Abstract

We investigate the question of whether investing in a child's development by having a parent stay at home when the child is very young (less than 6) has a causal impact on the child's adult outcomes. Specifically, do children with stay-at-home mothers have higher adult earnings or more labor force experience by a given age? Previous attempts to answer this question have been severely limited by the need to observe the child's outcome starting some 30 years after the initial investment. We overcome this data limitation by using administrative data that provides the complete earnings histories of parents and their children between 1951 and 2011. These long histories allow us to differentiate between different types of mothers and come closer to a causal impact of the mother's work history when the child is young. We show that when the endogeneity of the mother's history is ignored, we get the same result as many previous studies: the children of working mothers have lower salaries. However, when we control for endogeneity by using either the sibling estimator or an estimate of mother unobserved earnings heterogeneity, we find few significant differences between the adult earnings of children with stay-at-home mothers and those with working mothers. We do however see a significant positive effect of working on child labor force experience for families with low educated parents.

JEL Classifications: J13, J22, J24

Keywords: human capital, child development, female labor supply

The Impact of a Mother's Decision to Work on the Development of a Child's Human Capital*

Peter Gottschalk and Martha Stinson†

Nov 2012

1 Introduction

The sharp rise in the labor force participation rate of women, including women with young children, has led to a sharp increase in the proportion of children being raised by working parents. As a result, parents have increasingly relied on bought inputs, such as day care, as substitutes for the time they would have devoted to their children had they been stay at home parents. Whether the increased reliance on market inputs has had a long term effect on these children is the question we address in this paper.

A large body of literature has found a range of results on the impact of a mother working. Some studies have shown that the children of working mothers have lower mean outcomes on a variety of indicators of later success, such as reading and math scores in early elementary school (Ruhm (2004), Baker, Gruber, Milligan (2008), Waldfogel, Han, and Brooks-Gunn (2002)) while others have found limited or no effect (Blau and Grossberg (1992), Baker and Milligan (2010), Dustmann and Schonberg (2009)). But this evidence on the sign and magnitude of changes in indicators of later success, such as test scores in elementary school, is not evidence on the sign and magnitude of later success itself, as measured by adult outcomes, such as labor supply or earnings.

*We would like to thank Gary Benedetto, Graton Gathright, Christopher Wignall, and participants at the 2012 Census RDC Research Conference for helpful comments and suggestions. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau. All data used in this paper are confidential. The U.S. Census Bureau supports external researchers' use of some of these data through the Research Data Center network (www.census.gov/ces). For public-use data please visit www.sipp.census.gov/sipp/ and click "Access SIPP Synthetic Data."

†Peter Gottschalk is professor of economics at Boston College. Contact information: Economics Department, Boston College, Chestnut Hill MA 02467; peter.gottschalk@bc.edu. Martha Stinson is an economist at the U.S. Census Bureau, Social, Economic, and Housing Statistics Division. Contact information: 4600 Silver Hill Road, Washington, D.C. 20233; martha.stinson@census.gov.

Our paper fills this gap by using a longitudinal data set that is sufficiently long to allow us to measure the mother's work history before the child was born, during the formative years (0-5 years old), and after the child has left home. This data set also provides the child's earnings more than 30 years later. These rich data allow us to examine the relationship between the mother's work history and the child's adult labor market outcomes directly rather than having to rely on intermediary child outcomes as predictors of the child's adult outcome.

While we start by presenting descriptive statistics, such as the correlations between the mother's work history and the child's later outcomes, this is only a first step. The real challenge is to estimate the causal impact of the mother's work behavior on the child's adult outcomes. Does the mother's decision not to be a stay-at-home mom affect the child's outcome or is the correlation just spurious? This is the well known issue of endogeneity. High ability/high education mothers presumably command a higher wage in the labor market and hence are more likely to work. However, they might also be more likely to have high ability children. This will produce a spurious correlation between the mother′s work behavior when the child is young and the child's adult earnings.

We follow the literature by first conditioning on a rich set of observables, including family income of the parents, to see if this changes the estimated correlation. We then explore the role of unobservable characteristics of the parents by using our long earnings histories for both parents to estimate random effects which we then include in the child earnings equation in order to control for unobserved parental characteristics that might be correlated with work decisions and also produce high earning children[1]. Finally we estimate a model that exploits only the within family variation in the work behavior of the parents, where differences across families are not used. This within family estimator is the well-known sibling estimator that eliminates parent fixed effects since the siblings have the same parents.

We use a previously unexploited source of data, the Survey of Income and Program Participation (SIPP) linked to Social Security Administration (SSA) and Internal Revenue Service (IRS) lifetime earnings records, that enables us to account for a mother's and father's lifetime earnings as well as a mother's work patterns during the first five years of a child's life. This data set links parents and children living together at the time of the SIPP survey and provides annual earnings for all family members before, during, and after the survey, covering the period 1951-2010. Hence for those families that have been intact from the time of the birth of the child through the time of the survey, we are able to estimate total family earnings and mother work status for every year of the child's life until the time of the survey. We can then follow the child forward in time and observe earnings as an adult. We continue to observe parents' annual earnings forward in time and this allows us to gain precision in our estimate of unobservable factors, such as ability, that affect parent's earnings. These

[1] The random effects we use in this paper are not the typical random effects used by economists but rather are predicted values resulting from solutions to Henderson (1953) mixed effects equations. See Section 4.1 and Searle, Casella, and McCuloch (1992) for details. The appendix to this paper summarizes this literature.

data allow us to estimate the impact of a mother working on her child's labor market outcomes after age 30 while controlling for family income, mother and father education and unobserved ability. We find that the effect of a mother working varies by the education level of the parents but, after controlling for family characteristics and parental unobservables, we find substantial positive effects on labor supply that are statistically significant but few significant effects of a mother working on the child's adult wages.

Our paper proceeds as follows: first, we discuss the relevant literature and our contributions; second, we present a simple analytical model that focuses on the factors that influence the parent's decision whether to invest their own time in their child's future earnings or to rely more heavily on bought inputs; finally, we describe our data, present our statistical model, and discuss results. We conclude with an assessment of what we learn from this work.

2 Literature Review

As Almond and Currie (2010) describe, the last ten years have seen rapid growth in the number of studies on the impact of *in utero* or early childhood events on adult outcomes. The general idea has been that shocks to health in the early years of life may have a sustained impact on the human capital development of a child. Investments by parents or government may offset these effects, but not all investment are equal. Shocks or investments that happen in the first five years of life may have a differential impact than those which happen later. Our paper seeks to add to this literature by considering the investment of a mother's time in the first five years of a child's life and whether the effects of this investment depends on the education and ability level of the mother.

Existing studies have examined the impact of the mother working on near-term cognitive and behavioral outcomes for children. Most studies find either no effects or small negative effects only in early years. For example, Blau and Grossberg (1992) find a negative impact of the mother working during the first year of a child's life on cognitive development but find a positive impact of working from ages 2-3. In comparison, Ruhm (2004); Baker, Gruber, and Milligan (2008); Han, Waldfogel, and Brooks-Gunn (2001); and Waldfogel, Han, and Brooks-Gunn (2002) find small negative impacts of maternal employment on the cognitive development of children into their early elementary school years. Baker and Milligan (2010) and Dustmann and Schonberg (2008) find no significant impact of maternal employment on child outcomes.

Each of these papers recognizes the possible endogeneity of mother working. Blau and Grossberg adopt an instrumental variables approach for mother working,[2] Baker and Milligan, Dustmann and Schonberg, and Baker et al. use natural experiments stemming from changes in maternity leave laws, and Ruhm, Han et al., and Waldfogel et al use extensive mother, child, and family characteristics, including income, to control for observable characteristics.

[2] The authors use predicted values of mother labor supply generated from a two-limit Tobit analysis as instruments for actual labor supply in the first years of the child's life.

The two papers whose methodologies are most similar to ours, Ruhm (2004) and Blau and Grossberg (1992), both used the NLSY to provide a sample of linked mothers and children where both mother labor supply and child outcomes are observed, in this case test scores from age six or younger. Blau and Grossberg, initially controlling for only spouse's income and not the mother's earnings, find that children under age one with working mothers scored lower on a standardized vocabulary test that those with mothers who did not work. However, if the mother worked after the child's first year, the effect of working on test scores was positive. When the mother's earnings were controlled for, though, there was no significant effect of working after the first year. They conclude that the initial positive coefficients they found on working after the child had passed age one largely reflected the positive impact on children of having more income when the mother worked. In contrast, Ruhm finds negative and significant effects of longer hours of working during the first three years of a child's life on the math and reading scores of five and six year olds but no significant effect of income on test scores.

This literature on the effect of mother employment is complemented by the literature on the effect of income on child well-being and future outcomes. Blau (1999) models young child cognitive outcomes and treats income as endogenous. He finds that the effect of current income is small but that permanent income has a larger effect. Still he estimates that the significant effects in his models represent only modest impacts of changes in family income on the child outcomes he studies. He cites work from the literature on intergenerational income correlations to discuss how income impacts future child earnings. Early landmark studies in this area, such as Solon (1992)and Zimmerman (1992), find correlations between father and son earnings of approximately .4. Income correlation studies that control for family background measures, such as Haveman and Wolfe (1994), find almost no effect of income on child outcomes such as educational attainment and labor market participation. Blau concludes that family background and parent characteristics have a greater impact on the adult earnings of children than does family income.

The primary contribution of our study to this literature is that we are able to directly measure the correlation between mothers' labor market participation when their children are young and the adult outcomes of these children. This is in contrast to almost all other studies that examine the correlation between mothers' labor force participation and the childhood outcomes of their children, such as test scores when they enter school. These studies have to assume that these intervening child outcomes are themselves correlated with the future adult well-being. Dustmann and Schonberg (2008) is the only study that uses adult outcomes of the children. Like our paper, they use federal administrative data on wages and employment when the child is an adult (age 24 or 25) as outcome variables, although from Germany instead of the United States. They also use German state administrative data on child educational track choice. In contrast to our paper, they do not observe mothers linked to children in their administrative data and hence they rely on a change in German law in 1979 which extended maternity leave from two months to six months to identify the effect of maternal employment on child outcomes.

Our study also contributes to the literature by allowing a richer structure in this intergenerational correlation. We allow the impact of the mother's labor force participation on the child's adult outcomes to differ by observed characteristics of the mother, such as mother's education, and on the unobserved characteristics that affect the mother's and father's own wage rates. We are able to explore whether mothers with above average earnings who stay home with their children have a larger impact on their children's adult outcomes.

3 Analytical Framework

3.1 Basic Model

In this section we present a very simple model that can be used to analyze a parent's labor supply decision when the parent's utility depends not only on present consumption but also on the future labor market outcomes of the child. In order to focus on the key elements, we consider the work decisions of only one parent and ignore the other parent, assuming labor supply of the second parent is fixed. Adding a second parent would be straightforward but would raise a set of additional issues that would detract attention from the issues on which we focus. For ease of exposition we assume that the mother is the parent making the work decision[3].

Let $U(C_t, Y_t)$ be the household's utility function, where C_t is the household's consumption in time period t and Y_t is a $Kx1$ vector of the expected value of the K children's future outcomes. In our case Y_t is the present value of the future earnings of each of the K children who are less than six years old.

3.2 Externalities

The k^{th} child's future outcome depends on h_t^k, the number of hours the mother invests in the child's development, and on m_t^k, the market goods, such as day care and enrichment programs, that the mother buys as inputs into the child's development. Furthermore, if there is more than one child, then parental and bought inputs for one child may have externalities for the other child. For example, when the parent reads or tells a story, more than one child may benefit. This implies the following production function

$$Y_t = f(h_t^1, .., h_t^K, m_t^1, .., m_t^K, x, \theta^1, .., \theta^K) \tag{1}$$

where x is a vector of observable, exogenous variables and θ^k is a vector of time invariant unobservable traits that affect the k^{th} child's outcome, such as the parent's and child's natural intelligence.

The opportunity cost of staying home with a child is to work in the market where the mother can earn w_t^m. Her wage is determined by a standard human

[3] For the time period we study - children born in the 1970s - these stereotypic roles for men and women were still common.

capital production function that includes observables, x_{mt}, such as education and experience, and unobservables, θ^m, such as ambition and networking ability. If the mother works in the market, she buys day care and market goods at price P_m to substitute for her own input.

The household maximizes utility subject to an hours constraint, $H^m = h_t^m + \sum_k h_t^k$ and an intertemporal budget constraint $\sum_t \frac{w_t^m h_t^m}{d_t} = \sum_t P_{c_t} C_t / d_t + P_m \sum_k m_t^k$ where H^m is total hours available to the mother, h_t^m is the hours the mother works in the market, P_{c_t} is the price deflator for the consumption bundle, C_t is all other household consumption, and d_t is a discount factor.

Consider first the case of a family with only one child. The production function is: $Y_t^1 = f_1(h_t^1, m_t^1, x, \theta^1)$ and the first order conditions are familiar:

$$\frac{\partial U}{\partial Y_t}\frac{\partial Y_t}{\partial h_t^1} = \frac{\partial U}{\partial Y_t}\frac{\partial f_1}{\partial h_t^1} = \lambda w_t \tag{2}$$

$$\frac{\partial U}{\partial Y_t}\frac{\partial Y_t}{\partial m_t^1} = \frac{\partial U}{\partial Y_t}\frac{\partial f_1}{\partial m_t^1} = -\lambda P_m \tag{3}$$

$$\frac{\partial U}{\partial C_t} = -\lambda P_{c_t} \tag{4}$$

These imply the standard result that a mother will choose her hours of market work such that the marginal utility of the marginal increase in child human capital will equal the cost of that increase, her foregone wages. Likewise, market inputs into the child's human capital production function will be expanded until the marginal gain is equal to the marginal cost. In addition, the ratio of these two marginal utilities must equal the ratio of the costs of both inputs. Finally the ratio of the marginal utility of child human capital to the marginal utility of consumption must equal the ratio of the marginal cost of consumption to the marginal cost of the child human capital input. Thus mothers with equal ability to produce child human capital (i.e. the same production function), and equal enjoyment of child outcomes (i.e. the same utility function) will chose different amounts of time to spend with their children depending on the wage they face in the labor market. A mother will work until the marginal return of an additional hour would not buy adequate market goods to replace her time in the production of child human capital. Likewise, a mother will spend time with her child until the increase in child human capital is too small to compensate for the lost consumption. If a mother gets utility from spending time with her child, then the first order conditions would have an additional term $\frac{\partial U}{\partial h_t^1}$ and the wage required to compensate for the non-market time with the child would have to be higher.

