically different from that induced by long lived policy changes.

the estimate, more than doubling the standard error. Despite this, the estimate remains significant at the 5% level (t = -2.17). In our view the implausibly large fixed effects estimate probably reflects the difficulty in knowing just what fixed effects actually estimates in a context where (i) strict exogeneity seems implausible and (ii) changes in the variable of interest driven by transitory policy changes may have a very different impact from those driven by persistent changes.

An alternative way to address whether States where children had low skill endowments adopted stricter welfare reform is to look at the issue directly. In Appendix 5 we group States into those that adopted more vs. less strict approaches to welfare reform along five different dimensions. Then we compare average test scores in the pre-reform period between each group of States. In each case, there is no significant difference in average child test scores in the pre-reform period between States that subsequently adopted more vs. less strict welfare reform programs.[43] Thus, there is no clear evidence of cross-State correlation between the instruments and child skill endowments.

6.8. Robustness of the Results with Respect to the Instruments

It is well known that IV estimates can be sensitive to the instrument list, and that, given unobserved heterogeneity in treatment effects, what IV identifies depends on the instruments used. Thus, it is important examine the robustness of our results with respect to the instrument list used in the reduced form equations for the endogenous variables. Table 15 reports results using the baseline specification of the instrument list in column (1), and seven variants on that list in columns (2)-(8).

In column (2) we exclude Child Care Development Fund expenditures. This instrument may be of questionable validity from a theoretical point of view, given the arguments in Section 4, since it may shift a mother's budget constraint conditional on her work, income and child care usage decisions (by changing the price of child care). Excluding this instrument leads to a slightly stronger estimated day care effect (i.e., -3.2% vs. –2.9% per year) and little change in standard errors.

In column (3) we use only the main features of TANF as instruments: time limits, work requirements and earnings disregards.[44] This causes our estimate of the impact of day care to increase to –4.7% per year, and the standard error of the estimate increases by 43%. In column (4) in contrast, we drop the TANF related instruments (time limits, work requirements and disregards), using only other aspects of the policy and local demand environment to identify the day care effect. This reduces the estimated day care effect only slightly (-2.7%), and reduces the standard error.

[43] Based on point estimates, for 3 measures (time limit waivers, work requirements, exemptions) States that adopted stricter rules had *higher* pre-reform mean scores, while for 2 measures (work requirement waivers, time limit length) they had *lower* scores.

[44] In other words, we exclude CCDF and CSE expenditures, EITC rates, benefit amounts and local demand conditions. The time limits, work requirements and earnings disregards are still interacted with mother's education and AFQT.

In column (5) we drop all instruments specific to the welfare reforms of the 90s (e.g., TANF, CCDF, EITC), using only instruments that would have varied across States/time regardless. These are State welfare grant levels and local demand conditions (i.e., State unemployment rates and 20th percentile wage rate). This gives a 1st stage marginal R^2 of .056. The resulting estimate of the day care effect is very similar to our baseline specification (-2.8%). This is our most parsimonious model, in that it uses only 18 instruments. Given this relatively small instrument list, LIML and 2SLS estimates are rather similar; 2SLS gives -.0063 (.0024) compared to -.0071 (.0027) for LIML.

In our reduced form regression, we interact all policy and demand variables with mother's education and AFQT. This lets changes in policy/demand have different effects on different types of mothers (e.g., welfare rules are less important for the college educated). In column (6) we drop these interactions to see how important they are. This reduces the estimated day care effect to -1.9% per year - the smallest effect we obtain in any LIML specification – but it is significant at the 10% level.

Recall that some of our instruments are tailored to individuals based on ages of their children (e.g., whether a woman could have reached the time limit – see Section 3.1). In column (7) we drop these individual specific instruments, and rely purely on cross-State and over time variation to identify the child care effect. The resulting estimate is -3.4% per year, which is slightly larger than our baseline estimate. In column (8) we go further and also drop the interactions of the remaining instruments with mother's education and AFQT. This gives an estimate of -3.3% per year.

In summary, our result of a negative day care effect is robust to a wide range of alternative instrument sets, with the estimated effect ranging from –1.9% to –4.7% per year, with all but a couple estimates between -2.7% and -3.4%. We have experimented with a large number of other instrument sets (not reported) and continue to find results in this general range.

6.9. The Effects of Different Types of Child Care

So far we have reported on effects childcare *in general*, but it seems likely that the type of care matters. That is, *formal* center-based care by trained providers (e.g., daycare centers, pre-school) may have different effects from *informal* care provided by relatives (e.g., grandparents, siblings) or non-relatives. Thus, we estimated versions of equation (10) in which the effect of child care is allowed to vary by type of care.[45] First, Table 7 Panel B present results of the reduced form regressions for child care inputs of different types (i.e., formal, informal, etc.). Even at this more refined level, the welfare policy/demand condition variables are reasonably powerful instruments.

[45] Recall that we do not have direct measures of child care use in years 4 and 5, and we impute this using the procedure described in Section 5.2 and in Appendix 2. Now, having imputed child care use, we impute whether it was formal or informal by looking at the last available observation on type of care used.

The marginal R^2 for the excluded instruments in the reduced form regressions ranges from .086 to .098, and all the joint F-tests show the instruments are highly significant.

The reduced form regression coefficients (not reported) also appear reasonable. They show mothers are more likely to use formal relative to informal care if: (i) a State does *not* have a work requirement, (ii) it has young child or other work requirement exemptions, (iii) it has a longer work requirement time limit, (iv) work requirements were implemented more recently, (v) less time has elapsed since a time limit could have hit, (vi) remaining eligibility is greater, (vii) a State has higher CCDF spending, or (viii) earnings disregards are greater. Interestingly, work requirements raise the probability of using informal, but not formal care. And if a State has more exemptions, it reduces the probability of using child care in general, but that of using informal care is reduced much more. Education interactions are always opposite in sign to main effects, welfare rules have less influence on more educated women. A high AFQT score reinforces the effects that young child exemptions and higher disregards increase the likelihood of using formal care. Only for high AFQT women does the probability of using formal care increase in the time since time limits were implemented.

Strikingly, the LIML results in Table 16 indicate that formal (i.e., center-based) care does not have any adverse effect on cognitive outcomes. Only informal types of care lead to significant reductions in achievement. In particular, an additional year of informal childcare causes a 3.5% reduction in test scores. The estimated effect for formal care is actually positive, but insignificant.[46]

In column (4) of Table 16 we divide informal child care into that provided by relatives (most often grandparents) vs. non-relatives (e.g., family daycare). Here we find that only informal care by relatives has a negative effect. Note that informal care by relatives is the most common arrangement for the single mothers in our sample (60%). Informal care by non-relatives accounts for a little less than 20%, and formal center-based care accounts for a bit over 20%. This preponderance of informal care explains why our overall estimate of the effect of daycare is negative (i.e., -2.9% per year).

Our results here are basically consistent with prior work by Hansen and Hawkes (2007). Looking at Bracken school readiness scores in the Millenium Cohort Survey, they find a negative effect of grandmother care relative to formal center-based care. Similarly, Gregg et al (2005), using the Avon Longitudinal Survey, find that early maternal employment reduces subsequent child test scores only if children were placed in informal care (i.e., care by a relative or friend).

[46] We also tried interacting type of child care with mother's education. We consistently find a positive effect of formal care for low education mothers. But this effect is imprecisely estimated, and whether it is statistically significant is sensitive to the instrument set. Still, the result is at least consistent with prior results suggesting center-based care is beneficial for low SES children (see, e.g., Currie and Thomas (1995) on Head Start, or Pungello et al (2006) on the Perry Pre-School and Abecedarian experiments).

It may seem surprising that care by relatives – predominately grandparents – would lead to worse child outcomes, as grandparents presumably care a great deal about their grandchildren. There is a literature in sociology however, showing that grandparents often find caring for young children to be stressful and overly physically demanding (see Millwood (1998), Goodfellow and Laverty (2003)). Prior literature also suggests that center-based care has two advantages over informal care: (1) trained formal care providers may provide more cognitive simulation to children than do informal providers,[47] and (2) center-based care may provide more stimulating interaction with other children, and more educational activity, than informal care.

Ideally, we would also like to examine how effects of childcare differ by quality. But it is difficult to measure quality directly in the NLSY. Lacking a direct quality measure, we have instead differentiated between formal and informal care. But the notion that formal center-based care is superior is consistent with the evidence on who uses it. In Appendix 6 we present a logit for whether a mother uses formal or informal child care (conditional on child care use). The results show that more educated, urban women with fewer children are more likely to use formal care. This suggests that formal care is higher quality, as it is typically used by women who can afford more expensive care. Similarly, Appendix 6 also presents a logit for use of relatives vs. non-relatives (conditional on using informal care). The more educated, urban women with fewer children are more likely to use non-relatives, suggesting that non-relatives provide higher quality care than relatives.

6.10 Effects of Child Care at Different Child Ages

Developmental psychologists argue that the effects of home inputs on child development depend on the age at which the inputs are applied. Thus, in Table 17 we estimate how the effect of child care varies with child age. To let the child care effect differ by age, we replace the term $\varphi_2 \widehat{C}_t$ in equation (10) by the term $\sum_{\tau=1,t} (\varphi_{20} + \varphi_{21} \cdot t) \cdot C_t$. This lets the impact of child care be a linear function of child age. This expression may be rewritten $\varphi_{20}\ \widehat{C}_t + \varphi_{21} \cdot \sum_{\tau=1,t} (t \cdot C_t)$, where $\sum_{\tau=1,t} (t \cdot C_t)$ is an "age weighted cumulative child care" variable which we now add to the regression and treat as endogenous. In Table 7 Panel B we see that the incremental R^2 for the excluded instruments in the reduced form regressions for this variable is .105, with a F-statistic of 27.7.

