CHAPTER 15

Human Capital Development before Age Five☆

Douglas Almond, Janet Currie
Columbia University

Contents

1. Introduction 1316
2. Conceptual Framework 1322
   2.1. Complementarity 1323
   2.2. Fixed investments 1324
      2.2.1. Remediation *1325*
   2.3. Responsive investments 1326
3. Methods 1328
   3.1. Power 1333
   3.2. Data constraints 1336
      3.2.1. Leveraging existing datasets *1336*
      3.2.2. Improvements in the production of administrative data *1338*
      3.2.3. Additional issues *1339*
4. Empirical Literature: Evidence of Long Term Consequences 1340
   4.1. Prenatal environment 1341
      4.1.1. Maternal health *1341*
      4.1.2. Economic shocks *1367*
      4.1.3. Air pollution *1367*
   4.2. Early childhood environment 1370
      4.2.1. Infections *1371*
      4.2.2. Health status *1371*
      4.2.3. Home environment *1394*
      4.2.4. Toxic exposures *1396*
      4.2.5. Summary re: long term effects of fetal and early childhood environment *1396*
5. Empirical Literature: Policy Responses 1396
   5.1. Income enhancement 1397
   5.2. Near-cash programs 1405
   5.3. Early intervention programs 1419
      5.3.1. Home visiting *1425*
      5.3.2. US supplemental feeding program for women, infants, and children (WIC) *1426*
      5.3.3. Child care *1433*
      5.3.4. Health insurance *1453*

☆ We thank Maya Rossin and David Munroe for excellent research assistance, participants in the Berkeley Handbook of Labor Economics Conference in November 2009 for helpful comments, and Christine Pal and Hongyan Zhao for proofreading the equations.

E-mail addresses: da2152@columbia.edu (Douglas Almond), jc2663@columbia.edu (Janet Currie).

ISSN 0169-7218, DOI 10.1016/S0169-7218(11)02413-0

6. Discussion and Conclusions 1467
Appendix A 1468
Appendix B 1471
Appendix C 1472
Appendix D 1475
References 1476

Abstract

This chapter seeks to set out what economists have learned about the effects of early childhood influences on later life outcomes, and about ameliorating the effects of negative influences. We begin with a brief overview of the theory which illustrates that evidence of a causal relationship between a shock in early childhood and a future outcome says little about whether the relationship in question is biological or immutable. We then survey recent work which shows that events before five years old can have large long term impacts on adult outcomes. Child and family characteristics measured at school entry do as much to explain future outcomes as factors that labor economists have more traditionally focused on, such as years of education. Yet while children can be permanently damaged at this age, an important message is that the damage can often be remediated. We provide a brief overview of evidence regarding the effectiveness of different types of policies to provide remediation. We conclude with a list of some of the many outstanding questions for future research.

JEL classification: I12; I21; J13; J24; Q53

Keywords: Human capital; Early childhood; Health; Fetal origins

1. INTRODUCTION

The last decade has seen a blossoming of research on the long term effects of early childhood conditions across a range of disciplines. In economics, the focus is on how human capital accumulation responds to the early childhood environment. In 2000, there were no articles on this topic in the *Journal of Political Economy, Quarterly Journal of Economics,* or *the American Economic Review* (excluding the Papers and Proceedings), but there have been five or six per year in these journals since 2005. This work has been spurred by a growing realization that early life conditions can have persistent and profound impacts on later life. Table 1 summarizes several longitudinal studies which suggest that characteristics that are measured as of age 7 can explain a great deal of the variation in educational attainment, earnings as of the early 30s, and the probability of employment. For example, McLeod and Kaiser (2004) use data from the National Longitudinal Surveys and find that children's test scores and background variables measured as of ages 6 to 8 predict about 12% of the variation in the probability of high school completion and about 11% of the variation in the probability of college completion. Currie and Thomas (1999b) use data from the 1958 British Birth Cohort study and find that 4% to 5% of the variation in employment at age 33 can be predicted, and as much as 20% of the variation in wages. Cunha and Heckman (2008) and Cunha et al. (2010) estimate structural models in which initial endowments and investments feed

Table 1 How much of the differences in later outcomes (test scores, educational attainment, earnings) can be explained by early childhood factors?

Study and data	Inputs/intermediate variables	Outcomes	Results
Studies using National Longitudinal Survey of Youth-child Sample Data			
Childhood emotional and behavioral problems and educational attainment (McLeod and Kaiser, 2004) NLSY-child sample data on children 6-8 in 1986. Five waves through 2000 (when children 20-22) $n = 424$.	Emotional and behavioral problems at age 6-8 measured by BPI, mother emotional problems and delinquency, poverty status, mother's AFQT score, mother's education, mother's marital status and age, child's age, sex, and race, dummy for LBW.	Dummy for graduating high school by 2000, dummy for enrolling in college by 2000.	**R-squared for predicting HS graduation:** Only child emotional and behavioral problems: 0.046 Add child and mother demographics: 0.111 Add mother's emotional problems and delinquency: 0.124. **R-squared for predicting college enrollment:** Only child emotional and behavioral problems: 0.017 Add child and mother demographics: 0.093 Add mother's emotional problems and delinquency: 0.112.

(*continued on next page*)

Table 1 (continued)

Study and data	Inputs/intermediate variables	Outcomes	Results
Formulating, identifying, and estimating the technology of cognitive and noncognitive skill formation (Cunha and Heckman, 2008) NLSY-child sample data on white males. $n = 1053$.	Home environment measured by HOME score. Measures of parental investments: number of child's books, whether the child has a musical instrument, whether the family receives a daily newspaper, whether the child receives special lessons, how often the child goes to museums, and how often the child goes to the theater. Measures of child cognitive skills: PIAT test scores at various ages. Measures of child's non-cognitive skills: BPI at various ages.	Log earnings and likelihood of graduation from high school. Estimated a dynamic factor model that exploits cross-equation restrictions for identification. Estimated effects of cognitive and non-cognitive skills as well as parental investments at different stages of the child's lifecycle on log earnings and likelihood of graduation from high school.	A 10% increase in parental investments at ages 6-7 increases earnings by 24.9% (12.5% through cognitive skills, 12.4% through non-cognitive skills), and increases likelihood of graduating high school by 64.4% (54.8% through cognitive skills, 9.6% through non-cognitive skills).

Table 1 (continued)

Study and data	Inputs/intermediate variables	Outcomes	Results
Estimating the technology of cognitive and non-cognitive skill formation (Cunha et al., 2010) NLSY-child sample first-born white children. $N = 2207$.	Examples of measures of child cognitive skills: Motor-social development at birth; PIAT reading-comprehension at ages 5-6; PIAT math at ages 13-14. Examples of measures of child non-cognitive skills: BPI at various ages; "friendliness at birth". Examples of measures of parental investments: "How often child goes on outings during year of birth"; "How often mom reads to child during year of birth"; "Child has a CD player, ages 3-4"; How often child is praised, ages 13-14".	Main outcome variable: completed years of education by age 19. Multi-stage production functions estimated in which the productivity of later investments depends on early investments and on the stock of cognitive and non-cognitive skills and investments are endogenous. Identification based on nonlinear factor models with endogenous inputs. In preferred specification, used maximum-likelihood methods to estimate the production technology.	34% of variation in educational attainment is explained by measures of cognitive and non-cognitive capabilities. 16% is due to adolescent cognitive capabilities; 12% is due to adolescent non-cognitive capabilities. Parental endowments/investments account for 15% of variation in educational attainment.

(*continued on next page*)

Table 1 (continued)

Study and data	Inputs/intermediate variables	Outcomes	Results
Studies using National Child Development Study (1958 British Birth Cohort) Data			
Ability, family, education, and earnings in Britain (Dearden, 1998) Focus on individuals who participated in waves 4 and 5 of the survey in 1981 and 1991, who were employees in 1991. $n = 2597$ males, 2362 females.	Math and verbal ability (age 7), type of school, family characteristics—teacher's assessment of interest shown by parents in child's education at 7; type of school attended at 16 family's financial status at 11 and 16; region dummies; father's SES; parents' education levels.	Years of full-time education; Earnings at age 33.	**R-squared for education as outcome:** Including all explanatory variables: 0.33-0.34. **R-squared for earnings as outcome:** Baseline earnings equation including only years education: 0.15 (males), 0.25 (females) Add reading and math scores at age 7, school type and regional dummies: 0.26 (males), 0.31 (females) Including all explanatory variables:.029 (males), 0.41 for females.
Early test scores, socioeconomic status, and future outcomes (Currie and Thomas, 1999b)	Reading and math test scores at age 7, mother and father's SES and education, birth weight, other child background variables at age 7.	Number of O-level passes of exams by age 16; employed at age 23, 33; log wage at age 23, 33.	**R-squared for predicting age 16 exam passes:** Reading and math scores only: 0.21-0.22 Add other background variables: 0.31-0.32.

Table 1 (continued)

Study and data	Inputs/intermediate variables	Outcomes	Results
Full sample size (based on responses at 7): $n = 14,022$.			**R-squared for predicting employment at 33:** Reading and math test scores only: 0.01 Add other background variables: 0.04-0.05. **R-squared for predicting log wage at 33:** Reading and math test scores only: 0.08-0.09 Add other background variables: 0.18-0.20.
The lasting impact of childhood health and circumstance (Case et al., 2005) $n = 14,325$ (7016 men, 7039 women)	Mother's and father's education and SES, LBW, indicators for moderate, heavy, and varied maternal smoking during pregnancy, number of chronic conditions at age 7 and 16.	Number of O-level passes of exams by age 16; adult health status at age 42; part-time or full employment at age 42.	**R-squared for predicting age 16 exam passes/adult health/employment at 42.** Mother's education and SES: 0.062/0.082/0.076 Father's education and SES: 0.241/0.189/0.173 LBW and maternal smoking: 0.024/0.086/0.052.
Explaining intergenerational income persistence: non-cognitive skills, ability, and education (Blanden et al., 2006) **A.** NCDS: 1958 cohort Focus on 2163 males. **B.** British Cohort Study: 1970 cohort Focus on 3340 males.	**A.** Family income at age 16 reading and math test scores at age 11, scores for "behavioral syndromes" at age 11, O-level exam scores at age 16. **B.** Years of preschool education, birth weight, height at 5 and 10, emotional/behavioral scores at ages 5, 10, & 16, family income at ages 10 and 16, reading and math test scores at age 10, IQ at age 10, dummy for HS degree, exam scores at age 16.	**A.** Earnings at age 33. **B.** Earnings at age 30.	**R-squared Model A.** Birth weight, childhood health, and age 11 test scores only: 0.116 Including "behavioral syndromes" at age 11: 0.151 All variables: 0.263. **R-squared Model B.** Birth weight, childhood health, and age 10 test scores only: 0. 075. Including emotional/behavioral characteristics at age 10: 0.087. All variables: 0.222.

through to later outcomes; they arrive at estimates that are of a similar order of magnitude for education and wages. To put these results in context, labor economists generally feel that they are doing well if they can explain 30% of the variation in wages in a human capital earnings function.

This chapter seeks to set out what economists have learned about the importance of early childhood influences on later life outcomes, and about ameliorating the effects of negative influences. We begin with a brief overview of the theory which illustrates that evidence of a causal relationship between a shock in early childhood and a future outcome says little about whether the relationship in question is biological or immutable. Parental and social responses are likely to be extremely important in either magnifying or mitigating the effects of a shock. Given that this is the case, it can sometimes be difficult to interpret the wealth of empirical evidence that is accumulating in terms of an underlying structural framework.

The theoretical framework is laid out in Section 2 and followed by a brief discussion of methods in Section 3. We do not attempt to cover issues such as identification and instrumental variables methods, which are covered in some depth elsewhere (*cf* Angrist and Pischke (2009)). Instead, we focus on several issues that come up frequently in the early influences literature, including estimation using small samples and the potentially high return to better data.

Section 4 discusses the evidence for long-term effects of early life influences in greater detail, while Section 5 focuses on the evidence regarding remediation programs. The discussion of early life influences is divided into two subsections corresponding to *in utero* influences and after birth influences. The discussion of remediation programs starts from the most general sort of program, income transfers, and goes on to discuss interventions that are increasingly targeted at specific domains. In surveying the evidence we have attempted to focus on recent papers, and especially those that propose a plausible strategy for identifying causal effects. We have focused on papers that emphasize early childhood, but in instances in which only evidence regarding effects on older children is available, we have sometimes strayed from this rule. A summary of most of the papers discussed in these sections is presented in Tables 4 through 13. A list of acronyms used in the tables appears in Appendix A. We conclude with a summary and a discussion of outstanding questions for future research in Section 6.

2. CONCEPTUAL FRAMEWORK

Grossman (1972) models health as a stock variable that varies over time in response to investments and depreciation. Because some positive portion of the previous period's health stock vanishes in each period (e.g., age in years), the effect of the health stock and health investments further removed in time from the current period tends to fade out. As individuals age, the early childhood health stock and the prior health investments that it embodies become progressively less important.

In contrast, the "early influences" literature asks whether health and investments in early childhood have sustained effects on adult outcomes. The magnitude of these effects may persist or even increase as individuals age because childhood development occurs in distinct stages that are more or less influential of adult outcomes.

Defining h as health or human capital at the completion of childhood, we can retain the linearity of h in investments and the prior health stock as in Grossman (1972), but leave open whether there is indeed "fade out" (i.e. depreciation). For simplicity, we will consider a simple two-period childhood.[1] We can consider production of h:

$$h = A[\gamma I_1 + (1-\gamma) I_2], \tag{1}$$

where:

$$\begin{cases} I_1 \cong \text{investments during childhood through age 5} \\ I_2 \cong \text{investments during childhood after age 5.} \end{cases}$$

For a given level of total investments $I_1 + I_2$, the allocation of investments between period 1 and 2 will affect h for $\gamma \neq 0.5$. If $\gamma > 0.5$, then health at the end of period 1 is more important to h than investments in the second period, and if $\gamma A > 1$, h may respond more than one-for-one with I_1. Thus, (1) admits the possibility that certain childhood periods may exert a disproportionate effect on adult outcomes that does not necessarily decline monotonically with age. This functional form says more than "early life" matters; it suggests that early-childhood events may be more influential than later childhood events.

2.1. Complementarity

The assumption that inputs at different stages of childhood have linear effects is common in economics. While it opens the door to "early origins," perfect substitutability between first and second period investments in (1) is a strong assumption. The absence of complementarity implies that all investments should be concentrated in one period (up to a discount factor) and no investments should be made during the low-return period. In addition, with basic preference assumptions, perfect substitutability "hard-wires" the optimal investment response to early-life shocks to be compensatory, as seen in Section 2.3.

As suggested by Heckman (2007), a more flexible "developmental" technology is the constant elasticity of substitution (CES) function:

$$h = A\left[\gamma I_1^{\phi} + (1-\gamma) I_2^{\phi}\right]^{1/\phi}, \tag{2}$$

[1] See Zweifel et al. (2009) for a two period version of the Grossman (1972) model.

For a given total investment level $I_1 + I_2$, how the allocation between period 1 and 2 will also affect h depends on the elasticity of substitution, $1/(1-\phi)$, and the share parameter, γ. For $\phi = 1$ (perfect substitutability of investments), (2) reduces to (1).

Heckman (2007) highlights two features of "capacity formation" beyond those captured in (2). First, there may be "dynamic complementarities" which imply that investments in period t are more productive when there is a high level of capability in period $t-1$. For example, if the factor productivity term A in (2) were an increasing function of h_0, the health endowment immediately prior to period 1, this would raise the return to investments during childhood. Second, there may be "self-productivity" which implies that higher levels of capacity in one period create higher levels of capacity in future periods. This feature is especially noteworthy when h is multidimensional, as it would imply that "cross-effects" are positive, e.g. health in period 1 leads to higher cognitive ability in period 2. "Self-productivity" is more trivial in the unidimensional case like Grossman (1972)—even though the effect of earlier health stocks tends to fade out as the time passes, there is still memory as long as depreciation in each period is less than total (i.e., when $\delta < 1$ in Zweifel et al. (2009)).

Here, we will use the basic framework in (2) to consider the effect of exogenous shocks μ_g to health investments that occur during the first childhood period.[2] We begin with the simplest case, where investments do not respond to μ_g (and denote these investments $\bar{I}_1$ and $\bar{I}_2$). Net investments in the first period are:

$$\bar{I}_1 + \mu_g.$$

We assume that μ_g is independent of $\bar{I}_1$. While μ_g can be positive or negative, we assume $\bar{I}_1 + \mu_g > 0$. We will then relax the assumption of fixed investments, and consider endogenous responses to investments in the second period, i.e. $\delta I_2^*/\delta\mu_g$, and how this investment response may mediate the observed effect on h.

2.2. Fixed investments

Conceptually, we can trace out the effect of μ_g while holding other inputs fixed, i.e., we assume no investment response to this shock in either period. Albeit implicitly, most biomedical and epidemiological studies in the "early origins" literature aim to inform us about this *ceteris paribus*, "biological" relationship.

In this two-period CES production function adopted from Heckman (2007), the impact of an early-life shock on adult outcomes is:

$$\frac{\delta h}{\delta \mu_g} = \gamma A \left[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)\bar{I}_2^{\phi}\right]^{(1-\phi)/\phi} (\bar{I}_1 + \mu_g)^{\phi-1}. \tag{3}$$

[2] We include the subscript here because environmental influences at some aggregated geographic level g may provide exogenous variation in early childhood investments.

The simplest production technology is the perfect substitutability case where $\phi = 1$. In this case:

$$\frac{\delta h}{\delta \mu_g} = \gamma A.$$

Damage to adult human capital is proportional to the share parameter on period 1 investments, and is unrelated to the investment level $\bar{I}_1$.

For less than perfect substitutability between periods, there is diminishing marginal productivity of the investment inputs. Thus, shocks experienced at different baseline investment levels have heterogenous effects on h. Other things equal, those with higher baseline levels of investment will experience more muted effects in h than those where baseline investment is low. A recurring empirical finding is that long-term damage due to shocks is more likely among poorer families (Currie and Hyson, 1999). This is in part due to the fact that children in poorer families are subject to more or larger early-life shocks (Case et al., 2002; Currie and Stabile, 2003). However, it is also possible that the same shock will have a greater impact among children in poorer families if these children have lower period t investment levels to begin with. This occurs because they are on a steeper portion of the production function. *Ceteris paribus*, this would tend to accentuate the effect of an equivalent-sized μ_g shock on h among poor families.[3,4]

2.2.1. Remediation

Is it possible to alter "bad" early trajectories? In other words, what is the effect of a shock $\mu_g' > 0$ experienced during the second period on h? Remediation is of interest to the extent that (3) is substantially less than zero. However, large damage to h from μ_g *per se* says little about the potential effectiveness of remediation in the second period, as both initial damage and remediation are distinct functions of the three parameters A, γ, and ϕ.

The effectiveness of remediation relative to initial damage is:

$$\frac{\delta h/\delta \mu_g'}{\delta h/\delta \mu_g} = \frac{1-\gamma}{\gamma}\left(\frac{\bar{I}_1 + \mu_g}{\bar{I}_2 + \mu_g'}\right)^{1-\phi}. \tag{4}$$

Thus, for $\bar{I}_1 > \bar{I}_2$ and a given value of γ, a unit of remediation will be more effective at low elasticities of substitution—the lack of $\bar{I}_2$ was the more critical shortfall prior to the shock. If $\bar{I}_1 < \bar{I}_2$ high elasticities of substitution increase the effectiveness of remediation—adding to the existing abundance of $\bar{I}_2$ remains effective.

[3] I.e. $\delta^2 h/\delta \mu_g \delta I_1 < 0$. On the other hand, $\delta^2 h/\delta \mu_g \delta I_2 > 0$ so lower period two investments would tend to reduce damage to h from μ_g. The ratio of the former effect to the latter is proportional to $\gamma/(1-\gamma)$ (Chiang, 1984). Thus, damage from a period 1 shock is more likely to be concentrated among poor families when the period-1 share parameter (γ) is high.

[4] The cross-effect $\delta^2 h/\delta \mu_g \delta I_2$ is similar to dynamic complementarity, but Heckman (2007) reserves this term for the cross-partial between the stock and flow, i.e. $\delta^2 h/\delta h_0 \delta I_t$ for $t = 1, 2$ in the example of Section 2.1.

Fortunately, it is not necessary to observe investments and estimate all three parameters in order to assess the scope for remediation. In some cases, we merely need to observe how a shock in the second period, μ_g', affects h. Furthermore, this does not necessarily require a distinct shock in addition to μ_g. In an overlapping generations framework, the same shock, $\mu_g = \mu_g'$ could affect one cohort in the first childhood period (but not the second) and an older cohort in the second period (but not the first). For a small, "double-barreled" shock, we would have reduced form estimates of both the damage in (3) and the potential to alter trajectories in (4).[5] For example, in addition to observing how income during the prenatal period affects newborn health (Kehrer and Wolin, 1979), we might also be able to see how parental income affects the health of preschool age children to gain a sense of what opportunities there are to remediate negative income shocks experienced during pregnancy.

2.3. Responsive investments

Most analyses of "early origins" focus on estimating the reduced form effect, $\delta h/\delta \mu_g$. Whether this empirical relationship represents a purely biological effect or also includes the effect of responsive investments is an open question. In general, to the extent that "early origins" are important, so too will be any response of childhood investments to μ_g. For expositional purposes, we will consider $\mu_g < 0$ and responses that either magnify or attenuate initial damage.

Unless the investment response is costless, damage estimates that monetize $\delta h/\delta \mu_g$ alone will tend to understate total damage. In the extreme, investment responses could fully offset the effect of early-life shocks on h, but this would not mean that such shocks were costless (Deschnes and Greenstone, 2007). More generally, the damage from early-life shocks will be understated if we focus only on long-term effects and there are compensatory investments (i.e. investments that are negatively correlated with the early-life shock ($\delta I_2^*/\delta \mu_g < 0$)). The cost of investments that help remediate damage should be included. But even when the response is reinforcing ($\delta I_2^*/\delta \mu_g > 0$), total costs can still be understated by focussing on the reduced form damage to h alone (see below).

To consider correlated investment responses more formally, we assume parents observe μ_g at the end of the first period. The direction of the investment response—whether reinforcing or compensatory—will be shaped by how substitutable period 2 investments are for those in period 1. If substitutability is high, the optimal response will tend to be compensatory, and thereby help offset damage to h.

A compensatory response is readily seen in the case of perfect substitutability. Cunha and Heckman (2007) observed that economic models commonly assume that production at different stages of childhood are perfect substitutes. When $\phi = 1$, (2) reduces to:

$$h = A\left[\gamma(I_1 + \mu_g) + (1-\gamma)I_2\right]. \tag{5}$$

[5] How parameters of the production function might be recovered is discussed in Appendix D.

This linear production technology is akin to that used in a previous *Handbook* chapter on intergenerational mobility (Solon, 1999), which likewise considered parental investments in children's human capital. Further, Solon (1999) assumed parent's utility trades off own consumption against the child's human capital:

$$U_p = U(C, h), \tag{6}$$

where p denotes parents and C their consumption. The budget constraint is:

$$Y_p = C + I_1 + I_2/(1 + r). \tag{7}$$

With standard preferences, changes to h through μ_g will "unbalance" the marginal utilities in h versus C.[6] If μ_g is negative, the marginal utility of h becomes too high relative to that in consumption. The technology in (5) permits parents to convert some consumption into h at a constant rate. This will cause I_2^* to increase, which attenuates the effect of the μ_g damage. This attenuation comes at the cost of reduced parental utility. Similarly, if μ_g is positive parents will "spend the bounty" (at least in part), reduce I_2^* and increase consumption. Again, this will temper effects on h, leading to an understatement of biological effects in analyses that ignore investments (or parental utility). In either case, perfect substitutability hard-wires the response to be compensatory.

The polar opposite technology is perfect complementarity between childhood stages, i.e., a Leontieff production function. Here, a compensatory strategy would be completely ineffective in mitigating changes to h. As h is determined by the minimum of period 1 and period 2 investments, optimal period 2 investments should reinforce μ_g. If μ_g is negative, parents would seek to reduce I_2 and consume more. Despite higher consumption, parents' utility is reduced on net due to the shock (or this bundle of lower h and higher C would have been selected absent μ_g). Again, the full-cost of a negative μ_g shock is understated when parental utility is ignored.

The crossover between reinforcing and compensating responses of I_2^* will occur at an intermediate parameter value of substitutability. (The fixed investments case of Section 2.2 can be seen to reflect an optimized response at this point of balance between reinforcing and compensating responses). The value of ϕ at this point of balance will depend on the functional form of parental preferences in (6), as shown for CES utility in Appendix B.

To take a familiar example, assume a Cobb-Douglas utility function of the form:

$$U_p = (1 - \alpha)\log C + \alpha \log h. \tag{8}$$

[6] Obviously, the marginal utilities themselves will not be the same but equal subject to discount factor, preference parameters, and prices of C versus I, which have been ignored.

If the production technology is also Cobb-Douglas ($\phi = 0$), then no change to I_2^* is warranted. If instead substitution between period 1 and period 2 is relatively easy ($\phi > 0$), compensating for the shock is optimal. If substitution is relatively difficult ($\phi < 0$), then parents should "go with the flow" and reinforce. For this reason, whether conventional reduced form analyses under or over-state "biological" effects (effects with I_2 held fixed) depends on how easy it is to substitute the timing of investments across childhood. If the elasticity of substitution across periods is low, then it may be optimal for parents to reinforce the effect of a shock.

Tension between preferences and the production technology may also be relevant for within-family investment decisions. For example, Behrman et al. (1982) considered parental preferences that parameterize varying degrees of "inequality aversion" among (multiple) children. Depending on the strength of parents' inequality aversion relative to the production technology (as reflected by ϕ), parents may reinforce or compensate exogenous within-family differences in early-life health and human capital. If substitutability between periods of childhood is sufficiently difficult (low ϕ), reinforcement of sibling differences will be optimal. This reinforcement may be optimal even when the parents place a higher weight in their utility function on the accumulation of human capital by the less able sibling (see Appendix C). Thus, empirical evidence that some parents reinforce early-life shocks could reveal less about "human nature" than it would reveal about the developmental nature of the childhood production technology.

3. METHODS

As discussed above, we confine our discussion to methodological issues that seem particularly germane to the early influences literature. One of these is the question of when sibling fixed effects (or maternal fixed effects) estimation is appropriate. Fixed effects can be a powerful way to eliminate confounding from shared family background characteristics, even when these are not fully observed. This approach is particularly effective when the direction of unobserved sibling-specific confounders can be signed. For example, Currie and Thomas (1995) find that in the cross-section, children who were in Head Start do worse than children who were not. However, compared to their own siblings, Head Start children do better. Since there is little evidence that Head Start children are "favored" by parents or otherwise (on the contrary, in families where one child attends and the other does not, children who attended were more likely to have spent their preschool years in poverty), these contrasting results suggest that unobserved family characteristics are correlated both with Head Start attendance and poor child outcomes. When the effect of such characteristics is accounted for, the positive effects of Head Start are apparent.

However, fixed effects can not control for sibling-specific factors. The theory discussed above suggests that it may be optimal for parents to either reinforce or

compensate for the effects of early shocks by altering their own investment behaviors. Whether parents do or do not reinforce/compensate obviously has implications for the interpretation of models estimated using family fixed effects. If on average, families compensate, then fixed effects estimates will understate the total effect of the shock (when the compensation behavior is unobserved or otherwise not accounted for). In some circumstances, such a bias might be benign in the sense that any significant coefficient could then be interpreted as a lower bound on the total effect. It is likely to be more problematic if parents systematically reinforce shocks, because then any effect that is observed results from a combination of underlying effects and parental reactions rather than the shock itself. In the extreme, if parents seized on a characteristic that was unrelated to ability and systematically favored children who had that characteristic, then researchers might wrongly conclude that the characteristic was in fact linked to success even in the absence of parental responses.

The issue of how parents allocate resources between siblings has received a good deal of attention in economics, starting with Becker and Tomes (1976) and Behrman et al. (1982). Some empirical studies from developing countries find evidence of reinforcing behavior (see Rosenzweig and Paul Schultz (1982), Rosenzweig and Wolpin (1988) and Pitt et al. (1990)). Empirical tests of these theories in developed countries such as the United States and Britain generally use adult outcomes such as completed education as a proxy for parental investments (see for example, Griliches (1979), Behrman et al. (1994) and Ashenfelter and Rouse (1998)).

Several recent studies have used birth weight as a measure of the child's endowment and asked whether explicit measures of parental investments during early childhood are related to birth weight. For example, Datar et al. (2010) use data from the National Longitudinal Survey of Youth-Child and show that low birth weight children are less likely to be breastfed, have fewer well-baby visits, are less likely to be immunized, and are less likely to attend preschool than normal birth weight siblings. However, all of these differences could be due to poorer health among the low birth weight children. For example, if a child is receiving many visits for sick care, they may receive fewer visits for well care and this will not say anything about parental investment behaviors. Hence, Datar et al. (2010) also look at how the presence of low birth weight siblings in the household affects the investments received by normal birth weight children. They find no effect of having a low birth weight sibling on breastfeeding, immunizations, or preschool. The only statistically significant interaction is for well-baby care. This could however, be due to transactions costs. It may be the case that if the low birth weight sibling is getting a lot of medical care, it is less costly to bring the normal birth weight child in for care as well, for example.

Del Bono et al. (2008) estimate a model that allows endowments of other children to affect parental investments in the index child. They find, however, that the results from this dynamic model are remarkably similar to those of mother fixed effects models in most

cases. Moreover, although they find a positive effect of birth weight on breastfeeding, the effect is very small in magnitude.

We conducted our own investigation of this issue using data on twins from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B), using twin differences to control for potential confounders. At the same time, twins routinely have large differences in endowment in the form of birth weight. Table 2 presents estimates for all twins (with and without controls for gender), same sex twins, and identical twins. Overall, there are very few significant differences in the treatment of these twins: Parents seemed to be more concerned about whether the low birth weight twin was ready for school, and to delay introducing solid food (but this is only significant in the identical twin pairs). We see no evidence that parents are more likely to praise, caress, spank or otherwise treat children differently, and despite their worries about school readiness, parents have similar expectations regarding college for both twins. This table largely replicates the basic finding of Royer (2009), who also considered parental investments and birth weight differences in the ECLS-B data. In particular, Royer (2009) focussed on investments soon after birth, finding that breastfeeding, NICU admission, and other measures of neonatal medical care did not vary with within twin pair birth weight differences.