In a cross-sectional setting, it is difficult to estimate $\frac{\partial f_1}{\partial h_t^1}$ because it is difficult to control for θ^1, the unobservables that influence child human capital production. We also expect that θ^1 and θ^m are correlated, i.e. that a mother who is successful in the labor market may also be good at child rearing and that high ability mothers may also have high ability children. Considering how to hold unobserved ability constant is thus the thrust of any econometric specification that uses cross-sectional covariation in mother's work choices and eventual labor market outcomes of the child.

Now consider the implications of having externalities in the production of Y. In this case the first order conditions depend crucially on the number of children. If the family grows and there are now two young children to raise, then the production function incorporates these externalities :

$$Y_t = f_2(h_t^1, h_t^2, m_t^1, m_t^2, x, \theta^1, \theta^2)$$

and the first order conditions are:

$$\frac{\partial Y_t}{\partial h_t^1} = \frac{\partial f_2}{\partial h_t^1} + \frac{\partial f_2}{\partial h_t^2}\frac{\partial h_t^2}{\partial h_t^1} = \lambda w_t \tag{5}$$

$$\frac{\partial Y_t}{\partial m_t^1} = \frac{\partial f_2}{\partial m_t^1} + \frac{\partial f_2}{\partial m_t^2}\frac{\partial m_t^2}{\partial m_t^1} = -\lambda P_m \tag{6}$$

For example, consider the case of a mother with a high paying job who uses bought inputs to replace some of her time when she has only one child. The following year she has an additional child so there is the possibility that if the externalities in own time are sufficiently large then the mother will stop working in the market. In other words, the marginal utility of her time in child human capital production now exceeds her wage because time at home helps two children and so she chooses to stay home. Likewise, a mother who stays home with the first child may go to work in the market if the externalities to bought inputs are sufficiently large. For example, if the increase in the production of human capital by a child care provider is sufficiently high to outweigh the increase in the cost of the care, and the increase in earnings resulting from more work is sufficient to cover the cost of the child care, then the mother will return to work.

Parents who are highly productive in the labor market will tend to substitute away from own time spent with children towards purchased inputs which include day care arrangements and also extra activities such as SAT prep classes, music lessons, and academic enrichment activities. Other parents who earn less in the labor market but are equally productive in child care will stay at home with their children since their opportunity cost of staying home is low.

The child care decision will also depend on access to alternative sources of income. The most disadvantaged children may be those whose mothers are low-skilled in both paid employment and child human capital development and who stay home because they cannot find employment that will generate the income necessary to cover the day care costs. Alternatively, some of the most advantaged children are those of highly educated mothers and fathers. The high family income from either parent working allows them to buy an amount of market goods such that the marginal value of market goods is low. Hence, if the utility they receive from spending time with their children and from the future labor market outcomes of their children compensates for the foregone current income and consumption, the second parent may choose not to work in the paid labor market.

3.3 Identification

While the purpose of this analytical model is to provide a framework in which to analyze the impact of the parent's labor market decision, it also offers an identification strategy for dealing with a key endogeneity issue. Parents who work when their children are young are likely to have unobservable traits that lead to greater labor market activity. If the child shares these unobserved traits, then the parents' labor market outcomes are endogenous when included in the child's earnings equation. This problem can be addressed by identifying differences between some parents staying home and others working when this allocation mechanism is independent of the child's unobservables.

The analytical model in this section suggests that parents change their labor market behavior in response to the externalities from taking care of an additional child. Some parents gain greater externalities by expanding their own input into their children's future labor market outcomes than by expanding bought inputs. Hence working parents will decrease their labor market activity when the second child is born if the economies of scale for own inputs are sufficiently high relative to bought inputs.

In these models it is differences in economies of scale (such as the cost of child care for one versus two children) that determine the parent's investment of own time in their children's development. If these differences in economies of scale are independent of the child's ability then endogeneity is not a problem. Assignment is random with respect to the child's unobservables. If this is the case then one can use within family variation in parental inputs in their children to obtain consistent estimates of the parameters of the model.

4 Estimation

4.1 Statistical Model

Somewhat more formally, consider the following linearized version of the relationship between the child's adult outcome, Y_{it}, and mother labor force participation when the child is young, W_i :

$$Y_{it} = X_{it}\beta + \gamma W_i + \theta_i + \eta_{it} \tag{7}$$

where θ_i represents unobserved time-invariant characteristics of child and η_{it} captures time-period specific variation. The objective is to obtain consistent estimates of γ. If W_i is not independent of θ_i or η_{it} then this endogeneity must be modeled explicitly. Our assumption is that W_i is correlated with θ^m and θ^f, the unobserved time-invariant characteristics of the parents, which in turn are correlated with θ_i.

We estimate two models that address this endogeneity problem. The first model attempts to obtain direct estimates of mother and father unobserved ability and include these as controls in the child earnings regression. To do this, we first use a mixed-effects model to estimate earnings equations for mothers and fathers of the following form:

$$Y^m = X^m\beta + Z\theta^m + \eta \tag{8}$$
$$Y^f = X^f\beta + Z\theta^f + \eta \tag{9}$$

We describe mixed-effects models in detail in Appendix A, Section 9. These models are more general than the fixed effects or random effects models common in the econometrics literature. In particular they do not impose orthogonality between X and Z and they not only provide $\widehat{\beta}$ but also $\widetilde{\theta_i^m}$ and $\widetilde{\theta_i^f}$. We treat these predicted values of the random effects, $\widetilde{\theta_i^m}$ and $\widetilde{\theta_i^f}$, as measures of mother and father unobserved ability. By definition they are centered at zero and rank mothers and fathers relative to each other in terms of earnings that are not explained by observable characteristics such as education. We then estimate a simple OLS child-level earnings model which includes demographic characteristics of the child, the work decision of the mother during the first five years of the child's life, mother and father education, average total parent earnings over the first five years of the child's life, and $\widetilde{\theta_i^m}$ and $\widetilde{\theta_i^f}$. The actual specification of equation 7 that we estimate is

$$\begin{aligned} Y_{it} &= X_{it}\beta^c + \gamma W_i + \beta^{m1} MotherSomeCollege + \beta^{m2} MotherCollege \\ &\quad + \beta^{m3}\widetilde{\theta_i^m} + \beta^{f1} FatherSomeCollege + \beta^{f2} FatherCollege + \beta^{f3}\widetilde{\theta_i^f} + \eta_{it} \end{aligned} \tag{10}$$

where X_{it} includes the average combined earnings of both parents when the child was under the age of 5. The excluded education group for both parents is high school degree or less. The coefficient γ then becomes the effect of working for mothers with the same level of education, the same $\widetilde{\theta^m}$, and married to men with the same level of education and same $\widetilde{\theta^f}$, in families that have the same average total parent earnings. In this simple model, the mother working can affect child earnings by increasing formal child education, child labor force experience, or unobserved child skills, all of which affect child earnings. We do not attempt to distinguish the means of the impact but simply try to estimate whether there is an effect of the mother working, all else equal in early family circumstances.

We also estimate several versions of equation 10 with interactions between parent characteristics and mother work decisions. We first interact mother education and unobserved ability with mother working. Thus γ becomes a vector with three separate effects of working, one for each mother education group $\left(\gamma^{me1}, \gamma^{me2}, \gamma^{me3}\right)$, and a slope parameter, γ^{ma}, that tells how the effect of unobserved mother ability changes for working mothers. This specification can be written as:

$$
\begin{aligned}
Y_{it} &= X_{it}\beta^c + \gamma^{me1}W_i * MotherHS + \gamma^{me2}W_i * MotherSomeColl \quad (11)\\
&+\gamma^{me3}W_i * MotherColl + \gamma^{ma}W_i * \widetilde{\theta_i^m} \quad (12)\\
&+\beta^{m2}MotherSomeColl + \beta^{m3}MotherColl + \beta^{m4}\widetilde{\theta_i^m}\\
&+\beta^{f2}FatherSomeColl + \beta^{f3}FatherColl + \beta^{f4}\widetilde{\theta_i^f} + \eta_{it}
\end{aligned}
$$

We next interact mother and father education to create a nine-category parental education variable. We then create effects in our model for mother working and for the mother and father random effects interaction with mother work status that differ for each parental education category. This differentiates the effect of working for women with different education levels, different levels of unobserved ability and different types of husbands, again creating a vector γ that contains the effect of working for each mother-father education pair and slope parameters for each mother-father education pair that describes the marginal effect of mother and father ability for working mothers.

Our second model is a more traditional fixed effect model. Because for some families we observe multiple children per mother in our data, we ask what the effect of the mother working is when comparing siblings. This child earnings model uses variation across siblings to identify the effect. In this model, γ becomes the effect of working when comparing two children of the same mother, one of which had a working mom at age five and under and the other of which did not. We compare boys to their brothers and girls to their sisters in this mother fixed effect model so as to hold the gender of the child constant. Although the sample sizes become small because not all families have multiple children of the same gender, this specification provides a check for our OLS model with predicted parent random effects as controls.

We also use these same models to estimate the effect of the mother working on child labor force experience. We calculate total years a child has worked by the year 2011 and regress this on the same controls as for earnings, including mother and father education, average total parent earnings over the first five years of the child's life, and $\widetilde{\theta_i^m}$ and $\widetilde{\theta_i^f}$. We also estimate a mother fixed effect model using child labor force experience as the dependent variable.

5 Data

We now turn to a description of our data. Our data come from a linked survey-administrative database created by the Census Bureau using the Survey of Income and Program Participation (SIPP). This database is called the SIPP Gold Standard File (GSF) and contains all SIPP respondents from the 1984, 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008 SIPP panels. For a subset of the questions asked by the survey, consistent variables are created across all nine panels. For individuals where a validated Social Security Number is obtained, they are then linked to IRS and SSA administrative data on earnings, specifically data drawn from SSA's Master Earnings File (MEF).[4] This section describes how we chose children to include in our research sample, how Census obtained validated SSNs for these survey respondents, and how we created our analysis variables from the underlying administrative data.

To answer our question of the effect of early maternal employment on child outcomes, we needed children for whom we could observe their mothers' work patterns and total family income from birth and their own earnings as an adult. Due to concerns about the duration and timing of post-secondary schooling and the impact of frequent early-career job transitions on earnings, we chose to examine earnings for children from the year they turned 30 forward.[5] Since our last year of administrative data comes from 2011, this required all our children to be born in 1981 or earlier. This restriction determined the lower bound on the age of the children we chose from each panel. In the 1990 panel, for example, any child younger than 9 years old was not old enough to be included in our sample. We also chose children who were 18 or younger in each panel because children who live at home after 18 might have different earnings trajectories than children who leave home. This upper bound on the age of the children we selected imposes a lower bound on the birth year. For example in the 1990 panel, children could not have been born any earlier than 1972. In the 1984 panel, we imposed one additional restriction and only chose children who were at least 5 years old (i.e. birth year of 1979 or later). While younger children would have still turned 30 by 2011, we did not include these cases because of concerns that the household might have dissolved before the child turned 5 and our measure of family earnings would be inaccurate.

Concerns about the measurement of family earnings led to another sample restriction, namely that children live with both biological parents at the time of the survey. This restriction amounts to choosing to study families that have been intact from the birth of the child to the time of the survey, when the child is in his or her teenage years.[6] Because we cannot observe what happens to the family after the SIPP panel ends, it is important that the survey look retrospectively at the first five years of childhood. Otherwise, while we could continue to tell what the mother of a 4 year-old earned after the survey finished, we could not tell if both parents remained in the family together with the child, and hence we could not calculate total family earnings. Likewise if a family had broken up before the survey, and one parent was no longer living in the household, we would obtain no information about the identity of this parent and no past earnings history, leaving total family earnings when the child was young unknown.

Essentially the survey serves the purpose of identifying and linking together

[4]The SIPP GSF is the base file used to create the SIPP Synthetic Beta (SSB), a public-use product that uses data synthesis methods to protect the confidentiality of the linked data.

[5]Since the SIPP panels end before these children reach adulthood, we do not know educational outcomes for the children in our sample.

[6]We make the assumption that a biologically based family relationship observed when the child was surveyed in his or her teens years also existed when the child was under the age of 5 and do not explicitly require that the parents were married to each other at any point in time.

the parents and children and so provides the point of reference for choosing children of particular ages. Parent and child relationships are taken from the household roster compiled at the time of the interview.[7] From the survey, we also make use of the mother's self-reported fertility history to determine if the child was the oldest, youngest, or a middle child. We use self-reported race of both the parents and the children and the education of the parents. Unfortunately the SIPP panel is not long enough to observe complete educational outcomes for the children.

We rely on the administrative earnings histories to provide labor force participation and earnings information for both parents and children prior to, during, and after the survey was conducted. We obtain these histories by merging IRS/SSA records using each respondent's Social Security Number (SSN). For the panels we use, the Census Bureau asked each SIPP respondent at the time of the survey to provide an SSN. SSA then compared demographic information (name, sex, race, and date of birth) from the survey reports and the Numident, an administrative database containing demographic information collected upon issuance of the SSN. If a respondent's name and demographics were deemed to match between the two sources, the SSN was declared valid. This list of validated SSNs was the basis for creating a research extract from the MEF. Due to our reliance on the administrative data earnings history, we only include children in our sample when we obtain a valid SSN for them and both of their parents.