[47] McCartney (1984), Melhuish et al. (1992) and NICHD Early Child Care Research Network (2000) show that a key difference between high and low-quality care is the amount of language stimulation. Center-based teachers are more likely to have training in child development, and to be more educated in general, both of which are associated with more verbal stimulation. They also tend to provide more supportive, attentive and interactive care (see NICHD Early Child Care Research Network, 2000). Lamb (1998) and Smith (1998) show that children whose caregivers "provide ample verbal and cognitive stimulation … are more advanced in all realms of development compared with children who fail to receive these inputs…" (see National Research Council and Institutes of Medicine (2000, p. 315)).

The OLS results in Table 17 imply positive but insignificant effects of childcare at all ages. But the LIML results imply that age does matter. The estimate of φ_{21} is -.00308 with a standard error of .0015, implying child care effects grow more negative by 0.31% per quarter. The effect first turns significantly negative at quarter 10, when the combined effect $\varphi_{20} + \varphi_{21} \cdot t$ is -0.85% with a standard error of .30. Thus, our results imply that child care during the first two years has no detrimental effect on child cognitive outcomes, while child care at ages 2+ does have significant negative effects.[48] This is consistent with the idea that child-mother interactions are more valuable when the child is ready to engage in more challenging tasks like language learning, and less so during initial stages when the child requires more basic care.

Our results seem to contradict National Research Council and Institutes of Medicine (2000), who conclude there is clear evidence that maternal employment in the first year is a negative factor for infant development.[49] But our reading of the prior evidence is rather different. First, most of the cited studies fail to control for systematic differences between women who work in the first year after childbirth and those who don't. Second, results of prior literature are not really so consistent.

For instance, Ruhm (2000) finds that maternal employment in the first year after childbirth reduces PPVT scores at ages 3 and 4. But he also finds employment in the second and third years – and not the first year – are associated with lower PIAT math and reading scores at ages 5 to 6. Of particular note is Waldfogel et al (2002), the only study in the area that uses sibling fixed effects in an attempt to correct for endogeneity of maternal employment. They follow a sample of children of the NLSY up to ages 7 and 8. Their OLS estimates imply a negative effect of maternal employment in the first year on PPVT and PIAT scores, but their FE estimates show no negative effects. Thus, we do not view our failure to find a negative first year effect as inconsistent with prior literature.

6.11. The Effect of Welfare Reform on Child Test Scores

Our focus has been on using the welfare policy changes on the mid-90s as a source of exogenous variation to help identify the effect of childcare time vs. maternal time on child outcomes. But it is also interesting to examine how welfare reform itself affected child outcomes. We do this using the reduced form equation for child test scores, obtained by substituting all the excluded instruments listed in Table 3 for the three endogenous variables in equation (10). One change is that we include a quadratic time trend. While we found this made little difference to the

[48] We were unsuccessful at estimating more flexible specifications than a linear time trend – i.e., differentiating effects of child care more finely by age (i.e., year-by-year) – due to imprecision of the estimates of the annual effects.

[49] In particular, they cite Baydar and Brooks-Gunn (1991), Belsky and Eggenbeen (1991), Ruhm (2000), Desai et al. (1989), Vandell and Corasaniti (1988), Han et al. (2001) and Waldfogel et al (2002) as supporting this conclusion.

estimated day care effect, we felt it was important to control for possibly omitted time effects in the reduced form to avoid the risk of attributing the impact of these factors to welfare reform.

The 78 instruments are highly significant in the reduced form test score equation – the F-test for their joint significance is 3.42 compared to the 1% critical value of 1.41. In a simpler model that leaves out interactions with mother's education and AFQT, the remaining 26 instruments give an F-test of 2.52, compared to the 1% critical value of 1.76. These results show that changes in welfare rules did have a significant impact on child test scores. However, given the complexity of the set of variables that characterize the welfare rules, it is quite difficult to put a meaningful interpretation on the individual coefficients. Thus, we instead simulate the effect of changes of welfare rules by simulating test scores from the reduced form. We simulate average test scores under two scenarios: (i) using the policy variables that were actually in place and (ii) holding the policy variables fixed at a baseline level. A similar procedure was used by Fang and Keane (2004) to evaluate effects of changes in welfare rules on employment and welfare participation by single mothers.

Simulation of the reduced form model implies that changes in welfare rules had almost no impact on child test scores during the 1979-93 period. This is not surprising, as the rule changes during that period were modest. However, the rule changes began to reduce test scores after 1993. Our model implies that average test scores for children of single mothers in the 1994-99 period were 2.35% lower than they would have been had the rules not changed.

This figure is broadly consistent with our point estimate for the effect of child care. As we see in Table 6, the child care usage rate was about 10 percentage points higher in the post-93 period. Thus, by 1999 (6 years later) children would have had about 0.6 extra years of child care on average. Multiplying this by our estimated annual effect of -2.9%, we obtain -1.77%. This is slightly smaller than our simulated effect of -2.35%, but in the ballpark.

7. Conclusions

This paper evaluates effects of home inputs on child cognitive development using the sample of single mothers in the NLSY79. In particular, we assess the effects of childcare use and household income on child cognitive ability test scores at ages 3, 4, 5 and 6. Of course, an important selection problem arises when trying to assess the impact of maternal time and income on child development. To deal with this, we take advantage of (plausibly) exogenous variation in employment and child care choices of single mothers generated by differences in welfare rules across states and over time. This approach is motivated by the fact that the Welfare Reform of 1996, as well as earlier State level changes adopted under Section 1115 Welfare Waivers, generated substantial new incentives for single mothers to work and use child care. This event provides a good opportunity to extend the literature on effects of child care on child outcomes – a literature which has been limited by the difficulty of finding plausible instruments for child care usage. Thus, we construct a comprehensive set of welfare policy variables, and use these as instrumental variables to estimate child cognitive ability production functions. We also use local demand conditions as additional instruments.

Our main results indicate that the effect of child care on children's achievement is negative, significant and rather sizeable. In particular, one additional year of full time child care use is associated with a reduction of approximately 2.9% in cognitive ability test scores. This corresponds to 0.156 standard deviations, so it is a substantial effect.

But this general finding masks important differences across types of child care, child age, and types of mothers. What drives the negative estimate of day care effect is that most (i.e., about 75%) day care used by single mothers in our sample is informal (i.e., care by siblings, grandparents or other relatives, or by non-relatives in non-center based settings). Our estimates imply that a year of informal day care reduces child test scores by 3.5%. In contrast, we find that formal center-based care (e.g., center-based day care, pre-K programs, etc.), has no adverse effect on child outcomes.

Furthermore, child care only has an adverse affect for children who are 2+ years of age. We speculate this may be because maternal time inputs are more critical when the child is developing language skills. And, child care has a greater adverse affect for more educated mothers. This is not surprising, as education presumably increases the value of maternal time in child production.

We also provide estimates of how test scores at young ages are related to completed schooling. These imply, for example, that a 1% increase in PIAT math test scores at age 6, holding parental background variables like mother's education fixed, is associated with an increase in educational attainment (measured at age 18 or later) of approximately .019 years. For reading scores the figure is .025 years. Thus, for example, a 2.9% reduction in test scores induced by a years of full-time day care translates into roughly a .055 to .073 year reduction in completed schooling.

Finally, we find that the effect of household income since childbirth is quantitatively small, and statistically insignificant, given controls for mother's education and AFQT. In contrast, mother's education and AFQT are both highly significant in the child cognitive ability production function. This is consistent with a view that lifetime income is more important than transitory income in determining parental investment in children. But we do not disentangle whether this reflects genetic transmission of maternal ability to the child vs. the impact of permanent income on child investment.

References

Anderson, T.W. and H. Rubin, 1949, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46-63.

Anderson, T.W., N. Kunitomo and Y. Matsushita, 2005, "New Light from Old Wisdoms: Alternative Estimation Methods of Simultaneous Equations in Microeconomic Models," CIRJE Discussion paper #F-321, University of Tokyo.

Andrews, D.W. and J. Stock, 2006, "Inference with Weak Instruments," in Blundell R., W. Newey and T. Persson (eds.), Advances in Economics and Econometrics, Vol. 3, Ch. 6, Cambridge University Press.

Baydar, N. and J. Brooks-Gunn, 1991, "Effects of Maternal Employment and Child Care Arrangements on Preschoolers' cognitive and Behavioral Outcomes: Evidence from Children of the National Longitudinal Survey of Youth," *Developmental Psychology* 27(6), 932-945.

Ben-Porath, Y., 1967, "The production of human capital and the life-cycle of earnings," *Journal of Political Economy,* vol. 75, no. 4, part I, August, 352-265.

Bernal, R., 2006, "The Effect of Maternal Employment and Child Care on Children's Cognitive Development," manuscript, Northwestern University, January.

Bernal, R. and. M. P. Keane, 2007, "Quasi-Structural Estimation of a Model of Child Care Choices and Child Cognitive Ability Production," manuscript Northwestern University and Yale University.