The parental investment response has also been explored in the context of natural experiments. Kelly (2009) asked whether observed parental investments (e.g., time spent reading to child) were related to flu-induced damages to test scores in the 1958 British birth cohort study but did not detect an investment response.

In an interesting contribution to this literature, Hsin (2009) looks at the relationship between children's endowments, measured using birth weight, and maternal time use using data from the Child Supplement of the Panel Study of Income Dynamics. She finds that overall, there is little relationship between low birth weight and maternal time investments. However, she argues that this masks important differences by maternal socioeconomic status. In particular, she finds that in models with maternal fixed effects, less educated women spend less time with their low birth weight children, while more educated women spend more time. This finding is based on only 65 sibling pairs who had differences in the incidence of low birth weight, and so requires some corroboration. Still, one interpretation of this result in the context of the Section 2 framework is that the elasticity of substitution between C and h varies by socioeconomic status. In particular, if $\varphi_{\text{poor}} > \varphi_{\text{rich}}$, low income parents tend to view their consumption and children's h as relatively good substitutes. This would lead low-income parents to be more likely to reinforce a negative shock than high-income parents (assuming that the developmental technology, captured by γ and ϕ, does not vary by socioeconomic status). A second possible interpretation of the finding is that parents' responses may reflect their budget constraint more than their preferences. If parents would like to invest in both children, but have only enough resources to invest adequately in one, then they may be forced

Table 2 Estimated effect of birth weight on parental investments within twin pairs, estimates from the early childhood longitudinal study.

Outcome	All twins	Same sex twins	Identical twins
9 month survey			
1 if child was ever breastfed	0.0183	0.0187	0.0031
	[0.0238]	[0.0277]	[0.0355]
	1550	1000	350
1 if child is now being breastfed	0.0038	−0.0039	−0.0007
	[0.0126]	[0.0152]	[0.001]
	1550	1000	350
How long child was breastfed in months, given breastfed	−0.0753	−0.2165	−0.343
	[0.1752]	[0.204]	[0.3182]
	800	500	150
Age solid food was introduced in months, given introduced	−0.1802	−0.2478	−0.6660⋆
	[0.1523]	[0.1906]	[0.2914]
	1550	1000	350
Number of well-baby visits	0.283	0.3803	0.5797
	[0.1883]	[0.2414]	[0.5253]
	1550	1000	350
Number of well-baby visits only children in excellent or very good health	0.1956	0.2329	0.2668
	[0.1624]	[0.1944]	[0.3799]
	1500	950	300
1 if caregiver praises child	−0.0015	−0.051	0.096
	[0.0941]	[0.1189]	[0.2089]
	1250	800	250
1 if caregiver avoids negative comments	−0.0051	−0.0077	0
	[0.0055]	[0.0084]	[.]
	1250	800	250
1 if somewhat difficult or difficult to raise (caregiver report)	−0.0181	−0.0772	−0.0946
	[0.0583]	[0.0712]	[0.1395]
	1550	1000	350
1 if not at all difficult or not very difficult to raise (caregiver report)	0.1065	0.153	0.2237
	[0.0707]	[0.0812]	[0.1195]
	1550	1000	350
2-year survey			
1 if caress/kiss/hug child	0.0228	0.0055	0.0021
	[0.0266]	[0.0254]	[0.0049]
	1350	850	300

(*continued on next page*)

Table 2 (continued)

Outcome	All twins	Same sex twins	Identical twins
1 if spank/slap child	−0.0195	−0.0095	−0.0048
	[0.0249]	[0.0192]	[0.0316]
	1350	850	300
1 if time spent calming child > 1 hr usually	0.0317	−0.024	0.0719
	[0.0646]	[0.0759]	[0.093]
	1450	950	300
1 if somewhat difficult or difficult to raise (caregiver report)	−0.0432	−0.0901	−0.1412
	[0.0555]	[0.0621]	[0.086]
	1450	950	300
1 if not at all difficult or not very difficult to raise (caregiver report)	−0.0031	0.068	0.0527
	[0.0757]	[0.0869]	[0.1258]
	1450	950	300
Age when stopped feeding formula in months	−0.1903	−0.4504	−0.5903
	[0.255]	[0.3204]	[0.7844]
	1150	750	250
Age when stopped breastfeeding in months	−0.1492	−0.0267	−0.0422
	[0.5981]	[0.044]	[0.069]
	100	50	50
Preschool survey			
1 if parent expects child to enter kindergarten early	−0.0082	−0.0071	0
	[0.012]	[0.0102]	[0]
	1300	800	250
1 if parent concerned about child's kindergarten readiness	−0.1435**	−0.1299*	−0.1099
	[0.0554]	[0.0636]	[0.1253]
	1300	850	250
1 if expect child to get ≥ 4 yrs of college	−0.0073	0.0069	0.0228
	[0.0272]	[0.0327]	[0.0264]
	1350	850	300
Number of servings of milk in the past 7 days	−0.0598	−0.0577	0.0819
	[0.2074]	[0.2278]	[0.2489]
	1350	850	300
Number of servings of vegetables past 7 days	0.0632	0.2131	0.0871
	[0.2634]	[0.3027]	[0.4091]
	1350	850	300

Standard errors clustered on the mother are shown in brackets with sample sizes below. Twin pairs in which a child had a congenital anomaly are omitted. Birth weight measured in kilograms. Each entry is from a separate regression of the dependent variable on birth weight and a mother fixed effect. Models in column 1 also control for child gender. Sample sizes are rounded to the nearest multiple of 50.
Significance levels: $*p < 0.10$, $**p < 0.05$, $***p < 0.001$.

to choose the more well endowed child.[7] Interventions that relaxed resource constraints would have quite different effects in this case than in the case in which parents preferred to maximize the welfare of a favored child. More empirical work on this question seems warranted. For example, the PSID-CDS in 1997 and 2002 has time diary data for several thousand sibling pairs which have not been analyzed for this purpose.

Parent's choices are determined in part by the technologies they face, and these technologies may change over time, with implications for the potential biases in fixed effects estimates.[8] For example, Currie and Hyson (1999) asked whether the long term effects of low birth weight differed by various measures of parental socioeconomic status in the 1958 British birth cohort. They found little evidence that they did (except that low birth weight women from higher SES backgrounds were less likely to suffer from poor health as adults). But it is possible that this is because there were few effective interventions for low birth weight infants in 1958. In contrast, Currie and Moretti (2007) looked at Californian mothers born in the late 1960s and 70s and find that women born in low income zip codes were less educated and more likely to live in a low income zipcode than sisters born in better circumstances. Moreover, women who were low birth weight were more likely to transmit low birth weight to their own children if they were born in low income zip codes, suggesting that early disadvantage compounded the initial effects of low birth weight.

To the extent that behavioral responses to early-life shocks are important empirically, they will affect estimates of long-term effects whether family fixed effects are employed or not. Our conclusion is that users of fixed effects designs should consider any evidence that may be available about individual child-level characteristics and whether parents are reinforcing or compensating for the particular early childhood event at issue. This information will inform the appropriate interpretation of the estimates. There is little evidence at present that parents in developed countries systematically reinforce or compensate for early childhood events, but more research is needed on this question.

3.1. Power

Given that there are relatively few data sets with information about early childhood influences and future outcomes, economists may be tempted to make use of relatively small data sets that happen to have the requisite variables. Power calculations can be helpful in determining *ex ante* whether analysis of a particular data set is likely to yield any interesting findings. Table 3 provides two sample calculations. The first half of the table considers the relationship between birth weight and future educational attainment

[7] In the siblings model of Appendix C, this can be seen in the case of a Leontief production technology, where the second period investments generate increases in h only up to the level of first period investments. If parental income $\bar{Y}$ falls below the cost of maintaining the initial investment level $2\bar{I} + \mu_g$ in the second period, it may be optimal to invest fully in child b (i.e. $I^*_{2b} = \bar{I}$), but not in child a (i.e., $I^*_{2a} < \bar{I} + \mu_g$), who experiences the negative first-period shock.

[8] For example, the effectiveness of remedial investments would change over time if γ varied with the birth cohort. Remediation would be more effective for later cohorts if $\gamma_t > \gamma_{t+1}$ in Eq. (4).

Table 3 Sample power calculations.

Given a true population effect size, what is the power of a size alpha = 0.05 test against the null hypothesis that there is no effect for different sample sizes?

Basis study	Assumptions	Sample size	Power
Black et al. (2007)	True model: Prob(HSGRAD) =	100	0.097
Key result: a 1% increase in birth weight	$0.7 + 0.1*\ln(\text{birthweight}) + \text{error}$	300	0.167
increases the probability of high school	Calculation of error variance and SD:	500	0.263
completion by 0.09 percentage points.	Let y = Prob(HSGRAD), x = ln(birthweight),	600	0.298
Birth weight sample summary stats (twins):	e = error	700	0.351
mean = 2598 g, SD = 612 g.	Var(y) = 0.44^2 = 0.19	800	0.376
Probability of HS grad sample summary stats:	Var(x) = 0.26^2 = 0.07 (where $\text{SD}(x) = 0.26$,	900	0.409
mean = 0.73, SD = 0.44.	according to the distribution of ln(birthweight)).	1000	0.446
	If $y = 0.7 + 0.1x + e$, and x and e are independent,	1250	0.531
	Var(e) = Var(y) $-(0.1^2)*$Var(x)	1500	0.617
	$= 0.19 - (0.1^2)*(0.07) = 0.19$.	1620	0.660
	So, $\text{SD}(e) = \text{sqrt}(\text{Var}(e)) = 0.44$	2000	0.744
	Therefore, assume: birthweight ~N(2598, 612),	2200	0.750
	and take the natural log of birthweight.	2500	0.825
	error $\sim N(0, 0.44)$.	3000	0.892
		3500	0.928
		4000	0.962
		4500	0.975
		5000	0.982
		5500	0.993
		6000	0.994
		6500	0.996
		7000	0.999

Table 3 (continued)

Given a sample size, how large would the true effect size have to be in order to be able to detect it with reliable power using a test of size alpha = 0.05?

Basis study	Assumptions	True B1	Power
Conley et al. (2007)	Model: $y = B0 + B1 * x + \text{error}$	0.005	0.046
Sibling sample from PSID ($n = 1360$)	Assume: $z \sim N(2598, 612)$, $x = \ln(z)$	0.01	0.047
	error $\sim N(0, 0.44)$	0.02	0.077
	sample size = 1500	0.03	0.092
		0.04	0.146
		0.05	0.198
		0.06	0.274
		0.07	0.354
		0.08	0.461
		0.09	0.525
		0.1	0.631
		0.12	0.769
		0.15	0.926
		0.17	0.975
		0.2	0.99

Power calculations are based on Monte Carlo simulations with 1000 replications.

as in Black et al. (2007). Their key result was that a 1% increase in birth weight increased high school completion by 0.09 percentage points. The example shows that under reasonable assumptions about the distribution of birth weight and schooling attainment, it requires a sample of about 4000 children to be able to detect this effect in an OLS regression. We can also turn the question around and ask, given a sample of a certain size, how large would an effect have to be before we could be reasonably certain of finding it in our data? The second half of the table shows that if we were looking for an effect of birth weight on a particular outcome in a sample of 1300 children, the coefficient on (the log of) birth weight would have to be at least 0.15 before we could detect it with reasonable confidence. If we have reason to believe that the effect is smaller, then it is not likely to be useful to estimate the model without more data.

3.2. Data constraints

The lack of large-scale longitudinal data (i.e. data that follows the same persons over time) has been a frequent obstacle to evaluating the long-term impacts of early life influences. Nevertheless, the answer may not always be to undertake collection of new longitudinal data. Drawbacks include the high costs of data collection; the fact that long term outcomes cannot be assessed for some time; and the fact that limiting sample attrition is particularly costly. Unchecked, attrition in longitudinal data can pose challenges for inference.

3.2.1. Leveraging existing datasets

In many cases, existing cross-sectional microdata can serve as a platform for constructing longitudinal datasets. First, it may be possible to add retrospective questions to ongoing data collections. Second, it may be possible to merge new group-level information to existing data sets. Third, it may be possible to merge administrative data sets by individual in order to address previously unanswerable questions. The primary obstacle to implementing each of these data strategies is frequently data security. Depending on the approach adopted, there are different demands on data security, as described below.

Smith (2009) and Garces et al. (2002) are examples of adding retrospective questions to existing data collections. Smith had retrospective questions about health in childhood added to the Panel Study of Income Dynamics (PSID). The PSID began in the 1960s with a representative national sample, and has followed the original respondents and their family members every since. Using these data, Smith (2009) is able to show that adult respondents who were in poor health during childhood have lower earnings than their own siblings who were not in poor health. Such comparisons are possible because the PSID has data on large numbers of sibling pairs. Garces et al. (2002) added retrospective questions about Head Start participation to the PSID, and were able to show that young adults who had attended Head Start had higher educational attainment, and were less likely to have been booked or charged with a crime than siblings who had not attended.

While these approaches may enable analyses of long-term impacts even in the absence of suitable "off the shelf" longitudinal data, they have their drawbacks. First, retrospective data may be reported with error, although it may be possible to assess the extent of reporting error using data from other sources. Second, only outcomes that are already in the data can be assessed, so the need for serendipity remains. Still, the method is promising enough to suggest that on-going, government funded data collections should build-in mechanisms whereby researchers can propose the addition of questions to subsequent waves of the survey.

A second way to address long-term questions is to merge new information at the group level to existing data sets. The merge generally requires the use of geocoded data. For some purposes, such as exploring variations in policies across states, only a state identifier is required. For other purposes, such as examining the effects of traffic patterns on asthma, ideally the researcher would have access to exact latitude and longitude. There are many examples in which this approach has been successfully employed. For example, Ludwig and Miller (2007) study the long term effects of Head Start, which exploits the fact that the Office of Economic Opportunity initially offered the 300 poorest counties in the country assistance in applying for Head Start. They show, using data from the National Educational Longitudinal Surveys, that children who were in counties just poor enough to be eligible for assistance were much more likely to have attended Head Start than children in counties that were just ineligible. They go on to show that child mortality rates in the relevant age ranges were lower in counties whose Head Start enrollments were higher due to the OEO assistance. Using Census data they find that education is higher for people living in areas with higher former Head Start enrollment rates. Unfortunately, however, neither the decennial Census nor the American Community Survey collect county of birth, so they cannot identify people who were born in these counties (substantial measurement error is obviously introduced by using county of residence or county where someone went to school as a proxy for county of birth). An exciting crop of new research would be enabled by the addition of Census survey questions on county of birth, as well as county of residence at key developmental ages (e.g., ages 5 and 14).[9]

In addition to the observational approaches described above, an intriguing possibility is that participants in a completed randomized trial could be followed up. For instance, Rush et al. (1980) conducted a randomized intervention of a prenatal nutrition program in Harlem during the early 1970s. Following these children over time would

[9] In another example, Currie and Gruber (1996a) were able to examine the effects of the Medicaid expansions on the utilization of care among children by merging state-level information on Medicaid policy to data from the National Health Interview Survey (NHIS). At the time, this was only possible because one of the authors had access to the NHIS state codes through his work at the Treasury Department. It has since become easier to access geocoded health data either by traveling to Washington to work with the data, or by using it in one of the secure data centers that Census and the National Center for Health Statistics (NCHS) support. However, it remains a source of frustration to health researchers that NCHS does not make state codes and/or codes for large counties available on its public use data sets.

allow researchers to evaluate cognitive outcomes in secondary school, and it might be possible to collect retrospective data on parental investments during childhood, and to evaluate whether parental investments were affected by the randomization.

A third approach to leveraging existing data merges administrative records from multiple sources at the individual level, which obviously requires personal identifiers such as names and birth dates or social security numbers. Access to such identifiers is especially sensitive. Nevertheless, it constitutes a powerful way to address many questions of interest. Several important studies have successfully exploited this approach outside of the US. For example, Black et al. (2005) and Black et al. (2007) use Norwegian data on all twins born over 30 years to look at long-term effects of birth weight, birth order, and family size on educational attainment. Currie et al. (2010) use Canadian data on siblings to examine the effects of health shocks in childhood on future educational attainment and welfare use. Almond et al. (2009) use Swedish data to look at the long term effects of low-level radiation exposure from the Chernobyl disaster on children's educational attainment.

In the US, Doyle (2008) uses administrative data from child protective services and the criminal justice system in Illinois to examine the effects of foster care. He shows first that there is considerable variation between foster care case workers in whether or not a child will be sent to foster care. Moreover, whether a child is assigned to a particular worker is random, depending on who is on duty at the time a call is received. Using this variation, Doyle shows that the marginal child assigned to foster care is significantly more likely to be incarcerated in future. These examples exploit large sample sizes, objective indicators of outcomes, sibling or cohort comparisons, as well as a long follow up period. Some limitations of using existing data include the fact that administrative data sets often contain relatively little background information, and that outcomes are limited to those that are collected in the data bases. Finally, the application process to obtain individual-matched data is often protracted.

Looking forward, the major challenge to research that involves either merging new information to existing data sets, or merging administrative data sets to each other, is that privacy concerns are making it increasingly difficult to obtain data just as it is becoming more feasible to link them. In some cases, access to public use data has deteriorated. For example, for many years, individual level Vital Statistics Natality data from birth certificates included the state of birth, and the county (for counties with over 100,000 population). Since 2005, however, these data elements have been suppressed and it is now necessary to get special permission to obtain US Vital Statistics data with geocodes.

3.2.2. Improvements in the production of administrative data

There are several "first best" potential solutions to these problems. First, creators of large data sets need to be sensitive to the fact that their data may well be useful for addressing questions that they have not envisaged. In order to preserve the ability to use data to answer future questions, it is essential to retain information that can be used for

linkage. At a minimum, this should include geographic identifiers at the smallest level of disaggregation that is feasible (for example a Census tract). Ideally, personal identifiers would also be preserved.

Second, more effort needs to be expended in order to make sensitive data available to researchers. A range of mechanisms exist that protect privacy while enabling research:

1. Suppress small cells or merge small cells in public use data files. For example, NCHS data sets such as NHIS could be released with state identifiers for large states, and with identifiers for groups of smaller states.
2. Add small amounts of "noise" to public use data sets, or do data swapping in order to prevent identification of outliers. For example, Cornell University is coordinating the NSF-Census Bureau Synthetic Data Project which seeks to develop public-use "analytically valid synthetic data" from micro datasets customarily accessed at secure Census Research Data Centers.
3. Create model servers. In this approach, users login to estimate models using the true data, but get back output that does not allow individuals to be identified.
4. Data use agreements. The National Longitudinal Survey of Youth and the National Educational Longitudinal Survey have successfully employed data use agreements with qualified users for many years, and without any documented instances of data disclosure.
5. Creation of de-identified merged files. For example, Currie et al. (2009) asked the state of New Jersey to merge birth records with information about the location of pollution sources, and create a de-identified file. This allows them to study the effect of air pollution on infant health.
6. Secure data facilities. The Census Research Data Centers have facilitated access to much confidential data, although researchers who are not located close to the facilities may still face large costs of accessing them.

These approaches to data dissemination have been explored in the statistics literature for more than 20 years (see Dalenius and Reiss (1982)), and have been much discussed at Census (see for example, Reznek (2007)).

3.2.3. Additional issues

We conclude with two new and relatively unexplored data issues. First, how can economists make effective use of the burgeoning literature on biomarkers? These measures have recently been added to existing health surveys, such as the National Longitudinal Study of Adolescent Health data. Biomarkers include not only information about genetic variations but also hormones such as cortisol (which is often interpreted as a measure of stress). It is tempting to think of these markers as potential instrumental variables (Fletcher and Lehrer, 2009). For instance, if it was known that a particular gene was linked to alcoholism, then one might think of using the gene as an instrument for

alcoholism. The potential pitfall in this approach is clear if we consider using something like skin color as an instrument in a human capital earnings function–clearly, skin color may predict educational attainment, but it may also have a direct effect on earnings. Just because a variable is "biological" does not mean that it satisfies the criteria for a valid instrument.

A second issue is the evolving nature of what constitutes a "birth cohort." Improvements in neonatal medicine have meant that stillbirths and fetal deaths that would previously have been excluded from the Census of live births may be increasingly important, e.g. MacDorman et al. (2005). Such a compositional effect on live births may have first-order implications for program evaluation and the long-term effects literature. Indeed, both the right and left tails of the birth weight distribution have elongated over time—in 1970 there were many fewer live births with birthweight either less than 1500 g or over 4000 g. To date there has been little research exploring the implications of these compositional changes.

In summary, there are many secrets currently locked in existing data that researchers do not have access to. Economists have been skillful in navigating the many data challenges inherent in the analysis of long-term (and sometimes latent) effects. Nevertheless, we need to explore ways to make more of these data available, and to more researchers. In many cases, this will be a more cost effective and timely way to answer important questions than carrying out new data collections.

4. EMPIRICAL LITERATURE: EVIDENCE OF LONG TERM CONSEQUENCES

> *What is of importance is the year of birth of the generation or group of individuals under consideration. Each generation* after the age of 5 years *seems to carry along with it the same relative mortality throughout adult life, and even into extreme old age.*
>
> ***Kermack et al. (1934) in The Lancet (emphasis added).***

In this section, we summarize recent empirical research findings that experiences before five have persistent effects, shaping human capital in particular. A hallmark of this work is the attention paid to identification strategies that seek to isolate causal effects of the early childhood environment. An intriguing sub-current is the possibility that some of these effects may remain latent during childhood (at least from the researcher's perspective) until manifested in either adolescence or adulthood. Recently, economists have begun to ask how parents or other investors in human capital (e.g. school districts) *respond* to early-life shocks, as suggested by the conceptual framework in Section 2.3.

As the excerpt from Kermack et al. (1934) indicates, the idea that early childhood experiences may have important, persistent effects did not originate recently, nor did it first appear in economics. An extensive epidemiological literature has focussed on the early childhood environment, nutrition in particular, and its relationship to health

outcomes in adulthood. For a recent survey, see Gluckman and Hanson (2006). This literature has been criticized within epidemiology for credulous empirical comparisons (see, e.g. Rasmussen (2001) or editorial in *The Lancet* [2001]). Absent clearly-articulated identification strategies, health determinants that are difficult to observe and are therefore omitted from the analysis (e.g., parental concern) are presumably correlated with the treatment and can thereby generate the semblance of "fetal origins" linkages, even when such effects do not exist.

4.1. Prenatal environment

In the 1990s, David J. Barker popularized and developed the argument that disruptions to the prenatal environment presage chronic health conditions in adulthood, including heart disease and diabetes (Barker, 1992). Growth is most rapid prenatally and in early childhood. When growth is rapid, disruptions to development caused by the adverse environmental conditions may exert life-long health effects. Barker's "fetal origins" perspective contrasted with the view that pregnant mothers functioned as an effective buffer for the fetus against environmental insults.[10]

In Table 4, we categorize prenatal environmental exposures into three groups. Specifically, we differentiate among factors affecting maternal and thereby fetal health (e.g. nutrition and infection), economic shocks (e.g. recessions), and pollution (e.g. ambient lead).

4.1.1. Maternal health

Currie and Hyson (1999) broke ground in economics by exploring whether "fetal origins" (FO) effects were confined to chronic health conditions in adulthood, or might extend to human capital measures. Using the British National Child Development Survey, low birth weight children were more than 25% less likely to pass English and math O-level tests, and were also less likely to be employed. The finding that test scores were substantially affected was surprising, as epidemiologists routinely posited fetal "brain sparing" mechanisms, whereby adverse *in utero* conditions were parried through a placental triage that prioritized neural development over the body, see, e.g., Scherjon et al. (1996). Furthermore, Stein et al. (1975)'s influential study found no effect of prenatal exposure to the Dutch Hunger Winter on IQ.

Currie and Hyson (1999) were followed by a series of papers that exploited differences in birthweight among siblings and explored their relationship to sibling differences in completed schooling. In relatively small samples (approximately 800 families), Conley and Bennett (2001) found negative but imprecise effects of low birth weight on educational attainment. Statistically significant effects of low birth weight on educational attainment were found when birth weight was interacted with being poor, but in general

[10] For example, it has been argued that nausea and vomiting in early pregnancy (morning sickness) is an adaptive response to prevent maternal ingestion of foods that might be noxious to the fetus.

Table 4 Prenatal effects on later child and adult outcomes.

Study and data	Study design	Results
Effects of maternal health		
Is the impact of health shocks cushioned by economic status? The case of low birth weight. Currie and Hyson (1999) NCDS 1958 cohort. $N = 11{,}609$ at age 20, 10,267 at age 23, 9402 at age 33.	Multivariate regression with numerous background and demographic controls (including maternal grandfather's SES, birth order, and maternal smoking during pregnancy). Key explanatory variables are indicators for LBW, SES, and the interactions. Outcomes measured are the number of O-level passes at age 16 (transcripts collected at age 20), employment, wages and health status at ages 23 and 33. SES assigned using father's social class in 1958 (or mother's SES if father is missing).	LBW children are 38–44% less likely to pass Math O-level. LBW females are 25% less likely to pass English O-level tests. LBW females are 16% less likely to be employed full-time at age 23, LBW males are 9% less likely to be employed full-time at age 33. LBW females are 54% more likely to have fair/poor health at age 23, LBW males are 43% more likely to have fair/poor health at age 33. Few significant differences by SES.
Returns to birthweight (Behrman and Rosenzweig, 2004). Monozygotic female twins born 1936–1955 from the Minnesota Twins Registry for 1994 and the birth weights of their children. $N = 804$ twins, and 608 twin-mother pairs.	Twin fixed effects estimates compared to OLS. To explore generalizability of findings from twins sample, weighted the sample using the US singleton distribution of fetal growth rates.	**Results from twins sample:** 1 oz. per week of pregnancy increase in fetal growth leads to increases of 5%, in schooling attainment; 2% in height; 8% hourly wages. **Results from twins sample weighted using singleton distribution:** 1 oz. per week of pregnancy increase in fetal growth leads to increase of 5% in schooling attainment; 1.7% in height; no significant effect on wages. OLS underestimates effects of birth weight by 50%.

Table 4 (continued)

Study and data	Study design	Results
The costs of low birth weight (Almond et al., 2005). Linked birth and infant death files for US for 1983-85, 1989-91 and 1995-97. Hospital costs from healthcare cost and utilization project state inpatient database for 1995-2000 in New York and New Jersey. NCHS $N = 189{,}036$ twins, 497,139 singletons. HCUP $N = 44{,}410$.	Twin fixed effects to estimate effect of low birth weight (LBW) on hospital costs, health at birth, and infant mortality. Also estimated impact of maternal smoking during pregnancy on health among singleton births, controlling for numerous background and demographic characteristics using OLS and propensity score matching.	**Results using twin fixed effects:** 1 SD increase in birth weight leads to: 0.08 SD decrease in hospital costs for delivery and initial care 0.03 SD decrease in infant mortality rates 0.03 SD increase in Apgar scores 0.01 SD decrease in use of assisted ventilator after birth (OLS estimates w/out twin fixed effects are 0.51 SD, 0.41 SD, 0.51 SD, 0.25 SD). **Results from OLS on effects of maternal smoking:** Maternal smoking reduces birth weight by 200 g (6%); increases likelihood that infant is LBW (<2500 g) by more than 100% (mean = 0.061). No statistically significant effects on Apgar score, infant mortality rates, or use of assisted ventilator at birth.
The 1918 influenza pandemic and subsequent health outcomes (Almond and Mazumder, 2005). Data from SIPP for 1984-1996. $N = 25{,}169$.	Compare cohorts *in utero* before, during and after Oct. 1918 flu pandemic. Regressions estimate cohort effects including survey year dummies and quadratic in age interacted with survey year.	Individuals born in 1919 are 10% more likely to be in fair or poor health, also increases of 19%, 35%, 13%, 17% in trouble hearing, speaking, lifting and walking.
Estimating the impact of large cigarette tax hikes: the case of maternal smoking and infant birth weight (Lien and Evans, 2005). 1990-1997 US Natality files. Data on cigarette taxes from the tax burden on tobacco, various years.	IV using tax hikes in four states as instruments for maternal smoking during pregnancy. Controlled for state and month of conception and background characteristics.	Maternal smoking during pregnancy reduces birth weight by 5.4% and increases likelihood of low birth weight by ∼100%.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Long term effects of *in utero* exposure to the 1918 influenza pandemic in the post-1940 US population (Almond, 2006). 1960-1980 US Census data. 1917-1919 Vital Statistics data on mortality. For 1960 $n = 114{,}031$, for 1970 $n = 308{,}785$, for 1980 $n = 471{,}803$.	Estimated effects of *in utero* influenza exposure by comparing cohorts born immediately before, during, and after the 1918 pandemic and by employing the idiosyncratic geographic variation in intensity of exposure to conduct within-cohort analysis. Exposed cohort = those born in 1919. Surrounding cohorts = those born in 1918 and 1920. Used multivariate regression with dummy for birth cohort = 1919, and a quadratic cohort trend to measure departures in the 1919 birth cohort outcomes from the trend. For geographic comparison, used data on virus strength by week as well as data on epidemic timing by Census division to yield a measure of average pandemic virulence by division. Then estimated multivariate regression including the virulence measure, state and year of birth fixed effects, the infant mortality rate in state and year of birth, and the attrition of birth cohort in the Census data.	**Estimation results comparing birth cohorts:** The 1919 birth cohort, compared with the cohort trend: was 4%-5% (13%-15% among treated) less likely to complete high school received 0.6%-1.6% fewer years of education had 1%-3% less total income (for males only, 2005 dollars) had 1%-2% lower socioeconomic status index (Duncan index) was 1%-2% more likely to have a disability that limits work (for males only) had 12% higher average welfare payment (for women). Estimates are slightly larger for nonwhite subgroup. **Estimation results using geographic variation and state fixed effects:** (For the 1919 birth cohort, used average maternal infection rate = 1/3) Maternal infection: reduces schooling by 2.2% reduces probability of high school graduation by 0.05% decreases annual income by 6% reduces socioeconomic status index by 2%-3%.