Our data actually consist of two types of extracts. The first, called the Detailed Earnings Record (DER), contained uncapped earnings for all employment relationships of SIPP respondents that produced W-2 forms from 1978-2011. The second, called the Summary Earnings Record (SER), contained earnings capped at the FICA-taxable maximum from FICA-covered jobs from 1951-1977. We combine the SER and the DER to make a single administrative earnings history for each individual and create annual total earnings by imputing earnings above the taxable maximum for years covered only by the SER (prior to 1978) and summing across employers for the DER (after 1978).[8] This combined SER/DER work history has one serious drawback in that prior to 1978, only FICA-covered earnings are recorded. While today the vast majority of jobs are subject to FICA taxes, in the 1960s and 1970s, this was not necessarily the case. Thus we are concerned about missing parental, especially maternal, employment prior to 1978. To avoid this problem, we imposed one final restriction on the birth year of children, requiring them to be born in 1978 or later. This

[7] In the 1984 - 1993 SIPP panels, the relationship between a child and one parent is reported in the core data files by including the parent person number on the child's record. Most commonly this parent identifier links to the mother. The second parent's information must be obtained from the topical module that reports household relationships in a matrix form, i.e. person A's relation to person B, person C., etc. Beginning in 1996, links to both parents are included on a child's record in the core files and the topical module is not necessary for defining the parent-child relationship.

[8] The DER has both FICA and non-FICA covered earnings for each year as well as deferred wages. We sum all these amounts when creating annual total earnings. Both the DER and the SER include self-employment reports taken from Schedule C of income tax returns.

restriction, combined with the earlier requirements of turning 30 by 2011 and not being older than 18 at the time of the survey, leaves us with a sample of children born between 1978 and 1981, surveyed in the 1984, 1990, 1991, 1992, 1993, and 1996 SIPP panels, and age 30 to 33 in 2011.

We define a person as working in a given year if he or she has positive annual earnings, and we create years of labor force experience measures by summing these work indicators over time. Given our restriction to births happening in 1978 or later, we can create maternal work indicators for each year the child is under 5 and average combined mother and father earnings when the child is between 0 and 5 years old, solely from the DER. We also rely on the DER for child earnings at age 30 and beyond and for child labor force experience. We do use the combined DER/SER history when estimating the parent random effects so as to take account of all available earnings history data both before and after the birth of the child. We measure age for both children and parents using the birth date from the Numident. Mother's age at birth of the child is the difference between the mother and child administrative birth dates.

To arrive at our final sample, we drop children who themselves never had positive earnings after age 30, whose parents never had positive earnings between 1960 and 2011, or whose parents had no positive earnings in the years when the child was between the ages of 1 and 5. We also require non-missing education, race, and fertility variables from the SIPP. We analyze boys and girls in separate regressions due to the different nature of labor force participation between girls and boys. These sample restrictions focus this analysis strictly on children who are working as adults, had at least one working parent when they were under age 5, and whose mothers worked at some point either before or after the child's birth.

In Table 1A we show summary statistics for our sample of boys with overall means and standard deviations given in the far right column. Of the roughly 3,300 boys in our sample, 6% are black, 42% are the first-born child of their mother, and their mother's average age at the time of their birth was 26.7. Nearly half the sample had mothers with only high school degrees or less, and 42% had fathers with high school degrees or less. Average total mother and father earnings when the boy was age five or under was almost $63,000 in 2010 dollars.[9] At age 30, average total earnings of these boys was just over $46,000, in 2010 dollars, and average labor force experience was 13.6 years. The average predicted father random effects were zero on average whereas the predicted mother random effects were just slightly below zero on average. By design in the mixed effects model, the estimated mother random effects have zero mean when averaged across mothers. However in this table we are averaging across sons and hence the average changes because some mothers have multiple sons in our sample, effectively re-weighting the random effects. This result means that mothers with lower $\widetilde{\theta^m}$ had more sons.

[9] While this number for total family earnings may seem high, we remind the reader that our sample of boys comes from parents who remain together in the same household till the son is in his teens.

Overall, 75% of boys had mothers who worked at least some years when the son was under age 5, with 32% working all 5 years. To see whether there are observable differences between sons with working mothers and those whose mothers stay home, we report means by working status of the mother and present t-statistics testing the differences between sons with non-working mothers and those with mothers who work either some or all five years when the child was age five and under. Sons with working moms are more often black, more often oldest children, and have mothers with higher unobserved ability (i.e. higher mother RE). Sons with mothers who work every year the first five years of the son's life have higher levels of education. In contrast, sons whose mothers work some years have fathers with lower levels of education and both groups of sons with working mothers have fathers with lower unobserved ability. Thus it would seem that working moms are married to men who have lower earnings than the husbands of their non-working counterparts. Thus it will be important to control for both parents' characteristics as the bias introduced into the coefficient on mother working by unobserved parent qualities may go in opposite directions for mothers and fathers. Families where the mother works every year the son is under age 5 have significantly higher average total parent earnings during this time period. Families where the mother only works some years have similar average total parent earnings when the son is under age 5 to mothers who did not work at all. We do not see differences in earnings across the three groups of sons that are significantly different from zero but we do see such differences for labor force experience. Sons whose mothers were most strongly attached to the labor market have higher labor force experience themselves by age 30.

In Table 1B we report the same summary statistics for our sample of girls. The most noteworthy thing about this table is that these girls, restricted to have worked at least one year between age 30 and 33, look remarkably similar to the boys in Table 1A in almost every respect. Race, oldest child status, mother age at birth of daughter, parental education, average combined parent earnings when the daughter was under 5, and percentages of mothers who worked are very similar on average to boys. Similar trends hold that girls with working mothers also have more highly educated mothers (31% with college degree or more versus 18%); mothers with higher unobserved ability; slightly less educated fathers (31% with college degree or more versus 34%); and fathers with lower unobserved ability. The only major difference between the gender is that girls earn less on average than boys, despite having similar levels of work experience, a result probably due in part to unobserved labor supply differences. Girls may work fewer hours per week at age 30 than boys due to child care responsibilities. Unfortunately our administrative data do not contain labor supply measures so we cannot differentiate among various potential causes of lower earnings.

Given our selection of children interviewed in one of six different SIPP panels and coming from intact families, our sample is clearly not nationally representative of children born between 1978 and 1981, and there are currently no weights available to make our sample representative. In order to assess how different our sample is from the national population in this birth cohort, we compare our

SIPP boys to the universe of male workers born between 1978 and 1981 with W-2 records in 2010, to see how the W-2 earnings for our chosen survey respondents compared to the universe of W-2 records. We used decile cut-off points calculated from the W-2 universe earnings distribution to place each SIPP boy in a national decile, and in Figure 1 we report the percentage of our sample that falls into each decile. If our sample was a perfect random sample from the universe, we would expect close to 10% of the boys to fall into each decile as defined by the national distribution. As Figure 1 shows, our sample is slightly under-represented in the first five deciles, and slightly over-represented in the top three deciles. A similar pattern holds for girls as shown in Figure 2. We expect that some of the differences between the universe W-2 earnings distribution and our sample are driven by the fact that we chose only individuals from families that remain intact for a relatively long period of time, something we cannot control for in the universe distribution. Overall these figures give us some confidence that our sample can reasonably be used for estimating covariate relationships even though it cannot be used for calculating population totals.

While our data set clearly expands our knowledge by providing the long time series on earnings necessary to observe the mother's labor market attachment when the child is young and to observe the child's labor market outcomes 30 years later, like all data sets this one has limitations. Like many previous studies, we are also not able to control for quality of purchased child care services. Furthermore, we cannot determine what happens to family structure after the end of the survey nor observe any other child outcomes besides labor market participation and earnings. There are currently no weights and we have no sub-annual information available on labor supply. However, there is much to recommend these data. Our sample size is relatively large, we have long histories of earnings which are potentially more reliable than self-reports about earnings and work decisions from the far past, and we know a great deal about the history of the family over a time period that covers the important early years of a child's life.

6 Results

6.1 Labor Force Experience Outcomes

We begin with results for labor force experience for boys and show results from estimating equations 10 and 11 in Table 2A. In column 1 we include no controls for parent characteristics and then include parental education in column 2 and parental unobserved ability in column 3. Finally in column 4 we show the mother interaction model. We divide working mothers into two groups - those who worked some years when the child was under the age of 5 and those who worked all 5 years - and estimate separate effects of working for both groups.

Column 1 shows a positive relationship between mother working and total years of labor force experience of the son in the year 2011.[10] The effect of

[10] We chose this measure of labor force experience in order to use the highest observed level

the mother working some years is relatively modest (.05) and not significantly different from zero. If the mother worked all five of the first years of her son's life, the impact is larger (.22) and significant at the .05 level. These patterns stay the same as we move across the next two columns, adding first mother and father education indicators and then mother and father random effects. The stability of both coefficients on mother's work indicates that this variable is not simply capturing the indirect impact of these control variables.

In the fourth column, where we include interaction terms, we find that the lowest educated mothers are the only group with a statistically significant effect of working on sons' future labor force attachment. Mothers with a high school degree or less who worked every year between the birth of their sons and age 5, had sons with .38 years more adult labor force experience than sons of mothers did not work when their sons were age 5 and under. The point estimates for the other two education groups are of similar magnitude but not significantly different from zero. This implies that less educated mothers have at least as large an impact on their sons' future labor market experience as mothers with more education. Interestingly, the coefficient on the interaction between the mother random effect and working all 5 years is negative and significant, meaning that mothers with higher unobserved ability have a dampened positive impact on their son's labor market experience.

In order to consider the secondary effect of the mother working, namely additional income, we also include a cubic in average combined parent earnings over the first five years of the son's life in the regression. To facilitate the interpretation of these results, we report a marginal effect of earnings at the 25th, 50th, and 75th percentiles of the distribution. These effects are all statistically significant and range from .28 to .41, indicating diminishing marginal returns to parent earnings in terms of boosting long-run child labor force experience. Thus a mother working has potentially two positive effects on her son's future labor force experience. First, through a direct effect and second through increasing the family's income.

Table 2B shows results from the fully interacted model. Each of the nine categories representing possible pairings of mother and father education has a separate effect of mother working and within these nine categories, there is a separate slope coefficient that represents the effect of increasing unobserved ability. The results in Table 2B are for mothers who work all five years. Moving down the table changes the mother/father education combination. Moving across the table changes the mother's unobserved ability. Columns 1-3 show the total effects defined as $(\gamma^{me\{A\}fe\{B\}} + \gamma^{ma(me\{A\}fe\{B\})} * \theta^m)$ where $A = \{1, 2, 3\}$ and $B = \{1, 2, 3\}$ and represent the mother and father education levels respectively and θ^m is evaluated at the 25th, 50th, and 75th percentiles of the distribution. Column 4 reports separately the unobserved mother ability slope coefficient for that education category, $\gamma^{ma(me\{A\}fe\{B\})}$. Here we see that the only significant effects of working are for sons of mothers and fathers who both

for every child in our sample. We control for age of the child in the regression. Results using labor force experience at age 30 are not materially different.

have only a high school degree or less. The diminishing returns to working as mother unobserved ability rises are significantly different from zero for the sons of more educated mothers paired with fathers who hold at least a college degree.

Figure 3 displays these combinations of coefficients graphically. The diamond (blue) line represents the effect of the mother working all 5 years for families in which the father has no more than a high school degree. The first three points on the line represent the effects for mothers who have a high school degree or less, varying across the three ability levels (25th percentile, median, and 75th percentile of the unobserved heterogeneity distribution) and are equivalent to the first three columns of the first row in Table 2B. The slope of this line segment is equal to the parameter reported in the fourth column of the first row. The next three points show the effects for mothers with some college. The final three points show the effects for mothers with a college degree or more. Thus the diamond (blue) line shows how the effect of the mother working all 5 years changes as the mother's education level rises and as unobserved ability increases within education category. The square (red) line reports the same results for children with fathers with some college and the triangle (green) line for children with fathers with a college degree or more.[11]

The impact of mothers working when the sons are young is similar in families with fathers in the two lowest education categories where it is U-shaped across mother education levels. Mothers with low and high education appear to have more positive effects than mothers with middle education levels. However these differences are not statistically significant. Sons with more educated fathers show a different pattern. Here the least educated mothers have the lowest impact, with the middle and high educated mothers both having larger effects of working. Again, however these differences across mother education category are not statistically significant. Thus while for two of the three lines working mothers with a college degree or higher have bigger impacts when they work than mothers with a high school degree or less, our coefficients are not estimated with enough precision to say that the effect of working changes monotonically with the mother's education. Likewise for fathers, although for both low and high mother education the impact of working is lower when the father is highly educated, the lack of statistical significance for these differences prevents us from concluding that the effect of the mother working is monotonically decreasing with father's education.[12]

In Table 3 we show results from a traditional sibling model which is estimated with a mother fixed effect. This specification compares brothers and is identified from differences in the mother's labor supply when each son was an infant. Our initial unconditional specification finds no significant relationship between the mother's work pattern when each son was young and their future labor force experience at age 30. We then interact working with mother and father

[11] All these lines hold the father's unobserved ability constant at the median value, $\theta^f = 0$.

[12] F-stats for comparisons across mother and father education are shown in Appendix Table 2B.

education categories, although using fewer categories than in Table 2B due to sample size issues. We present these results in column 2 of Table 3 and find a significant impact of the mother working for mothers with at least some college and fathers with a high school degree or less. This is consistent with our results shown in Table 2B and Figure 3, where the mothers with college degrees who were more educated than their husbands tended to have larger positive impacts on their sons' labor force experience, although those results were not statistically significant.

In Table 4A, we show the same results for girls as were shown in Table 2A for boys. Without any parental controls, having a mother who works every year her daughter is under age 5 is associated with a .3 increase in the number of years of labor force experience for that girl by 2011. Adding education and mother and father ability controls does not have a large impact on the magnitude or significance of these coefficients. This is consistent with our findings for sons. When mother education is interacted with mother work status, we find a large positive significant effect of working all 5 years for mothers with a high school degree or less, a very similar effect to what we found for boys. Likewise the coefficient on the interaction of the mother random effect and work status is negative and significant while the marginal effect of average combined parent earnings is positive and diminishing over the income distribution.