Blau, D., 1999a, "The Effects of Child Care Characteristics on Child Development," *The Journal of Human Resources*, XXXIV, 4.

Blau, D., 1999b, "The Effect of Income on Child Development," *Review of Economics and Statistics,* 81(2): 261-276, May.

Blau, F. and A. Grossberg, 1992, "Maternal Labor Supply and Children's Cognitive Development," *The Review of Economics and Statistics* 74(3), 474-481.

Blau, D. and J. Currie, 2004, "Preschool, Day Care, and Afterschool Care: Who's Minding the Kids" in *Handbook of the Economics of Education*, Hanushek, E. and F. Welch eds. Elsevier.

Burchinal, M.R., S. Ramey, M. Reid and J. Jaccard, 1995, "Early Child Care Experiences and their Association with Family and Child Characteristics during Middle Childhood," *Early Childhood Research Quarterly,* 10: 33-61.

Cameron, S. and J. J. Heckman, 1998, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, University of Chicago Press, vol. 106(2), pages 262-333, April.

Carneiro, P. and J. Heckman, 2002, "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal*, 112: 705-34.

Caughy, M., J. DiPietro and D. Strobino, 1994, "Daycare Participation as a Protective Factor in the Cognitive Development of Low-Income Children," *Child Development* 65:457-471.

Chase-Lansdale, P.L., R. Moffitt, B. Lohman, A. Cherlin, R. Coley, L. Pittman, J. Roff, E. Votruba, 2003, "Mothers' Transitions from Welfare to Work and the Well-Being of Preschoolers and Adolescents," *Science,* 299(March):1548-1552.

Connolly, S., J. Micklewright and S. Nickell, 1992, "The Occupational Success of Young Men who Left School at Sixteen," *Oxford Economics Papers*, 44, pgs. 460-479.

Cragg, J.G. and S.G. Donald, 1993, "Testing Identifiability and Specification in Instrumental Variables Models, *Econometric Theory*, 9, 222-240.

Currie, J. and D. Thomas, 2005, "Does Head Start Make a Difference?," *American Economic Review,* 55:3, 341-364.

Currie, J. and D. Thomas, 2001, "Early Test Scores, School Quality and SES: Long Run Effects on Wage and Employment Outcomes," *Worker Wellbeing in a Changing Labor Market,* Vol. 20.

Desai, S., L. Chase-Lansdale and M. Robert, 1989, "Mother or Market? Effects of Maternal Employment on the Intellectual Ability of 4-year Old Children," *Demography*, Vol. 26, No. 4.

Duncan, G. and NICHD Early Child Care Research Network, 2003, "Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development," manuscript Northwestern University.

Fang, H. and M. Keane, 2004, "Assessing the Impact of Welfare Reform on Single Mothers," *Brookings Papers on Economic Activity* Vol. 1.

Goodfellow, J., & Laverty, J., 2003, Grandcaring: Insights into grandparents' experiences as regular child care providers. Canberra: Early Childhood Australia.

Gregg P., Washbrook E., Propper C., Burgess S., 2005, "The effects of a mother's return to work decision on child development in the UK," *The Economic Journal*, 115:F48-F80.

Han, W., Waldfogel, J., and J. Brooks-Gunn, 2001, "The Effects of Early Maternal Employment on Later Cognitive and Behavioral Outcomes," *Journal of Marriage and the Family*, 63(2): 336-354.

Hansen, C., J. Hausman and W. Newey, 2006, "Estimation with Many Weak Instruments," Working Paper, MIT Department of Economics.

Harmon, C. and I. Walker, 1998, "Selective Schooling and the Returns to the Quantity and Quality of Education," manuscript Department of Economics, University College Dublin.

Haveman, R. and B. Wolfe, 1994, Succeeding Generations: On the Effects of Investments in Children. New York: Russell Sage Foundation.

Howes, C. and J. Rubenstein, 1981, "Toddler Peer Behavior in Two Types of Day Care," *Infant Behavior and Development* 4:387-393.

______. 1985, "Determinants of Toddlers' Experiences in Day Care: Age of Entry and Quality of Setting," *Child Care Quarterly*, 14: 140-151.

Hutchinson, D., H. Prosser and P. Wedge, 1979, "The Prediciton of Educational Failure," *Educational Studies* 5(1).

James-Burdumy, S., 2005, "The Effect of Maternal Labor Force Participation on Child Development", *Journal of Labor Economics* Vol. 23(1): 177-211.

Keane, M.P. and K. Wolpin, 1997, "The Career Decisions of Young Men," *Journal of Political Economy*, University of Chicago Press, vol. 105(3), pages 473-522, June.

Keane, M.P. and K. Wolpin, 2001, "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, vol. 42(4), pages 1051-1103, November.

Lamb, M., 1996, "Effects of Nonparental Child Care on Child Development: An Update," *The Canadian Journal of Psychiatry*, 41:330-342.

Lamb, M., 1998, "Nonparental child care: Context, Quality, Correlates" in *Handbook of Child Psychology,* Vol. 4 W. Damon, I. Sigel and K. Renninger, eds. New York: John Wiley & Sons Inc.

Leibowitz, A., 1974, "Home investments in children," *Journal of Political Economy,* 82:2, Part I, 111-131.

Liu, H., T. Mroz and W. van der Klaauw, 2003, "Maternal Employment, Migration, and Child Development", manuscript University of North Carolina at Chapel Hill.

Love, J.M., P. Schochet and A. Meckstroth, 1996, "Are They in Real Danger? What Research Does -And Does not- Tell Us About Child Care Quality and Children's Well-being", Mathematica No. ED145 030.

McCartney, K., 1984, "The Effect of Quality of Day Care Environment upon Children's Language

Development", *Developmental Psychology* 20:244-260.

Melhuish, E., A. Mooney, E. Hennesy and S. Martin, 1992, "Characteristics of Child Care in Early Childhood and Child Development in Middle Childhood" presented at the Europen Conference on Developmental Psychology in Seville, Spain. September.

Millward, C., 1998, *Family Relationships and Intergenerational Exchange in Later Life*, WP #15, Australian Institute of Family Studies, Melbourne.

Moffitt, R., 1992, "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature,* 30(1), pp. 1-61.

Murnane, R. J., J. Willet and F. Levy, 1995, "The growing importance of cognitive skills in wage determination," *Review of Economics and Statistics,* LXXVII, pp. 251-266.

National Research Council and Institutes of Medicine, 2000, *From Neurons to Neighborhoods: The Science of Early Childhood Development*, Shonkoff, J. and D. Phillips, eds. Washington D.C. NAP.

Neal, D. and W. Johnson, 1996, "The Role of Premarket Factors in Black/White Wage Differences," *Journal or Political Economy,* 104(5), pgs. 869-895.

NICHD Early Child Care Research Network, 2000, "The Relation of Child Care to Cognitive and Language Development", *Child Development* 71(4): 958-978.

Parcel, T. and E. Menaghan, 1990, "Maternal Working Conditions and Children's Verbal Facility: Studying the Intergenerational Transmission of Inequality from Mothers to Young Children," *Social Psychology Quarterly* 53: 132-147.

Parcel, T. and E. Menaghan, 1994, "Early Parental Work, Family Social Capital, and Early Childhood Outcomes," *American Journal of Sociology* 99(4), 972-1009.

Peterson, C. and R. Peterson, 1986, "Parent-child Interaction and Day Care: Does Quality of Day Care Matter?," *Journal of Applied Developmental Psychology*, 7, 1-15.

Robertson, D. and J. Symons, 1996, "Do Peer Groups Matter? Peer Group Versus Schooling Effects on Academic Attainment?," Centre for Economic Performance, London School of Economics, WP No. 311.

Ruhm, C., 2000, "Parental Employment and Child Cognitive Development", NBER WP No. 7666.

Smith, S., 1998, "The Past Decade's Research on Child Care Quality and Children's Development: What We are Learning, Directions for the Future," Paper prepared for a meeting on 'Child Care in the New Policy Context' held in Bethesda, MD on May 1998.

Stock, J. and M. Yogo, 2004, "Testing for Weak Instruments in Linear IV Regression," Working paper, Harvard University Department of Economics.

Studer, M., 1992, "Quality of Center Care and Preschool Cognitive Outcomes: Differences by Family Income," *Sociological Studies of Child Development*, Vol. 5, P. Adler and P. Adler, eds.

Todd, P. and K. Wolpin, 2003, "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal* Vol. 113(485): 3-33.

Vandell, D. and J. Ramanan, 1992, "Effects of Early and Recent Maternal Employment on Children from Low-income Families," Child Development 63(4), 938-949.

Vandell, D. and M. Corasaniti, 1990, "Variations in Early Child Care: Do They Predict Subsequent Social, Emotional and Cognitive Differences?, *Early Childhood Research Quarterly 5:555-572.*

Waldfogel, J., W. Han and J. Brooks-Gunn, 2002, "The Effects of Early Maternal Employment on Child Cognitive Development," *Demography*, May 39(2): 369-392.

Zax, J. and D. Rees, 1998, "Environment, Ability, Effort and Earnings", Center for Research on Economic and Social Policy, Denver, Working Paper No. 9801.