Table 4 (continued)

Study and data	Study design	Results
Explaining sibling differences in achievement and behavioral outcomes: the importance of within- and between-family factors (Conley et al., 2007) Data from child development supplement of PSID. $N = 1360$.	Sibling fixed effects to examine effects of birth weight, birth order, and gender on later outcomes. Cognitive outcomes measured by the Woodcock- Johnson revised tests of achievement. Behavioral outcomes measured by the behavioral problems index (BPI). Control for family- and child-specific characteristics.	No statistically significant effects of birth weight on BPI or on cognitive assessments for whole sample. For blacks, positive effect of birth weight on cognitive assessments.
Twin differences in birth weight: the effects of genotype and prenatal environment on neonatal and postnatal mortality (Conley et al., 2006) Data on twin births from the 1995-97 matched multiple birth database. $N = 258{,}823$.	Twin fixed effects models of effects of birth weight on mortality for same-sex and mixed-sex pairs.	1 lb increase in birth weight leads to: 9% (10%) reduction in infant mortality for mixed-sex (same-sex) 7% (8%) reduction in neonatal mortality for mixed-sex (same-sex) 2% reduction in post-neonatal mortality for mixed-sex and same-sex. For full-term twins, mixed-sex effects much larger than same-sex effects.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Biology as destiny? Short- and long-run determinants of intergenerational transmission of birth weight (Currie and Moretti, 2007). California natality data for children born between 1989 and 2001 and their mothers (if born in CA) born between 1970 and 1974. Mothers who are sisters are matched using grandmother's name. $n = 638{,}497$ births.	Examine effect of mother's birth weight on child's birth weight in models with grandmother fixed effects. Examine interactions of mother's birth weight and grandmother's SES at time of mother's birth as proxied by income in zip code of mother's birth. Examine effect of maternal low birth weight on mother's SES at time she gives birth.	Mother's low birth weight increases likelihood that child is low birth weight by about 50%. The incidence of child low birth weight is 7% higher if mother was born into high poverty zip code than into low poverty zip code. Children born into poor households are 0.7% more likely to be low birth weight if their mothers were low birth weight; children born into non-poor households are 0.4% more likely to be low birth weight if their mothers were low birth weight (so, poverty raises the probability of transmission of low birth weight by 88%). Being low birth weight is associated with a loss of $110 in future income, on average, on a baseline income of $10,096 (in 1970 dollars). Being low birth weight increases the probability of living in a high-poverty neighborhood by 3% relative to the baseline. Being low birth weight reduces future educational attainment by 0.1 years.

Table 4 (continued)

Study and data	Study design	Results
From the cradle to the labor market: the effect of birth weight on adult outcomes (Black et al., 2007). Birth records from Norway for 1967-1997. Dropped congenital defects. Matched to registry data on education and labor market outcomes and to military records. $n = 33{,}366$ twin pairs.	Twin fixed effects, controlling for mother and birth-specific variables. Log(birth weight) is primary independent variable.	10% increase in birth weight: reduces 1-year mortality by 13% increases 5 min APGAR score by 0.3% increases probability of high school completion by 1.2% increases full-time earnings by 1%. **Male outcomes at age 18-20:** 10% increase in birth weight: increases height by 0.3% increases BMI by 0.5% increases IQ by 1.1% (scale of 1-9). **Effects of mother's birth weight on child's birth weight:** 10% increase in mother's birth weight: increases child's birth weight by 1.5%.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
The influence of early-life events on human capital, health status and labor market outcomes over the life course (Johnson and Schoeni, 2007) PSID. Adult sample born between 1951 and 1975. $N = 5160$ people and 1655 families. Child sample 0-12 years old in 1997. $N = 1127$ mothers and 2239 children.	Sibling fixed effects to estimate impact of maternal smoking during pregnancy on birth outcomes. Also estimated model using OLS, controlling for grandparent smoking in adolescence, maternal birth weight, and paternal smoking, among other background characteristics.	Note: means are not reported, so relative effects cannot be calculated. **Effects of income and mother's birth weight on child's birth weight:** An increase in income of $10,000 raises child's birth weight by 0.12 lbs if mother was low birth weight (by 0.02 lbs if mother was not low birth weight) for a family with income of $7500 (1997 dollars). Having no health insurance increases the probability of low birth weight by 10 pp. **Effects on child health outcomes:** (health index 1-100) Low birth weight siblings have a 1.67 point lower health index. Private health insurance increases health index by 1.02 points. A $10,000 increase in income for families with $15,000-50,000 income increases health index by 0.53 pp. **Effects on adult outcomes:** Low birth weight siblings: 4.7 pp more likely to be high school dropouts, have a 3.7 point lower health index (1-100 scale). Low birth weight brothers (sample of males only): are 4.3 pp more likely to have no positive earnings (sig at 10% level) have $2966 less annual earnings (sig at 10% level).

Table 4 (continued)

Study and data	Study design	Results
Maternal smoking during pregnancy and early child outcomes (Tominey, 2007) Children of NCDS mothers born between 1973 and 2000. $n = 2799$ mothers and 6291 sibling children.	Sibling fixed effects to estimate impact of maternal smoking during pregnancy on birth outcomes. Also estimated model using OLS, controlling for grandparent smoking in adolescence, maternal birth weight, and paternal smoking, among other background characteristics.	Note: only reporting results from sibling fixed effects regression here. Maternal smoking during pregnancy reduces birth weight by 1.7%. No statistically significant effect of maternal smoking during pregnancy on probability of having a low birth weight child, pre-term gestation, or weeks of gestation. No statistically significant effects of maternal smoking among mothers who quit by month 5 of pregnancy. Larger effects of maternal smoking on birth weight among low educated women.
Can a pint per day affect your child's pay? The effect of prenatal alcohol exposure on adult outcomes (Nilsson, 2008). Data from Swedish LOUISE database on first-born individuals born between 1964 and 1972. $n = 353{,}742$.	Swedish natural experiment in which alcohol availability in 2 treatment regions increased sharply as regular grocery stores were allowed to market strong beer for 6 months during 1967 with the minimum age for purchase being 16 (instead of 21). Difference-in-difference-in-difference comparing under-21 mothers with older mothers in treatment and control regions pre-, during, and post-experiment. Baseline estimations focused on children conceived prior to the experiment, but exposed *in utero* to the experiment. Controlled for quarter and county of birth fixed effects.	**DDD results:** Years of schooling: decreased by 0.27 (2.1%) years for whole sample; by 0.47 years for males; no statistically significant effect for females. HS graduation: decreased by 0.4% for whole sample; by 10% for males; no stat. significant effect for females. Graduation from higher education: decreased by 16% for whole sample; by 35% for males; no stat. significant effect for females. Earnings at age 32: decreased by 24.1% for whole sample; by 22.8% for males; by 17.7% for females. Probability no income at age 32: increased by 74% whole sample. Proportion on welfare: increased by 90% for whole sample; by 5.1pp for males.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Short-, medium- and long-term consequences of poor infant health (Oreopoulos et al., 2008). Data from the Manitoba Center for Health Policy matching provincial health insurance claims with birth records, educational records, and social assistance records. Includes all children born in Manitoba from 1979 to 1985. $n = 54{,}123$ siblings and 1742 twins.	Sibling and twin fixed effects. Used three measures of infant health: birth weight, 5-min Apgar score, and gestational length in weeks; used dummies for different categories of the variables to estimate nonlinear effects. Also estimated models using OLS for the whole sample, for the siblings sample, and for the twins sample without family fixed effects.	Note: only reporting results from regressions that included family fixed effects here. BW = birth weight. Relative to Apgar score = 10, $BW > 3500$ g, gestation 40-41 weeks lower values. **Effects on infant mortality:** increase infant morality. E.g. in sibling sample (infant mortality = 0.011). Apgar score <6 increases probability of infant mortality by 31.9 pp $BW < 1000$ g increases probability of infant mortality by 87.2 pp Gestation < 36 weeks increases probability of infant mortality by 11.9 pp. **Effect on language arts score (taken in grade 12):** Apgar score < 6 decreases test score by 0.1 of SD (sig. at 10% level) BW 2501-3000 g decreases test score by 0.04 of SD **Effect on probability of reaching grade 12 by age 17:** Apgar score < 6 decreases grade 12 probability by 4.1 pp (sig. at 10%) BW 1001-1500 g decreases grade 12 probability by 14.1 pp Gestation < 36 weeks decreases grade 12 probability by 4.0 pp. **Effect on social assistance take-up during ages 18-21.25:** $BW < 1000$ decreases probability of take-up by 21.5 pp.

Table 4 (continued)

Study and data	Study design	Results
Birth cohort and the black-white achievement gap: the role of health soon after birth (Chay et al., 2009). Data from the national assessment of educational progress long-term trends for 1971-2004. AFQT data from US military for 1976-2001 for male applicants 17-20. $n = 2{,}649{,}573$ white males and $n = 1{,}103{,}748$ black males. Hospital discharge rates from National Health Interview Survey.	Regression of test scores on year, age, and subject fixed effects (separately for blacks and whites) using NAEP-LTT and AFQT test score data. In AFQT test score data, correct for selection bias using inverse probability weighting from Natality and Census data. Estimated regressions separately for blacks and whites and for the North, the South, the Rustbelt, the deep South, and individual states within the South and North. Also estimated difference-in-difference-in-difference (DDD) models, comparing black and white test scores between cohorts born in 1960-62 and 1970-72 in the South relative to the Rustbelt, controlling for region-specific, race-specific age-by-time effects and race-by-region-by-time effects. Used post neonatal mortality rate (PNMR) as a proxy for infant health environment to assess impact of infant health on the test score gap.	**NAEP-LTT data:** The black-white test score gap declined from 1 SD to 0.6 of SD between 1971 and 2004. The convergence was primarily due to large increases in black test scores in the 1980s. Regression results indicate that the convergence was due to cohort effects, rather than time effects. **AFQT data:** The black-white gap in AFQT test scores declined by about 19% between the 1962-63 and 1972 birth cohorts. The decline in the gap is about 0.3 SDs greater in the South relative to the Rustbelt between 1960-62 and 1970-72 cohorts. Convergence in PNMR explains 52% of the variation across states in AFQT convergence. **Effects of access to hospitals:** A 30pp increase in black hospital birth rates from 1962-64 to 1968-70 increases cohort AFQT scores by 7.5 percentile points. A black child who gained admission to a hospital before age 4 had a 0.7-1 SD gain in AFQT score at age 17-18 relative to a black child who did not.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Separated at birth: US twin estimates of the effects of birth weight (Royer, 2009) California birth records. $n = 3028$ same-sex female twin births 1960-1982. Data from early childhood longitudinal study-birth cohort. $n = 1496$ twin births.	Twin fixed effects, allowing non-linear effects of birth weight.	**Twin fixed-effect results:** 250 g increase in birth weight leads to: 8.3% decrease in probability of infant mortality 0.82 day decrease in stay in hospital post-birth 0.02-0.04 of SD increase in mental/motor test score 0.2% increase in educ. attainment of mother at childbirth 0.5% increase in child's birth weight 11% decrease in pregnancy complications. No statistically significant effects of birth weight on adult health outcomes such as hypertension, anemia, and diabetes. No statistically significant effects of birth weight on income-related measures. 200 g increase in birth weight leads to a 0.5 pp increase in probability of being observed in adulthood—small selection bias. Found no statistically significant evidence of compensating or reinforcing parental investments when considering early medical care and breastfeeding.

Table 4 (continued)

Study and data	Study design	Results
The long-term economic impact of *in utero* and postnatal exposure to malaria (Barreca, 2010). Malaria mortality from US mortality statistics 1900-1936. $n = 1147$ state-year observations. Climate data $n = 1813$ obs. Adult outcomes from 1960 Census. All data merged at state/year of birth level.	IV using fraction of days in "malaria-ideal" temperature range as IV for malaria deaths in state and year.	Note: Only results from the IV regression are reported here. Exposure to 10 additional malaria deaths per 100,000 inhabitants causes 3.4% less years of schooling Exposure to malaria can account for approximately 25% of the difference in years of schooling between cohorts born in high and low malaria states.
Long-run longevity effects of a nutrition shock in early life: The Dutch Famine of 1846-47 (Van Den Berg et al., in press). Historical sample of the Netherlands. Exposed cohort born 9/1/1846-6/1/1848. Non-exposed born 9/1/1848-9/1/1855 and 9/1/1837-9/1/1944.	Key independent variable is exposure to famine at birth. Instrument access to food with variations in yearly average real market prices of rye and potatoes for three different regions. Controlled for macroeconomic conditions, infant mortality rates, and individual demographic and socio-economic characteristics.	**Results from nonparametric regression:** Residual life expectancy at age 50 is 3.1 years shorter for exposed men than for men born after the famine; 1.4 years shorter than for men born before famine. Kolmogorov-Smirnov test suggests that the survival curves after age 50 differ significantly for exposed and control men. Max difference in distribution = 0.15 at age 56. **Results from parametric survival models:** Exposure to famine at birth for men reduces residual life expectancy at age 50 by 4.2 years. No statistically significant results for women.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Maternal stress and child well-being: evidence from siblings (Aizer et al., 2009). National Collaborative Perinatal Project. Births 1959-65 in Providence and Boston. $N = 1103$ births to 915 women. 163 siblings in sample.	Sibling fixed effects. Estimated effect of deviations from mean cortisol levels (cortisol measures maternal stress) during pregnancy on outcomes.	No significant effects on birth weight or maternal postnatal investments. Exposure to top quartile of cortisol level distribution (relative to the middle) leads to 47% of SD decrease in verbal IQ at age 7 (sign. at 10% level). 1 SD deviation in cortisol levels leads to 26% of SD decrease in educational attainment. Exposure to top quartile of cortisol level distribution (relative to the middle) leads to 51% of SD decrease in educational attainment.
The scourge of Asian flu: the physical and cognitive development of a cohort of British children *in utero* during the Asian Influenza Pandemic of 1957, from birth until age 11 (Kelly, 2009). NCDS 1958 cohort. $N = 16{,}765$ at birth, 14,358 at 7, 14,069 at 11.	Effect of flu on each cohort member identified using variation in incidence of epidemic by local authority of birth. Epidemic peaked when cohort members were 17-23 weeks gestation. 1/3 of women of child-bearing age were infected, hence true treatment effect can be estimated by multiplying results by 3.	No statistically significant evidence of reinforcing or compensating parental investments. 1 SD increase in epidemic intensity decreases birth weight by 0.03-0.363 SD; decreases test scores by 0.067 of SD at age 7, by 0.043 of SD at age 11; increases detrimental effect of mother preeclampsia from -0.075 to -0.11 SD on birth weight; increases detrimental effect of mother smoking by 0.03 of SD on birth weight; increases detrimental effect of mother under 8 stone weight from -0.54 to -0.61 of SD on birth weight

Table 4 (continued)

Study and data	Study design	Results
Do lower birth weight babies have lower grades? Twin fixed effect and instrumental variables evidence from Taiwan (Lin and Liu, 2009). Birth Certificate data for Taiwan. High school entrance exam results from Committee of Basic Competence Test. $N = 118{,}658$. Twin sample $n = 7772$.	Twin fixed effects and IV using the public health budget and number of doctors in county where child was born as instruments for child's birth weight.	**Twin FE/IV for mothers w/ <9 yrs ed** & **<25 yrs old** Increase in birth weight of 100 g leads to: Increase Chinese score: 0.5%/8.2% Increase English score: 0.3%/8.1%⋆ Increase Math score: 0.7%/12.6% Increase Natural Science score: 0.8%/11.8% Increase Social Science score: 0.4%/4.0%⋆ ⋆ means significant at 10% level IV results for other subgroups show smaller/no effects.
Poor, hungry and stupid: numeracy and the impact of high food prices in industrializing Britain, 1750-1850 (Baten et al., 2007). Data on age heaping from 1951 to 1881 British Census. Information on poor relief from Boyer (1990).	Age heaping is rounding age to nearest 5 or 10. Whipple index (WI) = number of ages that are multiples relative to expected number given uniform age distribution. Regressions use WI as dependent variable. Wheat prices and poor relief measures are independent variables.	During the Revolutionary and Napoleonic Wars, the price of wheat almost doubled—and the number of erroneously reported ages at multiples of 5 doubled from 4% to 8%. Men and women born in decades with higher WI sorted into jobs that had lower intelligence requirements. 1 SD increase in the Whipple index associated with a 2.8% (relative to median earnings) decrease in earnings.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Impacts of economic shocks		
Birth weight and income: interactions across generations (Conley and Bennett, 2001) Data from PSID. Two samples: (1) children born 1986-1992, $n = 1654$; (2) reached 19 by 1992, $n = 1388$ individuals in 766 families.	Sibling fixed effects.	Low birth weight child who spent 6 years at poverty line is less likely to graduate high school than normal birth weight child who spent 6 years at poverty line. Low birth weight child who spent 6 years with income 5 times the poverty line is as likely to graduate high school as normal birth weight child of same income (significant at 10% level).
Economic conditions early in life and individual mortality (Van Den Berg et al., in press). Data from historical sample of the Netherlands on individuals born 1812-1903 with date of death observed by 2000. $n = 9276$. Merged with historical macro time series data.	Compared individuals born in booms with those born in the subsequent recessions. Controlled for wars and epidemics.	Comparing those born in boom of 1872-1876 with those born in recession of 1877-1881: Boom $T\|T > 2 = 66.0$ years, Recession $T\|T > 2 = 62.5$ years Boom $T\|T > 5 = 70.8$ years, Recession $T\|T > 5 = 67.5$ years **Regression results:** Boom $T\|T > 2$ is 1.58 years greater than Recession $T\|T > 2$ *Notation*: $T\|T > 2 =$ average lifetime given survival past age 2 $T\|T > 5 =$ average lifetime given survival past age 5

Table 4 (continued)

Study and data	Study design	Results
Evidence on early-life income and late-life health from America's Dustbowl era (Cutler et al., 2007). The Health and Retirement Study. $N = 8739$ people born between 1929 and 1941. Agricultural data from National Agricultural Statistics Service. Income data from Bureau of Economic Analysis.	Measured economic conditions *in utero* using income and yield from the same calendar year for those born in 3rd or 4th quarter of the year and from the previous calendar year for those born in 1st or 2nd quarter. Regressions include the *in utero* economic condition measure (log income, log yield), dummy for whether respondent's father was a farmer and interaction of the dummy with the economic conditions. Controlled for region and year of birth fixed effects, region-specific linear time trends, and region-year infant death and birth rates, and other demographic characteristics.	No statistically significant relationship between poor early-life economic conditions during the Dustbowl and late-life health outcomes such as heart conditions, stroke, diabetes, hypertension, arthritis, psychiatric conditions, etc.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Long-run health impacts of income shocks: wine and Phylloxera in 19th century France (Banerjee et al., forthcoming). See paper references for source of data on wine production, number of births, and infant mortality. Department level data on heights from military records 1872-1912. $n = 3485$ year-departments.	Regional variation in a large negative income shock caused by Phylloxera attacks on French vineyards between 1863 and 1890. Shock dummy equal to 0 pre Phylloxera and after 1890 (when grafting solution found). Dummy equal 1 when wine production < 80% of pre-level. Difference-in-difference comparing children born in affected and unaffected areas before and after.	Decrease in wine production was not compensated by an increase in other agricultural production (e.g. wheat), suggesting that this was truly a large negative income shock. **Main outcomes:** 3-5% decline in height at age 20 for those born in wine-growing families during the year that their region was affected by Phylloxera. Those born in Phylloxera-affected year are 0.35-0.38 pp more likely to be shorter than 1.56 cm. No statistically significant effects on other measures of health or life expectancy.

Table 4 (continued)

Study and data	Study design	Results
Impacts of environmental shocks		
Lifespan depends on the month of birth (Doblhammer and Vaupel, 2001) Data on populations of Denmark, Austria, and Australia. **Denmark:** Longitudinal data based on population registry for 1968-2000 n = 1,371,003. **Austria:** Data from death certificates for 1988-1996 n = 681,677 **Australia:** Data from death certificates from 1993-1997 n = 219,820 Also used data on individuals born in Britain who died in Australia: n = 43,074 Data on infant mortality for Denmark	Used t tests to perform pairwise comparisons between mean age at death by quarter of birth. To test whether the seasonal difference in the risk of death accounts for differences in adult life span by month of birth, calculated monthly deviations from annual death rates and used weighted least squares regression with dummies for month of birth, current month, age since last birthday in months, sex, and birth cohort. To test whether selective survival or debilitation during the 1st year of life explains differences in life expectancy at age 50, calculated monthly death rates during 1st year of life and monthly deviations from annual death rates during 1st year of life; then used a multivariate regression without controls for sex and birth cohort.	Note: Autumn-Spring = average difference in age at death between people born in Autumn (Oct-Dec) and in the Spring (Apr-Jun). **Denmark:** Mean remaining life expectancy at age 50 = 27.52 years 0.19 years shorter lifespans for those born in 2nd quarter 0.12 years longer lifespans for those born in 4th quarter Correlation between infant mortality at time of birth and adult mortality after age 50 = 0.87. **Austria:** Average age at death = 77.70 years 0.28 years shorter lifespans for those born in 2nd quarter **Austria:** average age at death = 77.70 years 0.28 years shorter lifespans for those born in 2nd quarter 0.32 years longer lifespans for those born in 4th quarter Additional results shown for specific causes of death. **Australia:** Mean age of death = 78.00 years for those born in 2nd quarter; mean age of death = 77.65 years for those born in 4th quarter British immigrants born Nov.-Jan. have age of death 0.36 years higher than natives. Those immigrants born Mar-May have age of death 0.26 years lower than Australian natives.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Air quality, infant mortality, and the Clean Air Act of 1970 (Chay and Greenstone, 2003a,b) County-level mortality and natality data 1969-1974. Annual monitor level data on total suspended particles (TSP) from EPA. $N = 501$ county-years.	Clean Air Act imposed regulations on polluters in counties with TSP concentrations exceeding federal ceilings. Used nonattainment status as an instrument for changes in TSP. Also used regression discontinuity methods to examine effect of TSP on infant mortality.	1% decline in TSP pollution results in 0.5% decline in infant mortality rate. Most effects driven by reduction of deaths occurring within 1 month of birth.
Air pollution and infant health: what can we learn from California's recent experience? (Currie and Neidell, 2005) Data on pollution from California EPA. Individual-level infant mortality data from California vital statistics 1989-2000. $n = 206{,}353$.	Estimated linear models that approximate hazard models, where the risk of death is defined over weeks of life, and length of life is controlled for with a flexible nonparametric spline. Controlled for prenatal and postnatal pollution exposure, weather, child's age, and numerous other child and family characteristics, as well as month, year, and zip code fixed effects. Examined effects of ozone (O3), carbon monoxide (CO), and particulate matter (PM10) on infant mortality.	1.1 unit reduction in postnatal exposure to CO (the actual reduction that occurred in CA in the 1990s) saved 991 infant lives, −4.6% decrease in infant mortality rate. No statistically significant effects of prenatal exposure to any of the pollutants.

Table 4 (continued)

Study and data	Study design	Results
Prenatal exposure to radioactive fallout (from Chernobyl) and school outcomes in Sweden (Almond et al., 2009). All Swedes born 1983-1988, $n = 562{,}637$. Data on radiation from Swedish Geological Survey at parish level.	Three empirical strategies: (1) Cohort comparisons—Compared cohorts *in utero* before, during, and after Chernobyl with particular focus on cohort 8-25 weeks post conception. (2) Within-cohort comparisons-Use geographic variation in levels of exposure. Define 4 regions, R0 (least exposure), R1, R2, R3. (3) Diff-in-diff sibling comparison—Compared those exposed to radiation *in utero* 8-25 to siblings who were not.	**Results from (1):** Probability of qualifying for high school reduced by 0.2% for cohort *in utero* during radiation; by 0.6% for cohort *in utero* 8-25. Grades reduced by 0.4% for cohort *in utero* during radiation; by 0.6% for cohort *in utero* 8-25. **Results from (2)**: (R3 relative to R0) Probability of qualifying for high school reduced by 3.6% for cohort *in utero* during radiation in R3; by 4% for cohort *in utero* 8-25 in R3. Grades reduced by 3% for cohort *in utero* during radiation in R3; by 5.2% for cohort *in utero* 8-25 in R3. **Results from (3)**: (R3 relative to R0) Difference in probability of qualifying for high school between siblings increased by 6% in R3. Difference in grades between siblings increased by 8% in R3.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Air pollution and infant health: lessons from New Jersey (Currie et al., 2009). Pollution data from EPA. Individual-level data on infant births and deaths from the New Jersey Department of Health 1989-2003. $n = 283{,}393$ for mother fixed effects models.	Mother fixed effects to estimate impact of exposure to pollution (during and after birth) on infant health outcomes. Air pollution measured from air quality monitors and assigned to each childbased on home address of mother. Also included interactions b/n variable for pollution exposure and maternal smoking as well as other maternal characteristics. Controlled for weather, pollution monitor locations, time trends, seasonal effects, and other background characteristics.	1 unit change in mean CO exposure during last trimester of pregnancy decreases average birth weight by 0.5%, increases likelihood of low birth weight by 8%, and decreases gestation by 0.2%. 1 unit change in mean CO exposure in first 2 weeks after birth increases likelihood of infant mortality by 2.5%. Estimated that a 1 unit decrease in mean CO exposure in first 2 weeks after birth would save 17.6 per 100,000 lives. Effects of CO exposure on infant health at birth are 2-6 times larger for smokers and mothers who are over age 35. Effects of PM10 (particulate matter) and ozone are not consistently significant across the specifications.

Table 4 (continued)

Study and data	Study design	Results
Fetal exposures to toxic releases and infant health (Currie et al., 2009) Toxic Release Inventory data at county-year level matched to county-year level natality and infant mortality data. $N = 5279$.	Multivariate regression of infant health outcomes on amount of toxic releases in each county and year, controlling for demographic and socio-economic characteristics, mother drinking or smoking during pregnancy, county employment, and county and year fixed effects. Compared effects of developmental toxins to other toxins, and "fugitive" air releases to "stack" air releases (since emissions that go up a smoke stack are more likely to be treated in some way, and hence will affect those in the vicinity less).	An additional thousand pounds per square mile of all toxic releases leads to: 0.02% decrease in length of gestation 0.04% decrease in birth weight 0.1% increase in probability of low birth weight 0.8% increase in probability of very low birth weight 1.3% increase in infant mortality Larger effects for developmental toxins and for fugitive air releases than for other toxins or for stack air releases. A 2-SD increase in lead releases decreases gestation by 0.02% and decreases birth weight by 0.05%. A 2-SD increase in cadmium releases decreases gestation by 0.03%, decreases birth weight by 0.07%, increases probability of low birth weight by 1.2%, increases probability of very low birth weight by 1.4%, and increases infant mortality by 5%. Similar results for toluene, epichlorohydrin. Reductions in releases over 1988–1999 can account for 3.9% of the reduction in infant mortality over the same time period.

(*continued on next page*)

Table 4 (continued)

Study and data	Study design	Results
Traffic congestion and infant health: evidence from E-ZPass (Currie and Walker, 2009) Birth records for Pennsylvania and New Jersey, 1994-2003. Data on housing prices for New Jersey. $N = 727{,}954$ for diff-in-diff. $N = 232{,}399$ for mother fixed effects.	Identification due to introduction of electronic toll collection (E-ZPass), which reduced traffic congestion and motor vehicle emissions in the vicinity of highway toll plazas. Difference-in-difference, comparing infants born to mothers living near toll plazas to infants born to mothers living near busy roadways (but away from toll plazas), before and after introduction of E-ZPass. Also estimated impacts of exposure to E-ZPass on infant health using mother fixed effects methods. Controlled for various background characteristics, year and month fixed effects, and plaza-specific time trends. Also estimated impact of E-ZPass introduction on housing prices near the toll plaza.	**Results from Difference-in-Difference Method:** Reductions in traffic congestion generated by introduction of E-ZPass reduced incidence of premature birth by 10.8% and low birth weight by 11.8% for children of mothers living w/in 2 km of a toll plaza. For those living w/in 3 km of a toll plaza, effects are 7.3% for premature births and 8.4% for low birth weight. Similar results using mother fixed effects. No effects of E-ZPass introduction on housing prices or demographic composition of mothers living near toll plazas.

Table 4 (continued)

Study and data	Study design	Results
Caution, Drivers! Children Present. Traffic, Pollution, and Infant Health (Knittel et al., 2009). Traffic data from the freeway performance measurement system. Pollution data from EPA. Birth data from California Dept. of Public Health 2002-2006. $N = 373{,}800$.	IV using traffic shocks (due to accidents or road closures) to instrument for air pollution. Preferred specifications include traffic flow, delays, and interactions between traffic and weather.	Note: only results from the IV regression are reported here. 1 unit decrease in exposure to particulate matter (PM10) leads to 5% decrease in infant mortality rate (saves 14 lives per 100,000 births).

Unless otherwise noted, only results significant at 5% level are reported.

sample size prevented detection of all but the largest effects (see Section 3.1). Using a comparable sample size, Behrman and Rosenzweig (2004) found the schooling of identical female twins was nearly one-third of a year longer for a pound increase in birth weight (454 grams), with relatively imprecise effects on adult BMI or wages.

In light of the above power concerns, Currie and Moretti (2007) matched mothers to their sisters in half a million birth records from California. Here, low birth weight was found to have statistically significant negative impacts on educational attainment and the likelihood of living in a wealthy neighborhood. However, the estimated magnitudes of the main effects were more modest: low birth weight increased the likelihood of living in a poor neighborhood by 3% and reduced educational attainment approximately one month on average. Like Conley and Bennett (2001), the relationship was substantially stronger for the interaction between low birth weight and being born in poor neighborhoods.

In a sample of Norwegian twins, Black et al. (2007) also found long-term effects of birth weight, but did not detect any heterogeneity in the strength of this relationship by parental socioeconomic status.[11] Oreopoulos et al. (2008) find similar results for Canada and Lin and Liu (2009) find positive long term effects of birth weight in Taiwan. Royer (2009) found long-term health and educational effects within California twin pairs, with a weaker effect of birth weight than several other studies, esp. Black et al. (2007). Responsive investments could account for this discrepancy if they differed between California and elsewhere (within twin pairs). Alternatively, there may be more homogeneity with respect to socioeconomic status in Scandinavia than in California. As described in Section 3, Royer (2009) analyzed investment measures directly with the ECLS-B data, concluding that there was no evidence of compensatory or reinforcing investments (see Section 2.2).