In Table 4B, we again examine the impact across education and ability of the father and mother. The only significant effects are for the lowest educated group of mothers and fathers, those where both parents have a high school degree or less and here the effect is positive. The effects for mothers with high school degrees or less and fathers with both some college and college degree or more are also positive but only statistically significant at the 10% level. In Figure 4, we see differences between daughters and sons in the apparent patterns of the effect of working across parent education levels. Generally the effect of working decreases as the mother's education rises rather than being U- shaped, while the effect increases with father education. This pattern is the reverse of what we saw in Figure 3 for boys. However none of these differences are statistically significant.[13]

In Table 5 we show results for the sibling model for girls. Unlike in Table 3 for boys, we find no significant effects of the mother working. Here it seems that our sample of girls who work between age 30 and 33 with sisters who are close in age and also work is perhaps too small to draw any conclusions.

6.2 Earnings Outcomes

We now turn to results from estimating equations 10 and 11 with earnings as the dependent variable. While ideally one would estimate a wage equation, our administrative data does not include any measures of hours or weeks worked, and hence we are constrained to use total annual earnings. Table 6A shows

[13] F-stats for comparisons across mother and father education are shown in Appendix Table 4B.

the results for boys. The point estimates of the mother working are negative when we do not control for parental education, ability or family earnings. The estimated impact of mothers working when their sons are 0 to 5 on the son's adult earnings becomes less negative as we control for parental education and unobserved ability.[14] However none of these effects are significant, a fact that remains true in column 4 when we interact working with mother education and the mother random effect. Rather than mother working, the strongest predictors of a son's earnings are father's education and unobserved ability (father RE), both of which are associated with statistically significant higher earnings. Having a father with a college degree is associated with an increase in real log annual earnings of 20%. Moving from the median of the unobserved father ability distribution to the 75th percentile (i.e. $\theta^f = .3$ instead of 0) is associated with a 10% rise in the son's earnings.

In Table 6B where we present results from the full interaction model. Here we find positive and significant effects of working for some groups of mothers, based on the education and ability of mothers and fathers, namely mothers with a college degree or more paired with either high or low educated fathers. While within these groups, sons of mothers who worked when they were young have higher earnings than the sons of stay at home moms, for most other groups the effect of working is not significantly different from zero. For mothers with a high school degree or less paired with a highly educated husband, the effect of working borders on being statistically significant and negative.

Figure 5 displays the point estimates from Table 6B and makes clear the trend that the effect of working becomes more positive as the mother's education increases. This is true for mothers paired with either a high or low educated father. For the diamond (blue) and triangle (green) lines, the difference between a mom with a high school degree or less and a mom with a college degree or more are statistically significant. For fathers the pattern is reversed. The impact of mothers working is higher for sons with less educated fathers. Again the highest returns are for sons in families in which the mother had more education than the father. While these point estimates are intriguing, these differences across father education are not statistically significant.[15]

Table 7 shows the earnings estimates from the sibling model with the mother fixed effects. While the point estimates are almost all positive, none of them are significant at the 5% level and only one is significant at the 10% level (lowest educated mothers and fathers when mother works some years). As in Tables 3 and 5, we again collapse education categories due to sample size constraints, and no clear pattern of the effects by parental education emerges.

[14] This result seems to contradict the hypothesis that working mothers are more able and also have more able sons and hence we would expect the effect of working to become more negative when we control for mother unobserved ability. However we are also including father education and ability in the regression and in our sample, father education and ability are negatively correlated with the mother working. Hence when we control for both father and mother unobserved ability, the net effect of the mother working coefficients is to move them towards zero.

[15] F-stats for comparisons across mother and father education are shown in Appendix Table 6B.

Finally turning to earnings results for girls, we see in the first three columns of Table 8A that the point estimate for the effect of the mother working when her daughter is zero to five is negative but moves towards zero as we control for more parent characteristics. However none of the mother working effects are statistically significant in this table, even when we include interactions with mother education and unobserved ability as shown in column 4. We do see in this column, though, that the point estimates are negative only for mothers with low education. Table 8B provides further evidence on the role of parental education and we see a significant negative effect of working for some low and high educated mothers paired with fathers who are low educated. Likewise we see some positive effects where both parents are highly educated, although these are only significant at the 10% level. In general the point estimates are only positive if the mother and father education levels match. Otherwise they are close to zero or negative.

We note with interest that the mother working when she has more education than the father does not have the same beneficial effects for daughters as it does for sons. As shown in Figure 6, the point estimates for women with college degrees or more are negative unless the husband is also highly educated. In comparison, for sons, the effect of working for highly educated mothers paired with low educated fathers was the highest positive point in Figure 5. In general the lines in Figure 6 trend down as opposed to up in Figure 5, meaning the effects decrease as mother education increases, although these changes are not significant. The only exception to this is the line for fathers with at least a college degree. Here the effect rises when the mother switches to the highest education category and this change is significantly different from zero. The effect of working across father education categories is generally increasing as well, with some exceptions, which is again opposite from the results for sons. This change is significant when the mother is highly educated (i.e. moving from the diamond (blue) line to the triangle (green) line).[16]

We finish with a siblings model for girls' earnings in Table 9 and again see similar trends in the effect of working across education groups but find no significant effects.

7 Conclusion

Overall, we find that children whose mothers worked when the child was very young were more likely to work themselves as adults than children with stay-at-home moms. This is true for both boys and girls and is consistent with a model in which parents pass on their values to their children and find alternative day-care arrangements that are close substitutes for what the parent could provide as a stay-at-home mom.

Paradoxically, we find weaker evidence for the proposition that a mother's decision to stay home when her children are very young has an impact on the

[16] F-stats for comparisons across mother and father education are shown in Appendix Table 8B.

child's earnings later in life. For the mother-father education pairings where child's years of work increase by a statistically significant amount when the mother works, annual earnings do not increase.

These two stylized facts imply that the returns to experience of children of working parents are actually lower than the returns to experience of children from families with stay-at-home moms. This is consistent with a model in which the parent's influence on the child's work-ethic is sufficiently strong that the children of working parents not only work more as adults than children raised by stay-at-home moms, but that these work-prone children grew up to become adults willing to accept jobs with lower wages that would be rejected by the offspring of the stay-at- home moms. While we do not claim to have tested this behavioral model, it is consistent with our reduced form findings. Identifying ways to test such a model against other alternative explanations is the next stage for this analysis.

Future analysis will also include subsetting our sample to include only first-born children in order to test the sensitivity of our results to birth-order. As additional years of administrative data become available, we will be able to increase our sample size and follow children further into adulthood, which may change the statistical significance of some of our results. Imputation of missing administrative records due to failure to confirm an SSN would also help with sample sizes.

Overall it does not appear that there are significant negative effects of the mother working on the adult labor market outcomes of either sons or daughters. Even if there is an negative impact of the mother working on early childhood test scores, such differences do not appear to translate into differential earnings later in life.

8 Bibliography

Almond. D. and J. Currie, "Human Capital Development before age 5," *Handbook of Labor Economics*, Vol. 4B, ed. Orley Ashenfelter and David Card, (2010): 1315-1486.

Baker, Gruber, Milligan, "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, Vol. 116, No. 4 (August 2008): 709-745.

Baker and Milligan, "Evidence from Maternity Leave Expansions of the Impact of Maternal Care on Early Child Development," *Journal of Human Resources,* Vol. 45, No. 1 (Winter 2010): 1-32.

Blau, D., "The Effect of Income on Child Development," *Review of Economics and Statistics,* Vol. 81, No. 2 (May 1999): 201-276.

Blau, F. and Grossberg, A., " Maternal Labor Supply and Children's Cognitive Develpment," *Review of Economics and Statistics, Vol. 74, No. 3* (1992): 474-481.

Han, Waldfogel, and Brooks-Gunn, "The Effects of Early Maternal Employment on Later Cognitive and Behavioral outcomes," *Journal of Family and*

Marriage, Vol. 63, No.2 (May 2001): 336-354.

Dustmann and Schonberg, "The Effect of Expansions in Maternity Leave Coverage on Children's Long-Term Outcomes," SSRN elibrary (2008).

Henderson, "Estimation of Variance and Covariance Components," *Biometrics*, Vol. 9 (1953), 226-252.

Ruhm, C., "Parental Employment and Child Cognitive Development," *Journal of Human Resources*, Vol. 39, No. 1 (Winter 2004): 155-192.

Searle, S., G. Casella, C. McCulloch, *Variance Components*, John Wiley & Sons, (1992).

Solon, Gary R., "Intergenerational Income Mobility in the United States," *American Economic Review*, Vol. 82, No. 3 (June 1992): 393-408.

Waldfogel, Han, and Brooks-Gunn, "The Effects of Early Maternal Employment on Child Cognitive Development," *Demography,* Vol. 39, No. 2 (May 2002): 369-393.

Zimmerman, David J., "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, Vol. 82, No. 3 (June 1992): 409-429.

9 Appendix A

In this Appendix we describe the estimation of the mother and father random effects that we use to control for unobserved heterogeneity of both parents in the child's equations. Our goal in estimating these effects is to exploit the long earnings history from the administrative data for each parent in order to create a measure of unobserved labor market heterogeneity, beyond what could be observed in terms of education and labor force experience of the parent. We treat these as random effects and estimate them using a mixed effects model. While such models are common in the statistics literature, especially biostatistics, they are not as common in the economics literature. Therefore, we begin by briefly presenting the mixed model and then we explain why the mixed effects model does not suffer from some of the same problems that economists typically associate with random effects. We end with a brief description of our estimation method.

9.1 Mixed Model

In their classic text on random and mixed effects models, Searle, Casella, and McCulloch begin by defining factor variables as information that classifies the data into categories. These factor variables have effects on variables of interest to a researcher and these effects can be either fixed or random. The authors define fixed effects as those which are "attributable to a finite set of levels of a factor that occur in the data." Random effects are unobserved factors with an infinite set of levels "of which only a random sample occur in the data." In each case there are multiple observations for each factor. For example, the data may be on housing prices which vary by city, neighborhood and block. Heterogeneity occurs at each level. The heterogeneity within blocks can be treated as random since the quality of homes has infinite support.

Note that the distinction in the statistics literature between random and fixed effects is based on whether the heterogeneity distribution is fully captured by covariates in the data (i.e. fixed effects) or whether the data only provides a sample of the heterogeneity distribution that has infinite support.

In our data, we treat mothers and fathers unobserved personal earnings heterogeneity, θ, as random because there is an infinite number of types of mothers and fathers– the support for the unobserved heterogeneity is infinite. Therefore, the heterogeneity for the group of mothers and fathers present in our data is only a finite sample of all possible values. In contrast, we treat the unobserved heterogeneity associated with different levels of education as fixed since there is a finite and relatively small number of levels of education, each with its own heterogeneity component. If the unobserved heterogeneity distribution is fully captured by the observed education then this form of heterogeneity is fixed. Note that the distinction between the parental heterogeneity, which is random, and the educational heterogeneity, which is fixed, does not require any assumption about the independence of the unobserved heterogeneity.

One particularly appealing characteristic of mixed effects models is that both fixed and random effects can be included. For example when estimating an earnings equation, one can include a set of dummies for a particular characteristic such as education that capture the mean of the heterogeneity distribution across time for individuals. These fixed effects control for time invariant attributes of the individual. A person random effect can also be included that captures the dispersion around this conditional means. This is in contrast with the standard fixed effects models where a person-level effect will soak up the effect of all time-invariant person characteristics.

9.2 Estimation

The models we first estimate are a set of parental earnings models with parental characteristics such as age, labor force experience, race, education, and year time dummies included as explanatory variables. We estimate separate models for mothers and fathers but they are not qualitatively different. To aid the flow of our description, we use mothers as our example in what follows. Everything can be equivalently applied to fathers. First let I be the total number of mothers in the sample with T observations each for a total of $N = I * T$ observations. Let Y_i be a $Tx1$ vector of annual earnings measures for mother i and let X_i be a Txk matrix of explanatory variables with coefficient vector β with dimensions $kx1$. Let d_i be a $1 \times I$ design matrix of the effects associated with mother i and θ be the $Ix1$ matrix of person effects such that $d_i\theta = \theta_i$. Finally let η_i be the $Tx1$ vector of residuals. The linear model for mother i is given by

$$Y_i = X_i\beta + d_i\theta + \eta_i$$

and then stacked across all mothers to become

$$Y = X\beta + Z\theta + \eta \tag{13}$$

$$Z = \begin{bmatrix} d_1 \\ ... \\ d_I \end{bmatrix} = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}, \theta = \begin{bmatrix} \theta_1 \\ ... \\ \theta_I \end{bmatrix}, Y = \begin{bmatrix} Y_1 \\ ... \\ Y_I \end{bmatrix}, X = \begin{bmatrix} X_1 \\ ... \\ X_I \end{bmatrix}, \eta = \begin{bmatrix} \eta_1 \\ ... \\ \eta_I \end{bmatrix}$$

Statisticians call Z the design matrix of the effects θ. It is merely a set of dummies that assign θ_i from the θ vector to the i^{th} mother.

This model described by 13can be treated as what Greene calls the least squares dummy variable (LSDV) model (page 466) with the following commonly made assumptions:[17]

$$\begin{aligned} \eta &\sim N(0, R) \\ R &= \sigma_\eta^2 I \end{aligned}$$

where β and θ are called fixed effects in the statistics literature if the unobservable and observable factors in the population (θ_1 ...θ_l and X_1 ...X_l) are finite and cover all possible values in the population.