Table 1
THE EFFECT OF MATERNAL EMPLOYMENT ON CHILDREN'S COGNITIVE ABILITY
(Studies that use NLSY data)

Author, year	Sample	Method	Effect of mother's employment
Mott, 1991	2387 1-4 yr olds	OLS	Negative effects
Harvey, 1999	3-12 yr olds	OLS	Negative effects
Ruhm, 2000	3-6 yr olds	OLS	Negative effects
Han et al., 2001	462 birth-8 yrs	OLS	Negative effects
Bernal, 2005	529 3 to 7 yr olds	Structural model	Negative effects
Liu, Mroz & Van der Klaauw, 2003	5 to 15 yr olds	Structural model	Negative effects
Vandel & Ramanan, 1992	1889 2nd graders	OLS	Positive effects
Parcel & Menaghan, 1994	768 3-6 yr olds	OLS	Positive effects
Greenstein, 1995	2040 4-6 yr olds	OLS	Insignificant effects
Moore & Driscoll, 1997	1154 5-14 yr olds	OLS	Insignificant effects
James-Burdumy, 2005	498 3-4 yr olds	FE and IV-FE[1]	Differing depending on test used
Waldfogel, et al., 2002	1872 birth-8 yrs	OLS and FE	Differing depending on group
Desai, et al., 1989	503 4 yr olds	OLS	Differing depending on group
Baydar & Brooks-Gunn, 1991	572 4 yr olds	OLS	Differing depending on timing
Blau & Grossberg, 1992	8784 3-4 yr olds	OLS and IV[2]	Differing depending on timing
Todd and Wolpin, 2004	6-13 yr olds	IV Child FE	Not reported

[1] Household FE, and instruments are local market conditions, e.g., county unemployment rate and percentage of the labor force in the services sector

[2] Work experience prior to childbirth is the instrument for maternal employment.

Table 2

THE EFFECT OF CHILD CARE ON CHILDREN'S COGNITIVE ABILITY

Author, year	Sample	Method	Effect of child care use
Baydar and Brooks-Gunn, 1991	572 4 yr olds	OLS	Negative effects (vary with timing)
Desai, et al., 1989	503 4 yr olds	OLS	Negative effects (only for boys)
Vandell & Corasaniti, 1990	236 8-year olds	OLS	Negative effects
Thornburg et al., 1990	835 kindergarten children	OLS	Insignificant effects
Ackerman-Ross and Khanna, 1989	3-yr olds, whites	OLS	Insignificant effects
Parcel and Menaghan, 1990	697 3-6 yr olds	OLS	Insignificant effects
Studer, 1992	95 children	OLS	Insignificant effects
Burchinal et al., 1995	6-12 yr olds	OLS	Insignificant effects
Blau, 1999	2000+ 3-5 yr olds	OLS and FE[1]	Differing depending on quality
Caughy, et al., 1994	867 5-6 year olds	OLS	Differing depending on background
Dunn, 1993	4-yr olds, middle-class	OLS	Differing depending on quality of daycare
Clarke-Stewart et al., 1994	2-4 yr olds, middle class	OLS	Differing depending on quality of daycare
Ruopp, et al., 1979	1600 preschool children	Experiment[2]	Differing depending on measure of quality
Duncan and Currie, 1995	3477 4+ yr olds	Siblings FE	Positive effects of Head Start
Peisner-Feinberg et al., 2001	773 4 - 8 yr olds	OLS	Positive effects (of high quality daycare)
NICHD Early Child Care Research Network, 2000	595 0-3 yr olds	OLS	Positive effects of center-based arrangements
Duncan-NICHD, 2003	1162 24-54 months old	OLS and FE[3]	Positive effects (of high quality daycare)

[1] Household fixed effects.

[2] The National Day Care Study randomly assigned children to classrooms with different staff-child ratios and to teachers with different levels of training. However, the 64 day care centers were not randomly selected.

[3] Child fixed effects.

Table 3

List of Instruments

Variable	Description
Time Limits	
TLI_{st}	Dummy for whether state s has time limit in place in period t.
TL_LENGTH_{st}	Length of time limit in state s in period t.
$ELAPSED\_TL_{st}$	Time (in months) elapsed since the implementation of time limit in state s.
TL_HIT_{ist}	Dummy variable indicating whether a woman could have hit time limit
$ELAPSED\_TL\_HIT_{ist}$	Time elapsed since woman i may potentially be subject to time limit
$REMAIN\_TL\_ELIG_{ist}$	Maximum potential remaining length of a woman's time limit, constructed: TL_LENGTH_{st}-min$\{AGE_OLDEST_CHILD_{ist}, ELAPSED_TL_{st}\}$
$REMAIN\_ELIG_{ist}$	Remaining length of time to be categorically eligible for welfare benefits: 18-$AGE\_YOUNGEST\_CHILD_{ist}$
$DCHILDBEN_{st}$	Dummy variable indicating whether the child portion of the welfare benefit continues after time limit exhaustion
Work Requirements	
DWR_{st}	Dummy for whether state s has work requirement in place in period t.
WR_LENGTH_{st}	Length (in months) of work requirement limit in state s in period t.
$ELAPSED\_WR_{st}$	Time (in months) elapsed since the implementation of work requirement in state s.
WR_HIT_{ist}	Indicator for whether a woman could be subject to a work requirement: =1 if[$WR\_LENGTH_{st} \leq \min\{AGE\_OLDEST\_CHILD_{ist}, ELAPSED\_WR_{st}\}$ & $AGE\_YOUNGEST\_CHILD_{ist} \geq AGE\_CHILD\_EXEM_{st}$]
$ELAPSED\_WR\_HIT_{ist}$	Time elapsed since woman i may be potentially subject to work requirement
$CHILD\_EXEM_{st}$	Dummy for whether state s has age of youngest child exemption in place at t
AGE_EXEM_{st}	Age of youngest child below which the mother will be exempted from work requirement in state s at time t.
$WR\_ULT\_SANC_{st}$	Dummy for whether state s has a full sanction for non-compliance of work requirement in state s at time t.
$EXEMP_{st}$	Number of work requirement exemptions in state s
Earnings Disregards	
$FLAT\_DISREGARD_{st}$	Flat amount of earnings disregarded in calculating the benefit amount.
$PERC\_DISREGARD_{st}$	Benefit reduction rate (Does not include phase-out)
Other Policy Variables	
$BEN(1)_{st}$	Real AFDC/TANF maximum benefits for a family with 1 child
$EITC(1)_{st}$	EITC phase in rate constructed from both the federal and state level for a family with 1 child
$BEN(2)_{st}$	Real AFDC/TANF maximum benefits for a family with 2 children
$EITC(2)_{st}$	EITC phase in rate for a family with 2 children
$CHILDCARE_{st}$	CCDF expenditure per single mother in state s at time t.
$ENFORCE_{st}$	Child support enforcement expenditure in state s at year t per single mother.
Local Demand Conditions	
UE_{st}	Unemployment rate in State s in period t
$SWAGE_{st}$	Hourly wage rate at the 20th percentile of the wage distribution in State s in period t.

The instruments used in our baseline specification also include these policy variables and local demand conditions interacted with mother's education and AFQT score. In addition, workbef, EXPBEF, urban and age of mother (see definitions in Table 4) are interacted with child's age.

Table 4
Control Variables in the Cognitive Ability Production Function

Variable	Description
Baseline Specification	
$I[AGE_i<20]$	Dummy for whether mother is younger than 20 years old
$I[AGE_i>=33]$	Dummy for whether mother is older than 33 years old
$EDUC_i$	Mother's educational attainment at childbirth
$AFQT_i$	Mother's AFQT score
$I[WORK\_BEF]_i$	Dummy for whether mother worked prior to childbirth
$I[WORK\_BEF]_i$ x $SKILL_i$	Work dummy interacted with mother's skill*
$EXPBEF_i$	Mother's total work experience (in number of years) prior to childbirth
$EXPBEF_i *age_i$	EXPBEF interacted with mother's age
$MARAFT_i$	Mother's marital status at time of child's test
$URBAN_i$	Urban/Rural residence at time of child's test
$NUMCHILD_i$	Number of children
$RACE_i$	Child's race (1 if black/hispanic, 0 otherwise)
$GENDER_i$	Child's gender (1 if male, 0 if female)
BW_i	Child's birthweight
$AGECHILD_i$	Child's age at assessment date
$dPPVT_i$	Dummy for whether the corresponding test is PPVT
$dMATH_i$	Dummy for whether the corresponding test is PIAT-MATH
Alternative specifications also include	
AGE_i	Age of the mother at childbirth
AGE_i^2	Age of the mother at childbirth squared
$NUMCHILD_{0\text{-}5\,i}$	Number of children 0-5 years of age
$NUMCHILD_{6\text{-}17\,i}$	Number of children 6-17 years of age
$C_{ti} *EDUC_i$	Cumulative child care use interacted with mother's education
$C_{ti} *AFQT_i$	Cumulative child care use interacted with mother's AFQT score
$C_{ti} *NUMCHILD_i$	Cumulative child care use interacted with number of children

*The variable "skill" is defined as the residual from a regression of mother's initial wage on age, age squared, education and race.