Following a literature in demography on seasonal health effects, Doblhammer and Vaupel (2001) and Costa and Lahey (2005) focused on the potential long-term health effects of birth season. A common finding is that in the northern hemisphere, people born in the last quarter of the year have longer life expectancies than those born in the second quarter. Both the availability of nutrients can vary seasonally (particularly historically), as does the likelihood of common infections (e.g., pneumonia). Therefore, either nutrition or infection could drive this observed pattern. Almond (2006) focused on prenatal exposure to the 1918 Influenza Pandemic, estimating that children of infected mothers were 15% less likely to graduate high school and wages were between 5 and 9% lower. Kelly (2009) found negative effects of prenatal exposure to 1957 "Asian flu" in Britain on test scores, though the estimated magnitudes were relatively modest. Interestingly, while birth weight was reduced by flu exposure, this effect appears to be independent of the test score effect. Finally, Field et al. (2009) found that prenatal iodine

[11] Royer (2009) notes that Black et al. (2007) find a "negligible effect of birth weight on high school completion for the 1967-1976 birth cohort, but for individuals born between 1977 and 1986, the estimate is nearly six times as large".

supplementation raised educational attainment in Tanzania by half a year of schooling, with larger impacts for girls.

4.1.2. Economic shocks

A second set of papers considers economic shocks around the time of birth. Here, health in adulthood tends to be the focus (not human capital), and findings are perhaps less consistent than in the studies of nutrition and infection described above. Van Den Berg et al. (2006)'s basic result is that adult survival in the Netherlands is reduced for those born during economic downturns. In contrast, Cutler et al. (2007) detected no long term morbidity effects in the Health and Retirement Survey data for cohorts born during the Dustbowl era of 1930s. Banerjee et al. (forthcoming) found that shocks to the productive capacity of French vineyards did not have detectable effects on life expectancy or health outcomes, but did reduce height in adulthood. Baten et al. (2007) related variations in grain prices in the decade of birth to numeracy using an ingenious measure based on "age heaping" in the British Censuses between 1851 and 1881. Persons who are more numerate are less likely to round their ages to multiples of 5 or 10. They find that children born in decades with high grain prices were less numerate by this index.

4.1.3. Air pollution

The third strand of the literature examines the effect of pollution on fetal health. Epidemiological studies have demonstrated links between very severe pollution episodes and mortality: one of the most famous focused on a "killer fog" in London, England and found dramatic increases in cardiopulmonary mortality (Logan and Glasg, 1953). Previous epidemiological research on the effects of moderate pollution levels on prenatal health suggests negative effects but have produced inconsistent results. Cross-sectional differences in ambient pollution are usually correlated with other determinants of fetal health, perhaps more systematically than with nutritional or disease exposures considered above. Many of the pollution studies have minimal (if any) controls for these potential confounders. Banzhaf and Walsh (2008) found that high-income families move out of polluted areas, while poor people in-migrate. These two groups are also likely to provide differing levels of (non-pollution) investments in their children, so that fetuses and infants exposed to lower levels of pollution may tend to receive, e.g., better quality prenatal care. If these factors are unaccounted for, this would lead to an upward bias in estimates. Alternatively, certain pollution emissions tend to be concentrated in urban areas, and individuals in urban areas may be more educated and have better access to health care, factors that may improve health. Omitting these factors would lead to a downward bias, suggesting the overall direction of bias from confounding is unclear.

Two studies by Chay and Greenstone (2003a,b) address the problem of omitted confounders by focusing on "natural experiments" provided by the implementation of the Clean Air Act of 1970 and the recession of the early 1980s. Both the Clean Air

Act and the recession induced sharper reductions in particulates in some counties than in others, and they use this exogenous variation in levels of pollution at the county-year level to identify its effects. They estimate that a one unit decline in particulates caused by the implementation of the Clean Air Act (recession) led to between five and eight (four and seven) fewer infant deaths per 100,000 live births. They also find some evidence that the decline in Total Suspended Particles (TSPs) led to reductions in the incidence of low birth weight. However, only TSPs were measured at that time, so that they could not study the effects of other pollutants. And the levels of particulates studied by Chay and Greenstone are much higher than those prevalent today; for example, PM10 (particulate matter of 10 microns or less) levels have fallen by nearly 50% from 1980 to 2000.

Several recent studies consider natural experiments at more recently-encountered pollution levels. For example, Currie et al. (2009) use data from birth certificates in New Jersey in which they know the exact location of the mothers residence, and births to the same mother can be linked. They focus on a sample of mothers who live near pollution monitors and show that variations in pollution from carbon monoxide (which comes largely from vehicle exhaust) reduces birth weight and gestation. Currie and Walker (2009) exploit a natural experiment having to do with introduction of electronic toll collection devices (E-ZPass) in New Jersey and Pennsylvania. Since much of the pollution produced by automobiles occurs when idling or accelerating back to highway speed, electronic toll collection greatly reduces auto emissions in the vicinity of a toll plaza. Currie and Walker (2009) compare mothers near toll plazas to those who live near busy roadways but further from toll plazas and find that E-ZPass increased birth weight and gestation. They show that they obtain similar estimates following mothers over time and estimating mother fixed effects models. These papers are notable in part because it has proven more difficult to demonstrate effects of pollution on fetal health than on infant health, as discussed further below. Hence, it appears that being *in utero* may be protective against at some forms of toxic exposure (such as particulates) but not others.

This literature on the effects of air pollution is closely related to that on smoking. Smoking is, after all, the most important source of indoor air pollution. Medical research has shown that nicotine constricts the oxygen supply to the fetus, so there is an obvious mechanism for smoking to affect infant health. Indeed, there is near unanimity in the medical literature that smoking is the most important preventable cause of low birth weight. Economists have focused on ways to address heterogeneity in other determinants of birth outcomes that are likely associated with smoking. Tominey (2007) found that relative to a conventional multivariate control specification, roughly one-third of the harm from smoking to birth weight is explained by unobservable traits of the mother. Moreover, the reduction in birth weight from smoking was substantially larger for low-SES mothers. In a much larger sample, Currie et al. (2009)

showed that smoking significantly reduced birth weight, even when comparisons are restricted to within-sibling differences. Moreover, Currie et al. (2009) document a significant interaction effect between exposure to carbon monoxide exposure and infant health in the production of low birth weight, which may help explain the heterogeneity in birth weight effects reported by Tominey (2007). Aizer and Stroud (2009) note that impacts of smoking on birth weight are generally much smaller in sibling comparisons than in OLS and matching-based estimates. Positing that attenuation bias is accentuated in the sibling comparisons, Aizer and Stroud (2009) use serum cotinine levels as an instrument for measurement error in smoking and find that sibling comparisons yield similar birth weight impacts (around 150 g). Lien and Evans (2005) use increases in state excise taxes as an instrument for smoking and find large effects of smoking on birth weight (182 g) as a result. Using propensity score matching, Almond et al. (2005) document a large decrease in birth weight from prenatal smoking (203 g), but argue that this weight decrease is weakly associated with alternative measures of infant health, such as prematurity, APGAR score, ventilator use, and infant mortality.

Some recently-released data will enable new research on smoking's short and long-term effects. In 2005, twelve states began using the new US Standard Certificate of Live Birth (2003 revision). Along with other new data elements (e.g., on surfactant replacement therapy), smoking behavior is reported *by trimester*. It will be useful to consider whether smoking's impact on birth weight varies by trimester, and also whether smoking is more closely tied to other measures of newborn health if it occurs early versus late in pregnancy. Second, there is relatively little research by economists on the long-term effects of prenatal exposure to smoking. Between 1990 and 2003, there were 113 increases in state excise taxes on cigarettes (Lien and Evans, 2005).[12] Since 2005, the American Community Survey records both state and quarter of birth, permitting linkage of these data to the changes in state excise taxes during pregnancy.

Almond et al. (2009) examine the effect of pollution from the Chernobyl disaster on the Swedish cohort that was *in utero* at the time of the disaster. Since the path of the radiation was very well measured, they can compare affected children to those who were not affected as well as to those born in the affected areas just prior to the disaster. They find that in the affected cohort those who suffered the greatest radiation exposure were 3% less likely to qualify for high school, and had 6% lower math grades (the measure closest to IQ). The estimated effects were much larger within families. A possible interpretation is that cognitive damage from Chernobyl was reinforced by parents.

To summarize, the recent "fetal origins" literature in economics finds substantial effects of prenatal health on subsequent human capital and health. As we discuss in

[12] Some states enacted earlier excise taxes: the "average state tax rate increased from 5.7 cents in 1964 to 15.5 cents in 1984" (Farrelly et al., 2003); high 1970s inflation can be an additional potential source of identification as excise taxes were set nominally.

Section 5, this suggests a positive role for policies that improve human capital by affecting the birth endowment. That is, despite being congenital (i.e. present from birth), this research indicates that the birth endowment is malleable in ways that shape human capital. This finding has potentially radical implications for public policy since it suggests that one of the more effective ways to improve children's long term outcomes might be to target women of child bearing age in addition to focusing on children after birth.

4.2. Early childhood environment

It would be surprising to find that a very severe shock in early childhood (e.g., a head injury, or emotional trauma) had no effect on an individual. Therefore, a more interesting question from the point of view of research is how developmental linkages operating at the individual level affect human capital formation in the aggregate. To answer this question, we need to know how many children are affected by negative early childhood experiences that could plausibly exert persistent effects? How big and long-lasting are the effects of less severe early childhood shocks relative to more severe shocks? Taken together, how much of the differences in adult attainments might be accounted for by things that happen to children between birth and age five? Furthermore, how are these linkages between shocks and outcomes mediated or moderated by third factors? For example, is the effect of childhood lead exposure on subsequent test scores stronger for families of lower socioeconomic status (i.e. is the interaction with SES an important one) and if so why? Alternatively, is the effect of injury mediated by health status, or is the causal pathway a direct one to cognition?

We might also wish to know how parents respond to early childhood shocks. To date, there has been less focus on this question in the early childhood period than in the prenatal period, perhaps because it seems less plausible to hope to uncover a "pure" biological effect of a childhood shock given that children are embedded in families and in society. However, this embeddedness opens the possibility that a richer set of behavioral responses—of the kind considered by economists—might be at play. Furthermore, early childhood admits a wider set of environmental influences than the prenatal period. For example, abuse in early childhood can be distinguished from malnutrition, a distinction more difficult for the *in utero* period, and these may have quite different effects.

We define early childhood as starting at birth and ending at age five. From an empirical standpoint, early childhood so defined offers advantages and disadvantages over analyses that focus on the prenatal period. Mortality is substantially lower during early childhood than *in utero*, which reduces the scope for selective attrition caused by environmental shocks to affect the composition of survivors. On the other hand, it is unlikely that environmental sensitivity during early childhood tapers discontinuously at any precise age (including age five). From a refutability perspective, we cannot make sharp temporal comparisons of a cohort "just exposed" to a shock during early childhood to a neighboring cohort "just unexposed" by virtue of its being too old to be sensitive.

Moreover, it will often be difficult to know *a priori* whether prenatal or postnatal exposure is more influential.[13] Thus, studies of early childhood exposures tend to emphasize cross-sectional sources of variation, including that at the geographic and individual level. The studies reviewed in this section focus on tracing out the relationships between events in early childhood and future outcomes, and are summarized in Table 5.

4.2.1. Infections

Insofar as specific health shocks are considered, infections are the most commonly studied. In epidemiology, long-term health effects of infections—and the inflammation response they trigger—has been explored extensively, e.g. Crimmins and Finch (2006). Outcomes analyzed by economists include height, health status, educational attainment, test scores, and labor market outcomes. The estimated impacts tend to be large. Using geographic differences in hookworm infection rates across the US South, Bleakley (2007) found that eradication after 1910 increased literacy rates but did not increase the amount of completed schooling, except for Black children. The literacy improvement was much larger among Blacks than Whites, and stronger among women then men. The return to education increased substantially, and Bleakley (2007) estimated that hookworm infection throughout childhood reduced wages in adulthood by as much as 40%. Case and Paxson (2009) focussed on reductions in US childhood mortality from typhoid, malaria, measles, influenza, and diarrhea during the first half of the 20th Century. They found that improvements in the disease environment in one's state of birth were mirrored by improved cognitive performance at older ages, but like Bleakley (2007), this effect did not seem to operate through increased years of schooling. However, the estimated cognitive impacts in Case and Paxson (2009) were not robust to the inclusion of state-specific time trends in their models.

Chay et al. (2009) found that reduced exposure to pneumonia and diarrhea in early childhood among Blacks during the late 1960s raised subsequent AFQT and NAEP scores towards those of Whites. Changes in postneonatal mortality rates (dominated by infections) explained between 50% and 80% of the (large) reduction in the Black-White AFQT gap. Finally, Bozzoli et al. (forthcoming) highlight that in developing countries, high average mortality rates cause the selection effect of early childhood mortality to overwhelm the "scarring" effect. Thus, the positive relationship between early childhood health and subsequent human capital may be absent in analyses that do not account for selective attrition in high mortality settings.

4.2.2. Health status

Many of the studies reviewed in Table 5 investigate the link between health in childhood and future cognitive or labor market outcomes. These studies can be viewed as a subset of

[13] For example, early postnatal exposure to Pandemic influenza apparently had a larger impact on hearing than did prenatal flu exposure (Heider, 1934).

Table 5 Impacts of early childhood shocks on later outcomes.

Study and data	Study design	Results
Childhood physical and mental health		
Mental health in childhood and human capital (Currie and Stabile, 2006). Data from NLSCY and NLSY. Canadian NLSCY: data on children aged 4-11 in 1994. Mental health screening in 1994; outcomes measured in 2002. $n = 5604$. US NLSY: data on children aged 4-11 in 1994. Mental health score averaged over 1990-1994. $n = 3758$.	Sibling fixed effects. Hyperactivity and aggression scores based on "screener" questions asked of all children. Estimated effect of hyperactivity on grade repetition, reading and math scores, special education, and delinquency. Controlled for individual background characteristics. Estimated same model omitting children with other learning disabilities besides those in the main explanatory variable.	**1 unit change in hyperactivity score:** increases probability of grade retention by 10-12% in both US and Canada decreases math scores by 0.04-0.07 SD in both US and Canada increases probability of being in special ed by 11% in US decreases reading scores by 0.05 SD in US **1 unit change in conduct disorder score (in the US only):** increases probability of grade retention by 10% decreases math scores by 0.02 SD decreases reading scores by 0.03 SD **1 unit change in aggression score (in Canada only)**: decreases probability that a youth aged 16-19 is in school by 4% high depression scores increase probability of grade retention by 10% in both US and Canada no significant effects of interaction between mental health scores and income or maternal education.

Table 5 (continued)

Study and data	Study design	Results
Disease and development: Evidence from hookworm eradication in the American South (Bleakley, 2007). Data on hookworm infection rates from the Rockefeller Sanitary Commission surveys for 1910-1914. Census data from IPUMS for 1880-1990. $n = 115$ state-years.	Identification due to different pre-eradication hookworm infection rates in different states. Compare individuals born before and after eradication campaigns funded by the Rockefeller Sanitary Commission (RSC). Active years of RSC were 1910-1915. Considered contemporaneous effects on children as well as long-term effects on adult wages and educational attainment. Controlled for geographic and year fixed effects, as well as some individual characteristics. In regressions for contemporaneous effects, geographic units are state economic areas. In regressions for long-term effects, geographic units are states of birth.	**Contemporaneous effects of infections on children:** 1 SD increase in lagged hookworm infection associated with: 0.18-0.25 SD decrease in school enrollment 0.21-0.28 SD decrease in full-time school enrollment 0.1 SD decrease in literacy Results robust to inclusion of state-year fixed effects, controlling for mean-reversion in schooling, and using state-level infection rates. Larger effects for blacks than for whites. No contemporaneous effects on adults. **Long-term effects of infections in childhood on adults:** Being infected with hookworm in one's childhood leads to a reduction in wages of 43% and a decrease in returns to schooling by 5%. 80% reduction in wages due to hookworm infections explained by reduced returns to schooling. Being infected with hookworm in one's childhood leads to a reduction in occupational income score by 23% and a decrease in Duncan's Socio-Economic Index by 42%.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Adult health and childhood disease (Bozzoli et al., forthcoming) Data on height from the European Community Household Panel, the Health Survey of England, and the National Health Interview Survey in the US for individuals born in 1950-1980. Data on post-neonatal mortality from the World Health Organization. $n = 316$ country-years. Also used data on women's heights from the international system of Demographic and Health Surveys on women aged 15-49 in more than 40 countries in the late 1990s-2000s. Data on infant mortality from the United Nations population division. $n = 1514$ country-years.	Analyzed relationship b/n post-neonatal mortality (PNM, death after 28 days and before first birthday) and adult height. PNM is a measure of childhood health environment. OLS regression with adult height as outcome variable, controlling for country and year fixed effects and a time trend. Also controlled for neonatal mortality rates and GDP in year and country of birth. Considered mortality from pneumonia, intestinal disease, congenital anomalies, and other causes separately.	Survivors are expected to be positively selected relative to those who died, but may still be stunted by illness. In poor countries, the selection effect dominates, whereas in rich countries (with low mortality rates) the stunting effect dominates. Overall correlation b/n PNM and average adult height of the same cohort in a given country $= -0.79$. In the US, PNM was 3x larger than in Sweden in 1970. This diff accounts for 20-30% of the 2-cm diff in average height b/n 30-yr-old Americans and Swedes in 2000. After controlling for PNM, no relationship b/n adult height and GDP in the year and country of birth. Biggest determinant of differences in PNM rates across countries is mortality from intestinal disease, followed by mortality from pneumonia. Out of the 4 determinants of PNM, only mortality from pneumonia has a significant negative effect on adult height.

Table 5 (continued)

Study and data	Study design	Results
Stature and status: height, ability and labor market outcomes (Case and Paxson, 2008a). Data from NCDS (on cohort of individuals born in week of 3/3/1958 in Britain), BCS, NLSY (on siblings in the US), and Fragile Families and Child Well-Being Study (on young children in the US). NCDS: $n = 9155$; BCS: $n = 9003$; NLSY: $n = 13{,}884$ (total, not siblings); Fragile Families: $n = 2150$.	Analyzed relationship b/n adult height and cognitive ability as a potential explanation for the significant and positive relationship b/n adult height and earnings that is observed. Adult height is a proxy for early childhood health conditions. OLS regression of cognitive test scores on height-for-age (HFA) z-scores,controlling for background and demographic variables (using NCDS, BCS and fragile families data). Same regression with mother fixed effects using NLSY79 data. OLS regression of log hourly earnings on adult height, controlling for cognitive test scores at ages 7, 10, and 11, and other background and demographic variables.	**Child's height and cognitive test scores at age 3:** 1 SD increase in HFA z-score linked to a 0.05-0.1 SD increase in PPVT score. **Child's height and cognitive test scores at ages 5-10:** (reporting only mother fixed effects results) 1 SD increase in HFA z-score linked to increased PIAT math score, PIAT reading recognition score, and PIAT reading comprehension score by 0.03 SD. Height does not explain differences in test scores across racial groups. **Adult height, earnings, and cognitive test scores:** Inclusion of cognitive test scores at ages 7, 10, and 11 makes the coefficient on adult height insignificant for predicting hourly wages. Height difference b/n men and women does not account for the difference in earnings.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Height, health and cognitive function at older ages (Case and Paxson, 2008b). Data from HRS on men and women aged 50 and older from 1996 to 2004. $n = 72{,}258$.	OLS regression of health and cognition measures on adult height, controlling for age, survey wave, race and sex. Adult height is a proxy for early childhood health and nutrition. Also estimated models including controls for childhood health (self-reported), completed education, and a dummy for employment in a white-collar occupation.	1 in increase in height associated with: 0.9% increase in delayed word recall score 0.3% increase in probability of being able to count backwards 0.3% increase in probability of knowing the date 1.4% decrease in depression score 0.4% decrease in health status scale (1 = excellent, 5 = poor) Childhood health, completed education, and employment in a white-collar occupation all positively related to adult cognitive function and health. Inclusion of these controls makes the coefficients on adult height insignificant except for the cases of delayed word recall and depression.
The role of childhood health for the intergenerational transmission of human capital: evidence from administrative data (Salm and Schunk, 2008). Administrative data from the department of health services in the city of Osnabrueck, Germany, collected during official school entrance medical examinations on children aged 6 between 2002-2005. Sibling sample: $n = 947$. 321 children had at least one parent w/ college degree.	Sibling fixed effects to estimate impact of health conditions on cognitive and verbal ability at school entrance. Controlled for individual background characteristics. Used the Oaxaca (1973) decomposition and a nonparametric decomposition to estimate how much of the gap b/n children of college-educated parents and children of less educated parents can be explained by chronic childhood health conditions.	Mental health conditions reduce cognitive ability by 10% for the whole sample, by 9% for college-educated sample, by 11% for less-educated sample. Asthma reduces cognitive ability by 8% for less-educated sample only. Mental health conditions reduce verbal ability by 11% for the whole sample, and by 13% for the less-educated sample. Health conditions explain 18% of the gap in cognitive ability and 65% of the gap in language ability b/n children of college-educated and less-educated parents.

Table 5 (continued)

Study and data	Study design	Results
Long-term economic costs of psychological problems during childhood (Smith and Smith, 2008) Data from PSID on siblings (with a special supplement designed by the authors in the 2007 wave, asking retrospective questions about childhood health). Sample consists of sibling children of the original participants who were at least 16 in 1968. Sibling children were at least 25 in 2005.	Sibling fixed effects to estimate the impact of reporting having had childhood psychological problems (measured by depression, drug or alcohol abuse or other psychological problems before age 17) on later life socio-economic outcomes. Controlled for individual childhood physical illnesses (asthma, diabetes, allergic conditions, and many others) as well as family background characteristics.	Having had psychological problems during childhood leads to: 20% reduction in adult earnings ($10,400 less per year; $17,534 less family assets) reduction in number of weeks worked by 5.76 weeks per year 11 pp reduction in likelihood of getting married 33.5 pp reduction in educational attainment (means not reported, so can't calculate relative effect sizes) **Costs of childhood psychological problems:** Lifetime cost: $300,000 loss in family income; $3.2 trillion cost for all those affected
Early life health and cognitive function in old age (Case and Paxson, 2009) Region-level historical data on mortality from infectious diseases, as well as total infant mortality. Data for 1900-1936 from Grant Miller's data archive on NBER website. Data for 1937-1950 from vital statistics documents. Data on later life outcomes from Health and Retirement Study for 1996-2004 on men and women aged 50-90. $n = 60{,}000$.	Identification due to variation in mortality rates across time and regions. OLS regression of later life outcomes on log mortality rates from various infectious diseases in year of birth and in 2nd year of life in region of birth. Controlled for age, sex, race, and current Census region of residence.	Decrease in infant mortality by half during 2nd year of life associated w/ an increase in delayed word recall score by 0.1 SD. Significant negative impacts of typhoid, influenza, and diarrhea mortality in 2nd year of life on delayed word recall score. (Means not reported so can't calculate relative effect sizes). Weaker associations b/n disease mortality and ability to count backwards. No significant impacts of disease mortality and overall mortality during year of birth, once mortality in 2nd year of life is included, on either of the later life outcomes. Results not robust to adding Census region-specific time trends.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Child health and young adult outcomes (Currie et al., 2010) Administrative data on public health insurance records from the Canadian province of Manitoba on children born b/n 1979 and 1987, followed until 2006. $n = 50{,}000$.	Sibling fixed effects of the long term effects of health problems at various child ages controlling for health at birth (with birth weight and congenital anomalies). Key explanatory variables examined are: asthma, major injuries, ADHD, conduct disorders, and other major health problems at ages 0-3, 4-8, 9-13, and 14-18. Key outcome variables: achievement on standardized test on language arts in grade 12, whether child took college-preparatory math courses in high school,whether child is in grade 12 by age 17, and welfare participation after becoming eligible at 18.	**Results reported below are only those that control for health problems at all age groups:** An additional major health condition at ages 0-3/4-8/14-18 increases the probability of being on welfare by 10%/9%/31%. Effects of major health conditions at younger ages on educational outcomes not significant when controlling for health at ages 14-18. **ADHD/conduct disorders diagnosis:** at ages 0-3/4-8/9-13/14-18 decreases probability of being in grade 12 by age 17 by 4%/10%/17%/19%; at ages 4-8/9-13/14-18 increases probability of being on welfare by 38%/44%/109%; at ages 4-8/9-13/14-18 decreases probability of taking a college prep math class by 11%/25%/35%; at ages 4-8/9-13/14-18 decreases literacy score by 0.15 SD/0.23 SD/0.27 SD. Effects of asthma at younger ages not statistically significant once health at ages 14-18 is controlled for. **Major injury:** at ages 0-3/14-18 increases probability of being on welfare by 7%/9%; at ages 9-13/14-18 decreases probability of being in grade 12 by age 17 by 2%/2%; at ages 9-13/14-18 decreases probability of taking a college prep math class by 6%/8%; at ages 9-13/14-18 decreases literacy score by 0.03 SD/0.03 SD. Children who have a major physical health condition and then recover do not have significant adverse outcomes. Children with mental health conditions, children with major health conditions at ages 14-18, and those with conditions that persist for multiple age periods suffer worse outcomes.

Table 5 (continued)

Study and data	Study design	Results
The impact of childhood health on adult labor market outcomes (Smith, 2009) Data on siblings from PSID. Childhood health measured as a self-reported retrospective health index regarding health at ages younger than 17. Adult outcomes measured in 1999. $n = 2248$.	Sibling fixed effects to estimate impact of self-reported retrospective childhood health status (on a 5-point scale) before age 16 on adult earnings, employment, education, marital status, etc., controlling for demographics and family background.	Better health during childhood increases adult family income by 24%, increases adult family wealth by 200% (relative to mean = \$2000), and increases adult earnings by 25%. Better health during childhood increases probability of having worked the year before by 5.4 pp (mean not reported). About 2/3 of overall impact of poor childhood health on adult family income is present at age 25; the remaining 1/3 is due to a slower growth path after age 25 due to poor childhood health. About 1/2 of overall impact of poor childhood health on individual earnings is present at age 25; the remaining 1/2 is due to a slower growth path after age 25 due to poor childhood health. No statistically significant impacts of childhood health on educational attainment.
The effect of childhood conduct disorder on human capital (Vujic et al., 2008) Data from the Australian Twin Register on twins born between 1964 and 1971. Data collected in 1989-1990 and 1996-2000 $n = 5322$ twins; 2250 identical twins.	Twin fixed effects to estimate impact of childhood conduct disorder (measured by various indicators) based on diagnostic criteria from psychiatry on educational attainment and criminal behavior in adulthood. Controlled for birth weight, timing of the onset of conduct disorders, and other family and individual background characteristics. Conducted separate analyses for all twins and identical twins.	Childhood conduct disorders lead to: 5-16% decrease in likelihood of high school graduation 100-228% increase in likelihood of being arrested (mean =0.07) 6-68% increase in likelihood of grade retention 20-60% increase in likelihood of having 3+ job quits (not significant in identical twin sample) 50-325% increase in likelihood of telling lies (mean = 0.04) 37-526% increase in likelihood of going to jail (mean = 0.019) Earlier occurrence of conduct disorder has larger negative effects than later occurrence.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Causes and consequences of early life health (Case and Paxson, 2010a) Data from NCDS, BCS, PSID, Whitehall II Study (longitudinal study of British civil servants b/n ages 34 and 71, collected from 1985 to 2001), HRS, and NLSY79. See Case and Paxson (2008a,b) for more info on the data sets. $n = 11{,}648$ (NCDS), 11,181 (BCS), 63,995 (PSID), 29,774 (Whitehall II), 66,269 (HRS), $n = 3200$-46,000 (NLSY79).	Multivariate regression of educational attainment, employment, log earnings, and self-reported health status and cognitive function on adult height in inches (proxy for early childhood health) using 5 longitudinal data sets. Controlled for age, ethnicity, sex, and survey wave. Used sibling fixed effects in NLSY79 to understand what aspects of early life health adult height captures—regressed children's test scores, grade level, self-perception in school, and childhood health outcomes on children's HAZ.	Results from NCDS, BCS, PSID, Whitehall II, and HRS: One inch of height is associated with: 0.05-0.16 more years of schooling; 0.2-0.6 pp increase in likelihood of employment; 0.012-0.028 increase in average hourly earnings for men; 0.007-0.027 increase in average hourly earnings for women. A 4-in increase in height leads to: 8% decrease in probability of long-standing illness; 40% decrease in probability of disability; 4% of SD decrease in depression score. Results from NLSY79 (reporting sibling fixed effects only) indicate that 1 pt. increase in HAZ leads to: 0.3% increase in PIAT math score; 0.1% increase in PIAT reading recognition score; 0.2% increase in PIAT reading comprehension score; 0.9% increase in digital span test score; 0.4% increase in PPVT score. 2% decrease in likelihood that child is in appropriate grade level for age (ages 6-14). 1% increase in child doing school work quickly; 1% increase in child remembering things easily; 0.9% increase in total scholastic competence score. 8% decrease in likelihood that child has limiting emotional/neurological condition; 0.105 increase in birth weight z-score (mean not reported).

Table 5 (continued)

Study and data	Study design	Results
The long reach of childhood health and circumstance: evidence from the Whitehall II Study (Case and Paxson, 2010b) Data from the Whitehall II Study (see above for description). $n = 10{,}308$.	Since most Whitehall II study participants belong to the highest occupational classes, estimated the selection effect using data from NCDS and BCS for comparison. OLS regressions of initial placement and promotion in Whitehall on height, family background characteristics, educational attainment, and other individual characteristics. Also estimated individual fixed effects and first-differenced regressions of future promotions on self-reported health, and of future health on grade in Whitehall.	Selectivity of Whitehall II study attenuates the impact of father's social class on various outcomes relative estimates that would be obtained in the full population. OLS regressions indicate significant correlations between family background (father's social class), initial placement and likelihood of promotion. Significant relationship between initial placement and promotion, height, and self-reported health. Although some of the effects of health on placement and promotion are mediated by educational attainment, there is an independent effect of health. Individual fixed effects estimates indicate that cohort members who report "excellent" or "very good" health are 13% more likely to be promoted (relative to lowest grade). No effect of grade in Whitehall on future health.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Home environment		
The impact of maternal alcohol and illicit drug use on children's behavior problems: evidence from the children of the National Longitudinal Survey of Youth (Chatterji and Markowitz, 2000) Data from the 1988, 1992, and 1994 waves of NLSY-Child surveys on children who were 4-14 yrs old during the survey years. $n = 6194$ children in total sample; $n = 2498$ mothers who have siblings in sample; $n = 7546$ obs for children who have 2 or 3 observations in sample.	Estimated impact of maternal substance abuse on children's behavior problems using OLS, IV, child-specific & maternal family fixed effects models. In IV, used alcohol and illicit drug prices as instruments for maternal substance abuse. Controlled for numerous demographic and background variables. Behavior problems measured by the Behavior Problems Index (higher BPI means more problems).	No statistically significant relationship between maternal substance abuse and child behavior in IV regressions, but 1st stage relationship is weak. Fixed effects suggest that 1 more drink consumed by mother in past month increases child BPI by 1%. Maternal marijuana use in past month increases child BPI by 7-8%. Maternal cocaine use in past month increases child BPI by 19%.