The standard normal equations for the OLS estimator are

$$\begin{bmatrix} Z'Z & Z'X \\ X'Z & X'X \end{bmatrix} \begin{bmatrix} \theta \\ \beta \end{bmatrix} = \begin{bmatrix} Z'Y \\ X'Y \end{bmatrix} \tag{14}$$

which are familiar to most economists. These can be solved to yield

$$\begin{aligned} \beta &= \left[X' \left(I - Z(Z'Z)^{-1}Z'\right) X\right]^{-1} \left[X' \left(I - Z(Z'Z)^{-1}Z'\right) Y\right] \\ \theta &= \left[Z' \left(I - X(X'X)^{-1}X'\right) Z\right]^{-1} \left[Z' \left(I - X(X'X)^{-1}X'\right) Y\right] \end{aligned}$$

using the general rules for obtaining solutions for partitioned regressions (Greene page 179). One characteristics of the LSDV method is that the solutions for (θ, β) do not impose orthogonality between Z and X. In the terms used in the econometrics literature, one does not need to assume that the time invariant unobservables are independent of the $X's$.

The term "random effects" has a different meaning in the econometrics literature where unobserved heterogeneity is treated as a random effect in the

[17] In all our descriptions here we will assume that the variance structure of the model error, η, is defined as $R = \sigma_\eta^2$ but this assumption can be changed to a more complicated variance structure without substantially changing the model descriptions presented here.

following sense:

$$
\begin{aligned}
Y &= X\beta + \theta + \eta \\
\eta &\sim N(0, \sigma_\eta^2 I) \\
\theta &\sim N(0, \sigma_\theta^2 I) \\
cov(\theta, \eta) &= 0, cov(X, \eta) = 0, cov(X, \theta) = 0 \\
\Omega &= var(y_i) = \begin{bmatrix} \sigma_\eta^2 + \sigma_\theta^2 & \sigma_\theta^2 & ... & \sigma_\theta^2 \\ \sigma_\theta^2 & \sigma_\eta^2 + \sigma_\theta^2 & ... & \sigma_\theta^2 \\ ... & ... & ... & ... \\ \sigma_\theta^2 & \sigma_\theta^2 & ... & \sigma_\eta^2 + \sigma_\theta^2 \end{bmatrix} \\
R &= \begin{bmatrix} \Omega & 0 & ... & 0 \\ 0 & \Omega & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \Omega \end{bmatrix}
\end{aligned}
$$

In this model, the random effect is merely treated as a portion of the error term. The identity of the mother imposes additional structure on the variance/covariance matrix of the error term. This type of model does not estimate θ directly but rather estimates σ_θ^2. The solution for fixed effects, β, is

$$\beta = (X'R^{-1}X)^{-1}X'R^{-1}Y$$

which is the standard GLS estimator. There is no $X'Z$ term in this model because of the assumption of orthogonality between the random effects and the observed characteristics in the X vector.[18]

In contrast to these two methods, mixed effects models allow θ to be treated as a random effect but also allow $\widehat{\theta_i}$ to be estimated for each mother in the sample. These methods were pioneered by Henderson, a biostatistician interested in estimating genetic models that predicted milk production of cows as a function of the identity of their sires and dames. The goal of his models was to be able to predict parent effects for the milk production of a child cow, with the intent of identifying which bulls sired the best milk-producing daughters. He began with the same model as above, namely,

$$Y = X\beta + Z\theta + \eta$$

along with the assumptions (Searle, Casella, McCulloch page 275)

$$\begin{bmatrix} \theta \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ X\beta \end{bmatrix}, \begin{bmatrix} G & GZ' \\ ZG & V \end{bmatrix} \right)$$

$$
\begin{aligned}
var(Y) &= V = ZGZ' + R \\
R &= \sigma_\eta^2 I \\
var(\theta) &= G = \sigma_\theta^2 I \\
cov(Y, \theta) &= ZG
\end{aligned}
$$

[18] The widely-used Hausman test is in fact a test of whether $X'Z = 0$ and the frequent rejection of this hypothesis has left most economists skeptical of using random effects.

Henderson shows that the pdf of the joint distribution is given by

$$\begin{aligned} f(y, \theta) &= f(y \mid \theta) f(\theta) \\ &= \frac{\exp\left\{-\frac{1}{2}\left[(y - X\beta - Z\theta)' R^{-1}(y - X\beta - Z\theta) + \theta' G^{-1}\theta\right]\right\}}{(2\pi)^{1/2(N+I)} \left|R\right|^{1/2} \left|G\right|^{1/2}} \end{aligned} \tag{15}$$

By taking partial derivatives of 15 with respect to β and θ, Henderson arrived at what are now known as the mixed model equations (MME) (Searle, Casella, McCulloch page 276).

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\beta} \\ \widehat{\theta} \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Z'R^{-1}Y \end{bmatrix} \tag{16}$$

The important thing to notice in these equations is that $X'R^{-1}Z \neq 0$, and hence the standard economist concern about imposing orthogonality between the characteristics in X and the design of the random effects matrix is no longer an issue.

It is also informative to compare equation 16 to equation 14, the normal equations for the LSDV model. Without G^{-1} in the bottom right cell, the MME are simply the maximum likelihood versions of the normal equations for the LSDV model. As $|G| \to \infty$, the MME converge to the normal equations. Thus the LSDV model is a special case of the mixed effect model.

In estimating our mixed effect model we use Restricted Maximum Likelihood (REML). The basic concept of REML estimation is to maximize a marginal likelihood. A set of linear error contrast equations are created that do not include β and these are used to create a likelihood function that contains only σ_η^2 and σ_θ^2 from the variance matrices G and R (Searle, Casella, and McCulloch (1992)). These parameters are called variance components and are estimated by maximizing this marginal likelihood. Using these estimates of G and R, the mixed model equations are solved to give estimates for the fixed effects, $\widehat{\beta}$, and then the predicted random effects, $\widehat{\theta}$. For samples of our size and earnings equations with simple random effects, the Stata version of REML for mixed effects models (xtmixed) is sufficient to generate $\widehat{\theta}_i$ for each parent in our sample in a computationally feasible amount of time.

Table 1A: Summary Statistics for Sons

	Overall		Mother - no work	Mother - some work			Mother - all work		
	mean	st. dev.	mean	mean		t-stat	mean		t-stat
N	3356		842	1452		for diff	1062		for diff
black	0.06	0.24	0.04	0.06	*	-2.06	0.09	***	-4.55
oldest child	0.42	0.49	0.32	0.42	***	-4.73	0.49	***	-7.48
age of mother at birth	26.66	4.75	27.27	26.05	***	5.78	27.01		1.16
Mother Educ. Indicators									
high school or less	0.48	0.50	0.53	0.52		0.6	0.37	***	6.95
some college	0.29	0.46	0.26	0.27		-0.77	0.35	***	-4.11
college or more	0.23	0.42	0.21	0.21		0.1	0.28	***	-3.64
Father Educ. Indicators									
high school or less	0.42	0.49	0.40	0.44	*	-2.05	0.39		0.2
some college	0.27	0.45	0.26	0.28		-0.73	0.28		-0.54
college or more	0.31	0.46	0.34	0.28	**	2.88	0.33		0.3
Mother random effect	-0.01	0.58	-0.07	-0.12	*	2.05	0.19	***	-9.9
Father random effect	0.00	0.49	0.04	0.00	*	1.8	-0.03	***	3.35
Average combined parent earnings when son was ages 1-5	62,629	80,791	57,596	58,078		-0.15	72,840	***	-6.10
Real earnings at age 30	46,214	53,355	45,674	44,933		0.36	48,362		-1.17
Labor force experience age 30	13.57	2.04	13.44	13.52		-0.93	13.74	***	-3.17
mom no work ages 1-5	0.25	0.43							
mom some work ages 1-5	0.43	0.50							
mom work all years ages 1-5	0.32	0.47							
average %total parent earnings due to mom, son ages 1-5	0.21	0.24							

Sample is boys who: 1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011 5. worked themselves at least one year between 30 and 33
T-stats are from two sample tests with unequal variances comparing means for working moms, all or some years,
to non-working moms. * p<0.05, ** p<0.01, *** p<0.001

Table 1B: Summary Statistics for Daughters

	Overall		Mother - no work	Mother - some work			Mother - all work		
	mean	st. dev.	mean	mean		t-stat	mean		t-stat
N	**2910**		**728**	**1256**		**for diff**	**926**		**for diff**
black	0.07	0.26	0.03	0.07	***	-4.05	0.11	***	-6.63
oldest child	0.41	0.49	0.33	0.41	***	-3.75	0.47	***	-5.75
age of mother at birth	26.67	4.73	27.54	26.00	***	6.9	26.89	**	2.78
Mother Educ. Indicators									
high school or less	0.47	0.50	0.55	0.48	**	2.93	0.41	***	5.7
some college	0.29	0.45	0.27	0.30		-1.46	0.29		-0.53
college or more	0.24	0.42	0.18	0.22	*	-2.03	0.31	***	-6.17
Father Educ. Indicators									
high school or less	0.43	0.49	0.43	0.43		0.08	0.42		0.35
some college	0.27	0.44	0.23	0.29	**	-2.77	0.27	*	-1.83
college or more	0.31	0.46	0.34	0.29		2.49	0.31		1.32
Mother random effect	-0.00	0.58	-0.02	-0.12	***	3.57	0.18	***	-7.21
Father random effect	-0.00	0.51	0.07	-0.02	***	3.43	-0.04	***	4.14
Average combined parent earnings when son was ages 1-5	62,747	51,893	64,868	53,702	***	4.51	73,347	**	-2.81
Real earnings at age 30	35,405	27,487	35,502	33,456		1.60	37,974	*	-1.73
Labor force experience age 30	13.48	2.10	13.29	13.46		-1.61	13.66	***	-3.53
mom no work ages 1-5	0.25	0.43							
mom some work ages 1-5	0.43	0.50							
mom work all years ages 1-5	0.32	0.47							
average %total parent earnings due to mom, son ages 1-5	0.21	0.24							

Sample is girls who: 1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011 5. worked themselves at least one year between 30 and 33
T-stats are from two sample tests with unequal variances comparing means for working moms, all or some years,
to non-working moms. * p<0.05, ** p<0.01, *** p<0.001

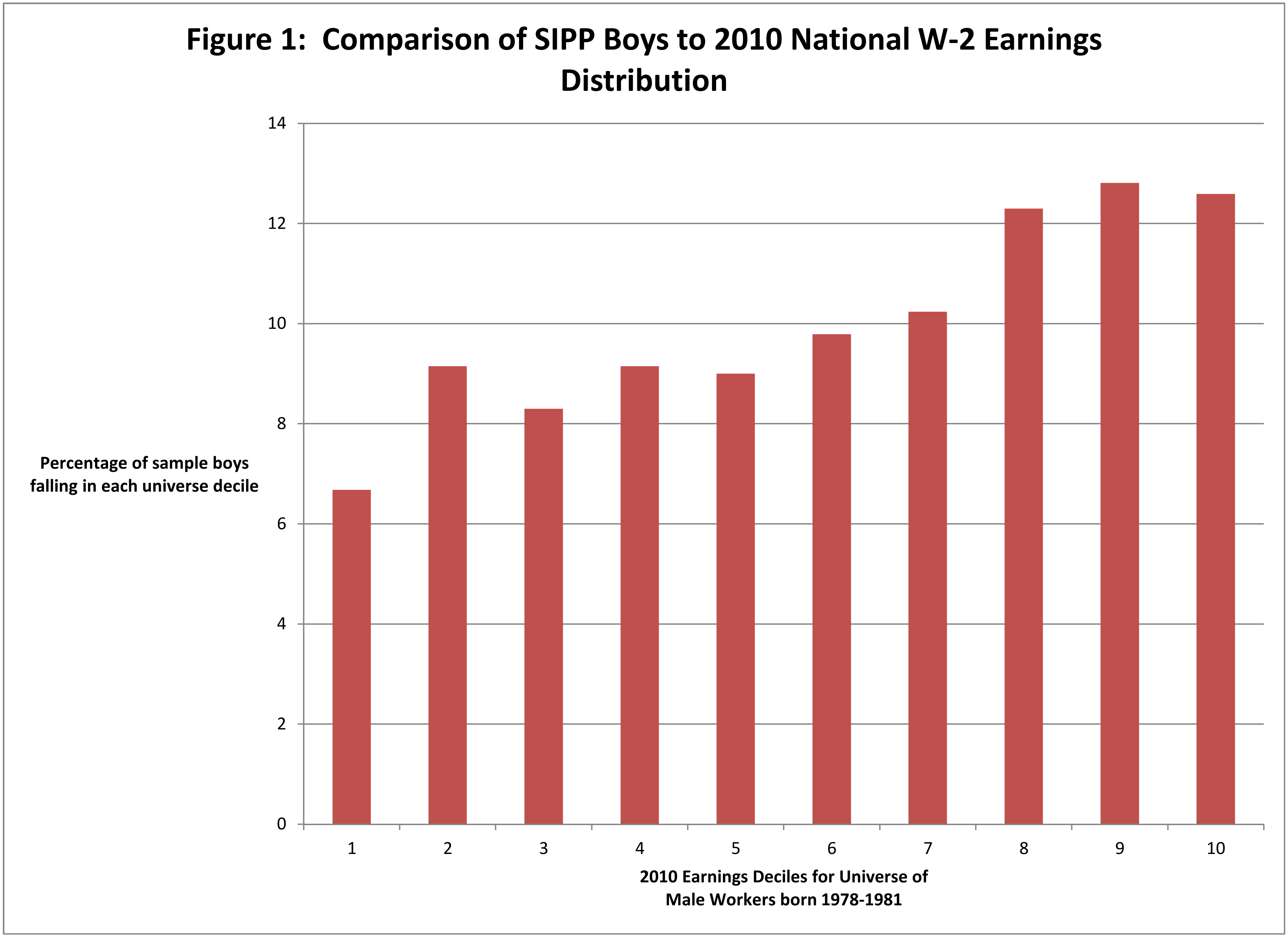
Figure 1: Comparison of SIPP Boys to 2010 National W-2 Earnings Distribution
Percentage of sample boys falling in each universe decile
0
2
4
6
8
10
12
14
1
2
3
4
5
6
7
8
9
10
2010 Earnings Deciles for Universe of Male Workers born 1978-1981