Table 5
Mean Characteristics of Mothers in the Sample

Description	All mothers in NLSY	Single mothers at childbirth only	Single mothers for 5 yrs after childbirth	Our Sample
Mother's age at childbirth	24.8 (5.56)	23.56 (5.07)	23.80 (5.15)	23.13 (4.59)
Mother's education at childbirth (in years)	12.0 (2.475)	11.3 (1.920)	11.3 (1.917)	11.2 (1.909)
Mother's AFQT score	37.9 (27.23)	21.7 (20.09)	19.9 (19.11)	19.3 (18.30)
Hispanic or Black	0.47 (0.499)	0.73 (0.445)	0.79 (0.404)	0.83 (0.379)
Hourly wage before childbirth (first child)	6.32 (7.71)	4.74 (8.23)	4.90 (9.85)	4.39 (2.01)
Total number of children of mother	2.9 (1.37)	3.1 (1.57)	3.1 (1.61)	3.1 (1.53)
Father present at birth	0.55 (0.004)	-	-	-
Observations	4,814	2,528	1,820	1,464
Cases with wages at childbirth observed	2,622	1,208	753	670

Our sample screens are (1) The mother does not have a husband/partner for 5 years after childbirth and (2) The child has at least one test score observation.

Table 6
Summary of Variables used in the Empirical Analysis

Variable	Mean (standard error)	Variable	Mean (standard error)
log(Test Score)	4.49855 (0.1861)*	Urban	0.8189 (0.3851)
Mother's education	11.208 (1.8972)	Average yearly income (Thousands)	10.9274 (13.568)
Mother's age	23.136 (4.5820)	Cumulative income (Thousands)	51.1787 (67.415)
Boys (Children of single mothers)	0.4976 (0.5001)	Units of child care per quarter[1]	0.3546 (0.3064)
Hispanic or Black	0.8262 (0.3790)	Cumulative child care use (Quarters)	7.0923 (6.1273)
Birthweight	111.97 (21.976)	Labor participation rate (avg 1979-1993)	48.23 (6.5)
Mother worked before giving birth	0.6431 (0.4792)	Labor participation rate (avg 1994-1999)	60.40 (5.4)
Wage rate prior to giving birth	4.3938 (2.0075)	Welfare participation rate (avg 1979-1993)	58.93 (4.1)
Accumulated work experience prior to giving birth (number of years)	4.7202 (6.0088)	Welfare participation rate (avg 1994-1999)	44.65 (12.1)
Never married after childbirth	0.7215 (0.4483)	Childcare use rate (avg 1979-1993)[2]	59.05 (5.0)
Separated after childbirth	0.1540 (0.3611)	Childcare use rate (avg 1994-1999)	69.27 (5.9)
Divorced after childbirth	0.1158 (0.3201)		

*Standard error of log(test score) calculated after taking out the test-specific means of the three tests, i.e., the standard error of the residuals from a regression of log(test score) on test dummies PPVT and PIAT Math.

[1] One quarter of full-time child care use is 1 unit and one quarter of part-time child care use is 1/2 unit.

[2] It is equal to 1 if child care is used (either full-time or part-time).

Table 7
Explanatory Power of the Instruments

Input	Partial correlation squared	Shea partial correlation squared	R^2 with exogenous variables only	Incremental R^2	F-statistic	P-value
A. Endogenous variables in the baseline model						
Cumulative Child Care Use	0.1749	0.1495	0.4755	0.0917	15.560	0.0000
Current Child Care Use	0.1316	0.0988	0.3360	0.0825	10.910	0.0000
Cumulative Income	0.1111	0.1161	0.2158	0.0871	26.150	0.0000
Current Income	0.0791	0.0864	0.1128	0.0766	3.4900	0.0000
Number of Children	0.3942	0.3254	0.2468	0.2970	25.360	0.0000
B. Other endogenous variables in additional models						
Cumulative Formal Child Care	0.0944	0.1004	0.0898	0.0860	50.590	0.0000
Cumulative Informal Child Care	0.1393	0.1455	0.3381	0.0922	17.160	0.0000
Cumulative Child Care by Nonrelatives	0.0958	0.1008	0.0884	0.0874	1.8100	0.0001
Cumulative Child Care by Relatives	0.1267	0.1460	0.2254	0.0982	14.370	0.0000
Age Weighted Cumulative Child Care*	0.1928	0.1117	0.4532	0.1054	27.740	0.0000

Instruments are: variables in the main equation (see Table 4) plus mother´s age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and mother's AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence and both interacted with mother's education and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother (see definitions in Table 4).

Cumulative child care, cumulative income and number of children are predicted using lags and current values of the instruments listed above.

Current child care and current income are predicted using current values of the instruments listed above.

F-stat is cluster robust.

* Weights are the age of the child (in number of quarters)

Critical value at 1% is 1.47 (78 d.f. in the numerator and 1463 in the denominator)

Table 8
Do maternal time inputs matter for children's achievement?

Dependent Variable -> Log(Test Score)

	(1)	(2)	(3)	(4)
Cumulative Child Care	-0.00727		0.00080	
	(0.0030) **		(0.0008)	
Current Child Care		-0.00520		0.00440
		(0.0191)		(0.0027) *
Log(Cumulative Income)	0.00694	-0.00003	-0.00301	-0.00338
	(0.0266)	(0.0266)	(0.0057)	(0.0056)
Mother's education	0.01394	0.01133	0.01058	0.01040
	(0.0032) **	(0.0034) **	(0.0027) **	(0.0026) **
Mother's AFQT	0.00143	0.00136	0.00134	0.00134
	(0.0003) **	(0.0003) **	(0.0002) **	(0.0002) **
Estimation Method	LIML	LIML	OLS	OLS
Number of Observations	3,787	3,787	3,787	3,787
R-squared	0.3532	0.3755	0.3782	0.3786
k #	1.047	1.050		
Weak/Many Instruments Test	5.81	4.96		

Instruments are: variables in the main equation (see Table 4) plus mother´s age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and mother's AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence, and both interacted with mother's educations and AFQT and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother (see definitions in Table 4).

Cumulative child care is predicted using lags and current values of the instruments listed above. Current child care in (2) is predicted using current values of the instruments listed above.

Robust standard errors (Huber-White) by child clusters.

k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

Table 9
Heterogeneity in Effect of Maternal Time Inputs

Dependent Variable -> Log(Score)

	(1)		(2)		(3)	
Cumulative Child Care	-0.00620		-0.00629		-0.00732	
	(0.0031)	**	(0.0030)	**	(0.0030)	**
$\widetilde{\text{Education}}$*(Cum. Child Care)	-0.00174					
	(0.0010)	*				
$\widetilde{\text{AFQT}}$*(Cum. Child Care)			-0.00019			
			(0.00011)	*		
($\widetilde{\text{Number of Children}}$)*(Cum. Child Care)					-0.00114	
					(0.0032)	
Log(Cumulative Income)	0.01153		0.01389		0.00784	
	(0.0271)		(0.0282)		(0.0263)	
Mother's Education	0.02553		0.01279		0.01355	
	(0.0077)	**	(0.0033)	**	(0.0032)	**
Mother's AFQT score	0.00141		0.00315		0.00141	
	(0.0003)	**	(0.0010)	**	(0.0003)	**
No. of observations	3,787		3,787		3,787	
Estimation Method	LIML		LIML		LIML	
R-squared	0.3550		0.3617		0.3555	
k [#]	1.0453		1.0451		1.0470	
Weak/Many Instruments Test	5.82		5.62		4.80	

$\widetilde{\text{Education}}$=Education-$\overline{\text{Education}}$, where $\overline{\text{Education}}$ is the mean (same for number of children and mother's AFQT score)
Instruments are: variables in the main equation (see Table 4) plus mother's age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence, and both of these interacted with mother's educations and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother (see definitions in Table 4).
Robust standard errors (Huber-White) by child clusters.

[#] k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

Table 10
Comparison by Estimation Method

Dependent Variable -> Log(Test Score)

	(1)		(2)		(3)		(4)		(5)	
Cumulative Child Care	0.00080		-0.00447		-0.00477		-0.00725		-0.00727	
	(0.0008)		(0.0015)	**	(0.0021)	**	(0.0030)	**	(0.0030)	**
Log(Cumulative Income)	-0.00301		0.01430		0.00551		0.00694		0.00694	
	(0.0057)		(0.0090)		(0.0165)		(0.0265)		(0.0266)	
Mother's education	0.01058		0.01162		0.01289		0.01394		0.01394	
	(0.0027)	**	(0.0023)	**	(0.0030)	**	(0.0032)	**	(0.0032)	**
Mother's AFQT	0.00134		0.00121		0.00139		0.00143		0.00143	
	(0.0002)	**	(0.0002)	**	(0.0003)	**	(0.0003)	**	(0.0003)	**
Estimation Method	OLS		GMM		2SLS		FULLER[&]		LIML	
Number of Observations	3,787		3,787		3,787		3,787		3,787	
R-squared	0.3782		0.3659		0.3664		0.3534		0.3532	
k [#]							1.047		1.047	
Weak/Many Instruments Test			5.81		5.81		5.81		5.81	

Instruments are: variables in the main equation (see Table 4) plus mother´s age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and mother's AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence, and both interacted with mother's educations and AFQT and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother (see definitions in Table 4).

Note: Residuals in the score equation depart modestly from normality, exhibiting some skewness and excess kurtosis. Neither consistency nor asymtotic normality of LIML depend on normality, but small sample properties are presumably improved by the approximation being reasonably accurate.

Robust standard errors (Huber-White) by child clusters.