Table 5 (continued)

Study and data	Study design	Results
Parental employment and child cognitive development (Ruhm, 2004). Data from the Children of NLSY79 for 1986-1996 (mothers aged 29-38 at the end of 1995). $n = 3042$.	OLS to estimate the impact of maternal work in the year prior to child's birth, and in the first 4 years of child's life on child cognitive test scores at ages 3-4 and 5-6. Used a rich set of controls for maternal background, demographic, and socio-economic characteristics, as well as assessment year fixed effects. Maternal work measured both by hours and weeks worked during the year. Controlled for family income.	**20 more hrs. worked each week by mother during child's 1st year:** Decrease PPVT score by 0.06-0.10 SD (ages 3-4) **20 more hrs. worked each week by mother during child's age 2-3:** Decrease PIAT-Reading score by 0.06-0.08 SD (ages 5-6) Decrease PIAT-Math score by 0.05-0.06 SD (ages 5-6) Results robust to alternative specifications (measuring employment by weeks worked or part-time vs. no employment). More negative effects on PPVT and PIAT-Math scores for boys than girls. More negative effects on PIAT-Reading scores for girls than boys. More negative effects on PPVT an PIAT-Reading scores for whites than blacks. More negative effects on PIAT-Math scores for blacks than whites.
Maternity leave, early maternal employment and child health and development in the US (Berger et al., 2005) Data from the Children of NLSY79 on children born b/n 1988 and 1996. $n = 1678$.	OLS and propensity score matching to estimate impact of maternal return to work within 12 weeks of birth on child health and developmental outcomes. Models control for numerous background characteristics as well as state and year of birth fixed effects. Sample limited to mothers who worked pre-birth.	**Mother returning to work within 12 weeks of giving birth leads to:** 2.5% decrease in likelihood of well-baby visit 12.8% decrease in likelihood of any breastfeeding 40.4% decrease in number of weeks breastfed 4.3% decrease in child getting all DPT/Polio immunizations Results from OLS consistent with results from propensity score matching methods. Larger negative effects of returning to work full-time within 12 weeks of giving birth compared to any work at all.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Evidence from maternity leave expansions of the impact of maternal care on early child development (Baker and Milligan, 2010) Data from the NLSCY in Canada on children up to age 29 months and their mothers in 1998-2003.	Identification due to large expansions in parental leave policies in Canada on December 31, 2000 (from 25 to 50 weeks of leave). OLS regression. Key explanatory variable is a dummy for birth after December 31, 2000. Outcomes include child development indicators, maternal employment, and use of child care. Controlled for various individual and family background characteristics as well as labor market variables. Observations from Quebec omitted due to changes in childcare policies over the same time period. Observations from single-parent families omitted b/c of expansions in benefits to those families. Subsample analysis for women who returned to work within 1 year of child's birth.	Expansions in parental leave policies led to a 48-58% increase in months of maternal care during the 1st year of life and a 25-29 pp decrease in non-parental care. No statistically significant effects of parental leave expansions on child development indicators (motor/social development, behavior, physical ability, cognitive development).

Table 5 (continued)

Study and data	Study design	Results
The effect of expansions in maternity leave coverage on children's long-term outcomes (Dustmann and Schönberg, 2009) Administrative data on students in Germany attending public schools in Hesse, Bavaria, and Schleswig-Holstein (3 states in Germany) for 2002-03 to 2005-06. $n = 101{,}257$. Administrative data on social security records for German individuals born b/n July 1977 and June 1980. $n = 140{,}387$	Identification due to 3 policy reforms in Germany that expanded unpaid and paid maternity leave to estimate impact on adult wages, employment, and educational outcomes. First reform in 1979 increased paid leave from 2 to 6 months. Second reform in 1986 increased paid leave from 6 to 10 months.Third reform in 1992 increased unpaid leave from 18 to 36 months. OLS regression, comparing children born one month before and after each reform,controlling for background variables and state fixed effects. Also difference-in-difference, comparing children born before and after leave expansions with children born in the same months the year before (hence not affected by expansions).	Despite significant delays in maternal return to work due to policy reforms, there are no statistically significant impacts of expansions in maternity leave policies (paid or unpaid) on any long-run outcomes (wages, employment, selective high school attendance, grade retention and grade attendance).

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Child protection and adult crime: using investigator assignment to estimate causal effects of foster care (Doyle, 2008) Data from computerized criminal history system from the Illinois State Police on arrests and imprisonment up to age 31 for 2000-2005, linked with child abuse investigation data on individuals who are 18-35 in 2005. $n = 23{,}254$.	IV models of the effect of foster care placement on measures of criminal activity. Identification due to the fact that child protection cases are randomized to investigators, and investigators influence whether a child is placed in foster care or remains at home. Model allows for treatment effect heterogeneity—used random coefficient model. IV where the instrument is the investigator's probability of using foster placement relative to the other investigators.	Children on the margin for placement into foster care are 1.5 times more likely to be arrested, 2.68 times more likely to have a sentence of guilty/withheld, 3.41 times more likely to be sentenced to prison if they are placed into foster care. Children on the margin are likely to include African-Americans, girls, and young adolescents.

Table 5 (continued)

Study and data	Study design	Results
Long-term consequences of child abuse and neglect on adult economic well-being (Currie and Widom, 2009). Sample of abused/neglected children based on court-substantiated cases of childhood physical and sexual abuse in 1967-1971 in one Midwestern metropolitan county. Maltreated children matched to controls on the basis of sex, race, and elementary school class/hospital of birth. Cases restricted to children 11 yrs or younger. Adult outcomes measured at mean age 41. Used info collected in 1989-95 and 2003-04 interviews. $n = 1195$ in 1989-95; $n = 807$ in 2003-04. Matched sample (both members of matched pair interviewed during the 2 waves): $n = 358$.	Multivariate regression with key explanatory variable being a dummy for having been maltreated. Controlled for demographic and family background characteristics, as well as quarter of year at time of interview. Also estimated models separately for males and females and for the subsample of participants whose families received food stamps or welfare when they were children.	**Individuals who were maltreated as children:** complete 4.3% less years of schooling (1989-95) score 5.3% lower on IQ test (1989-95) have 24% lower imputed earnings (2003-04) are 0.52 times as likely to have a skilled job (1989-95) are 0.46 times as likely to be employed (2003-04) are 0.58 times as likely to own a vehicle (2003-04) For males, the only significant effects of maltreatment are for years of schooling and having a skilled job. For females, significant negative effects of maltreatment on years of schooling, IQ test scores, imputed earnings, being employed, owning a bank account, owning a stock, owning a vehicle, and owning a home. Effects larger for females than for males.

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Child maltreatment and crime (Currie and Tekin, 2006) Data from the National Longitudinal Study of Adolescent Health (AddHealth). First wave: 1994-95; last wave: 2001-2002. $n = 13{,}509$ (full sample), $n = 928$ (twins sample).	OLS, propensity score matching, and twin fixed effects methods. Estimate the impact of maltreatment on criminal activity. Controlled for numerous demographic and background characteristics as well as state fixed effects. Also controlled for parental reports of child being bad-tempered or having a learning disability.	Only results that are significant in all 3 main specifications (OLS, propensity scores, and twin fixed effects) are reported. **Children who experienced any maltreatment are:** 99-134% more likely to commit any non-drug related crime (mean = 0.109) 288-489% more likely to commit a burglary (mean = 0.009) 113-181% more likely to damage property (mean = 0.052) 106-131% more likely to commit an assault (mean = 0.049) 183-222% more likely to commit a theft > \$50 (mean = 0.018) 76-101% more likely to commit any hard-drug related crime (mean = 0.085) 96-103% more likely to be a crime victim (mean = 0.077) Probability of crime increases if a person suffers multiple forms of maltreatment. Being a victim of maltreatment doubles the probability that an individual is convicted as a juvenile. **Cost-Benefit:** Estimated costs of crime induced by abuse = \$8.8-68.6 billion/year Estimated costs of home visiting programs (that have been shown to reduce cases of maltreatment by 50%) = \$14 billion/year.

Table 5 (continued)

Study and data	Study design	Results
The effect of maternal depression and substance abuse on child human capital development (Frank and Meara, 2009) Data from children of NLSY79 on children aged 1-5 in 1987. $n = 1587$.	OLS regression of the impact of maternal depression and substance abuse on maternal inputs and children's outcomes in grades 1-5 and 6-9. Controlled for a rich set of demographic and background variables including family measures of parent and sibling behavior and health. For robustness, estimated models using mother fixed effects and propensity scores.	**Effects on maternal inputs:** Maternal depression leads to a 0.2 SD decrease in emotional stimulation sub-component of the HOME score (ages 7-10, 11-14). Maternal substance abuse leads to a 0.23 SD decrease in emotional stimulation sub-component of the HOME score (ages 11-14). **Effects on child outcomes:** Maternal depression leads to a 0.46 SD increase in behavioral problems index (ages 7-10, 11-14). Maternal alcohol abuse leads to a 0.31 SD decrease in child's PIAT math score (ages 11-14); 0.29 SD increase in behavioral problems index (ages 7-10); 377% decrease in likelihood of ever being suspended or expelled at any age (mean = 0.22).

(*continued on next page*)

Table 5 (continued)

Study and data	Study design	Results
Environmental shocks		
Environmental policy as social policy? The impact of childhood lead exposure on crime (Reyes, 2007). Data on lead in gasoline for 1950-1990 on a state-by-state level from Yearly Report of Gasoline Sales by State, Petroleum Marketing Annual, and Petroleum Products Survey. Data on air lead exposure from the EPA's Aerometric Information Retrieval System. Data on per capita crime rates from the Uniform Crime Reports compiled by the Federal Bureau of Investigation for 1985-2002. Data on individual blood lead levels from NHANES for 1976-1980 only.	Used state-by-year variation in the decline of lead exposure from gasoline between 1975 and 1985 due to enforcement of the Clean Air Act to identify the impacts of childhood lead exposure on violent crime in young adulthood. Also estimated the impact of lead exposure on levels of lead in blood during childhood for robustness. Primary measure of lead exposure is the grams of lead per gallon of gasoline in automobile sources. Calculated air lead exposure (measure of average concentration of lead in the air in each state and year). Calculated effective lead exposure relative to each crime in each state and year as the weighted average of lead exposure b/n ages 0 and 3 for all cohorts of arrestees. Regression of per capita crime rate on effective lead exposure, controlling for state and year fixed effects and other state-year characteristics that could affect crime rates.	Elasticity of violent crime with respect to lead exposure = 0.8. Changes in childhood lead exposure in the 1970s are responsible for a 56% decrease in violent crime in the 1990s (while abortion is responsible for a 29% drop in violent crime). Weak evidence for link b/n lead exposure and murder or property crimes. **Cost-Benefit Analysis:** Cumulative social cost of switching to unleaded gasoline: $15-65 billion. Cumulative social benefit of reduced crime from reduced lead exposure: $1.2 trillion. Note, due to data limitations, it is not possible to make a direct connection between blood lead levels, lead in gasoline, and criminal activity at the individual level.

Table 5 (continued)

Study and data	Study design	Results
The long-term effects of early childhood lead exposure: evidence from sharp changes in local air lead levels induced by the phase-out of leaded gasoline (Nilsson, 2009) Data on air lead exposure from Swedish Environmental Protection Agency that uses a nationwide grid of moss samples collected in 1000 locations evenly spread across Sweden for average lead exposure levels for 1972-74, 1977-79, and 1982-84. Individual outcome data from the Educational database at the Institute for Labor Market Policy Evaluation in Uppsala, Sweden on those born in 1972-74, 1977-79, and 1982-84. $n = 797{,}889$	Identification due to geographical variation in childhood lead exposure due to a policy in Sweden that induced a phase-out of leaded gasoline between 1973 and 1981. Difference-in-difference comparing children born in municipalities that experienced large reductions in lead exposure to those born in municipalities with relatively no change, before and after policy went into effect, to estimate impact of childhood lead exposure on later life outcomes. Also used mother fixed effect methods, comparing siblings who experienced different levels of lead exposure in childhood. Controlled for various background characteristics as well as year of birth and municipality of birth fixed effects.	A decrease in blood lead levels by 30 mg/kg leads to: 3% decrease in likelihood of being in the lower end of the GPA distribution 0.6 pp increase in average IQ for males 0.05 yr increase in schooling attainment 0.7 pp increase in likelihood of graduating high school 0.5 pp decrease in likelihood of receiving welfare in adulthood Effects of lead are stronger at lower end of ability/skills distribution and for children from poorer households. Effects of lead are stronger when exposure is at ages 0-2 than when exposure is at ages 5-7. No statistically significant difference in effects by gender. Nonlinear effects—significant and negative effects of lead exposure above 75th percentile (>50 mg/kg). No statistically significant effects of exposure to lead below 75th percentile levels. Estimated that there exists a blood lead threshold of 60 mg/L below which reductions in lead exposure have no effect on outcomes. Similar results using mother fixed effects. Note: means not reported, so relative effect sizes can't be calculated.

Unless otherwise noted, only results significant at the 5% level are reported. Percent changes reported relative to the mean.

a broader literature asking whether income affects health, and how health affects income. For example, using cross-sectional US data, Case et al. (2002) find a striking relationship between family income and a child's reported health status, which becomes stronger as children age. Their motivation for looking at children is that the child's health is unlikely to have a large direct effect on family income, so that the direction of causality is relatively clear. Currie and Stabile (2003) investigate this relationship using Canadian panel data and argue that one reason the relationship between income and child health increases over time is that poorer children are subject to many more negative health shocks. In fact, in Canada, this is the dominant mechanism driving the relationship (which is not surprising given that all Canadian children have public health insurance so that gaps in treatment rates are small).[14]

The question we focus on here is how much poor health in childhood, in turn, affects future outcomes. One of the chief ambiguities in answering this question is what we mean by health in childhood. While it has become conventional to measure fetal health using birth weight (though there may be better measures, see Almond, Chay and Lee, 2005) there are a wide variety of different possible measures of child health, ranging from maternal reports about the child's general health status through questions about diagnoses of specific chronic conditions, to the occurrence of "adverse events".

Case and Paxson (2008a,b, 2010a,b) do not have a direct measure of child health, but argue that adult height is a good proxy for early child health. This is a useful observation given that most surveys of adults have no direct information on child health. Height at age 5 is affected by a range of early health shocks including fetal conditions, poor nutrition, and illness. In turn, height at age 5 is strongly predictive of adult height. And like birthweight, it is predictive of many shorter and longer range outcomes.

A second problem is that it is often unclear whether the ill health dates from a particular period (e.g. an injury) or whether it might reflect a continuing, perhaps a congenital, condition. For example, Smith (2009) uses data from the Panel Study of Income Dynamics which asked young adults a retrospective question about their health status before age 16. In models with sibling fixed effects, he finds that the sib with the worse health had significantly lower earnings, although educational attainment was not significantly affected. He also finds using data from the Health and Retirement Survey that reports of general poor health in childhood do tend to be correlated in the expected way with the presence of specific health conditions. However, it is not possible to ascertain that the negative effects are due to poor health at any particular "critical" window. Salm and Schunk (2008) attempt to deal with this problem using detailed health

[14] Condliffe and Link (2008) argue that in the US, differential access to care also plays a role in the steepening of the relationship between income and child health with age. A number of studies have investigated this relationship, dubbed "the gradient", in other countries (*cf* Currie et al. (2007), Case et al. (2008), Doyle et al. (2005), Khanam et al. (2009)). Propper et al. (2007) and Khanam et al. (2009) are particularly interesting because they find that when maternal mental health is controlled, the relationship between family income and child health disappears, suggesting that it is mediated largely by factors that affect maternal mental health.

information from a medical examination of young German children entering school. In models with sibling fixed effects, they find a significant relationship between poor mental health and asthma on the one hand, and measures of cognitive functioning on the other. They control for the child's birth weight in an effort to distinguish between the effects of health at birth and health after birth (though to the extent that birth weight is an imperfect measure of health at birth, it is possible that the other health measures partly capture congenital conditions).

Currie et al. (2010) use administrative data from the Canadian province of Manitoba's public health insurance system to follow children from birth through young adulthood. Using information about all contacts with medical providers, they construct measures of whether children suffered injuries, asthma, mental health problems or other health problems at ages 0 to 3, 4 to 8, 9 to 13, and 14 to 18. It is interesting that even in a large sample, there were relatively few children with specific health problems other than injuries, asthma, or mental health problems, so that it was necessary to group the remaining problems together. They then look at the relationship between health at various ages, educational attainment, and use of social assistance as a young adult in sibling fixed effects models that also control flexibly for birth weight and for the presence of congenital anomalies. The results are perhaps surprising in view of the conceptual framework developed in Section 3. When entered by themselves, early childhood health conditions (at age 0-3 and at age 4-8) are predictive of future outcomes, conditional on health at birth. However, when early physical health conditions are entered along with later ones, generally only the later ones matter. This result suggests that physical health in early childhood affects future outcomes largely because it affects future health (i.e., subsequent health mediates the relationship), and not because there is a direct link between early physical health status and cognition. In contrast, mental health conditions at early ages seem to have significant negative effects on future outcomes even if there are no intermediate report of a mental health condition. This result suggests that common mental health problems such as Attention Deficit Hyperactivity Disorder (ADHD, also called Attention Deficit Disorder or ADD) or Conduct Disorders (i.e. disorders usually involving abnormal aggression and anti-social behavior) may impair the process of human capital accumulation even if they do not lead to diagnoses of mental health disorders in adulthood.

Several recent papers focus specifically on measures of mental health conditions. Currie and Stabile (2006) use questions similar to those on mental health "screeners" which were administered to large samples of children in the US and Canada in two national surveys. They find that children whose scores indicated mental health problems in 1994 had worse outcomes as of 2002-4 than siblings without such problems. They controlled for birth weight (among other variables) and estimated models with and without including children with diagnosed learning disabilities. In all specifications, they found negative effects of high ADHD scores on test scores on schooling attainment.

Smith and Smith (2008) report similar results using data from the PSID which includes retrospective questions about mental health problems before age 16. Like Smith (2009) and Currie et al. (2010) they estimate models with sibling fixed effects, and find significant long term effects of mental health conditions which are much larger than those of physical health conditions. Vujic et al. (2008) focus on conduct disorders using a panel of Australian twins and find that conduct disorder before age 18 has strong negative effects on the probability of high school graduation as well as positive effects on the probability of criminal activity. None of these three papers focus specifically on measures of mental health conditions before age 5 but "externalizing" mental health conditions such as ADHD and Conduct Disorder typically manifest themselves at early ages. Finally, although they are conceptually distinct, many survey measures of mental health resemble measures of "non-cognitive skills". Hence, one might interpret evidence that non-cognitive skills in childhood are important determinants of future outcomes as further evidence of the importance of early mental health conditions (Blanden et al., 2006; Heckman and Rubinstein, 2001).

4.2.3. Home environment

The home is one of the most important environments affecting a young child and there is a vast literature in related disciplines investigating the relationship between different aspects of the home environment and child outcomes. We do not attempt to summarize this literature here, but pick three aspects that may be most salient: Maternal mental health and/or substance abuse, maternal employment, and child abuse/foster care (which may be considered to be an extreme result of bad parenting). Given the importance of child mental health and non-cognitive skills, it is interesting to ask how maternal mental health affects child outcomes? Frank and Meara (2009) examine this question using data from the National Longitudinal Survey of Youth. They include a rich set of control variables (mother's cognitive test score, grandparent's substance abuse, permanent income) and estimate models with mother fixed effects and models with propensity scores. Their estimates suggest large effects (relative to the effects of income) of contemporaneous maternal depression on the quality of the home environment and on children's behavioral problems, but little effect on math and reading scores. Estimates of the effects of maternal substance abuse are mixed, which echoes the findings of Chatterji and Markowitz (2000) using the same data. Unfortunately, the authors are not able to look at the long term effects of maternal depression experienced by children aged 0 to 5 because the depression questions in the NLSY have been added only recently. As these panel data are extended in time, further investigation of this issue is warranted.

There is also a large literature, including some papers by economists, examining the effect of maternal employment at early ages on child outcomes. Much of this literature suffers from the lack of an appropriate conceptual framework. If we think of child outcomes being produced via some combination of inputs, then the important question is how maternal employment affects the inputs chosen? This will evidently depend on

how much her employment income relaxes the household budget constraint, and the price and quality of the child care alternatives and other inputs that are available. Some of the literature on maternal employment seems to implicitly assume that the mother's time is such an important and unique input that no purchased input can adequately replace it. This may possibly be the case but is a strong assumption. If the mother's time is replaceable at some price, then one might expect maternal employment to have quite different effects on women with different levels of household income (moreover, mother's time may not all be of equal quality, so that it is easier to replace some mother's time than others with the market). This argument suggests that it is extremely important to consider explicitly the quality of the mother's time inputs and the availability of potential substitute inputs in models of maternal employment, something that is difficult to do in most available data sets. Studies that rely on regression methods and propensity score matching (see Berger et al. (2005) and Ruhm (2004)) often find small negative effects of maternal employment (especially in the first year) on children's cognitive development. However, two recent studies using variation in maternity leave provisions find that while more generous maternity leave policies are associated with increased maternal employment, there is little effect on children's outcomes (Baker and Milligan, 2010; Dustmann and Schönberg, 2009). Dustmann and Schönberg (2009) have data that permit cohorts affected by expansions in German maternity leave laws to be followed for many years. They see no effect of maternal employment on educational attainment or wages.

Finally, there are a few papers examining the effects of child abuse/foster care on child outcomes. This is a difficult area to investigate because it is hard to imagine that abuse (or neglect) can be divorced from other characteristics of the household. Currie and Widom (2009) use data from a prospective longitudinal study in which abused children (the treatments) were matched to controls. After following these children until their mid 40s, they found that the abused children were less likely to be employed, had lower earnings, and fewer assets, and that these patterns were particularly pronounced among women. It is possible that these results are driven by unobserved differences between the treatments and controls, although focusing on various subsets of the data (e.g. children whose mothers were on welfare; children of single mothers) produced similar results. Currie and Tekin (2006) use data from the National Longitudinal Study of Adolescent Health to examine the effect of having been abused before age 7 on the propensity to commit crime. They find strong effects which are quite similar in OLS, sibling, and twin fixed effects models. It is possible that these results reflect a characteristic of an individual child (such as difficult temperament) which makes it both more likely that they will be abused and more likely that they will commit crime. However, controlling directly for measures of temperament and genetic endowments does not alter the results. The Doyle (2008) study of the effects of foster care on the marginal child is also summarized in Table 5.

4.2.4. Toxic exposures

Epidemiological studies of postnatal pollution exposure and infant mortality have yielded mixed results and many are likely to suffer from omitted variables bias. Currie and Neidell (2005) examine the effect of more recent (lower) levels of pollution on infant health, along with the role of specific pollutants in addition to particulates (only TSPs were measured during the time periods analyzed by Chay and Greenstone (2003a,b)). Using within-zip code variation in pollution levels, they find that a one unit reduction in carbon monoxide over the 1990s in California saved 18 infant lives per 100,000 live births. However, unlike Currie et al. (2009) they were unable to find any consistent evidence of pollution effects on health at birth, probably because of the crudeness of their measure of maternal location.

Reyes (2007) found large effects of banning leaded gasoline on crime in the US, but results were not robust to state-specific time trends despite a relatively long panel of state-level lead measurements. Nilsson (2009) considered reductions in ambient lead levels in Sweden following the banning of lead in gasoline and measures possible exposures using the concentrations of lead in 1000 moss (bryophyte) collection sites that have been maintained by the Swedish environmental protection agency since the early 1970s. Nilsson (2009) found that early childhood exposure reduced human capital, as reflected by both grades and graduation rates. These effects persisted when comparisons were restricted within siblings, and were substantially larger for low-income families.

4.2.5. Summary re: long term effects of fetal and early childhood environment

The last 10 years have seen an upsurge of empirical work on the long-term effects of early childhood. As a result, much has been learned. We can state fairly definitively that at least some things that happen before age five have long-term consequences for health and human capital. Moreover, these effects are sufficiently large and general to shape outcomes at the population level. On balance, effects of fetal exposure tend to be somewhat larger than postnatal effects, but there are important exceptions. Mental health is a prime example. Mental health conditions and non-cognitive skills seem to have large, persistent effects independent of those captured by measures of child health at birth.

5. EMPIRICAL LITERATURE: POLICY RESPONSES

The evidence discussed above indicates that prenatal and early childhood often have a critical influence on later life outcomes. However, by itself this evidence says little about the effectiveness of remediation. Hence, this section discusses evidence about whether remediation in the zero to five period can be effective in shaping future outcomes. In so doing, we take a step away from explicit consideration of an early-childhood shock u_g as in Section 2. Instead, we focus on the specific public policies that may be able to alter developmental trajectories, often in disadvantaged sub-populations. We begin with programs that raise income, and then move on to programs that target specific domains.

5.1. Income enhancement

In the model sketched above, there are many ways for poverty to affect child outcomes. Even with identical preferences, poorer parents will make different investment choices than richer ones. In particular, poor families will optimize at lower investment (and consumption) levels and thereby have children with lower health and human capital, other things equal. Further, poor parents may face different input prices for certain goods, or have access to different production technologies. Providing cash transfers addresses the budgetary problems without necessarily changing the production technology. Hence, it is of interest to see whether cash transfers, in and of themselves, can improve outcomes. It is however, remarkably difficult to find examples of policies that increase incomes without potentially having a direct effect on outcomes. For example, many studies of cash welfare programs have demonstrated that children who are or have been on welfare remain worse off on average than other children. This does not necessarily mean, however, that welfare has failed them. Without welfare, their situation might have been even worse. Berger et al. (2009) explore the relationship between family income, home environments, child mental health outcomes, and cognitive test scores using data from the Fragile Families and Child Well-being Study which follows a cohort of five thousand children born in several large US cities between 1998 and 2000. They show that all of the measures of the home environment they examine (which include measures of parenting skills as well as physical aspects of the home) are highly related to income and that controlling for these measures reduces the effects of income on outcomes considerably.

Levine and Zimmerman (2000) showed that children who spent time on welfare scored lower than other children on a range of tests, but that this difference disappeared when the test scores of their mothers were controlled for, suggesting that welfare had little effect either positive or negative. Similarly, Levine and Zimmerman (2000) argue that children of welfare mothers were more likely to grow up to be welfare mothers, mainly because of other characteristics of the household they grew up in.

Currie and Cole (1993) compare siblings in families where the mother received welfare while one child was *in utero*, but not while the other child was *in utero*, and find no difference in the birth weight of the siblings. Given that research has shown little evidence of positive effects of cash welfare on children, it is not surprising that the literature evaluating welfare reform in the United States has produced similarly null findings. The National Research Council (Smolensky and Gootman, 2003) concluded that "no strong trends have emerged, either negative or positive, in indicators of parent well-being or child development across the years just preceding and following the implementation of [welfare reform]". However, US welfare reform was a complex intervention that changed many parameters of daily life by, for example, imposing work requirements on recipients.

Conditional tax credits represent an alternative approach to providing income to poor families, and hence to poor children. The early years of the Clinton administration in the

United States saw a huge expansion of the Earned Income Tax Credit (EITC), while in the UK, the Working Families Tax Credit approximately doubled in 1999. These are tax credits available to poor working families. Their essential feature is that they are "refundable"—in other words, a family whose credit exceeds its taxes receives the difference in cash. The tax credits are like welfare in that they give cash payments to poor families. But like welfare reform, the tax credits are a complex intervention in that recipients need to work and file tax returns in order to be eligible, and a great deal of research has shown that such tax credits affect maternal labor supply and marriage patterns (Eissa and Liebman, 1996; Meyer and Rosenbaum, 2001; Blundell, 2006). This is because the size of the payment increases with earnings up to a maximum level before being phased out, so that it creates an incentive to work among the poorest households but a work disincentive for households in the phase-out range. In the US, the number of recipients grew from 12.5 million families in 1990 to 19.8 million in 2003, and the maximum credit grew from $953 to $4204. The rapid expansion of this formerly obscure program run through the tax system has resulted in cash transfers to low-income families that were much larger than those that were available under welfare. Gundersen and Ziliak (2004) estimate that the EITC accounted for half of the reduction in after-tax poverty that occurred over the 1990s (the other half being mainly accounted for by strong economic growth).

Table 6 provides an overview of some of the research on the effects of income on children. Dahl and Lochner (2005) use variation in the amount of the EITC households are eligible for over time and household type to identify the effects of household income and find that each $1000 of income improves childrens' test scores by 0.02% to 0.04% of a standard deviation. An attractive feature of the changes in the EITC is that households may well have regarded them as permanent, so this experiment may approximate the effects of changes in permanent rather than transitory income. Their result implies, though, that it would take on the order of a $10,000 transfer to having an educationally meaningful effect on test scores.

Milligan and Stabile (2008) take advantage of a natural experiment resulting from changes in Canadian child benefits. These benefits vary across provinces and were reformed at different times. An advantage of their research is that the changes in income were not tied to other changes in family behavior, in contrast to programs like the EITC. They find that an extra $1000 of child benefits leads to an increase of about 0.07 of a standard deviation in the math scores and in the Peabody Picture Vocabulary Test, a standardized test of language ability for four to six year old children. If we think of a change of a third or a half a standard deviation in test scores as a meaningful educational effect, then these results suggest that an increase of as little as $5000 in family income has a meaningful effect. Milligan and Stabile (2008) go beyond Dahl and Lochner by examining effects on other indicators. They find that higher child benefits lower aggression in children and decrease depression scores for mothers. They do not find much

Table 6 Effects of income on birth and early childhood outcomes: evidence from the US and around the world.