Figure 2: Comparison of SIPP Girls to 2010 National W-2 Earnings Distribution

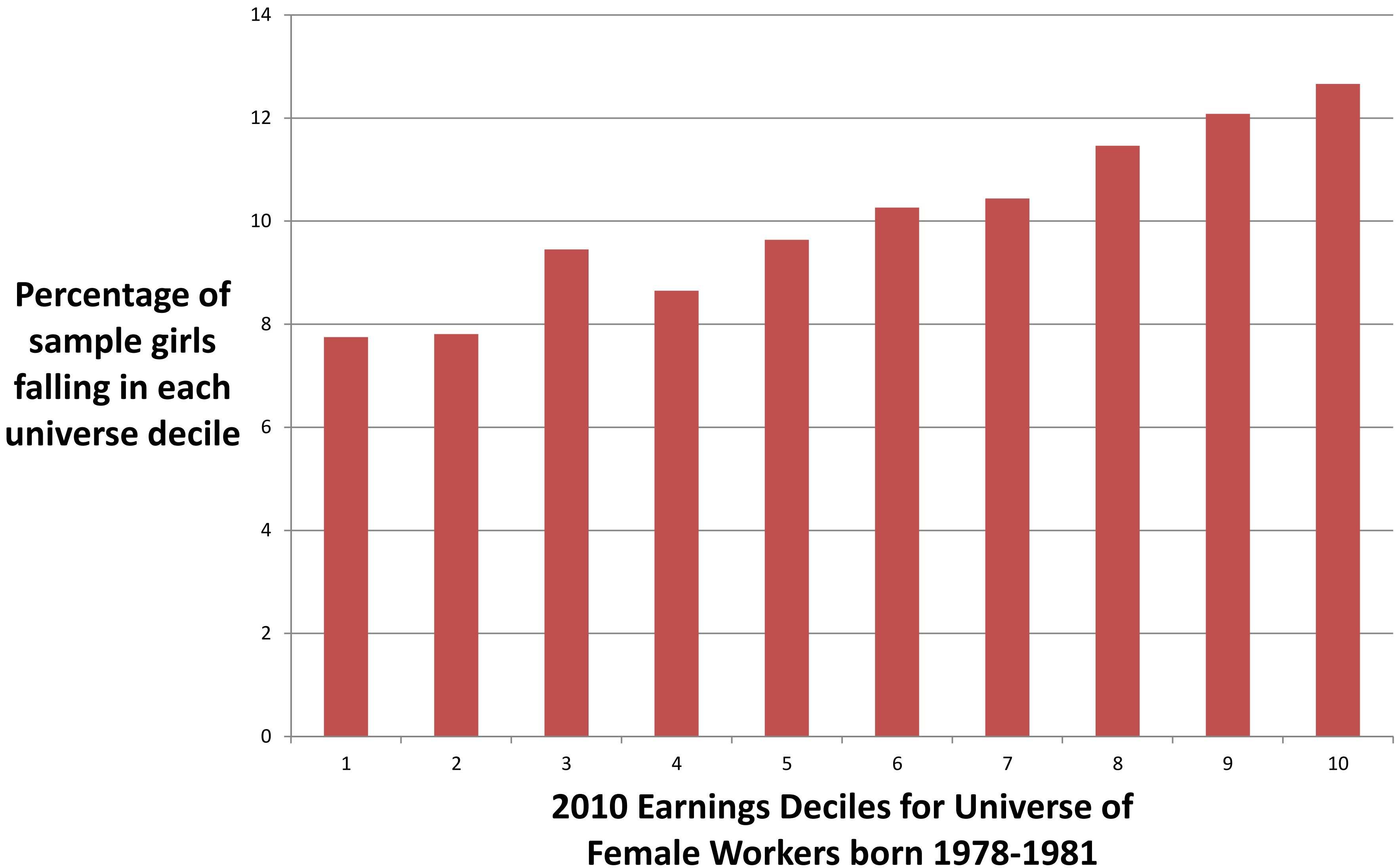

Table 2A: Labor Force Experience of Sons in 2011

		1	2	3	4
Mom work		Obs=3356			
all educ levels	some years	0.053	0.053	0.044	
		(0.094)	(0.094)	(0.094)	
all educ levels	all 5 years	0.220*	0.224*	0.244*	
		(0.103)	(0.103)	(0.107)	
Mom work Interactions					
HS or less	some years				0.038
					(0.129)
HS or less	all 5 years				0.377*
					(0.154)
Some college	some years				0.149
					(0.182)
Some college	all 5 years				0.205
					(0.188)
College +	some years				0.008
					(0.204)
College +	all 5 years				0.334
					(0.207)
Mom RE	some years				0.129
					(0.159)
Mom RE	all 5 years				-0.492**
					(0.178)
Parental Controls					
Mother	Some College		0.001	-0.006	0.005
			(0.096)	(0.096)	(0.183)
Mother	College +		-0.134	-0.146	-0.142
			(0.124)	(0.124)	(0.207)
Father	Some College		0.310**	0.300**	0.298**
			(0.099)	(0.099)	(0.099)
Father	College +		-0.014	-0.032	-0.043
			(0.117)	(0.119)	(0.119)
Mother	Random Effect			-0.128	-0.052
				(0.067)	(0.124)
Father	Random Effect			-0.034	-0.055
				(0.091)	(0.091)

Column 4 (cont)		
Marginal Effect of Average Parents' Earnings in years son was ages 1-5		
25th Pct.	50th Pct.	75th Pct.
0.405***	0.342***	0.281**
(0.083)	(.085)	(.092)

Sample is boys who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Regressions also included controls for age, age squared, age of mother at birth of son, black, indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 2B: Labor Force Experience of Sons in 2011: Full interaction Model

Mom work - all 5 years		Obs=3356			
Mother Educ	Father Educ	Low Ability	Average Ability	High Ability	Slope
HS or less	HS or less	0.514*	0.430*	0.346+	-0.279
		(0.222)	(0.190)	(0.201)	(0.311)
Some college	HS or less	-0.056	-0.024	0.007	0.104
		(0.516)	(0.473)	(0.472)	(0.464)
College +	HS or less	0.859	0.967	1.075	0.359
		(0.903)	(0.819)	(0.820)	(0.898)
HS or less	Some college	0.752+	0.651+	0.551+	-0.335
		(0.418)	(0.337)	(0.331)	(0.565)
Some college	Some college	0.087	0.062	0.037	-0.083
		(0.339)	(0.294)	(0.292)	(0.395)
College +	Some college	0.535	0.385	0.235	-0.500
		(0.636)	(0.571)	(0.567)	(0.644)
HS or less	College +	-0.444	-0.512	-0.580	-0.227
		(0.631)	(0.532)	(0.488)	(0.626)
Some college	College +	0.530	0.192	-0.147	-1.129*
		(0.389)	(0.337)	(0.339)	(0.465)
College +	College +	0.411	0.211	0.011	-0.667*
		(0.274)	(0.243)	(0.251)	(0.335)

Sample is boys who:
1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Full interactions included:
nine category mother/father education interacted with three category mother work
mother and father random effects interacted with nine category educ and three category mother work
Regressions also included controls for age, age squared, age of mother at birth of son, black, indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5
Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001, + p<.10

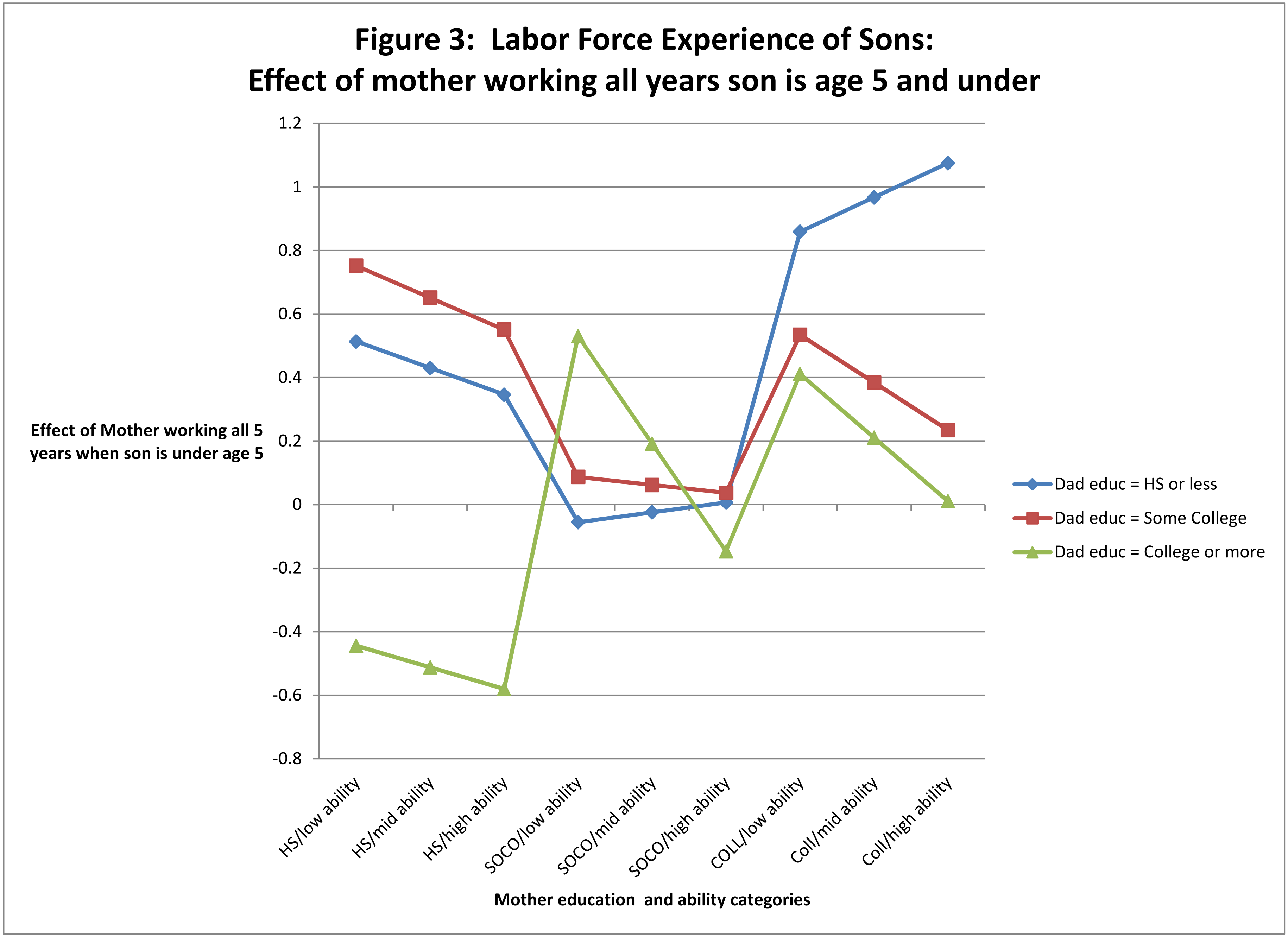

Figure 3: Labor Force Experience of Sons:
Effect of mother working all years son is age 5 and under
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Effect of Mother working all 5 years when son is under age 5
HS/low ability
HS/mid ability
HS/high ability
SOCO/low ability
SOCO/mid ability
SOCO/high ability
COLL/low ability
Coll/mid ability
Coll/high ability
Mother education and ability categories
Dad educ = HS or less
Dad educ = Some College
Dad educ = College or more

Table 3: Labor Force Experience of Sons at age 30 - Mother Fixed Effect/Sibling Model

obs=523 boys, 258 mothers			1	2
Mom work				
all educ levels		some years	0.254	
			(0.403)	
all educ levels		all 5 years	-0.150	
			(0.633)	
Mom work Interactions				
Mother educ	Father educ	Mom work		
HS or less	HS or less	some years		0.030
				(0.709)
HS or less	HS or less	all 5 years		0.010
				(1.309)
HS or less	Some College +	some years		1.005
				(0.884)
HS or less	Some College +	all 5 years		0.053
				(1.396)
Some College +	HS or less	some years		2.373*
				(1.106)
Some College +	HS or less	all 5 years		-0.211
				(1.911)
Some College +	Some College +	some years		-0.858
				(0.688)
Some College +	Some College +	all 5 years		-0.830
				(0.962)

Marginal Effect of Average Parents' Earnings in years son was ages 1-5 (Column 2 cont.)		
25th Pct.	50th Pct.	75th Pct.
0.244	0.193	0.168
(0.673)	(0.762)	(0.847)

Sample is boys who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33
6. had a brother who met criteria #1-#5

Regression also included controls for year dummies (2008, 2009, 2010, base year of 2011), indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 4A: Labor Force Experience of Daughters in 2011

		1	2	3	4
Mom work		Obs=2910			
all educ levels	some years	0.166	0.175	0.173	
		(0.104)	(0.104)	(0.104)	
all educ levels	all 5 years	0.310**	0.376***	0.325**	
		(0.113)	(0.113)	(0.116)	
Mom work Interactions					
HS or less	some years				0.174
					(0.143)
HS or less	all 5 years				0.566***
					(0.162)
Some college	some years				-0.002
					(0.194)
Some college	all 5 years				0.141
					(0.210)
College +	some years				0.277
					(0.236)
College +	all 5 years				0.165
					(0.237)
Mom RE	some years				-0.391*
					(0.173)
Mom RE	all 5 years				-0.439*
					(0.200)
Parental Controls					
Mother	Some College		0.258*	0.253*	0.458*
			(0.104)	(0.104)	(0.196)
Mother	College +		-0.334*	-0.343*	-0.233
			(0.134)	(0.134)	(0.240)
Father	Some College		0.263*	0.257*	0.258*
			(0.108)	(0.108)	(0.108)
Father	College +		0.487***	0.465***	0.447***
			(0.127)	(0.129)	(0.129)
Mother	Random Effect			0.090	0.385**
				(0.073)	(0.137)
Father	Random Effect			-0.144	-0.164
				(0.095)	(0.095)