[#] k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

Table 11
A. Factor Analysis of the Instruments

Dependent Variable -> Log(Test Score)

	(1)	(2)	(3)	(4)	(5)
Cumulative Child Care	-0.00727	-0.00727	-0.00724	-0.00658	-0.00543
	(0.0030) **	(0.0026) **	(0.0026) **	(0.0024) **	(0.0023) **
Log(Cumulative Income)	0.00694	0.02809	0.02800	0.02527	0.02269
	(0.0266)	(0.0246)	(0.0245)	(0.0215)	(0.0200)
Mother's education	0.01394	0.01282	0.01281	0.01261	0.01220
	(0.0032) **	(0.0031) **	(0.0031) **	(0.0030) **	(0.0029) **
Mother's AFQT	0.00143	0.00129	0.00129	0.00129	0.00125
	(0.0003) **	(0.0003) **	(0.0003) **	(0.0003) **	(0.0003) **
Estimation Method	LIML	LIML	FULLER[&]	2SLS	GMM
Number of Observations	3,787	3,787	3,787	3,787	3,787
R-squared	0.3532	0.3507	0.3509	0.3552	0.3610
k [#]	1.047	1.009	1.008	-	-
Weak/Many Instruments Test	5.81	16.45	16.45	16.45	16.45

Instruments used in column (1) are described in the footnote in Table 10.

Instruments in columns (2) to (5) are 16 factors derived from the factor analysis of our original 78 instruments described in the footnote in Table 10.

[&] Fuller parameter=1. Robust standard errors (Huber-White) by child clusters.

[#] k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

B. Explanatory Power of Instruments in First Stage Regressions for Childcare (Instruments in Table 11A)

Dep var-> Cum. Childcare

Instruments listed in footnotes in Table 11A	Partial correlation squared	Shea partial correlation squared	Incremental R^2	F-statistic	P-value
(1)	0.1749	0.1495	0.0917	15.560	0.000
(2)-(5)	0.1123	0.1014	0.0589	15.020	0.000

R^2 of first stage regression with only exogenous variables=0.4755

Table 12
Robustness with respect to the Specification of the Main Equation

Dependent Variable -> Log(Test Score)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)**
Cumulative Child Care	-0.00678	-0.00743	-0.00766	-0.00727	-0.00764	-0.00730	-0.01468	-0.00738	-0.00675	-0.50584
	(0.0030) **	(0.0030) **	(0.0030) **	(0.0030) **	(0.0030) **	(0.0033) **	(0.0080) *	(0.0030) **	(0.0030) **	(0.2302) **
Log(Cumulative Income)	0.01072	0.02215	0.01009	0.00694	0.00910	0.06204	0.02996	0.00842	0.01824	0.26586
	(0.0264)	(0.0267)	(0.0288)	(0.0266)	(0.0289)	(0.0270) **	(0.0403)	(0.0269)	(0.0272)	(2.0502)
Mother's education	0.01171	0.01368	0.01510	0.01394	0.01517	0.01605	0.01607	0.01361	0.01434	1.07187
	(0.0029) **	(0.0030) **	(0.0034) **	(0.0032) **	(0.0034) **	(0.0037) **	(0.0043) **	(0.0034) **	(0.0032) **	(0.2548) **
Mother's AFQT score	0.00145	0.00132	0.00139	0.00143	0.00139		0.00106	0.00143	0.00127	0.12715
	(0.0003) **	(0.0003) **	(0.0003) **	(0.0003) **	(0.0003) **		(0.0005) **	(0.0003) **	(0.0003) **	(0.0251) **
Child's age	0.04162	0.03797	0.04170	0.04258	0.04190	0.02468	0.04058	0.04274	0.03954	3.64182
	(0.0133) **	(0.0137) **	(0.0141) **	(0.0133) **	(0.0141) **	(0.0139) **	(0.0148) **	(0.0133) **	(0.0134) **	(1.0007) **
Mother's age		-0.00322	-0.01966		-0.01479	-0.00558				
		(0.0016) *	(0.0118) *		(0.0157)	(0.0166)				
Mother's age squared			0.00035		0.00026	0.00004				
			(0.0002)		(0.0003)	(0.0003)				
I[age of mother_i<20]				0.02626	0.01061	0.01140	0.02452	0.02283	0.0191	1.92435
				(0.0116) **	(0.0154)	(0.0157)	(0.0123) *	(0.0149)	(0.0118)	(0.9259) *
I[age of mother_i>=33]				0.0057	-0.0023	0.0061	-0.0208	0.0048	0.0079	0.3154
				(0.0258)	(0.0323)	(0.0343)	(0.0395)	(0.0260)	(0.0261)	(2.2042)
Maternal employment[&]							0.2971			
							(0.3062)			
Number of children	-0.03321	-0.02683	-0.02523	-0.02811	-0.02511	-0.02301	-0.02030		-0.0231	-2.25421
	(0.0060) **	(0.0070) **	(0.0069) **	(0.0065) **	(0.0069) **	(0.0075) **	(0.0111) *		(0.0068) **	(0.5146) **
Number of children 0-5								-0.0267		
								(0.0070) **		
Number of children 6-17								-0.0311		
								(0.0105) **		
Year (at time of test)									-0.0031	
									(0.0014) **	
Estimation Method	LIML	LIML	LIML	LIML	LIML	LIML	LIML	LIML	LIML	LIML
Number of Observations	3787	3787	3787	3,787	3787	3787	3787	3787	3787	3787
R-squared	0.3529	0.352	0.3522	0.3532	0.3524	0.3284	0.3077	0.3519	0.3583	0.3757
k [#]	1.049	1.047	1.046	1.047	1.046	1.050	1.045	1.047	1.048	1.044
Weak/Many Instruments Test	5.85	5.70	5.30	5.81	5.42	6.71	3.09	5.79	5.88	5.81

Instruments are: variables in the main equation (see Table 4) plus mother´s age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and mother's AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence, and both interacted with mother's educations and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother

[&] Equals 1 if mother always worked since childbirth, 0.5 if she worked at least one quarter and 0 if she did not work at all (also instrumented for).

**Column (10) uses test scores in levels as the dependent variable. The mean score is 91.9, so the point estimate implies a day care effect of -0.55 per quarter, or -2.2% per year, similar to the log results.

Robust standard errors (Huber-White) by child clusters.

[#] k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

Table 13
Robustness with respect to mother's age

Dependent Variable -> Log(Test Score)

	(1)		(2)	
Cumulative Child Care	-0.00882		-0.00998	
	(0.0046)	*	(0.0052)	*
Log(Cumulative Income)	0.01715		0.02186	
	(0.0248)		(0.0261)	
Mother's education	0.01210		0.01230	
	(0.0050)	**	(0.0050)	**
Mother's AFQT	0.00154		0.00158	
	(0.0004)	**	(0.0005)	**
Estimation Method	LIML		LIML	
Number of Observations	1,643		1,680	
R-squared	0.3398		0.3323	
k [#]	1.080		1.086	
Weak/Many Instruments Test	4.63		4.39	

Instruments are: see footnote in Table 11
(1) Mothers 24 to 34 years old at childbirth
(2) Mothers 24 + years old at childbirth
Robust standard errors (Huber-White) by child clusters.
[#] k is the parameter of the k-class estimator.
** Significant at 5%, * Significant at 10%

Table 14
Robustness with respect to State Fixed Effects

Dependent Variable -> Log(Test Score)

	No State Fixed Effects		With State F.E.		State F.E. + year dummies		State F.E. + year dummies[&]	
Cumulative Child Care	-0.00727		-0.01430		-0.01267		-0.01403	
	(0.0030)	**	(0.0066)	**	(0.0069)	*	(0.0061)	**
Log(Cumulative Income)	0.00694		0.02656		0.01119		0.06675	
	(0.0266)		(0.0339)		(0.0467)		(0.0472)	
Mother's education	0.01394		0.01660		0.01796		0.01569	
	(0.0032)	**	(0.0040)	**	(0.0041)		(0.0039)	**
Mother's AFQT	0.00143		0.00149		0.00150		0.00115	
	(0.0003)	**	(0.0003)	**	(0.0004)		(0.0004)	**
Estimation Method	LIML		LIML		LIML		LIML	
Number of Observations	3,787		3,787		3,787		3,787	
R-squared	0.3532		0.3061		0.3294		0.3011	
k [#]	1.047		1.047		1.049		1.007	
Weak/Many Instruments Test	5.81		4.30		4.11		7.47	

Instruments are: see footnote in Table 11.
[&] Instruments are 16 factors derived from factor analysis of our original 78 instruments.
Robust standard errors (Huber-White) by child clusters.
[#] k is the parameter of the k-class estimator.
** Significant at 5%; * Significant at 10%

Table 15
A. Robustness with respect to the Instrument List

Dependent Variable -> Log(Test Score)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative Child Care	-0.00727	-0.00800	-0.01176	-0.00673	-0.00706	-0.00471	-0.00842	-0.00828
	(0.0030) **	(0.0029) **	(0.0043) **	(0.0026) **	(0.0027) **	(0.0028) *	(0.0026) **	(0.0027) **
Log(Cumulative Income)	0.00694	0.00820	0.02911	0.01946	0.01424	-0.00855	0.00003	0.00115
	(0.0266)	(0.0274)	(0.0440)	(0.0361)	(0.0492)	(0.0242)	(0.0270)	(0.0399)
Mother's education	0.01394	0.01425	0.01501	0.01115	0.01192	0.01279	0.01381	0.01218
	(0.0032) **	(0.0033) **	(0.0036) **	(0.0038) **	(0.0041) **	(0.0030) **	(0.0035) **	(0.0038) **
Mother's AFQT	0.00143	0.00143	0.00137	0.00134	0.00138	0.00149	0.00150	0.00150
	(0.0003) **	(0.0003) **	(0.0004) **	(0.0004) **	(0.0004) **	(0.0003) **	(0.0003) **	(0.0004) **
Estimation Method	LIML	LIML	LIML	LIML	LIML	LIML	LIML	LIML
Number of Observations	3,787	3,787	3,787	3,787	3,787	3,787	3,787	3,787
R-squared	0.3532	0.3486	0.3172	0.3491	0.3496	0.3635	0.3414	0.3352
k [#]	1.047	1.046	1.041	1.011	1.009	1.008	1.032	1.010
Weak/Many Instruments Test	5.81	5.82	5.20	7.03	5.67	9.09	6.28	6.98
Number of instruments	78	75	58	27	18	26	63	25