Study and data	Study design	Results
The link between AFDC participation and birthweight (Currie and Cole, 1993). Data from NLSY merged with state and county level data. Data on children born between 1979-1988 ($n = 5000$).	IV: using parameters of AFDC, FSP and Medicaid as instruments, controlling for child characteristics and region fixed effects. Separate regressions for poor black and poor white women. Sibling FE: comparing sibs where mother on AFDC during pregnancy to others.	OLS results suggest that AFDC has negative effects. Sibling fixed effects indicate no significant effects. IV results suggest that AFDC during pregnancy causes large increases in mean birth weight among poor whites only. Hence, evidence suggests that AFDC has neutral or positive rather than negative effects. Negative estimates driven by selection.
Does money really matter? estimating impacts of family income on children's achievement with data from random-assignment experiments (Morris et al., 2004). Data from four studies that evaluated 8 welfare and anti-poverty programs with randomized designs: Connecticut's Jobs First, the Los Angeles Jobs First GAIN, the New Brunswick and British Columbia sites of the Canadian Self-Sufficiency Project, and the Atlanta, GA, Grand Rapids, MI, and Riverside, CA sites of the National Evaluation of Welfare to Work Strategies. $n = 18,471$ children aged 2-15 at the time of random assignment.	IV models use random assignment as instruments for income, welfare receipt, and employment to estimate impact of income on children's cognitive achievement. Included dummies for sites in both stages of the regression analysis. Cognitive achievement measured with test scores and parent/teacher reports. Controlled for various baseline background family characteristics.	A \$1000 increase in annual income raises cognitive achievement by 0.06 of SD for children aged 2-5. No statistically significant effects of income on children aged 6-9 or 10-15. A 3-SD increase in the proportion of quarters that welfare is received leads to a 1.5 SD decrease in cognitive achievement among 10-15 year-olds.

(*continued on next page*)

Table 6 (continued)

Study and data	Study design	Results
Who benefits from child benefit? (Blow et al., 2005). Data from UK Family Expenditure Survey 1980-2000. $n = 9811$ two-parent households; $n = 2920$ one-parent households.	Identification from variation in real value of Child Benefit (CB) 1980-2000 due to inflation and policy changes. Calculated anticipated and unanticipated changes in CB payments. Examine the propensity to spend CB on child goods (children's clothing) and adult goods (alcohol, tobacco and adult clothing).	**Spending out of unanticipated vs. anticipated change in CB:** 15x greater for alcohol in 2-parent households; 10x greater for adult clothing in lone-parent households. Results suggest that parents fully insure children against income shocks so that unanticipated changes only affect spending on adult goods.
Effects of EITC on birth and infant health outcomes. (Baker, 2008). US Vital Statistics Natality data file 1989-96. ($n = 781{,}535$). Exclude observations from 1994. Used data from 1997 March CPS to compute proportion of treatment group that is eligible for EITC.	Difference-in-Difference-in-Differences (DDD). Exploit the large expansion of EITC in 1993 (phased in 1994-96) that increased benefits to families with 2 or more children. "Treatment" group = mothers giving birth to 3rd or higher child. Control 1 = mothers giving birth to 1st or 2nd child. Control 2 = mothers giving birth to 1st child. Used mothers with less than High school as a proxy for EITC eligibility. Effect of interest is interaction between <HS⋆treatment⋆after EITC expansion.	Birth weight: Increased by 0.4% for all women Incidence of low birth weight: Decreased by 3.7% for all No statistically significant effects on # prenatal visits or maternal smoking during pregnancy.

Table 6 (continued)

Study and data	Study design	Results
The impact of family income on child achievement: evidence from the earned income tax credit (Dahl and Lochner, 2008). Data from NLSY on children and their mothers for 1988-2000. $n = 4720$ children (2527 mothers).	Identification based on EITC expansions in 1987-1993 and 1993-1997 that increased benefits for low- and middle-income. Used simulated instrumental variables (IV): IV is predicted change in EITC income. Controlled for family background and demographic variables and for time-varying state-specific policies that might affect child outcomes. Examine contemporaneous and lagged income. Also estimate OLS and child fixed effects models.	A $1000 increase in income raises combined math and reading scores by 0.06 SD. Larger gains from contemporaneous income for children aged 5 to 10 than for those aged 11-15. Larger gains for children from disadvantaged families.
Do child tax benefits affect the well-being of children? Evidence from Canadian child benefit expansions (Milligan and Stabile, 2008). Data from survey of labor and income dynamics on families with children for 1999-2004 for benefit simulation. Data from Canadian National Longitudinal Study of Children & Youth for 1994-2005 for children 10 and under ($n = 56{,}000$).	Instrumental variables. Used variation in child benefits across time, provinces, and number of children per family to develop a measure of benefit income as IV for child benefits in regressions of the effect of child benefits on child outcomes. Controlled for province-year fixed effects. Also reduced forms for simulated benefits' effect on child outcomes.	An increase in $ 1000 in simulated benefits leads to: 1.5% reduction in likelihood of learning disability (if mom < HS); 3.6% decline conduct disorder/aggression score 4-10; 4.3% of SD decline maternal depression; 11.6% SD decline in maternal depression (if mom < HS); 1.1% decline in child ever experiencing hunger (if mom < HS) Larger positive impacts on education and physical health for boys than girls; larger positive impacts on mental health for girls. Effects on math and vocabulary scores, behavioral outcomes, maternal depression and likelihood of ever experiencing hunger persist at least 4 years.

(*continued on next page*)

Table 6 (continued)

Study and data	Study design	Results
South African Child Support Grant (CSG): Unconditional cash transfer program with payments made to child's primary caregiver. Intended to cover the country's poorest 30% of children. Study of effects on child nutrition (Agüero et al., 2009). Data from KwaZulu-Natal Income Dynamics Study. Main outcome measure: Height-for-Age z-score (HAZ). $T = 245$, $C = 886$.	Exploit varying lengths of exposure to CSG among children aged 0-3 due to the timing of the rollout of CSG. Fit a quadratic OLS model to application delay (number of days between program creation and child's enrollment) using data for children born ≤ 2 years prior to survey date. Calculated an expected application delay variable for each child conditional on birth date, location and family characteristics, defined as % deviation in application delay from expected delay. Conditional on family characteristics, deviation in delay, and observables, the extent of CSG treatment should be random.Use generalized propensity scores, MLE.	No gains in HAZ for treatments covering $\leq 20\%$ of the 0-3 window. HAZ 0.20 higher for treatment covering $2/3$ of window than for a child receiving treatment covering 1% of window. **Benefit-cost ratio of CSG:** Between 1.06 and 1.48 (estimating lifetime earnings gains from gains in HAZ, using annual 5% discount rate, and assuming unemployed 33% of time). Results robust to checking for age cohort effects and location/spacial effects.

Table 6 (continued)

Study and data	Study design	Results
Conditional Cash Transfer Programs (CCT)		
Mexican "Progresa" program (now "Oportunidades"): Conditional cash transfers of 20-30% of household income every 2 months. Families must take young children to health clinics, immunizations, get adequate prenatal care, & receive nutrition supplement. (Families also to keep older children in school) (Gertler, 2004). Data from survey collected from experimental villages. Sample sizes: 7703 children under age 3 at baseline; 1501 newborns; Height analysis: T = 1049, C = 503; Anemia analysis: T = 1404, C = 608.	Program phased in to 320 treatment villages and 185 control villages randomly selected. Control villages received benefits 2 years after start of program. They did not know this would be the case. Investigate 0/1 treatment dummy as well as different lengths of program exposure.	**Treatment effects:** Newborns 25.3% less likely to be ill in past month. Children 0-3 22.3% less likely to be ill in past month. Height (12-36 mos.) 1.2% higher. Probability of being anemic (12-48 mos.) 25.5% less.

(*continued on next page*)

Table 6 (continued)

Study and data	Study design	Results
"Atención a Crisis" in rural Nicaragua. Conditional cash transfer ∼15% average per capita expenditures. Women receive payments every 2 months if preschool children taken for regular visits to health clinics. (Older children must also be enrolled in school) (Macours et al., 2008). Sample sizes: $T = 3002$, $C = 1019$. Additional data from the 2001 Nicaragua Living Standards Measurement Study.	Randomized experimental design. 4 groups: CCT, CCT + vocational training, CCT + productive investment grant, control. Subgroup analysis by gender and age. Analyzed various transmission mechanisms by estimating expenditures on different foods (to measure nutrient intake), and differences in indicators for early childhood development.	**Treatment effects:** Developmental screener (DDST social-personal): increase by 0.13 SDs. DDST language: increase by 0.17 SDs. Receptive vocabulary (TVIP): increase by 0.22 SDs. (Children made up approx. 1.5 months of delay on TVIP.) **Effects by age:** DDST language: up 0.06 SDs 0-35 mos., up 0.20 SDs 60-83 mos. TVIP: up 0.05 SDs 36-59 mos., up 0.36 SDs 60-83 mos. (Oldest children made up 2.4 months of delay on TVIP) No statistically significant differences by gender. Transmission mechanisms: Better diet; more likely to have books, paper, pencils;more likely to be read to; more likely to have checkup, vitamins, de-worming drugs; more likely to be enrolled; more likely to see doctor when ill.

Unless otherwise noted, only results significant at 5% level are reported.
Percent changes (denoted by % instead of "pp") are reported relative to the mean, if means were reported in the paper.
"T" = Treatment group, "C" = Control group.
Outcomes written as "$T > C$" mean that treatment outcome greater than control outcome at 5% significance level.
Outcomes written as "$T = C$" mean that there is no statistically significant difference in outcomes between 2 groups at 5% level.

impact on physical health measures, though they do find a decrease in families reporting that their children went hungry. There is some evidence of gender differences, with girls showing greater responsiveness to income on the mental health and behavioral scores while boys show greater responsiveness on test scores.

These findings are extremely intriguing, but raise several questions. First, do the effects of income vary depending on the child's age? Morris et al. (2004) argue that income is more important at younger ages, though persistent poverty is worst of all. Second, are there really gender effects in the impact of income, and if so why? Third, the effects that Milligan and Stabile find are roughly twice those found by Dahl and Lochner. Is this because the former study a pure income transfer while the latter study a tied transfer? Fourth, will the effects last, or will they be subject to "fade out" as the children grow older?

Table 6 also includes examples from a growing literature analyzing "conditional cash transfer programs" (CCTs). These are programs that tie transfers to specific behavior on the part of the family. For example, the parents may be required to make sure that the children attend school or get medical care in return for the transfer. These programs have become increasingly popular in developing countries, and have also been implemented to a limited extent in rich countries (for example, there is a program in New York City which is being evaluated by Manpower Development Research Corporation). By their nature, CCTs are complex programs that cannot tell us about the pure impacts of income. Still, these programs have attracted attention because randomized controlled trials have shown at least short-term results. It is difficult however to compare across programs, given that they all tend to focus on different outcomes.

Given this positive evidence about the effects of income, it is a puzzle why so much aid to poor families is transferred in kind. Currie and Gahvari (2008) survey the many reasons for this phenomena that have been offered in the economics literature and conclude that the most likely reasons aid is offered in kind are agency problems, paternalism, and politics. In a nutshell, policy makers and the voters they represent may be more concerned with ensuring that children have medical care than with maximizing their parent's utility, even if the parent's utility is assumed to be affected by the children's access to health care. Politics come in because coherent lobby groups (such as doctors, teachers, or farmers) may have incentives to advocate for various types of in kind programs. In any case, in kind programs are an important feature of aid policies in all Organization for Economic Cooperation and Development states, accounting for over 10% of GDP if health care and educational programs are included. In what follows, we first discuss "near cash programs" and then programs whose benefits are less fungible with cash.

5.2. Near-cash programs

Programs such as the US Food Stamp Program (FSP, now renamed the Supplemental Nutrition Assistance Program, or SNAP) and housing assistance are often referred to

as "near cash" programs because they typically offer households benefits that are worth less than what the household would have spent on food or housing in any case. Hence, canonical microeconomic theory suggests that households should think of them as equivalent to cash and that they should have the same impact as the equivalent cash transfer would have. In the case of food stamps, it has proven difficult to test this prediction because the program parameters are set largely at the national level, so that there is only time series variation.

Currie (2003) provides an overview of the program, and the research on its effects that had been conducted up to that point. Schanzenbach (forthcoming) uses data from a food stamp cash out experiment to examine the effect on food spending. She finds that a minority of households actually received more in food stamps than they would otherwise spend on food. In these constrained households, families did spend more on food than they would have otherwise, while in other households, food stamps had the same effect as cash. Unfortunately, there is little evidence that constrained households bought foods that were likely to have beneficial effects; they seem, for example, to have spent some of the "extra" food money on products such as soda.

Hoynes and Schanzenbach (2009) use variation from the introduction of the FSP to identify its effects on food spending. The FSP began as a small pilot program in 1961, and gradually expanded over the next 13 years: In 1971, national eligibility standards were established, and all states were required to inform eligible households about the program. In 1974, states were required to extend the program statewide if any areas of the state participated. Using data from the PSID, the introduction of the FSP was associated with an 18% increase in food expenditures in the full sample, with somewhat larger effects in the most disadvantaged households. They find that the marginal propensity to consume (MPC) food out of food stamp income was 0.16 compared to 0.09 for cash income. Thus, it does seem that many households were constrained to spend more on food than they otherwise would have (or alternatively, that the person receiving the food stamps had a stronger preference for food than the person controlling cash income in the household). From a policy maker's point of view, this means that the FSP has a bigger impact on food spending than an equivalent cash transfer. Still, it is a leaky bucket if only 16 cents of every dollar transferred goes to food.

Bingley and Walker (2007) conduct an investigation of the Welfare Milk Program in the UK. They identify the effect of the program on household milk expenditures using a large change in eligibility for the program that had differential effects by household type. They find that about 80% of a transfer of free milk is crowded out by reductions in milk purchases by the household. This estimate is quite similar to that of Hoynes and Schanzenbach, though it still suggests that the in kind transfer is having some effect on the composition of spending. Details of these two studies are shown in Table 7.

Given that these programs appear to have some effect on food expenditures, it is reasonable to ask what effect they have on child outcomes. There is a substantial older

Table 7 Impacts of Food Stamps on birth and early childhood outcomes.

Study	Study design	Results
Consumption responses to in-kind transfers: evidence from the introduction of the Food Stamp program (Hoynes and Schanzenbach, 2009). Data from PSID for 1968-78 and from the 1960, 1970, and 1980 decennial censuses. $n = 39{,}623$ person-year obs.	Difference-in-difference using variation in county-level implementation of FSP to estimate impact of FSP on food consumption and labor supply. Controlled for county and year fixed effects as well as state linear time trends. Included trends interacted w/ pre-treatment characteristics and three measures of annual per capita county transfer payments.	Introduction of FSP is associated with: 18% increase in total food expenditures (whole sample); 26-28% increase in total food expenditures for female-headed HHs; 6-13% increase in total food exp. for non-white female-headed HHs. No significant effect on meals out and cash expenditures on food at home. Elasticity of food spending with respect to income = 0.30. MPC for food out of cash income = 0.09 (for whole sample); MPC for food out of cash income = 0.111 (<$25,000 income); MPC for food out of FSP income = 0.16 (for whole sample); MPC for food out of FSP income = 0.238 (<$25,000 income). Decrease in whether the HH head reports any work by 21%.
There's no such thing as a free lunch: Altruistic parents and the response of household food expenditures to nutrition program reforms (Bingley and Walker, 2007). Data from UK Family Expenditure Surveys for 1981-1992. $n = 29, 222$.	Analyzed 3 nutrition programs in the UK: free school lunch for children from poor HHs, free milk to poor HHs w/ pre-school children, and free milk at day care for pre-schoolers in attendance regardless of income. For identification, exploited 1988 reform that ended eligibility for poor HHs with working parents. Difference-in-difference (DD). Also did DD using the fact that free school lunches available only during term time, and summer holidays begin earlier in Scotland.	Free school lunch reduces food expenditure by 15% of the purchase price of the lunch. Free pint of milk reduces milk expenditure by 80%.

(*continued on next page*)

Table 7 (continued)

Study	Study design	Results
Impact of Food Stamp Program (FSP) on birthweight, neonatal mortality, and fertility (Almond et al., 2009). Birth and death micro data from the National Center for Health Statistics merged with county-level data for 1968-77. FSP data from USDA. County characteristics from 1960 City and County Data Book. Data on government transfers and per-capita income from REIS. Participation rates calculated using CPS. $n = 97,785$ whites; 27,274 blacks.	Difference-in-difference, using the fact that FSP was introduced to different counties at different times due to available funding and policy changes. Key policy/treatment variable is the month and year that each county implemented FSP. Estimated the impact of FSP on county-level birth outcomes, using county and time fixed effects. Main outcomes concerned with availability of FSP during 3rd trimester of pregnancy.	Introduction of FSP in 3rd trimester led to: 0.06-0.08% (0.1-0.2%) increase in birth weight for whites (blacks); a 1% (0.7-1.5%) decrease in fraction of low birth weight for whites (blacks). Insignificant impacts of exposure to FSP during earlier trimesters. Results robust to adding county & time fixed-effect and other controls, as well as various time trends to the analysis. Results robust to conducting an event study analysis.
Effects of FSP benefits on weight gained by expectant mothers (Baum, 2008). Data from the NLSY. Limited to low-income black and Hispanic women w/ pregnancy information in the surveys. $n = 1477$ pregnancy-level obs.	Random effects models using Heckman and Singer method to model unobserved heterogeneity. Dependent variable is whether women gain correct amount of weight during pregnancy based on pre-pregnancy BMI. Assume state variation in FSP eligibility rules, and program administration affects FSP takeup but not weight gain. Control for gestation, pre-pregnancy FSP, WIC.	Increasing average monthly FSP benefits from $0 to $100 decreases probability of gaining too little weight by 11.8-13.7%. No effect on probability of gaining too much weight. No statistically significant difference in effects of FSP on weight gain between first-time and non-first-time mothers.

Table 7 (continued)

Study	Study design	Results
Impact of FSP on birth outcomes in California (Currie and Moretti, 2008). Data on FSP from state records and REIS. Data from birth records in CA for 1960-74. Aggregated data into cells defined using county, race, year of birth, maternal age group, parity, and the third of the year. $n = 38{,}475$ cells.	Difference-in-difference using county-level variation in timing of FSP introduction. FSP measured using dummy (= 1 if FSP introduced), log expenditures, or log participation. FSP dummy refers to 9 months prior to birth. County fixed effects and county-specific time trends included. Examined teenage mothers and LA county separately.	FSP introduction led to a 10% increase in number of first births to white teen mothers (only in Los Angeles); a 24% increase in number of first births to black teen mothers; a 12% increase in number of first births to all blacks; a 0.1% increase in probability infant 1500-2000 g survives for whites; a 4% decrease in probability infant <3000 g survives for blacks; a 4% increase in probability of low birth weight among white teens.

Unless otherwise noted, all reported results are statistically significant at 5% level.
Percent changes (denoted by % instead of "pp") are reported relative to the mean.

literature examining this question (see Currie (2003) for a summary). The modal study compares eligible participants to eligible non-participants using a multiple regression model. The main problem with drawing inferences about the efficacy of the FSP from this exercise is that participants are likely to differ from eligible non-participants in ways that are not observed by the researcher. Thus, for example, Basiotis et al. (1998) and Butler and Raymond (1996) both find that participation in the FSP reduces consumption of some important nutrients. Since it is hard to imagine how giving people food coupons could do this, one suspects that these results are driven by negative selection into the FSP program.

Several recent papers examining the effects of the FSP on young children are summarized in Table 7. Currie and Moretti (2008) were the first to try to use variation in the timing of the introduction of the Food Stamp program to look at effects on birth outcomes. Using Vital Statistics Natality data from California, they find that the introduction of the FSP increased the number of births, particularly in Los Angeles County. They also find some evidence that the FSP increased the probability of fetal survival among the lightest white infants, but the effect is very small, and only detectable in Los Angeles. Notably, the FSP increased (rather than decreased) the probability of low birth weight but the estimated effect is small, and concentrated among teenagers giving birth for the first time. Thus, it appears that in California, the FSP increased fertility and infant survival (in some groups) with overall zero or negative effects on the distribution of birth weight.

Almond et al. (forthcoming) examine the same question using national data, and focus on receipt of the FSP during the third trimester, when the fetus typically puts on most of the weight the baby will have at birth. In contrast to Currie and Moretti, they find that the introduction of the FSP increased birth weights for whites and had even larger effects on blacks. The percentage reductions in the incidence of low birth weight were greater than the percentage increases in mean birth weight, suggesting that the FSP had its largest effects at the bottom of the birth weight distribution. Almond et al. find no effect of Food Stamp receipt in the first trimester of pregnancy and much weaker evidence for effects of receipt in the second. This suggests that one reason for the contrast between their results and those of Currie and Moretti is that the latter did not focus narrowly enough on the relevant part of pregnancy. Moreover, Almond et al. find larger effects in the South than in other regions, raising the possibility that overall effects were smaller in California than in other regions. Finally, it is possible that the effects in California are obscured by the substantial in-migration that the state experienced over this period.

Baum (2008) examines the effects of the FSP on weight gain among pregnant women, with particular attention to whether women gained either less than the recommended amount or greater than the recommended amount given their pre-pregnancy body mass index. He estimates a simultaneous equations model in which weight gain and FSP participation are jointly determined. FSP participation is assumed to be affected by

various state-level rules about eligibility, outreach and so on. One difficulty is that these rules may be affected by other characteristics of states (such as overall generosity of social programs) which have direct effects on weight gain (e.g. through superior access to health care during pregnancy). Baum finds that FSP participation reduces the probability that women experienced inadequate weight gain during pregnancy, but has no effect on the probability that they gained too much weight. Since inadequate maternal weight gain is an important risk factor for low birth weight, it is likely that FSP had a positive effect on birth weights among affected mothers.

As discussed above, the other large category of "near cash" offer subsidized housing. Many OECD countries have large housing assistance programs, but their effects on families are understudied. In fact, we were able to find only one paper that examined the effects of housing programs on the outcomes of children less than five, and only a handful that examined effects on children at all. These studies are summarized in Table 8.

Since by design, families receiving housing assistance are among the poorest of the poor, it is clearly important to address the endogeneity of program receipt. Currie and Yelowitz (2000) look at the effects of living in a public housing project in families with two children. They combine information from the Census and from the Survey of Income and Program Participation in a two-sample instrumental variables framework where the instrument for receipt of housing assistance is the sex composition of the siblings (families with a boy and a girl are entitled to larger apartments, and so are more likely to take up housing benefits). They find that families living in projects are less likely to be subject to overcrowding and that the children are much less likely to have been held back in school. The latter effect is three times bigger for boys (who are more likely to be held back in any case) than for girls. Since most "holding back" occurs at younger ages (Kindergarten and grade 1), this suggests that this type of assistance is in fact beneficial for young children.

Goux and Maurin (2005) focus on the effect of overcrowding in France using a similar instrumental variables strategy: They argue that children in families in which the two eldest children are the same sex are more likely to live in crowded conditions in childhood. They also propose an alternative strategy in which crowding is instrumented with whether or not the parent was born in an urban area—parents who are from urban areas are more likely to live in crowded conditions. They find evidence consistent with Currie and Yelowitz in that crowding has a large and significant effect on the probability that a child falls behind in school and eventually drops out.

Fertig and Reingold (2007) examine the effect of receipt of public housing assistance using data from the Fragile Families Study and three instruments: the gender composition of children in the household, the supply of public housing in each location, and the length of waiting lists in each location. They find mixed estimates of effects on maternal health and little evidence of an effect on child health, though their samples are quite

Table 8 Effects of housing and neighborhoods on child outcomes.

Study and data	Study design	Results
Housing and child outcomes		
Are public housing projects good for kids? (Currie and Yelowitz, 2000). Data from SIPP for 1992-1993, March CPS for 1990-1995, and US Census for 1990 on families w/ 2 children (under 18) and income $<$ \$50,000. $n = 279{,}129$.	Two-sample instrumental variable (TSIV) of the effects of living in public housing projects on child's education, housing conditions. Outcome variables are in SIPP and Census. Endogenous regressor is in CPS. Instrument is a dummy for siblings being of different sexes since families with different sex children get larger apartments in public housing than families with same sex children. Controlled for per-capita availability of projects, vouchers, Section 8 subsidies, as well as other neighborhood and family background characteristics.	First stage results: Having siblings of different sex increases likelihood of family living in project by 24%. Second stage results: Families who live in projects are 16 pp (mean = 0.04) less likely to be overcrowded; and 12pp (mean = 0.02) less likely to live in high-density housing. Children who live in projects are 111% less likely to have been held back in school (larger effects for boys than for girls). Black children who live in projects are 19% less likely to have been held back in school. No statistically significant effect for white children.

Table 8 (continued)

Study and data	Study design	Results
The long-term effects of public housing on self-sufficiency (Newman and Harkness, 2002) Data from PSID—Assisted housing database on cohorts born b/n 1957 and 1967. Sample limited to individuals whose families were eligible for public housing b/n ages 10 and 16. $n = 1183$. Adult outcomes measured at ages 20-27.	Amemiya generalized least squares regression where the instrument was the vector of residuals from a regression of the number of public housing units per eligible family in the area on demographic characteristics of the area. Instrumented for whether child lived in public housing to estimate impact on various adult outcomes. Controlled for numerous background characteristics and state and year fixed effects.	Living in public housing as a child leads to an increase in the probability of any employment b/n ages 25 and 27 of 7.8%; an increase in annual earnings b/n ages 20 and 27 of 14.3%; an increase in the number of years not on welfare b/n ages 20 and 27 of 11.3%. No statistically significant effect on household earnings relative to the poverty line. Note: above results significant at 6% level.
Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago (Jacob, 2004) Administrative data from the Chicago Housing Authority and Chicago Public Schools for 1991-2002 on students who lived in high-rise public housing for at least one semester. $n = 10{,}556$.	Instrumental variables regression of the effect of living in a public housing project on student outcomes. Identification from variation in timing of housing demolitions in Chicago in the 1990s. Instrumented for living in public housing with a dummy for whether the student lived in a unit scheduled for demolition at the time of the closure announcement. Controls for background characteristics and project and year fixed effects.	No statistically significant effect of living in high-rise projects on student outcomes. Living in a building that was demolished leads to an 8.2% increase in probability that students>14 yrs drop out of school within 3 years relative to those that lived in buildings scheduled for demolition that had not yet been demolished (12% more likely for girls; 4% more likely for boys).

(*continued on next page*)

Table 8 (continued)

Study and data	Study design	Results
Public housing, health and health behaviors: is there a connection? (Fertig and Reingold, 2007). Data from the fragile families and child well-being study on children born in 20 US cities between 1998 and 2000. Sample 1: $T = 422, C = 2,055$; Sample 2: $T = 323, C = 1,999$; Sample 3: $T = 150, C = 919$.	Instrumental variables regression. Three instruments: gender composition of household, the supply of public housing in each city, and length of waiting list. Three subsamples. "Control" group is mothers who have below 80% of median income in their area but are not in public housing. "Treatment 1" is mothers who live in public housing at time of first interview immediately after childbirth. "Treatment 2" is mothers who have moved into public housing between childbirth and second interview at child's age 1 ("move-in" subsample). "Treatment 3" is the "'move-in" subsample limited to mothers with two or three children at second interview. Controlled for family, background, neighborhood, and other demographic and socio-economic characteristics.	Note: only reporting IV estimates with complete combination of instruments here. **Effects of moving into public housing between childbirth and one-year interview:** 1.63 point increase in mother's health index (5 point scale, mean not reported, can't calculate effect size) 55 pp decrease in likelihood that mother has limiting health condition (mean = 0.08) 41 pp decrease in likelihood that mother is underweight (mean = 0.07) 18% increase in mother's BMI Decrease in domestic violence (imprecise coefficient estimate) No statistically significant effects on child's birth weight or PPVT scores.
Neighborhood characteristics and child outcomes		
The long-run consequences of living in a poor neighborhood (Oreopoulos, 2003) Data from the intergenerational income data base for 1978-1999 (taken from income tax files) on individuals living in Toronto, born b/n 1963 and 1970. $n = 4060$.	Among families who apply for housing projects, assignment to a particular project is approximately random—based on 1st available unit. Compared housing-project means of various adult outcomes across neighborhood quality (measured by density of housing, total size of project, proportion of the Census tract below the low-income cutoff, and whether the project is all high-rises).	The quality of housing project or neighborhood has no statistically significant impact on total income, annual earnings, or number of years on welfare in adulthood. Family characteristics account for up to 30% of total variation in adult outcomes.

Table 8 (continued)

Study and data	Study design	Results
Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment (Kling et al., 2005) Data from Moving to Opportunities (MTO) experiment. Families living in public housing projects in Boston, Los Angeles, New York City, Baltimore, and Chicago were randomly chosen to stay in their current home, to receive a Section 8 housing voucher, or to receive a Section 8 housing voucher restricted to use in neighborhoods w/ poverty rate less than 10%. Surveys conducted in 2002. Data on youths aged 15-25. Data on arrest records for MTO states. $n = 4475$.	OLS regression of crime outcomes on dummy for whether family used a voucher (instrumented by whether family was assigned to treatment group) or on voucher type, and on an age-treatment interaction variable. Instrument necessary since not all families who were assigned a voucher chose to use it - some remained in their initial housing.	Assignment to a restricted voucher leads to a 31% decrease in violent crime arrests for female youths. Use of restricted voucher leads to a 76% reduction in violent crime arrests among female youths. No statistically significant effects on males for violent crime arrests. Assignment to a restricted voucher leads to a 33% decrease in property crime arrests for female youths. Use of restricted voucher leads to a 85% decrease in property crime arrests for female youths. Assignment to a restricted voucher leads to a 32% increase in property crime arrests for male youths. Use of restricted voucher leads to a 77% increase in property crime arrests for male youths.