Column 4 (cont)		
Marginal Effect of Average Parents' Earnings in years daughter was ages 1-5		
25th Pct.	50th Pct.	75th Pct.
.333***	.273**	.212*
(0.088)	(0.093)	(0.106)

Sample is girls who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Regressions also included controls for age, age squared, age of mother at birth of daughter, black, indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 4B: Labor Force Experience of Girls in 2011: Full interaction Model

		1	2	3	4
Mom work - all 5 years		Obs=2910			
Mother Educ	Father Educ	Low Ability	Average Ability	High Ability	Slope
HS or less	HS or less	0.625*	0.575**	0.524*	-0.168
		(0.247)	(0.201)	(0.210)	(0.369)
Some college	HS or less	0.257	0.297	0.337	0.133
		(0.469)	(0.409)	(0.421)	(0.587)
College +	HS or less	-2.301	-2.542	-2.784	-0.805
		(1.991)	(1.90)	(1.868)	(1.191)
HS or less	Some college	0.722+	0.559	0.396	-0.544
		(0.390)	(0.339)	(0.356)	(0.520)
Some college	Some college	0.474	0.393	0.312	-0.269
		(0.413)	(0.357)	(0.356)	(0.482)
College +	Some college	-0.049	-0.304	-0.560	-0.852
		(0.735)	(0.688)	(0.677)	(0.536)
HS or less	College +	1.271+	0.926	0.582	-1.148
		(0.681)	(0.589)	(0.571)	(0.731)
Some college	College +	-0.009	-0.306	-0.604	-0.991
		(0.617)	(0.454)	(0.393)	(0.824)
College +	College +	0.499	0.320	0.141	-0.597
		(0.311)	(0.282)	(0.295)	(0.371)

Sample is girls who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Full interactions included:

nine category mother/father education interacted with three category mother work

mother and father random effects interacted with nine category educ and three mother work categories

Regressions also included controls for age, age squared, age of mother at birth of daughter, black, indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001, + p<.10

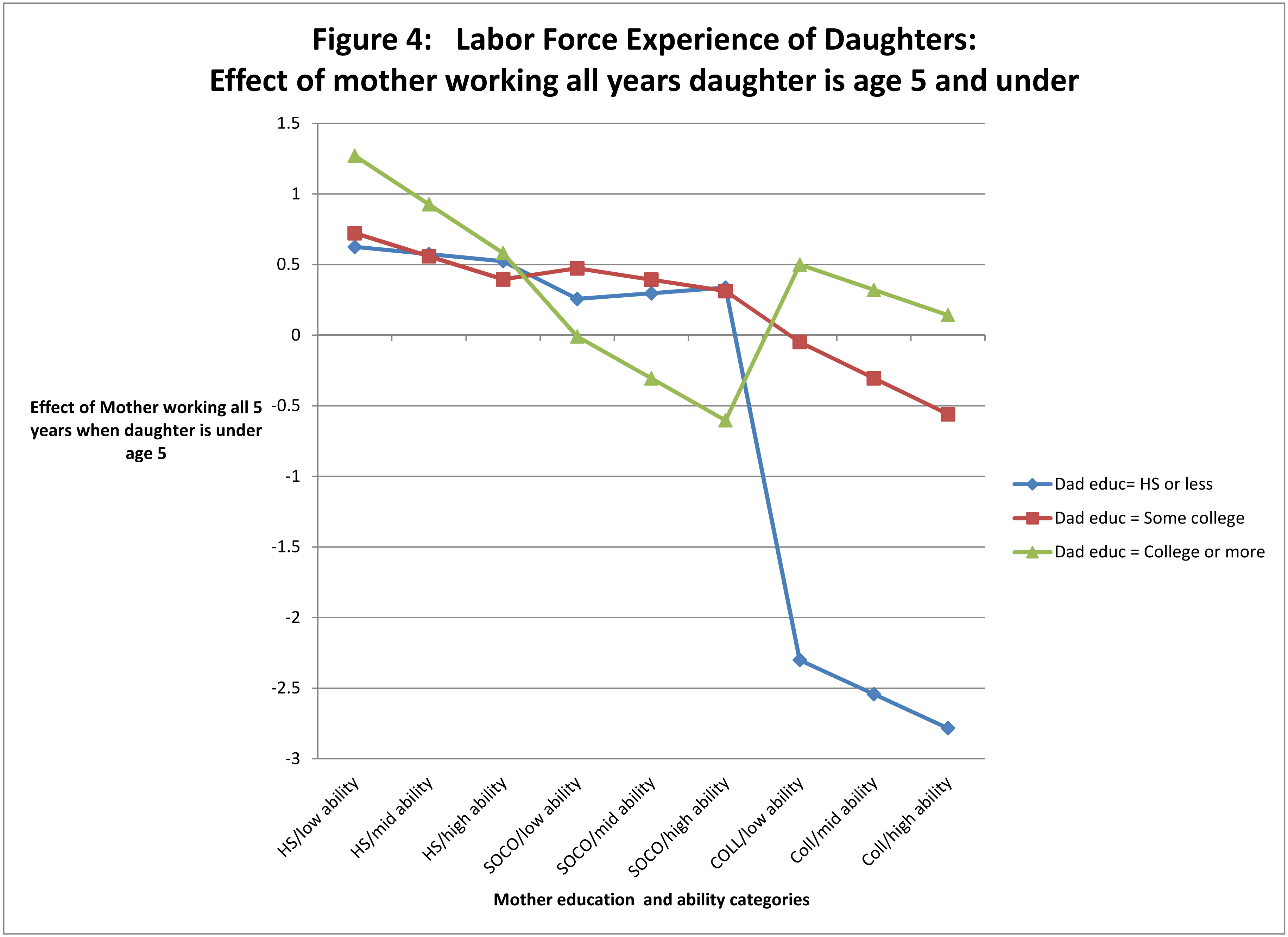

Figure 4: Labor Force Experience of Daughters:
Effect of mother working all years daughter is age 5 and under
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
Effect of Mother working all 5 years when daughter is under age 5
HS/low ability
HS/mid ability
HS/high ability
SOCO/low ability
SOCO/mid ability
SOCO/high ability
COLL/low ability
Coll/mid ability
Coll/high ability
Mother education and ability categories
Dad educ= HS or less
Dad educ = Some college
Dad educ = College or more

Table 5: Labor Force Experience of Daughters at age 30 - Mother Fixed Effect/Sibling Model

obs=375 girls, 184 mothers			1	2
Mom work				
	all educ levels	some years	0.077	
			(0.587)	
	all educ levels	all 5 years	0.701	
			(0.872)	
Mom work Interactions				
Mother educ	Father educ	Mom work		
HS or less	HS or less	some years		-1.442
				(0.956)
HS or less	HS or less	all 5 years		-0.638
				(1.248)
HS or less	Some College +	any years		0.468
				(1.736)
Some College +	HS or less	some years		1.119
				(2.465)
Some College +	HS or less	all 5 years		-0.144
				(3.462)
Some College +	Some College +	some years		1.112
				(0.869)
Some College +	Some College +	all 5 years		1.766
				(1.503)

Marginal Effect of Average Parents' Earnings in years daughter was ages 1-5 (Column 2 cont.)		
25th Pct.	50th Pct.	75th Pct.
0.585	0.665	0.768
(1.135)	(1.081)	(1.164)

Sample is girls who:
1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33
6. had a sister who met criteria #1-#5

Regression also included controls for year dummies (2008, 2009, 2010, base year of 2011), indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 6A: Earnings of Sons Ages 30-33

		1	2	3	4
Mom work		Obs=7781 son-years, 3356 sons			
all educ levels	some years	-0.049	-0.044	-0.017	
		(0.045)	(0.045)	(0.045)	
all educ levels	all 5 years	-0.074	-0.052	0.002	
		(0.050)	(0.051)	(0.050)	
Mom work Interactions					
HS or less	some years				-0.083
					(0.059)
HS or less	all 5 years				-0.080
					(0.073)
Some college	some years				0.019
					(0.084)
Some college	all 5 years				0.009
					(0.086)
College +	some years				0.121
					(0.108)
College +	all 5 years				0.187
					(0.105)
Mom RE	some years				0.038
					(0.079)
Mom RE	all 5 years				-0.009
					(0.084)
Parental Controls					
Mother	Some College		0.009	0.026	-0.042
			(0.044)	(0.044)	(0.084)
Mother	College +		0.074	0.112	-0.062
			(0.059)	(0.058)	(0.101)
Father	Some College		0.161***	0.200***	0.196***
			(0.045)	(0.045)	(0.045)
Father	College +		0.216***	0.289***	0.290***
			(0.054)	(0.054)	(0.054)
Mother	Random Effect			0.125***	0.111
				(0.034)	(0.057)
Father	Random Effect			0.330***	0.333***
				(0.045)	(0.045)

Column 4 (cont)		
Marginal Effect of Average Parents' Earnings in years son was ages 1-5		
25th Pct.	50th Pct.	75th Pct.
0.042	0.05	0.055
(.038)	(.038)	(.041)

Sample is boys who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Regressions also included controls for age, age squared, age of mother at birth of son, black, indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5 and year dummies (2008, 2009, 2010, 2011 base year); clustered standard errors

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 6B: Earnings of Sons Ages 30-33 Full interaction Model

		1	2	3	4
Mom work - all 5 years		Obs=7781 son-years, 3356 sons			
Mother Educ	Father Educ	Low Ability	Average Ability	High Ability	Slope
HS or less	HS or less	-0.107	-0.094	-0.082	0.041
		(0.108)	(0.094)	(0.099)	(0.151)
Some college	HS or less	0.449	0.355	0.260	-0.316
		(0.275)	(0.261)	(0.262)	(0.210)
College +	HS or less	0.663*	0.527+	0.392	-0.451*
		(0.274)	(0.280)	(0.300)	(0.203)
HS or less	Some college	0.045	0.025	0.005	-0.066
		(0.181)	(0.142)	(0.146)	(0.277)
Some college	Some college	-0.073	-0.015	0.043	0.194
		(0.161)	(0.138)	(0.133)	(0.182)
College +	Some college	-0.187	-0.069	0.049	0.394
		(0.301)	(0.264)	(0.271)	(0.372)
HS or less	College +	-0.303	-0.337+	-0.371	-0.113
		(0.262)	(0.218)	(0.208)	(0.309)
Some college	College +	-0.009	-0.057	-0.104	-0.158
		(0.158)	(0.142)	(0.156)	(0.221)
College +	College +	0.285*	0.236*	0.187	-0.162
		(0.128)	(0.116)	(0.127)	(0.175)

Sample is boys who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Full interactions included:

nine category mother/father education interacted with three category mother work

mother and father random effects interacted with nine category educ and three mother work categories

Regressions also included controls for age, age squared, age of mother at birth of son, black,
indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5
and year dummies (2008, 2009, 2010, 2011 base year); clustered standard errors

Standard errors in parentheses   * p<0.05, ** p<0.01, *** p<0.001, + p<.10

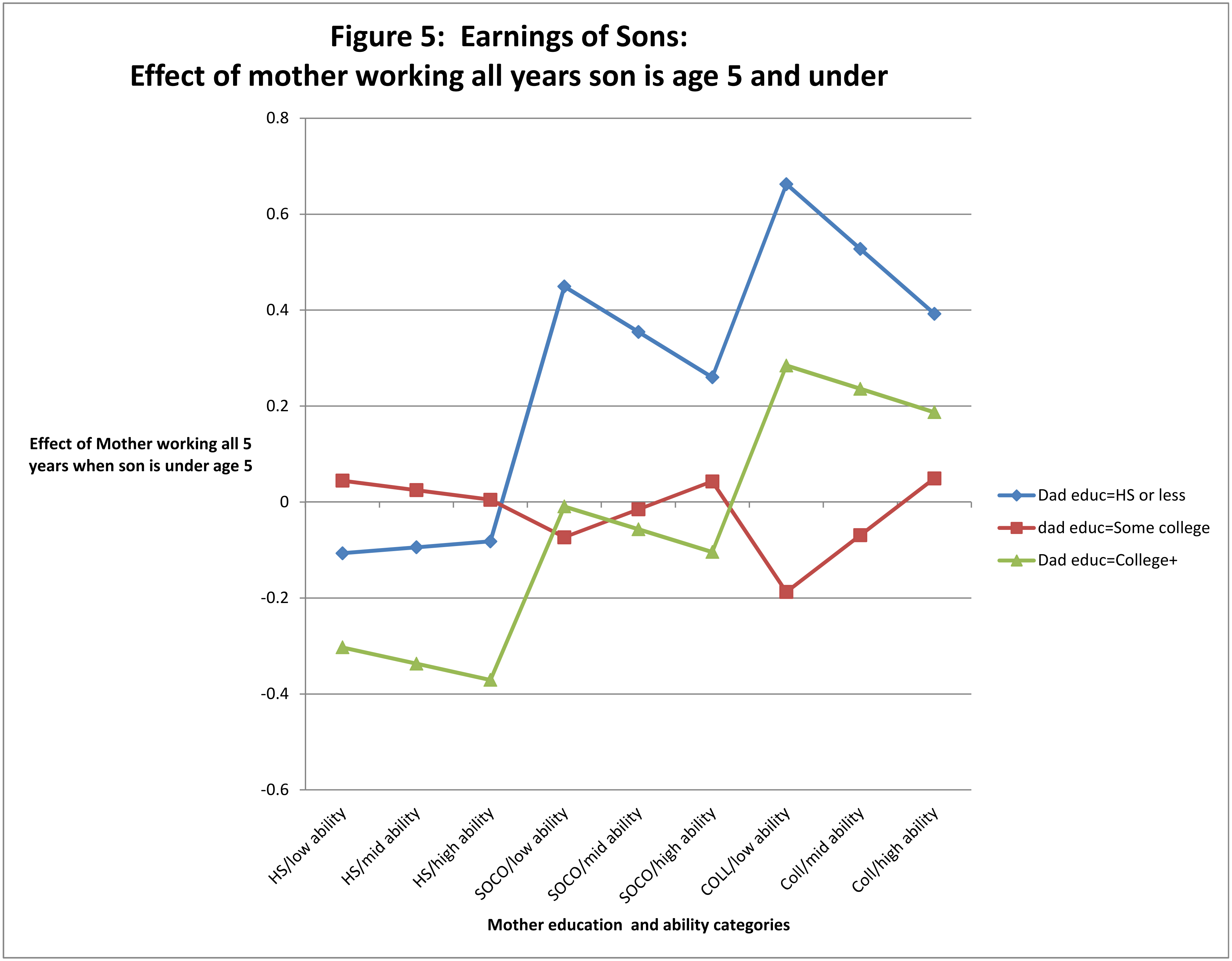
Figure 5: Earnings of Sons:
Effect of mother working all years son is age 5 and under
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Effect of Mother working all 5 years when son is under age 5
HS/low ability
HS/mid ability
HS/high ability
SOCO/low ability
SOCO/mid ability
SOCO/high ability
COLL/low ability
Coll/mid ability
Coll/high ability
Mother education and ability categories
Dad educ=HS or less
dad educ=Some college
Dad educ=College+