(1) All policy variables and local demand conditions in Table 3 (+ policy variables and local demand conditions interacted with mother's education and AFQT and and child's age interacted with workbef, EXPBEF, urban and age of mother). Unless otherwise noted, the specifications below still include these interaction terms.
(2) Excluding childcare expenditures (CCDF)
(3) Only work requirements, time limits and earnings disregards (excludes CCDF expenditures, CSE expenditures, EITC, local demand conditions and benefit amounts).
(4) Excluding time limits, work requirements and earnings disregards.
(5) Only benefit amounts and local demand conditions (unemployment rate and 20th percentile wage distribution).
(6) Instruments in (1) without all interactions.
(7) Includes only State-specific instruments (i.e., excludes all individual-specific welfare rules).
(8) Only State-specific instruments and excludes all interactions.
Robust standard errors (Huber-White) by child clusters.
[#] k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.
** Significant at 5%; * Significant at 10%

B. Explanatory Power of Instruments in First Stage Regressions for Childcare (Instruments in Table 15A)

Dep var-> Cum. Childcare

Instruments listed in footnotes in Table 14A	Partial correlation squared	Shea partial correlation squared	Incremental R^2	F-statistic	P-value
(1)	0.1749	0.1495	0.0917	15.560	0.000
(2)	0.1689	0.1438	0.0886	15.890	0.000
(3)	0.1321	0.0955	0.0693	19.240	0.000
(4)	0.1166	0.1108	0.0611	13.070	0.000
(5)	0.1060	0.0999	0.0556	17.970	0.000
(6)	0.1328	0.1072	0.0697	11.920	0.000
(7)	0.1457	0.1404	0.0764	10.140	0.000
(8)	0.1203	0.1158	0.0631	16.990	0.000

R^2 of first stage regression with only exogenous variables=0.4755

Table 16
Child Care Effects by Type of Care

Dependent Variable -> Log(Score)

	Mean (sd error)	(1)	(2)	(3)	(4)
Cumulative Informal child care	5.8533	0.00049	-0.00878		
	(5.873)	(0.0008)	(0.0032) **		
Relatives	5.00766			0.00021	-0.00981
	(5.7360)			(0.0008)	(0.0032) **
Nonrelatives	1.14537			0.00142	0.00661
	(3.3549)			(0.0011)	(0.0079)
Cumulative Formal child care	1.2229	0.00283	0.00245	0.00286	0.00442
(Daycare, Nursery, Pre-K, Other)	(3.055)	(0.0011) **	(0.0077)	(0.0011) **	(0.0076)
Log(Cumulative Income)	3.6332	-0.00316	0.00312	-0.00309	0.00538
	(0.730)	(0.0057)	(0.0256)	(0.0057)	(0.0254)
No. of observations		3787	3787	3787	3787
Method of Estimation		OLS	LIML	OLS	LIML
R-squared		0.3791	0.3405	0.3794	0.3081
k #			1.0450		1.0409
Weak/Many Instruments Instruments			4.49		4.47

Instruments are: variables in the main equation (see Table 4) plus mother's age and age squared, all policy variables (see Table 3), policy variables interacted with mother's education and mother's AFQT, unemployment rate in State of residence and average hourly wage at 20th percentile of wage distribution in State of residence, and both interacted with mother's educations and AFQT and child's age interacted with workbef, EXPBEF, urban and age of mother (see definitions in Table 4).

Robust standard errors (Huber-White) by child clusters.

k is the parameter of the k-class estimator, which equals 1 for 2SLS and exceeds 1 for LIML.

** Significant at 5%; * Significant at 10%

Table 17
Child Care Effects by Age

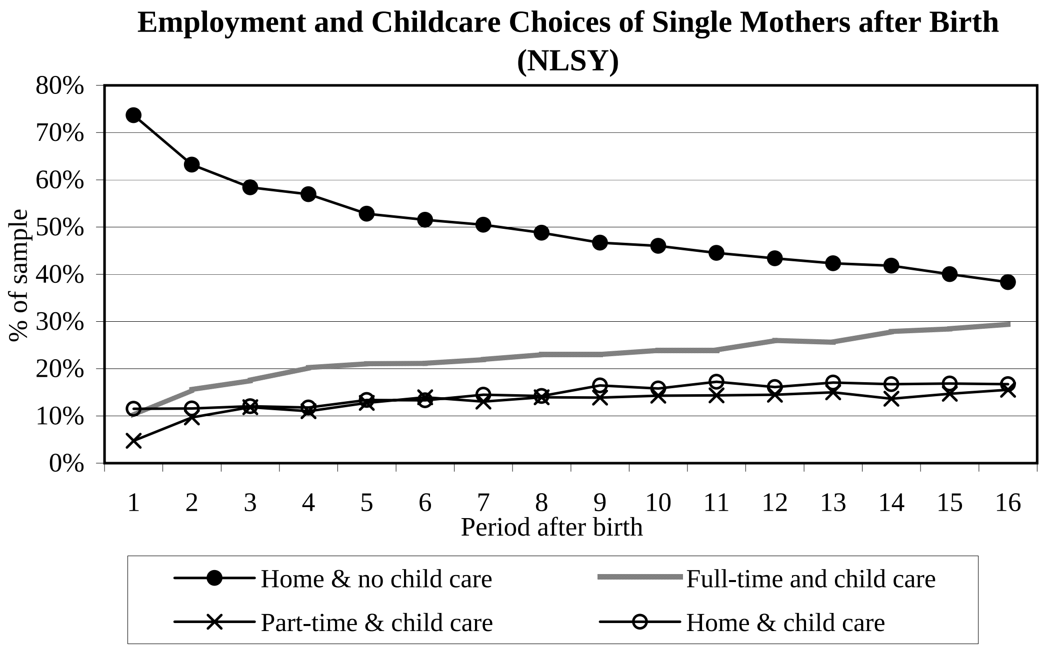
Dependent Variable -> Log(Score)

	(1)	(2)
Cumulative child care	0.00262	0.02234
	(0.0027)	(0.0151)
Age weighted cumulative child care&	-0.00020	-0.00308
	(0.0003)	(0.0015) **
Log(Cumulative Income)	-0.00318	0.00908
	(0.0057)	(0.0262)
No. of observations	3787	3787
Method of Estimation	OLS	LIML
R-squared	0.3783	0.3365
k #		1.04
Weak/Many Instruments Instruments		5.18

& Weights are the age of the child in number of quarters

k is the parameter of the k-class estimator.

Figure 1



Note: The first 12 periods correspond to quarters (3-months) and the last four correspond to semesters.

Appendix 1
Effect of early cognitive ability test scores on highest grade completed by 2000 (sample=young adults 18 years or older)

Dependent Variable -> Highest grade completed by 2000

	PPVT at age 4		PIAT Math at age 5		PIAT Reading at age 5		PIAT Math at age 6		PIAT Reading at age 6	
Test score	0.00819	0.01574	0.00633	0.01627	0.00960	0.02092	0.01908	0.03166	0.02493	0.03744
(Test and age in column heading)	(0.0041) **	(0.0035) **	(0.0046)	(0.0044) **	(0.0048) **	(0.0045) **	(0.0049) **	(0.0045) **	(0.0056) **	(0.0055) **
Age of completed education measure#	0.16449	0.06336	0.69629	0.68394	0.69097	0.66007	0.45305	0.41675	0.45629	0.40288
	(0.1563)	(0.1575)	(0.0752) **	(0.0717) **	(0.0758) **	(0.0723) **	(0.0438) **	(0.0409) **	(0.0439) **	(0.0411) **
Highest grade completed by mother	0.09231		0.05216		0.04901		0.09646		0.10179	
	(0.0403) **		(0.0348) *		(0.0343)		(0.0270) **		(0.0268) **	
Highest grade completed by father	0.02147		0.02069		0.01948		0.00833		0.01065	
	(0.0083) **		(0.0076) **		(0.0075) **		(0.0064)		(0.0064) *	
Number of siblings	-0.14160		-0.14066		-0.12912		-0.08883		-0.08942	
	(0.0586) **		(0.0543) *		(0.0535) **		(0.0428) *		(0.0424) **	
Birthorder	-0.11146		-0.13111		-0.09435		-0.11223		-0.07853	
	(0.0979)		(0.0957)		(0.0946)		(0.0754)		(0.0751)	
Race (1=Non-white)	0.06958		0.08739		0.06939		-0.06182		-0.21639	
	(0.1751)		(0.1523)		(0.1496)		(0.1258)		(0.1243) *	
Gender (1=Male)	-0.36024		-0.42114		-0.42716		-0.39478		-0.37505	
	(0.1380) **		(0.1236) **		(0.1228) **		(0.1011) **		(0.1008) **	
Mother's age at child's birth	-0.03878		-0.01219		-0.02523		0.02586		0.03390	
	(0.0387)		(0.0336)		(0.0331)		(0.0282)		(0.0280)	
Mother's AFQT score	0.00389		0.00378		0.00450		0.00128		-0.00030	
	(0.0038)		(0.0033)		(0.0033)		(0.0029)		(0.0028)	
Constant	7.2531	8.3078	-2.8778	-3.8171	-2.9770	-4.0097	-0.6925	-0.1622	-1.5869	-0.6049
	(3.1866) **	(2.9599) **	(1.8248)	(1.4088) **	(1.7977)	(1.3892) **	(1.2501)	-(0.1622) **	(1.2644)	(0.9295) **
No. of observations	363	363	451	451	446	446	747	747	739	739
Pseudo R-squared	0.1578	0.0537	0.2791	0.2014	0.2912	0.2209	0.2365	0.1761	0.2457	0.1760