(*continued on next page*)

Table 8 (continued)

Study and data	Study design	Results
Neighborhoods and Academic Achievement (Sanbonmatsu et al., 2006) Data from MTO experiment surveys for 2002 (see above) on youths aged 6-20. $n = 5074$.	Same regression as Kling et al. (2005) for educational outcomes.	Children in families assigned to a restricted voucher attend better quality schools: the mean rank of schools on state exams is 30% greater than the control group mean, and the mean proportion of school-lunch-eligible children is 8% lower. No statistically significant effects of either voucher on child educational outcomes (school attendance, whether child does homework, child's behavior in class, whether child attended class for gifted students or class for special help, and Woodcock-Johnson Revised reading and math test scores).
Close Neighborhoods Matter: Neighborhood Effects on Early Performance at School (Goux and Maurin, 2005). Data from the French Labor Force Survey for 1991-2002 on 16-year-old youths. $n = 13116$.	Authors assume that a child's birthdate will impact his/her own educational outcomes, but not the outcomes of his/her neighbors. (Proportion of 15-yr-olds held back a grade is 15 pp higher among those born at the end of the year relative to those born at the beginning of the year). IV regression, where first stages are: regression of being held back at age 15 on timing of birthday and regression of proportion of neighborhood youths held back at age 15 on their birthdays. Second stage is regression for being held back at age 16 on having been held back at age 15 and proportion of neighborhood youths held back at age 15. Controlled for year fixed effects, gender, and nationality.	1 SD increase in the proportion of neighboring youth that were held back in school at age 15 increases the probability of being held back at age 16 by 0.2 SD.

Table 8 (continued)

Study and data	Study design	Results
Experimental analysis of neighborhood effects (Kling et al., 2007). Data from MTO experiment surveys for 2002 (see above) on youths aged 15-20. $n = 1807$.	Same regression as (Kling et al., 2005) for health and behavior outcomes.	Assignment to an unrestricted voucher leads to a decrease in likelihood of experiencing anxiety symptoms by 62%/91% among females/males; a decrease in marijuana use by 44% among females; an increase in the likelihood of non-sports injury by 130% among males; an increase in incidence of smoking by 121% among males. Assignment to restricted voucher leads to a decrease in likelihood of experiencing psychological distress by 100% among females; a decrease in likelihood of experiencing anxiety by 57% among females; a decrease in marijuana use by 50% among females; an increase in likelihood of non-sports injury by 140% among males; an increase in incidence of smoking by 82% among males.

Unless otherwise noted, only results significant at 5% level are reported here.

small. Newman and Harkness (2002) use data from the PSID to examine the effect of living in public housing as a child on future earnings and employment. Living in public housing is instrumented using the residual from a regression of local housing supply on the demographic characteristics of the area. They find that public housing is associated with increases in the probability of any employment (from 88% to about 95%) and increases in annual earnings (by $1861 from a mean of $11,210). While all of these instrumental variables strategies are subject to caveats (is gender composition really uncorrelated with sibling's outcomes? Are characteristics of local housing markets associated with unobserved factors such as the quality of schools that might also affect child outcomes?) they certainly all point in a similar direction.

An important question is whether public housing assistance benefits children more than the equivalent cash transfer. It is difficult to answer this question given the available data. However, it is possible to eliminate some possible channels through which public housing programs might have different effects. One is that public housing programs may constrain the recipient's choice of neighborhoods, with either positive or negative effects. Jacob (2004) studies students displaced by demolitions of the most notorious Chicago high-rise projects. The US Congress passed a law in 1996 that required local housing authorities to destroy units if the cost of renovating and maintaining them was greater than the cost of providing a voucher for 20 years. Jacobs argues that the order in which doomed buildings were destroyed was approximately random. For example, in January 1999, the pipes froze in some buildings in the Robert Taylor Homes, which meant that those buildings were demolished before others in the same complex. By comparing children who stayed in buildings scheduled to be demolished to others who had already been displaced by demolitions, he obtains a measure of the effect of living in high-rise public housing. Despite the fact that the high rises in Jacob's study were among the most notorious public housing projects in the country, he finds very little effect of relocation on children's educational outcomes. However, this may be because for the most part, children stayed in the same neighborhoods and in the same schools.

The most exhaustive examination of the effects of giving vouchers to project residents is an ongoing experiment called "Moving to Opportunity" (MTO). MTO was inspired by the Gautreaux program in Chicago, which resulted from a consent decree designed to desegregate Chicago's public housing by relocating some black inner-city residents to white suburbs. MTO is a large-scale social experiment that is being conducted in Chicago, New York, Los Angeles, Boston and Baltimore (see Orr et al. (2003), Kling et al. (2005) and Sanbonmatsu et al. (2006)). Between 1994 and 1998, volunteers from public housing projects were assigned by lottery to one of three groups. The first group received a voucher that could only be used to rent housing in a low-poverty area (a Census tract with a poverty rate less than 10%). This group also received help locating a suitable apartment (referred to here as the "MTO group"). The second group received

a voucher which they could use to rent an apartment in any neighborhood. The third group was the control and received no vouchers or assistance although they were eligible to remain in their project apartment. Families in the first group did move to lower poverty neighborhoods and the new neighborhoods of the MTO group were considerably safer. The move to new neighborhoods had positive effects on the mental health and schooling attainment of girls, and negative effects on the probability that they were ever arrested. But surprisingly, MTO either had no effect, or negative effects, on boys. Boys in the experimental group were 13% more likely than controls to have ever been arrested. This increase was due largely to increases in property crimes. These boys also report more risky behaviors such as drug and alcohol use. And boys in the MTO and voucher groups were more likely to suffer injuries. These differences between boys and girls are apparent even within families (Orr et al., 2003).

It remains to be seen how the long-term outcomes of the MTO children will differ from controls. Oreopoulos (2003) uses data from Canadian income tax records to examine the earnings of adults who lived in public housing projects in Toronto as children. There are large differences between projects in Toronto, both in terms of the density of the projects, and in terms of the poverty of the neighborhoods. Oreopoulos argues that the type of project a family lives in is approximately randomly assigned because the family is offered whatever happens to be available when they get to the top of the waiting list. Oreopoulos finds that once the characteristics of the family are controlled, the neighborhood has no effect on future earnings or on the likelihood that someone works.

The findings on near cash programs can be summarized as follows. There is credible evidence that the FSP may improve birth weight. More work remains to be done to determine whether it has positive effects on the nutrition of children after birth, whether similar programs in other countries have positive effects, and whether this particular type of in kind program has effects that are different than cash subsidies to poor households. The evidence regarding housing programs also suggests that they can be beneficial to families, but offers little guidance about the important question of whether housing programs matter primarily because they subsidize family incomes or operate through some other mechanism. It seems doubtful, given the available evidence, that housing programs benefit child outcomes primarily by improving their neighborhoods (especially since many housing projects are located in less desirable neighborhoods).

5.3. Early intervention programs

Many programs specifically seek to intervene in the lives of poor children in order to improve their outcomes. Three interventions that have been shown to be effective are nurse home visiting programs, nutritional supplementation for pregnant women, and quality early childhood education programs. Table 9 summarizes some recent evidence about home visiting programs.

Table 9 Randomized trials of home visiting programs.

Study/program name	Data, program description, and study design	Results
The Comprehensive Child Development Program (CCDP) (St. Pierre and Layzer, 1999)	Biweekly visits starting 0-1, ending at 5 years. Population served: 43% African-American, 26% Hispanic; all below poverty. Background of Home Visitors: paraprofessionals Sample sizes: T = 2, 213, C = 2, 197 Evaluation Sites: 21 sites throughout the US. Age of children at last follow-up: 5 years old	Developmental Checklist: T = 57.93, C = 57.51 Found no significant effects on wide range of outcomes including: Development and Behavior scores, medical care, mortality, HOME scores, maternal depression, welfare use, maternal income, education or employment, maternal substance use. Total cost per participant: $ 37,488. Total benefit per participant: $91. Net present value = −$37,397.
Healthy start (Duggan et al., 1999, 2004; Harding et al., 2007; DuMont et al., 2006)	Weekly visits, fading to quarterly age of participation: birth to 5 years. population served: low-income, at-risk families of newborns recruited through an HSP screening and assessment protocol. All English-speaking. Background of home visitors: paraprofessionals. Sample sizes: Alaska: T = 179, C = 185; Hawaii: T = 373, C = 270, 6 sites; Virginia: T = 422, C = 197, 2 sites. 19 additional sites discussed in (Harding et al., 2007). Age of children at last follow-up: 2 years old (3 in San Diego)	Some positive effects on parenting practices and negative effects on domestic violence in some sites. E.g. Hawaii partner violence: T = 16%, C = 24%. Less corporal/verbal punishment T < C (odds ratio 0.59). Health effects in some sites but not others, e.g. VA: birth complications T = 0.2, C = 0.48; New York: low birth weight T = 3.3%, C = 8.3%; maternal depression T = 23%, C = 38%. Increases in child Bayley Scale for Infant Development in San Diego, Arkansas. Increases in maternal education and decreases in serious child abuse only in New York (T 12.5% less than C). All sites tested for a wide range of possible effects, with generally insignificant effects on other measures of child well being and child abuse.

Table 9 (continued)

Study/program name	Data, program description, and study design	Results
The nurse-family partnership program (Olds et al., 1999)	Weekly visits, fading to monthly, prenatal to 2 yrs. Population served: disadvantaged first-time mothers less than 30 weeks pregnant (62% unmarried, 47% teenage, 23% poor, unmarried and teenage). Background of home visitors: Nurses. Sample sizes: C1 = 90, C2 = 94, T3 = 100 T4 = 116 (see below for description of groups). Evaluation site: Elmira, New York. C1 = screening; C2 = screening & transportation; T3 = screening, transportation, & prenatal home visits; T4 = screening, transportation, prenatal and postnatal home visits. Prenatal analysis: T = T3 + T4, C = C1 + C2; Postnatal analysis: T = T4, C = C1 + C2. Age of children at last follow-up: 15 years old.	Pre-term births for women who smoked more than 4 cigarettes per day: T = 2.08%, C = 9.81% (mothers also less likely to smoke during pregnancy, better nutrition during pregnancy, more likely to use WIC). For children: fewer emergency room visits at 0-12, 12-24 months. **AT 15-yr follow-up** Mother's number of months receiving AFDC: T4 = 60.4, C = 90.3; Mother's substance use impairments: T4 = 0.41, C = 0.73; Mother's arrests: T4 = 0.18, C = 0.58; Convictions: T4 = 0.06, C = 0.28; Substantiated reports of child abuse and neglect, 0 to 15 yrs: T4 = 0.29, C = 0.54; Child's incidence of arrests: T4 = 0.20, C = 0.45; Child's convictions and probation violations : T4 = 0.09, C = 0.47; Child's number of sex partners: T4 = 0.92, C = 2.48; Child's number of days drank alcohol: T4 = 1.09, C = 2.49. **Cost-benefit analysis:** Total cost per child: $10,300 ; Total benefit per child: $ 30,000. Net present value = +$19,700. Most benefits due to reduced crime on part of the child and reductions in child abuse on part of parent.

(*continued on next page*)

Table 9 (continued)

Study/program name	Data, program description, and study design	Results
The nurse-family partnership program (Olds et al., 2007)	Weekly visits, fading to monthly, prenatal to 2 yrs. Population served: first-time mothers less than 29 weeks pregnant and at least 2 of the following: unmarried, less than 12 yrs of education, or unemployed; 92% African-American, 98% unmarried, 64% 18 yrs of age or younger, 85% at or below poverty level. Background of home visitors: nurses. Sample sizes: C1 = 166, C2 = 515, T3 = 230, T4 = 228 (see below for description of groups). Evaluation sites: Memphis, Tennessee. C1 = transportation; C2 = screening & transportation; T3 = screening, transportation, prenatal home visits, one visit postpartum in hospital, one postpartum visit at home; T4 = T3 plus home visits through child's 2nd birthday. Prenatal analysis: T = T3 + T4, C = C1 + C2; Postnatal analysis: T = T4, C = C2. Age of children at last follow-up: 9 years old.	During first 2 years of child's life: Number of health encounters for injuries/ingestions: T = 0.43, C = 0.56 Number of outpatient visits for injuries/ingestions: T = 0.11, C = 0.20 Number of days hospitalized for injuries/ingestions: T = 0.04, C = 0.18 Mother attempted breast-feeding: T = 26%, C = 16% Subsequent live births: T = 22%, C = 31% **AT 9-yr follow-up** Child GPA (low-resource only): T = 2.68, C = 2.44 Reading and math achievement (low-resource only): T = 44.89, C = 35.72 Mother's # months with current partner: T = 51.89, C = 44.48 Number of months on AFDC/TANF per year: T = 5.21, C = 5.92 Number of months on food stamps per year: T = 6.98, C = 7.80 Maternal mastery: T = 101.03, C = 99.50 No. of months with employed partner: T = 46.04, C = 48.43 No significant effects on ER visits for injuries in first 2 years, health at birth, use of medical care, maternal health.

Table 9 (continued)

Study/program name	Data, program description, and study design	Results
Early start (Fergusson et al., 2006)	Weekly visits for 1st month, then varying age of participation: prenatal to 3 years. Population Served: families recruited through the same screening process as in Hawaii Healthy Start. Background of home visitors: "family support workers" with nursing or social work qualifications. Sample sizes: T = 220, C = 223. 1 site in New Zealand. Age of children at last follow-up: 3 years old.	Average number of doctor's visits 0-36 mo: T > C by 0.24 SD. Percentage of up-to-date well-child checks 0-36 mo: T > C by 0.25 SD. Percentage enrolled w/ dentist 0-36 mo: T > C by 0.20 SD. Percentage attended hospital for accident/injury or accidental poisoning 0-36 mo: T < C by 0.22 SD. Mean duration of early childhood education: T > C by 0.22 SD. Mean number of community service contacts: T > C by 0.31 SD. Positive parenting attitudes score: T > C by 0.26 SD. Non-punitive parenting attitudes score: T > C by 0.22 SD. Overall parenting score: T > C by 0.27 SD. Percentage of parental report of severe physical assault: T < C by 0.26 SD. Child internalizing (negative) behavior score★: T < C by 0.26 SD. Child total negative behavior score★: T < C by 0.24 SD.

(*continued on next page*)

Table 9 (continued)

Study/program name	Data, program description, and study design	Results
Effectiveness of home visitation by public health nurses in prevention of the recurrence of child physical abuse and neglect: a randomized controlled trial (MacMillan et al., 2005)	Weekly visits for first 6 months, then biweekly Age of participation: Entry: 0-13 yrs old, program lasted 3 years. Population served: parents who have a reported episode of physical abuse or neglect in the 3 months prior to joining program. Background of home visitors: public health nurses. Sample sizes: T = 73, C = 66. 1 site in Hamilton, Canada. Note: Control group received standard services from the child protection agency CPA). Treatment group also received the standard services in addition to the home visiting. Age of children at last follow-up: Varied (3-year follow-up).	Recurrence of physical abuse or neglect based on hospital records: T = 24%, C = 11% Also tested, but found no statistically significant effects on: Recurrence of abuse or neglect based on CPA records; HOME score; child's behavioral, anxiety, attention problems, aggression or conduct disorder scores; parenting behavior scores or CAP score.
Economic evaluation of an intensive home visiting programme for vulnerable families: A cost-effectiveness analysis of a public health intervention (McIntosh et al., 2009)	Weekly visits Age of Participation: prenatal to 18 months old. Population Served: Pregnant women identified by community midwives as being at risk for child abuse or neglect. Background of home visitors: paraprofessionals who received program training. Sample Sizes: T = 67, C = 64. 1 site in the UK. Age of Children at Last Follow-up: 1 year old	Maternal sensitivity (CARE Index): T > C by 13%. Infant cooperativeness (CARE Index): T > C by 18%. Likelihood infant placed on child protection register: T > C (1.35 times more likely). Proportion of children removed from home: T = 6%, C = 0%.

Some cost-benefit analysis from Aos et al. (2004), Technical Appendix.
"T" refers to treatment group, "C"; refers to control group. "T = C" means no discernable difference between groups at 5% significance level.

5.3.1. Home visiting

Unlike many social programs, home visiting has been subject to numerous evaluations using randomized control trials. A recent survey appears in Howard and Brooks-Gunn (2009). David Olds and collaborators have developed a particular model for home visiting and conducted randomized controlled trials in a number of settings (Olds et al., 1999, 2007) to evaluate it. Olds' programs focus on families that are at risk because the mother is young, poor, uneducated and/or unmarried, and involve home visits by trained public health nurses from the prenatal period up to two years postpartum. The evaluations have shown many positive effects on maternal behavior, and on child outcomes. As of two years of age, children in the Elmira New York were much less likely to have been seen in a hospital emergency room for unintentional injuries or ingestion of poisonous substances, although this finding was not replicated at other study sites. As of age 15, children of visited mothers were less likely to have been arrested or to have run away from home, had fewer sexual partners, and smoked and drank less. The children were also less likely to have been involved in verified incidents of child maltreatment. This finding is important given the high incidence of maltreatment among US children (and especially among poor children), and the negative outcomes of maltreated children discussed above. There was little evidence of effects on cognition at four years of age (except among children of initially heavy smokers), though one might expect the documented reduction in delinquent behavior among the teens to be associated with improvements in eventual schooling attainment.

In Olds' model, using nurses as home visitors is viewed as key to getting good results. This may be because nurse home visitors are more acceptable to parents than social workers or community workers because families may want medical services. A randomized trial of nurses versus trained paraprofessionals (Olds et al., 2002) suggests that the effects that can be obtained by paraprofessionals are smaller. Also, the Olds programs are strongly targeted at families considered to be at risk and so they do not shed light on the cost-effectiveness of the universal home visiting programs for pregnant women and/or newborns that exist in many countries.

Olds' positive results do not imply that all home visiting programs are equally effective. In fact, Table 9 suggests that the average home visiting program has relatively small effects. They often improve parenting in subtle ways and may result in some improvements in specific health outcomes. However, these may not be sufficient to justify the cost of a large scale program (Aos et al. (2004) offers a cost benefit analysis of several programs). Home visiting programs can be viewed as a type of parenting program—presumably the reason why Olds' home visitors improved outcomes is because they taught mothers to be better parents. Since parents are so important to children, programs that seek to improve parenting practices are perennially popular. Yet studies of these programs suggest that it is remarkably difficult to change parents' behavior and that many attempted interventions are unsuccessful. The most successful parenting programs

are those that combine parent education with some other intervention that parents want, such as visits by nurses (as in Olds case) or child care (Brooks-Gunn and Markham, 2005).

5.3.2. US supplemental feeding program for women, infants, and children (WIC)

A second type of early intervention program that has been extensively studied is the US Supplemental Feeding Program for Women, Infants, and Children (WIC). As its name implies, WIC is a program targeted at pregnant and lactating women, infants, and children up to age 5. Participants receive vouchers that can be redeemed for particular types of food at participating retailers. Participants must generally go to the WIC office to receive the vouchers, and generally receive nutrition education services at that time. Many WIC offices are run out of clinics and may also facilitate access to medical care. Dozens of studies (many of them reviewed in Currie (2003)) have shown that participation in WIC during pregnancy is associated with longer gestations, higher birth weights, and generally healthier infants, and that the effects tend to be largest for children born to the most disadvantaged mothers. Economists have critiqued these studies, on the grounds that there may be unobservable variables that are correlated with WIC participation among eligibles and also with better birth outcomes. Moreover, it may be implausible to expect WIC to have an effect on pre-term birth. A recent Institute of Medicine report on the subject reviewed the evidence and concluded that randomized trials of many different interventions with women at risk of pre-term birth had failed to find effects (Behrman and Butler, 2007). So it might be surprising to find an effect for WIC, when more specific and intensive interventions aimed at preventing pre-term birth have generally failed.

A number of new studies have attempted to deal with various aspects of this critique, as shown in Table 10. Bitler and Currie (2005) look at data from the Pregnancy Risk Monitoring System, which contains very detailed data from new mothers obtained by combining data from birth records and survey data taken from women before and after pregnancy. They directly address the question of selection bias by examining the population of mothers eligible for Medicaid (all of whom are adjunctively eligible for WIC) and asking how participants differ from non-participants along observable dimensions. They find that the WIC women are more disadvantaged than the non-participants along all observables. This finding does not prove that WIC women are also negatively selected in terms of unobservable variables, but it does mean that women who were very negatively selected in terms of education, health, family relationships and so on would have to have other attributes that were systematically correlated with positive outcomes. Like previous studies, Bitler and Currie also find that WIC participation is associated with higher maternal weight gain, longer gestation, and higher birth weight, particularly among women on public assistance, high school dropouts, teen mothers, and single mothers.

Joyce et al. (2004) adopt a similar strategy with regard to selection, and focus on a sample of first births to women who initiated prenatal care in the first four months of

Table 10 Selected studies of special supplemental feeding program for women, infants, and children (WIC).

Study	Study design	Results
Impact on birth outcomes		
Bitler and Currie (2005). Data from Pregnancy Risk Assessment Monitoring System (PRAMS), ($n = 60,731$).	Compared WIC participants and non-participants in the sample of women whose deliveries were paid by Medicaid. Addressed selection bias by comparing a broad range of observable characteristics between eligible WIC participants and non-participants.	WIC mothers are negatively selected into the program relative to all Medicaid recipients. WIC participants are 1.4-1.5 times more likely to have had prenatal care in 1st trimester; 0.7 times as likely to give birth to low birth weight infant; 0.9 times as likely to give birth to infants who are below the 25th percentile of birth weight given gestational age; 0.9 times as likely to have their infant admitted to the Intensive Care Unit. WIC associated with increases in maternal weight gain, gestation, and birth weight. Larger impact for more disadvantaged women (such as those who received public assistance, high school drop-outs, teen mothers, single mothers).

(*continued on next page*)

Table 10 (continued)

Study	Study design	Results
Figlio et al. (2009). Matched data on births in Florida during 1997-2001 with school records of older siblings of the infants to identify whether the older child receives free lunch or reduced-price lunch though the NSLP (those on NSLP are WIC eligible). Marginally eligible: $n = 2530$; Marginally ineligible: $n = 1744$.	Instrumental variables comparing marginally eligible and ineligible WIC women. Marginally ineligible for WIC = families not participating in NLSP during pregnancy, but participating during either the year before or after. Marginally eligible = families that received NSLP during the pregnancy, but did not receive the lunches at least one adjacent year. A federal policy change that increased income reporting requirements for WIC eligibility in September 1999, made it more difficult for eligibles to obtain WIC. The instrument for WIC participation is an interaction between an indicator for the policy change and eligibility.	WIC participation reduces the likelihood of low birth weight by 12.9 pp (imprecise estimate). No statistically significant effect of WIC on gestational age or likelihood of prematurity.
Gueorguieva et al. (2009). Data on mother-infant pairs from birth files on all singleton births in Florida hospitals between 1996 and 2004. Merged with Medicaid eligibility and WIC participation data from the Florida Department of Health. ($n = 369{,}535$).	Adjusted for selection using propensity scores. "Treatment" is percent days on WIC during pregnancy. Main outcome is SGA ("small for gestational age"). Separate analyses for full-term, late pre-term, very pre-term, and extremely pre-term births.	A 10% increase in percent of time during pregnancy in WIC associated with a 2.5% decrease in probability of a full-term and SGA infant; 2.0% decrease in probability of a late pre-term and SGA infant; 3.7% decrease in probability of a very pre-term and SGA infant.

Table 10 (continued)

Study	Study design	Results
(Joyce et al., 2004). Data from birth certificate files in New York City between 1988 and 2001 on women who were on Medicaid and/or WIC during pregnancy. ($n = 35{,}415$ in 1988-1990; $n = 50{,}659$ in 1994-1996; $n = 52{,}608$ in 1999-2001).	Multivariate regression with dummies for WIC participation, interacted with year of birth. To address selection bias, limited analysis to women with first births who initiated prenatal care in first 4 months of pregnancy. Estimated separate models by race, ethnicity, nativity, and parity. Estimated same model comparing twin births.	Among US-born blacks, WIC participants are 2.4 pp less likely to experience a low birth weight birth than non-WIC participants in 1988-1991. No statistically significant effects of WIC participation for foreign-born Hispanic women. In twin analysis, for US-born blacks, WIC participation is associated with a 55 g increase in birth weight adjusted for gestation, and a 3.9 pp decrease in SGA rates. Effects biggest for US-born black women under age 25. (Note, means for subgroups not reported, so effect sizes can't be calculated).
Joyce et al. (2007). Pregnancy Nutrition Surveillance System (PNSS) data ($n = 2{,}870{,}031$). Included all women who were enrolled in WIC during pregnancy and re-enrolled postpartum. Comparison group is women who enrolled in WIC after delivery but were not exposed to WIC during pregnancy.	Multivariate regression with dummies for WIC in each trimester. To deal with selection bias, estimated separate models by race/ethnicity. Also analyzed subgroups whose pre-pregnancy characteristics put women at high risk for anemia, low weight gain, and intrauterine growth retardation. Finally, analyzed subgroup of first-births.	Conditional on gestational age, mean birth weight is 40 g greater among prenatal enrollees than postpartum enrollees. Rates of SGA are 1.7 pp (14%) less, and rates of term low birth weight are 0.7 pp (30%) less. Difference between 1st and 3rd trimester enrollees in mean birth weight is 13.5 g.

(*continued on next page*)

Table 10 (continued)

Study	Study design	Results
Kowaleski-Jones and Duncan (2002). NLSY Mother-Child data. 2000 children 1990-96. 104 sibling pairs, 71 pairs in which one child participated and one didn't.	Sibling fixed effects.	Increase of 7 ounces in mean birthweight. Positive effect on temperament score. No effect on social or motor skills test scores.
(Hoynes et al., 2009). Data on county-level WIC availability in 1971-1975 and 1978-1982 from lists of local agencies that provided WIC services. Data on birth weight from vital statistics records. Data on population of women by county-year from the CANCER-SEER population data. Data on various control variables from 1970 IPUMS. ($n = 18{,}517$ county-year cells)	Difference-in-difference, using variation in the timing of WIC implementation across different counties in 1972-1979, to estimate impact of WIC availability on birth weight. Controlled for three measures of per capita government transfers, an indicator for Food Stamp program availability, other demographics, and county and year fixed effects, and state-year fixed effects. Scaled estimates by WIC participation rates. Also conducted sub-group analysis in county-year-maternal education level cells.	If WIC is available in a county by the third trimester, average birth weight increases by 0.1% (estimate not scaled by participation rate). Among women with low levels of education, WIC increases average birth weight by 10% (estimate scaled by participation rate) and reduces the fraction of births classified as low birth weight by 11% (estimate scaled by participation rate). No effects of WIC on fertility—so results not driven by selection into birth.
Impact on breast feeding and infant feeding practices		
Chatterji et al. (2002). NLSY Mother-Child file. 1282 children born 1991-95. 970 siblings born 1989-95.	IV with WIC state program characteristics as instruments. Sibling fixed effects.	OLS and IV indicate WIC reduces breastfeeding initiation, but no effect on duration. Fixed effect suggests reductions in length breastfeeding.

Table 10 (continued)

Study	Study design	Results
Impact on nutrition and health outcomes of children		
Black et al. (2004) Data from surveys administered by The Children's Sentinel Nutritional Assessment Program in a multi-site study at urban medical centers in 5 states and Washington, DC in 1998-2001 ($n = 5923$: 5395 received WIC assistance, 528 did not).	Compared WIC-eligible families who participated in WIC with those who did not due to self-reported access problems. Multivariate regression including a participation dummy, controlling for relevant background characteristics.	Compared to infants who received WIC assistance, those who did not receive WIC assistance were more likely to be underweight (weight-for-age z-score $= -0.23$ vs. 0.009), short (length-for-age z-score $= -0.23$ vs. -0.02), and perceived as having fair or poor health (adjusted odds ratio $= 1.92$). No statistically significant differences in rates of food insecurity.
Lee and Mackey-Bilaver (2007). Data from IL Integrated Database on Children's Services. Includes info on Medicaid, FSP, WIC enrollment. All IL children born between 1990 and 1996 who entered Medicaid within first month. Tracked to 2001. Total sample = 252,246. Sibling FE sample = 36,277.	Multivariate regression with sibling fixed-effects, using participation dummies for FSP, WIC, and FSP-WIC jointly. (Note, sibling FE only make sense for WIC, effects of FSP estimated using OLS. Also, most families on WIC also receive FSP).	Effect of WIC: Abuse or neglect rates decrease by 84% (mean = 0.10). Incidence of anemia decrease by 41% (mean = 0.195). Failure to thrive decrease by 78% (mean = 0.128). Nutritional deficiency decrease by 115% (mean = 0.038).

Unless otherwise noted, only results significant at 5% level are reported here.
Percent changes (denoted by % instead of "pp") are reported relative to the mean.

pregnancy in order to ensure that participants and non-participants were more likely to be similar in terms of unobservables. In their sample of women giving birth in New York City, they find positive effects of WIC among US born black women, but not in other groups. Joyce et al. (2007) use a national sample of women, compare women who enrolled in WIC pre and post delivery, and focus on whether infants are small for gestational age (SGA). If one does not believe that WIC can affect gestation, then focusing on SGA is appropriate because it is not affected by gestational age. They find that the incidence of SGA is lower for the prenatal enrollees than for the postpartum enrollees. Gueorguieva et al. (2009) use a large sample of births from Florida and try to deal with potential selection using propensity score matching. They side step the issue of whether WIC affects gestation by presenting separate analyses for pregnancies of different length, and focusing on SGA. They find that longer participation in WIC is associated with reductions in the incidence of SGA. Kowaleski-Jones and Duncan (2002) examine sibling pairs from the NLSY and find that WIC participation is associated with an increase of seven ounces in birth weight. However, the number of pairs in which one child participated and one did not is quite small, so it would be useful to try to replicate this finding in a larger sample of siblings.

Figlio et al. (2009) present an innovative instrumental variables strategy using a large sample of births from Florida that have been merged to school records of older siblings. While the characteristics of WIC programs vary across states, they do not show a lot of variation over time, and previous analyses have demonstrated that these characteristics are weak instruments (Bitler and Currie, 2005). Figlio et al. (2009) first try to select participant and non-participant groups who are very similar. They do this by defining "marginally ineligible" families as those who participated in the National School Lunch Program (NLSP) in the year before or after the birth, but did not participate in the birth year. Thus, the study focuses on families whose incomes hover around the eligibility threshold for NSLP, which is the same as the eligibility threshold for WIC. The instrument is a change in income reporting requirements for WIC in Sept. 1999 which made it more difficult for eligible families to receive benefits. Figlio et al. (2009) find that WIC participation reduces the probability of low birth weight, but find no significant effect on gestational age or prematurity.