Table 7: Earnings of Sons at age 30 - Mother Fixed Effect/Sibling Model

obs=506 boys, 258 mothers			1	2
Mom work				
all educ levels		some years	0.542	
			(0.298)	
all educ levels		all 5 years	0.371	
			(0.463)	
Mom work Interactions				
Mother educ	Father educ	Mom work		
HS or less	HS or less	some years		0.930
				(0.513)
HS or less	HS or less	all 5 years		0.774
				(0.948)
HS or less	Some College +	some years		0.442
				(0.639)
HS or less	Some College +	all 5 years		-0.669
				(1.010)
Some College +	HS or less	some years		0.503
				(0.802)
Some College +	HS or less	all 5 years		0.306
				(1.384)
Some College +	Some College +	some years		0.180
				(0.558)
Some College +	Some College +	all 5 years		0.392
				(0.749)

Marginal Effect of Average Parents' Earnings in years son was ages 1-5 (column 2 cont.)		
25th Pct.	50th Pct.	75th Pct.
-0.142	-0.031	0.026
(.494)	(.567)	(.640)

Sample is boys who:

1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33
6. had a brother who met criteria #1-#5

Regression also included controls for year dummies (2008, 2009, 2010, base year of 2011), indicator for oldest child, constant, cubic in average parents' earnings when son was age 1-5

Standard errors in parentheses * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 8A: Earnings of Daughters Ages 30-33

		1	2	3	4
Mom work		Obs=6777 daughter-years, 2910 daughters			
all educ levels	some years	-0.051	-0.078	-0.045	
		(0.053)	(0.053)	(0.053)	
all educ levels	all 5 years	-0.033	-0.023	-0.019	
		(0.058)	(0.058)	(0.059)	
Mom work Interactions					
HS or less	some years				-0.132
					(0.067)
HS or less	all 5 years				-0.129
					(0.081)
Some college	some years				0.071
					(0.102)
Some college	all 5 years				0.064
					(0.107)
College +	some years				0.038
					(0.132)
College +	all 5 years				0.147
					(0.137)
Mom RE	some years				-0.033
					(0.088)
Mom RE	all 5 years				-0.074
					(0.103)
Parental Controls					
Mother	Some College		0.156**	0.168***	0.022
			(0.051)	(0.051)	(0.099)
Mother	College +		0.119	0.145*	-0.027
			(0.065)	(0.064)	(0.130)
Father	Some College		0.190***	0.228***	0.227***
			(0.053)	(0.053)	(0.053)
Father	College +		0.335***	0.391***	0.398***
			(0.059)	(0.059)	(0.060)
Mother	Random Effect			0.237***	0.268***
				(0.036)	(0.071)
Father	Random Effect			0.171***	0.173***
				(0.047)	(0.047)

Column 4 (cont)		
Marginal Effect of Average Parents' Earnings in years daughter was ages 1-5		
25th Pct.	50th Pct.	75th Pct.
0.068	0.03	-0.004
(.042)	(.048)	(.061)

Sample is girls who:
1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Regressions also included controls for age, age squared, age of mother at birth of daughter, black, indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5 and year dummies (2008, 2009, 2010, 2011 base year); clustered standard errors

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 8B: Earnings of Daughters Ages 30-33 Full interaction Model

		1	2	3	4
Mom work - all 5 years		Obs=6777 daughter-years, 2910 daughters			
Mother Educ	Father Educ	Low Ability	Average Ability	High Ability	Slope
HS or less	HS or less	-0.135	-0.174	-0.213*	-0.130
		(0.144)	(0.110)	(0.103)	(0.201)
Some college	HS or less	0.085	0.012	-0.062	-0.244
		(0.272)	(0.245)	(0.250)	(0.298)
College +	HS or less	-0.631	-0.643*	-0.655*	-0.041
		(0.422)	(0.289)	(0.194)	(0.515)
HS or less	Some college	-0.010	-0.003	0.004	0.023
		(0.192)	(0.154)	(0.145)	(0.246)
Some college	Some college	0.352	0.239	0.127	-0.374
		(0.249)	(0.203)	(0.188)	(0.292)
College +	Some college	-0.345	-0.385	-0.425+	-0.133
		(0.284)	(0.246)	(0.230)	(0.268)
HS or less	College +	0.000	-0.074	-0.148	-0.246
		(0.267)	(0.229)	(0.213)	(0.262)
Some college	College +	-0.242	-0.240	-0.238	0.007
		(0.232)	(0.177)	(0.156)	(0.293)
College +	College +	0.340+	0.276+	0.213	-0.212
		(0.175)	(.163)	(0.167)	(0.166)

Sample is girls who:
1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33

Full interactions included:
nine category mother/father education interacted with three category mother work
mother and father random effects interacted with nine category educ and three mother work categories
Regressions also included controls for age, age squared, age of mother at birth of daughter, black,
indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5
and year dummies (2008, 2009, 2010, 2011 base year); clustered standard errors
Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001, + p<.10

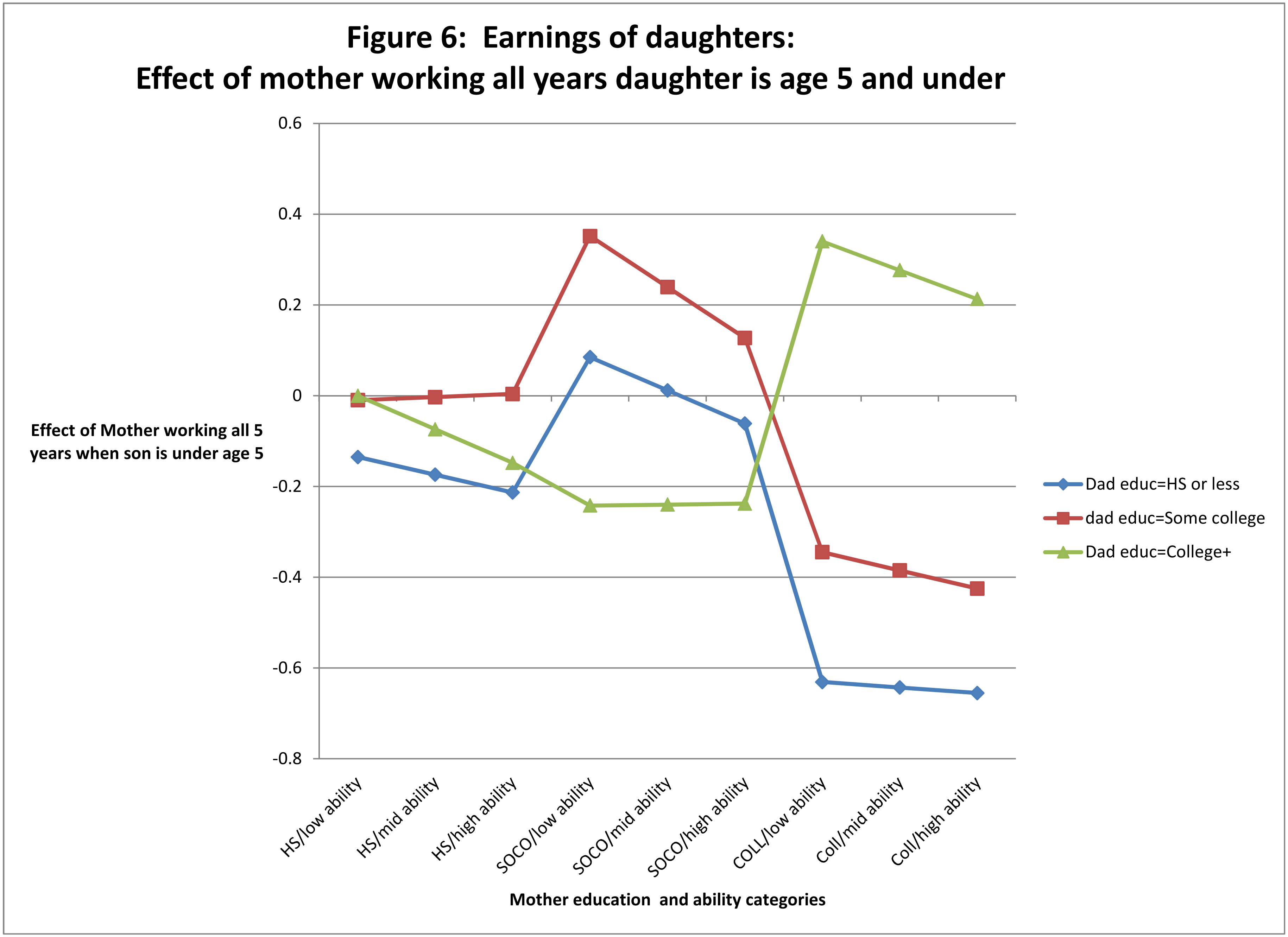
Figure 6: Earnings of daughters:
Effect of mother working all years daughter is age 5 and under
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Effect of Mother working all 5 years when son is under age 5
HS/low ability
HS/mid ability
HS/high ability
SOCO/low ability
SOCO/mid ability
SOCO/high ability
COLL/low ability
Coll/mid ability
Coll/high ability
Mother education and ability categories
Dad educ=HS or less
dad educ=Some college
Dad educ=College+

Table 9: Earnings of Daughters at age 30 - Mother Fixed Effect/Sibling Model

obs= 364 girls, 184 mothers			1	2
Mom work				
	all educ levels	some years	0.674	
			(0.381)	
	all educ levels	all 5 years	-0.301	
			(0.594)	
Mom work Interactions				
Mother educ	Father educ	Mom work		
HS or less	HS or less	some years		0.733
				(0.590)
HS or less	HS or less	all 5 years		-0.476
				(0.830)
HS or less	Some College +	any years		0.334
				(1.525)
Some College +	HS or less	some years		-0.248
				(1.522)
Some College +	HS or less	all 5 years		-0.773
				(2.134)
Some College +	Some College +	some years		0.773
				(0.578)
Some College +	Some College +	all 5 years		0.186
				(1.044)

Marginal Effect of Average Parents' Earnings in years daughter was ages 1-5 (column 2 cont.)		
25th Pct.	50th Pct.	75th Pct.
0.109	-0.055	-0.158
(.722)	(.687)	(.733)

Sample is girls who:
1. were born between 1978 and 1981 and surveyed by SIPP in 1984, 1990-1993, 1996 panels
2. were living with both biological parents at time of survey
3. had valid SSN to be used in linking to admin. earnings data, both parents had valid SSN
4. had mom who worked at least one year between 1978-2011
5. worked themselves at least one year between 30 and 33
6. had a sister who met criteria #1-#5

Regression also included controls for year dummies (2008, 2009, 2010, base year of 2011), indicator for oldest child, constant, cubic in average parents' earnings when daughter was age 1-5

Standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Appendix Table 2B: Labor Force Experience of Sons in 2011: Full interaction Model
F-statistics for differences in mother working effect between parent educational groups

	Mother education: Comparison across categories		
Father Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	1.05	0.14	0.78
Some college	1.57	0.08	0.39
College	1.75	1.57	0.07

	Father education: Comparison across categories		
Mother Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.26	2.08	2.53
Some college	0.05	0.84	0.76
College	0.09	0.23	0.03

Appendix Table 4B: Labor Force Experience of Daughters in 2011: Full interaction Model
F-statistics for differences in mother working effect between parent educational groups

	Mother education: Comparison across categories		
Father Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.51	2.14	1.57
Some college	0.20	0.87	0.39
College	1.97	1.08	0.55

	Father education: Comparison across categories		
Mother Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.05	0.81	0.5
Some college	0.12	0.12	0.43
College	1.13	1.94	0.48

Appendix Table 6B: Earnings of Sons Ages 30-33: Full interaction Model

F-statistics for differences in mother working effect between parent educational groups

	Mother education: Comparison across categories		
Father Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	3.64	6.99	0.31
Some college	0.24	0.44	0.11
College	0.94	4.14	2.21

	Father education: Comparison across categories		
Mother Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.53	0.49	1.21
Some college	2.74	2.14	0.08
College	4.42	1.60	2.13

Appendix Table 8B: Earnings of Daughters Ages 30-33: Full interaction Model

F-statistics for differences in mother working effect between parent educational groups

	Mother education: Comparison across categories		
Father Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.53	1.28	2.06
Some college	1.35	0.98	3.46
College	0.48	1.16	4.12

	Father education: Comparison across categories		
Mother Educ	HS or less v. Soco	HS or less v. College+	Soco v. College+
HS or less	0.29	0.21	0.00
Some college	0.53	0.85	3.10
College	0.32	4.60	4.32