All estimated by OLS. ** indicates significance at 5% and * at 10%.
The age of the young adult by 2000 if she is older than 18 years old. The average age is 21.8.
PPVT: Peabody Picture Vocabulary Test; PIAT: Peabody Individual Achievement Test

Appendix 2
Probit to predict child care choices of non-working women in years 4 and 5 after childbirth

Dependent Variable-> Pr(using child care in *t*)		
Whether worked before giving birth	0.5920	
	(0.208)	**
(Whether worked before) x (Avg. wage before)	-0.0642	
	(0.040)	*
Total work experience (prior to giving birth)	-0.0060	
	(0.019)	
Child's race	-0.0874	
	(0.170)	
Child's gender	0.0497	
	(0.120)	
Mother's education	0.0821	
	(0.038)	**
Total work experience since child birth	-0.3983	
	(0.070)	**
Total child care use since child birth	0.2226	
	(0.053)	**
Whether used child care or not in *t-1*	1.7801	
	(0.164)	**
Estimation	Probit	
Number of observations	867	
Pseudo-R^2	0.4585	

*Additional controls: Marital status at child birth (never married, separated,divorced, widowed),urban/rural residence and mother's age at birth.

**For women who reported working full-time in a given period after the third year, we imputed a child care value equal to 1; if the mother reported working part-time, we imputed a child care value equal to 0.5. Finally, if the mother does not work in a given period, we imputed a child care value of 0.5 if thepredicted probability of child care use based on this model exceeds 0.65.We choose this threshold to obtain a smooth trend of child care use since childbirth and until the end of the fifth year.

Appendix 3
Cognitive Ability Tests in our NLSY sample
Descriptive Statistics

	PPVT			PIAT - Math		PIAT-Reading	
Child's Age	3	4	5	5	6	5	6
Log(test score) in our sample	4.367	4.2689	4.402	4.539	4.543	4.633	4.606
	(0.191)	(0.295)	(0.239)	(0.152)	(0.128)	(0.152)	(0.095)
Test Scores in our sample	80.263	74.334	83.767	94.719	94.802	104.089	100.585
	(14.952)	(19.512)	(17.504)	(14.329)	(11.727)	(15.319)	(9.462)
Non-whites	78.007	70.836	82.135	93.836	94.247	103.358	100.482
	(14.169)	(17.958)	(16.889)	(14.289)	(11.685)	(15.454)	(9.269)
Whites	92.167	89.299	93.852	99.576	97.657	108.100	101.112
	(13.348)	(18.885)	(18.001)	(13.634)	(11.578)	(13.970)	(10.422)
Maternal education (12 yrs+)	82.820	78.748	88.743	97.084	96.823	106.755	102.265
	(14.369)	(18.917)	(17.648)	(14.178)	(11.663)	(15.131)	(9.425)
Maternal education (<12 yrs)	76.301	68.748	79.508	91.767	92.751	100.697	98.847
	(15.025)	(18.847)	(16.245)	(13.991)	(11.449)	(14.909)	(9.197)
Male	79.753	72.242	83.035	93.726	93.710	102.557	99.232
	(14.664)	(20.048)	(18.143)	(14.307)	(12.292)	(15.563)	(9.404)
Female	80.707	76.299	84.569	95.739	95.827	105.685	101.838
	(15.225)	(18.820)	(16.783)	(14.305)	(11.091)	(14.922)	(9.357)
No. of observations	339	512	438	598	663	584	653

PPVT: Peabody Picture Vocabulary Test
PIAT: Peabody Individual Achievement Test
Standard errors in parenthesis.

Appendix 4

Let S_3, S_4, S_5 and S_6 be the child's test scores at ages 3 though 6, respectively.[1] For example, S_3 can be the PPVT score at age 3.[2] In addition, let Y_3, Y_4 and Y_5 represent the endogenous variables that appear in the test score equation (10) in year 3, 4 and 5 after childbirth respectively. For example, Y_5 would include cumulative child care use up through age 5. Finally, let $R_1, R_2, ..., R_5$ represent vectors of instruments that are relevant for the mother's decisions in years 1 through 5 after the birth of the child. Thus, for example, R_5 would include welfare policy rules operative in the state of residence of the mother in year 5, and Y_5 is potentially influenced by R_1 through R_5. The first stage regressions in the 2SLS procedure will thus look like:

$$\begin{aligned}
Y_{3i} &= \alpha_0 + \alpha_1 R_{1i} + \alpha_2 R_{2i} + \alpha_3 R_{3i} + \alpha_4 \cdot 0 \;\; + \alpha_5 \cdot 0 \;\; + \underline{\alpha_6} X_{3i} + \varepsilon_i \\
Y_{4i} &= \alpha_0 + \alpha_1 R_{1i} + \alpha_2 R_{2i} + \alpha_3 R_{3i} + \alpha_4 . R_{4i} + \alpha_5 \cdot 0 \;\; + \underline{\alpha_6} X_{4i} + \varepsilon_i \\
Y_{5i} &= \alpha_0 + \alpha_1 R_{1i} + \alpha_2 R_{2i} + \alpha_3 R_{3i} + \alpha_4 . R_{4i} + \alpha_5 R_{5i} + \underline{\alpha_6} X_{5i} + \varepsilon_i
\end{aligned} \tag{A4.1}$$

where X_{ti} is a vector of exogenous characteristics of mothers and children that include all variables described in Table 4, and $\underline{\alpha_6}$ is an associated parameter vector. Notice that R_1 through R_t all enter the equation for Y_t. From A4.1, we obtain the fitted values $\hat{Y}_{ti}$ for t=3, 4, 5.

Finally, the second stage regressions in the 2SLS procedure would look like:

$$\begin{aligned}
S_{3i} &= \beta_0 + \beta_1 \hat{Y}_{3i} + \underline{\beta_2} X_{3i} + \xi_i \\
S_{4i} &= \beta_0 + \beta_1 \hat{Y}_{4i} + \underline{\beta_2} X_{4i} + \xi_i \\
S_{5i} &= \beta_0 + \beta_1 \hat{Y}_{5i} + \underline{\beta_2} X_{5i} + \xi_i \\
S_{6i} &= \beta_0 + \beta_1 \hat{Y}_{5i} + \underline{\beta_2} X_{6i} + \xi_i
\end{aligned} \tag{A4.2}$$

where β_1 is the parameter of interest. Notice that the test score at age 6 is only influenced by the endogenous variable dated at t=5 (i.e., cumulative day care use up through age 5), since at age 6 the child is of school age so day care is no longer necessary.

In the baseline specification, in order to avoid proliferation of parameters, we estimate a constrained version of the 1st stage regressions A4.1 where we assume that effects of the instruments on the endogenous variable Y_{ti} are the same in every year after birth, i.e., $\alpha_1 = \alpha_2 = ... = \alpha_5$.

---

[1] Recall that cognitive ability test scores are available as early as age 3 in the NLSY.

[2] Since we have quarterly data, a test score at 3 literally means a test score in the 12th quarter after the birth of the child.

Appendix 5
Average Test Scores for Children born prior to 1990

	Average	St. Dev	ttest
States that implemented TL waivers	93.34	(1.82)	-0.46
States that did not implement TL waivers	92.42	(1.08)	
States that implemented WR waivers	89.77	(1.35)	1.56
States that did not implement WR waivers	93.45	(1.09)	
States with TL lower than 3 years	90.2	(2.46)	0.87
States with TL higher than 3 years	93.02	(1.00)	
States with immediate WRs	93.48	(1.81)	-0.66
States with WRs of at least 1 month	92.20	(0.95)	
States with Age of Youngest child exemption < 6 months	93.40	(2.20)	-0.51
States with Age of Youngest child exemption > 6 months	92.38	(0.84)	

Source: NLSY, sample of single mothers

Appendix 6
Who is using formal child care and care provided by non-relatives?

Dependent Variable ->	1 if formal childcare used (0 if informal)	1 if care provided by non-relative (0 if relative)
Mother's education	0.12126	0.11247
	(0.0149) **	(0.0167) **
Mother's age at birth	-0.01140	0.02024
	(0.0056) *	(0.0061) **
Number of children	-0.08925	-0.04318
	(0.0191) **	(0.0213) **
Urban/rural	0.17590	0.63539
	(0.0637) **	(0.0798) **
No. of observations	12,167	9,471
Method of Estimation	Logit	Logit
Pseudo R-squared	0.0116	0.0209