There has been much less study of the effects of WIC on other outcomes, or other groups of participants. A couple of studies that make some attempt to deal with the selection issue are summarized in Table 10. One problem with WIC is that it subsidizes baby formula, which is likely to discourage breast-feeding. Chatterji and Brooks-Gunn (2004) use the NLSY Mother-Child file and estimate both sibling fixed effects models and instrumental variables models using characteristics of state programs as instruments. They find that WIC reduces breast feeding initiation and the length of breastfeeding. However, these results are subject to the caveats above (i.e. small samples and possibly weak instruments). Turning to the effects of WIC on older children, Black

et al. (2004) compare WIC eligible participants and those who did not participate due to "access problems". These problems were assessed based on the families own reports about why they were not participating. They found that infants who received WIC were less likely to be underweight, short, or perceived by their parents to be in fair or poor health. Lee and Mackey-Bilaver (2007) use a large data base from Illinois that integrates administrative data from several sources. Using sibling fixed effects models, they find that siblings who received WIC were less likely to be anemic, to have exhibited failure-to-thrive, or other nutritional deficiencies, and that the infants were less likely to be abused or neglected. As discussed above, one issue in the interpretation of these findings is why one infant would receive WIC while the other did not?

In one of the most interesting recent studies, Hoynes et al. (2009) use the initial roll-out of the WIC program in the 1970s to identify its effects. They find that the implementation of WIC increased average birth weight by 10% and decreased the fraction of low birth weight births. They did not find any evidence of changes in fertility.

In summary, the latest group of studies of WIC during pregnancy largely support the findings of earlier studies which consistently found beneficial effects on infant health. The finding is remarkable because WIC benefits are relatively modest (often amounting to about $40 per month) and Americans are generally well fed (if not overfed at least in terms of total calories). Research that attempted to peer into the "black box" and shed light on why the program is effective would be extremely interesting. Another question that cries out for future research is whether WIC benefits infants and children (i.e. children who participate after birth)? While a few studies suggest that it does, the effects of WIC in this population has been subject to much less scrutiny than the effects on newborns.

5.3.3. Child care

There have been many evaluations of early intervention programs delivered through the provision of child care. One reason for focusing on early intervention through the provision of quality child care is that the majority of young children are likely to be placed in some form of care. In 2008, 64% of women with children under 3 worked for pay (US Bureau of Labor Statistics, 2009). While the US may be an outlier in this respect, labor force participation among women with children is high and rising in many other economies. Blau and Currie (2006) provide an overview of the literature on early intervention through child care. Many studies concern experimental evaluations of model programs that serve relatively small numbers of children and involve intensive services delivered by well-trained and well-supervised staff. These studies generally find that early intervention has long-lasting effects on schooling attainment and other outcomes such as teen pregnancy and crime, even if it does not result in any lasting increase in cognitive test scores. These results point to the tremendous importance of "non-cognitive skills" (*cf* Heckman and Rubinstein (2001)) or alternatively, to the importance of mental as well as physical health in the production of good child outcomes (Currie and Stabile, 2006).

A few of the most notable model programs are summarized in Table 11. Two studies of "model" early intervention child care programs stand out because they randomly assigned children to treatment and control group, had low dropout rates, and followed children over many years. They are the Carolina Abecedarian Project and the Perry Preschool Project. Both found positive effects on schooling. A recent cost-benefit analysis of the Abcedarian data through age 21 found that each dollar spent on Abecedarian saved tax payers four dollars. And by focusing only on cost savings, this calculation does not even include the value of higher achievement to the individual children and society (Masse and Barnett, 2002). Each dollar spent on Perry Preschool has been estimated to have saved up to seven dollars in social costs (Karoly et al., 1998), although this high benefit-cost ratio is driven largely by the effect of the intervention on crime, which in turn depends on a handful of individuals.

Anderson (2008) conducts a re-analysis of the Perry Preschool and Abcedarian data (and a third intervention called the Early Training Project) and finds that like the MTO public housing experiment, the significant effects of the intervention were largely concentrated among girls. In addition to analyzing the data by gender, Anderson pays careful attention to the idea that there may be a reporting bias in the published studies of early intervention experiments; that is, researchers who found largely null effects of the experiment might still be able to publish results focusing on one or two positive outcomes out of many outcomes investigated. Conversely, if all effects tended in the same direction, but there was insufficient power to detect significant effects on each outcome, it might be possible to detect a significant effect on an index of the outcomes. Anderson finds positive effects (for girls) on a summary index of effects, and the effects are quite large at about a half a standard deviation. This study highlights an interesting question, which is whether it is generally easier to intervene with girls than with boys, and why that might be the case?

The fact that special interventions like Perry Preschool or Abcedarian had an effect on at least some target children does not prove that the types of programs typically available to poor inner-city children will do so. Head Start is a preschool program for disadvantaged 3, 4, and 5 year olds which currently serves about 800,000 children each year. It is funded as a federal-local matching grant program and over time, federal funding has increased from $96 million when the program began in 1965 to about $7 billion in 2009 (plus additional "stimulus" funds). Head Start is not of the same quality as the model interventions, and the quality varies from center to center. But Head Start centers have historically been of higher average quality than other preschool programs available to low income people. This is because, in contrast to the private child care market, there are few very low-quality Head Start programs (see Blau and Currie (2006) for an overview of preschool quality issues).

An experimental evaluation of Head Start has recently been conducted (Puma et al., 2010). The evaluation compares Head Start children to peers who may or may not be in

Table 11 Selected recent evaluations of early childhood programs with randomized designs.

Study/program name[a]	Data, program description, and study design	Results
Carolina Abecedarian follow-up and cost-benefit analysis at 21 years of age (Barnett and Masse, 2007)	Preschoolers: full-day child care School age: parent program Sample sizes: Initial: T = 57, C = 54 Age 8: T = 48, C = 42 Age 15: T = 48, C = 44 Age 21: T = 53, C = 51 Age of participation in program: Entry: 6 weeks to 3 months old Exit: 5 to 8 years	**Follow-up at 21 years of age:** Grade retention: T = 34%, C = 65%, age 21 Special education: T = 31%, C = 49%, age 21 High school dropout: T = 33%, C = 49%, age 21 College attendance: T = 36%, C = 13%, age 21 Crime rate: T = C, age 21 Employment status: T = C at age 21 Average age first child born: T > C at age 21 **Cost-Benefit Analysis:** (using 5% discount rate, $2002) Net cost per child = $34,599 Net benefit of program = $72,591 per participant
Infant Health and Development Project Follow-up at 18 years of age (McCormick et al., 2006)	Home visits, full-day child care Sample sizes: Initial: T = 377, C = 608 Followup at age 8: T = 336, C = 538 Followup at age 18: T = 254, C = 381 (divided in 2 groups: lighter low birth weight (LLBW) and heavier low birth weight (HLBW)) Age of participation in program: Entry: birth (home visits), 1 year (care). Exit: 3 years	Math achievement: T > C by 6.8%, age 18 HLBW Reading achievement: T > C by 5.6%, age 18 HLBW Risky behaviors: T > C by 23.3%, age 18 HLBW IQ: T = C, age 18 HLBW Note: For all outcomes: T = C, age 18 LLBW

(*continued on next page*)

Table 11 (continued)

Study/program name[a]	Data, program description, and study design	Results
A reevaluation of early childhood intervention—Abecedarian, Perry Preschool and Early Training Project— with emphasis on gender differences and multiple inference (Anderson, 2008)	Abecedarian: T = 57, C = 54 Perry: T = 58, C = 65 ETP: T = 44, C = 21 Ages of entry: Abecedarian/Perry/ETP: 4.4 mo./3 yrs./3-4 yrs	Outcomes include: IQ, grade repetition, special ed., high school, college attendance, employment, earnings, receipt transfers, arrests, convictions, drug use, teen pregnancy, marriage. Summary index pools multiple outcomes for a single test. Separate tests by gender. Effects on summary index for girls 5-12: ABC/Perry: increase by 0.45/0.54 SDs. Effects on summary index for girls 13-19: ABC/Perry/ETP: increase by 0.42/0.61/0.46 SDs. Effects on summary index for women over 21-40: ABC/Perry: increase by 0.45/0.36 SDs. No statistically significant effects on males of any age.
High/Scope Perry Preschool project follow-up and cost-benefit analysis at 40 years of age (Barnett et al., 2006)	Home visits, Preschool program Sample sizes: Initial: T = 58, C = 65 Age 40: T = 56, C = 63 Age of participation in program: Entry: 3 to 4 years, Exit: 5 years	Arrests: T = 32%, C = 48%; In jail: T = 6%, C = 17%; Report of stopping work for health reasons: T = 43%, C = 55%; Hard drug use: T = 22%, C = 29%; Abortions: T = 17%, C = 32% **Cost-benefit analysis:** Main result: Benefit of $12.90 for each $1 cost Most benefits due to reduced crime rates for males Cost: $15,827 ($2000) per student Total Net Private Benefit = $17,730 per participant Total Net Public Benefit = $180,455 per participant

Table 11 (continued)

Study/program name[a]	Data, program description, and study design	Results
National evaluation of Early Head Start (Administration on Children, Youth and Families, 2002 & Love et al., 2005) Cost-Benefit (Aos et al., 2004).	17 Early Head Start sites selected to reflect EHS programs funded in 1995-96. Random assignment within site (possible given wait lists). Sample: T=1513, C=1488. Age of participation in program: Entry at 0-1 year, exit 3 years	Mental Development Index (MDI): T > C by 0.12 SD PPVT-III Vocabulary score: T > C by 0.13 SD Percentage with PPVT score <85 pts: T < C by 0.12 SD Aggressive behavior: T < C by 0.11 SD Supportiveness during parent-child play: T > C by 0.15 SD HOME score: T > C by 0.11 SD Index of severity of discipline: T < C by 0.11 SD No statistically significant effects on parental mental or physical health or on measures of family functioning. **Cost-benefit analysis:** Total Cost per child: $20,972 Total Benefit per child: $4768, NPV: −$16,203
Head Start impact study (US Department of Health and Human Services, 2010)	Congressionally-mandated study of Head Start. Children from wait lists randomly assigned to one of 383 randomly selected Head Start centers across 23 different states. Baseline data collected in fall 2002; annual spring follow-ups through spring 2006. Sample Sizes: T = 2783, C = 1884. Entry: 3-4 years old; Exit: 4-5 years old	Summary of effects for 4-year-old entry cohort at end of 1st grade: PPVT: T > C by 0.09 SD; Withdrawn behavior: T < C by 0.13 SD; Shy/socially reticent: T > C by 0.19 SD; Problems with teacher interaction: T > C by 0.13 SD. No statistically significant impacts at age 4, kindergarten, or 1st grade on: math scores, Spanish language tests, oral comprehension, and several parent- and teacher-reported measures of emotional and behavioral outcomes. No statistically significant impacts at kindergarten or 1st grade on: school accomplishments, promotion, language and literacy ability, math ability, and social studies and science ability. Summary of effects for 3-year-old entry cohort as of 1st grade (selected results): Oral comprehension: T > C by 0.08 SD. No significant effect on other outcomes.

(*continued on next page*)

Table 11 (continued)

Study/program name[a]	Data, program description, and study design	Results
The rates of return to the High/Scope Perry Preschool Program (Heckman et al., 2009)	See Barnett et al. (2006) entry for sample sizes and information about the Perry Preschool Program. Randomization was compromised because of reassignment after initial randomization. Standard errors on the rate of return estimates adjusted for failure of randomization using bootstrapping and Monte Carlo methods. Also, adjusted for the deadweight loss due to taxation (assuming 0%, 50%, and 100% deadweight losses). Used a wide variety of methods for within-sample imputation of missing earnings. Used local data on costs of education, welfare participation, and crime instead of national data, wherever possible. Used several methods to extrapolate benefits and costs beyond age 40 (after last follow-up). Used several measures of the statistical cost of life to estimate costs of murder.	Estimated social rates of return to the Perry Preschool Program are 7%–10%. Estimated benefit-cost ratio = 2.2 to 31.5 (depends on discount rate used, and the measure of cost of murder).

Throughout the table, "T" refers to treatment group and "C" refers to control or comparison group.
Outcomes listed as T > C or C > T were statistically significant at the 5% level.
[a] Programs are grouped such that those enrolling children younger than three years old appear first, followed by those enrolling children after age three.

some other form of preschool (including state-funded preschools modeled in Head Start). In fact, the majority of children who did not attend Head Start did end up attending some other preschool program. Even relative to this baseline, initial results show that Head Start children make gains, particularly in terms of language ability. But children are followed only into the first grade, and so this evaluation did not address the important issue of whether Head Start has longer term effects. This example illustrates one of the limitations of experiments for the study of longer-term effects, which is that one may have to wait a long time for evidence to accumulate. There has also been a federal evaluation of Early Head Start (EHS), a version of the program geared to infants and toddlers under three years old. As Table 11 shows, EHS has small positive effects on cognitive test scores and some measures of behavior though Aos et al. (2004) conclude that it does not pass a cost-benefit test.

Table 12 summarizes notable non-experimental evaluations of Head Start and other public preschool programs. In a series of studies Currie and Thomas use national publicly-available survey data to try to measure the effect of Head Start. In most of these studies, they compare the outcomes of children who attended Head Start to siblings who did not attend. As discussed above, sibling fixed effects control for many shared characteristics of children, but are not a panacea. However, careful examination of differences between participant and non-participant children within families suggested that the Head Start sibling typically attends when the family is relatively disadvantaged. For example, a young single mother might have her first child attend Head Start. If she then marries, her next child will enjoy higher income and be ineligible for Head Start. Currie and Thomas found no within-family differences in birth weight or other individual characteristics of the children. They also investigated spillover effects, which as discussed above, can bias the estimated effect of Head Start. They found some evidence (Garces et al., 2002) that having an older sibling attend Head Start had positive effects on younger siblings. In all, it seems likely that sibling fixed effects models understate the true effect of Head Start.

Nevertheless, they found significant positive effects of Head Start on educational attainments among white youths, and reductions in the probabilities of being booked or charged with a crime among black youths (Garces et al., 2002). Test score gains for blacks and whites were initially the same, but these gains tended to fade out more quickly for black than white students, perhaps because black former Head Start students typically attend worse schools than other students (Currie and Thomas, 1995). Effects were especially large for Hispanic children (Garces et al., 2002).

More recently, Deming (2009) replicates the results of Currie and Thomas (1995) using the same cohort of NLSY children observed at older ages. Like Anderson, he focuses on an index of outcomes (although he also reports results for separate outcomes) and finds that Head Start results in an increase of 0.23 standard deviations, which is equivalent to about 1/3 of the gap between Head Start and other children. He notes

Table 12 Selected studies of large-scale public early childhood programs.

Study/program name and data	Study design	Results
Evaluations of Head Start		
Does Head Start make a difference? Does Head Start help Hispanic children? (Currie and Thomas, 1995, 1999a). NLSCM. Sample size: T = 896, C = 911 Hispanic study: T = 182, C = 568 Entry: 3 to 5 years; Exit: 5 to 6 years.	Estimate sibling fixed effects models of effects of Head Start and other preschool attendance on various outcomes. Examine differences between siblings that might potentially explain differential attendance by siblings.	**Achievement tests:**T > C (1/3 SD whites and Hispanics only) Grade retention: T < C (∼50% whites and Hispanics only) Immunization rates: T > C (8%-11%) Child height-for-age: T = C
Long term effects of Head Start (Garces et al., 2002). PSID, Sample size: T = 583, C = 3502. Entry: 3 to 4 years; Exit 5 to 6 years.	Compared Head Start participants to their own siblings who did not participate. Outcomes measured between ages 18 and 31. Retrospective reports on Head Start participation.	High school graduation: T > C (∼25% for whites only) Arrests T < C (∼50% for African-Americans only) College T > C (∼25% for Whites) Teen pregnancy T = C Welfare T = C
Effect of Head Start on health and schooling (Ludwig and Miller, 2005). Vital Statistics Compressed Mortality Files 1973-83; Individual data from NELS, where T = 649, C = 674.	Regression discontinuity around cutoff at which counties were eligible for assistance in applying for Head Start in 1965. T = 300 poorest counties in 1965, C = 301-600th poorest counties. 80% of treatment counties received funding vs. 43% of all counties nationwide.	Effects of participation in Head Start: Mortality, age 5-9: T < C by 35%-79% High school completion rates: T > C by 5.2%-8.5% Some college+: 16.2%-22.4% for oldest cohort only

Table 12 (continued)

Study/program name and data	Study design	Results
Head Start Participation and Childhood Obesity (Frisvold, 2006). PSID Child Development Supplement. Sample size = 1332.	Estimated the effect of Head Start on likelihood of a child being overweight or obese. Assume that # of spaces available in a community is a valid instrument for Head Start participation.	Head Start reduces probability of obesity at ages 5-10 among blacks. No effect in full sample of children or in children over 10. Estimates are large relative to sample means implying ∼100% reductions in overweight/obesity.
Evidence from Head Start On Lifecycle Skill Development (Deming, 2009). Data from NLSY Children for cohort enrolled in Head Start between 1984 and 1990. Children in study at least 5 years old in 1990. Sample size: 3415 total.	Sibling fixed effects estimates of benefits of Head Start.	Test scores: **T** > **C** by 0.145 SD ages 5-6, by 0.133 SD ages 7-10, by 0.055 SD ages 11-14. Noncognitive school-age outcomes index: **T** > **C** by 0.265 SD. Long-term effect on index⋆ of young adult outcomes: **T** > **C** by 0.228 SD. Large fade-out in test scores of African Americans, none for whites or Hispanics. No effects on criminal activity. Summary: Head Start increases index of long term outcomes by 0.23 SD (∼1/3 of gap attendees and others). Projecting wages implies that benefits (∼$1500 in greater earnings per year) exceed program costs of ∼$6000. ⋆ index includes: graduate high school, complete 1 yr college idle (no job, not in school), poor health.

(*continued on next page*)

Table 12 (continued)

Study/program name and data	Study design	Results
Preventing behavior problems in childhood and adolescence: Evidence from Head Start (Carneiro and Ginja, 2008). NLSY Children data. Sample size = 1786 males. Behavior problems, grade rep. and obesity at 12-13. Depression, crime, and obesity at 16-17. Oldest in data—born in 1974; youngest in data—born in 1992.	Head Start eligibility rules create discontinuities in income eligibility. Compare families above and below the cutoff. Identification strategy requires that families are not able to strategically locate above or below cutoff.	**USING REDUCED FORM ESTIMATION** **Head Start participation impacts on 12-13 year-old males:** behavioral problems index decreases by 38% probability of grade retention decreases by 33.3% probability of obesity decreases by ∼100% for blacks only **Head Start participation impacts on 16-17 year-old males:** Depression (CESD) decreases by 23.4% probability of obesity decreases by 57.9% probability of being sentenced decreases by >100% for blacks only **USING STRUCTURAL EQUATIONS** **Head Start participation impacts on 12-13 year-old males:** probability of grade retention decreases by >100% **Head Start participation impacts on 16-17 year-old males:** probability of being sentenced decreases by >100% probability of obesity decreases by >100% probability of being sentenced decreases by >100% for blacks only probability of obesity decreases by >100% for blacks only Note: baseline means are for sample of children with incomes between 5% and 195% of Head Start eligibility cut-off.

Table 12 (continued)

Study/program name and data	Study design	Results
Investing in health: the long-term impact of Head Start on smoking (Anderson et al., 2009). Data from the PSID. Used smoking data from 1999 and 2003 on individuals aged 21-36 in 1999. $n = 922$ in 1999; $n = 1005$ in 2003.	Compared smoking of siblings who did and did not attend Head Start or any preschool using sibling fixed effects. Controlled for family background characteristics specific to the age children were eligible for Head Start. Examined sibling differences that might be predictive of Head Start attendance. Examined spillover effects by including interactions between Head Start and birth order.	**Results from 1999 data:** Head Start participants are 58% less likely to smoke than siblings. **Results from 2003 data:** Head Start participants are 65% less likely to smoke than siblings. Including control for educational attainment makes results statistically insignificant. **Cost-Benefit:** PV reduction in smoking is $9967 per participant (using 3% discount rate, accounting for medical expenses and productivity losses) Average cost per Head Start participant in 2003 is $7092. Depending on discount rate used, the value of reduction in smoking is associated with 36-141% of program costs.
Expanding exposure: can increasing the daily exposure to Head Start reduce childhood obesity? (Frisvold and Lumeng, 2009) Administrative data from a Michigan Head Start for 2002-2006. $n = 1833$ obs. (from 1532 children, since some attend for multiple years) Full-day class sample = 424 obs. Half-day class sample = 1409 obs.	Estimated the effect of full-day vs. half-day Head Start on obesity at end of school year. Identification via elimination of a grant which led to a decrease in the # of full- day classes from 16 classrooms in 2002 to 4 classrooms in 2003. (IV = % full-day funded slots). Controls for observable family characteristics.	**First Stage Results:** 10 pp increase in percentage of full-day slots increases likelihood of full-day attendance by 85% (relative to baseline = 11% enrollment in full-day slots in 2003). **Second Stage Results:** Full-day enrollment in Head Start decreases likelihood of obesity by 143%. This implies that children who attended full-day classes would have been almost 3 lbs heavier had they attended half-day classes. Simulation results suggest that the 143% change in the likelihood of obesity can be explained by a change in caloric intake of 75 calories per day with no change in physical activity.

(*continued on next page*)

Table 12 (continued)

Study/program name and data	Study design	Results
Studies of public pre-K/K and child care programs		
Impact of early childhood care and education on children's preschool cognitive development: Canadian results from a large quasi-experiment (Lefebvre et al., 2006) Data from Canada's NLSCY on 4- and 5-yr old children from 5 consecutive cycles. $n = 15,546$	Identification from changes in Quebec's child care subsidies. On Sep. 1, 1997, child care facilities offered $5-per-full-day services to children who were 4 yrs old by Sep. 30th. In the following years, age cutoffs decreased and the number of spaces increased. No similar policies in other provinces of Canada in 1994-2003. Difference-in-difference (DD) and Difference-in-difference-in-difference (DDD) designs, comparing Quebec's preschool children with children of similar ages from other provinces using the fact that different cohorts of children were exposed to different numbers of treatment years.	Subsidies increase the number of hours in child care by 5.5-7.4 hours per week for children aged 4-5. Effect larger for mothers in highest educational group. No effects on 4-year-old children's PPVT scores. Decrease in PPVT scores of 0.33 SD for 5-year-old children.

Table 12 (continued)

Study/program name and data	Study design	Results
Promoting school readiness in Oklahoma: An evaluation of Tulsa's Pre-K Program (Gormley and Gayer, 2006)[a]. Data from Tulsa Public Schools (TPS) on test scores of Pre-K and Kindergarten children from test administered in Aug. 2001. T = 1112, C = 1284 (T = children who just completed Pre-K, C = children who are about to begin Pre-K). Entry: 4 years old; Exit: 5 years old.	Regression discontinuity design arising from cutoff of Sept. 1 to enroll in Pre-K in a given year. Compare kindergarten children who just completed Pre-K with slightly younger children who were ineligible to attend. Used quadratic polynomial to fit underlying age/test score relationship.	Cognitive/knowledge score: T > C by 0.39 SD Motor skills score: T > C by 0.24 SD Language score: T > C by 0.38 SD Largest impacts for Hispanics, followed by blacks, little impact for whites. Children who qualify for free school lunch have larger impacts than other children.
Does Prekindergarten Improve School Preparation and Performance? (Magnuson et al., 2007) Data from ECLS-K. $n = 10,224$.	Primary method is a multivariate regression to estimate the impact of Pre-K attendance on various outcomes. Robustness checks using teacher fixed effects, propensity score matching, and instrumental variables (IV). IV is different measures of access to pre-K in a given state. Dependent variables are measured in the fall of kindergarten and in the spring of first grade to assess any lasting impacts of Pre-K.	Pre-K attendance: increases reading scores at school entry by 0.86 SD (IV); increases aggression at school entry by 0.69 SD (IV). No effect on math scores or self control in IV. Effect sizes for all outcomes are larger for Pre-K than for other forms of child care, but Pre-K children have different characteristics than other children. Among children attending Pre-K in the same public school as their kindergarten, higher reading scores are not accompanied by increased behavioral problems. For disadvantaged children, cognitive gains are more lasting than in the whole sample. Effect sizes for cognitive outcomes much lower in spring of 1st grade than at school entry. Effect sizes for behavioral outcomes are the same.

(*continued on next page*)

Table 12 (continued)

Study/program name and data	Study design	Results
The Effects of Oklahoma's Pre-K Program on Hispanic Children (Gormley, 2008) Tests administered by Tulsa public schools in Aug. 2006. T = 194, C = 295. (T = children who just completed Pre-K, C = children who are about to begin Pre-K). Entry: 4 years old; Exit: 5 years old.	See Gormley and Gayer (2006) entry.	Letter-Word Identification Test score: T > C by 0.846 SD Spelling score: T > C by 0.52 SD Applied Problems score: T > C by 0.38 SD Significant effects only for Hispanic students whose primary language at home is Spanish.
Universal Child Care, Maternal Labor Supply, and Family Well-Being (Baker et al., 2008). Canadian NLSCY (1994-2003) includes only married women and their children. Average of 2000 children at each age per yr. Primary sample ages 0-4, robustness checks ages 8-11.	Compare outcomes in Quebec, which began $5 per day daycare for 4 year olds in 1997, extended program to 3 year olds in 1998, 2 year olds in 1999, and all children <2 in 2000, to the rest of Canada. Difference in differences.	Increase in use of any child care/institutional care/mothers working by 35%/>100%/14.5%. Crowding out of other informal child care. Increase in emotional disorder anxiety score (physical aggression and opposition) by 12% (9%) for 2-3 yr olds. Decrease in standardized motor and social development score by 1.7%. Increase in mother depression score by 10%. 40% of the cost of the child care subsidy is offset by increased taxes on extra labor supply.

Table 12 (continued)

Study/program name and data	Study design	Results
Impacts of New Mexico Pre-K on Children's School Readiness at Kindergarten Entry: Results from the Second Year of a Growing Initiative (Hustedt et al., 2008). Data on children who participated in the 2nd year of the Pre-K program during 2006-2007 and entered kindergarten in Fall 2007. T = 405, C = 519 (T = children who just completed Pre-K, C = children who are about to begin Pre-K). Entry: 4 years old; Exit: 5 years old.	Regression discontinuity design due to a birthday eligibility cut-off of Aug. 31st to enroll in Pre-K in a given year. Compared "young" kindergarten children who just completed Pre-K with slightly younger children who are about to being Pre-K. Used linear model for vocabulary score as dependent variable, quadratic model for early literacy score as dependent variable, cubic model for math score as dependent variable.	Vocabulary (PPVT) score: T > C by 0.25 SD Math score: T > C by 0.50 SD Early literacy score: T > C by 0.59 SD No statistically significant difference in effects between Pre-K programs funded by the Public Education Department and those funded by the Children, Youth and Families Department.

(*continued on next page*)

Table 12 (continued)

Study/program name and data	Study design	Results
An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs (Wong et al., 2008). Data on test scores from fall 2004 from Michigan, New Jersey, Oklahoma, South Carolina and West Virginia. Sample sizes: T = 485, C = 386 (MI); T = 1177, C = 895 (NJ); T = 431, C = 407 (OK); T = 353, C = 424 (SC); T = 379, C = 341 (WV). (T = children who just completed Pre-K, C = children who are about to begin Pre-K). Entry: 4 years old; Exit: 5 years old.	Regression discontinuity design due to a strict birthday eligibility cut-off. Looked at effect size differences due to programmatic variation between states. Used a polynomial approximation to the continuous function on the assignment variable in the regressions.	**Intent-to-Treat Results:** MI: T > C by 0.53 SD for math; 1.09 SD for Print Awareness. NJ: T > C by 0.36 SD for PPVT, 0.23 SD for math, 0.32 SD for Print Awareness. OK: T > C by 0.28 SD for PPVT, 0.78 SD for Print Awareness. WV: T > C by 0.92 SD for Print Awareness. **Treatment-on-Treated Results:** MI: T > C by 0.47 SD for math, 0.96 SD for Print Awareness. NJ: T > C by 0.36 SD for PPVT, 0.23 SD for math, 0.50 SD for Print Awareness. OK: T > C by 0.29 SD for PPVT. SC: T > C by 0.79 SD for Print Awareness. No statistically significant results for WV. No clear relationship between state funding for Pre-K programs and effect sizes. State Pre-K programs have larger effect sizes than Head Start.

Table 12 (continued)

Study/program name and data	Study design	Results
Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools (Cascio, 2009). Data from 4 Decennial censuses for 1970, 1980, and 2000 from the Public Use Microdata Samples. $n = 840$ whites, 425 blacks. Data from PSID on Head Start enrollment: $n = 174$ whites, 126 blacks.	Analyzed effect of expansion of public kindergarten on long-term outcomes. Identification from the variation in the timing of funding initiatives among treated states. Event study model, comparing individuals aged 5 before and after the initiatives were implemented. Included dummies for cohorts interacted with dummies for 3 different groups of treated states defined on the basis of average education expenditure per pupil in the early 1960s. Also controlled for cohort-by-region-of-birth fixed effects and state fixed effects. Units of observation are cohort-state cells.	White children aged 5 after the typical state reform are 2.5% less likely to be high school drop-outs and 22% less likely to be institutionalized as adults. No significant effects on grade retention, college attendance, employment, or earnings. No significant effects for blacks, despite comparable increases in enrollment in public kindergartens post reform. Potential explanation is that state funding for public kindergartens reduced the likelihood that a black 5-year-old attended Head Start by ~100%. Reduction in Head Start attendance may account for 16% of the 1.13 pp increase in the black-white gap in high school drop-out rates. Difference in effects on educational attainment between whites and blacks are driven by females.

(*continued on next page*)

Table 12 (continued)

Study/program name and data	Study design	Results
No Child Left Behind: Universal Child Care and Children's Long-Run Outcomes (Havnes and Mogstad, 2009). Data from Statistics Norway on individuals from 1967 to 2006. Household information from the Central Population Register. Administrative data on child care institutions and their locations for 1972-1996. Restricted sample to individuals whose mothers were married at the end of 1975. $n = 499{,}026$ children; 318,367 families. Adult outcomes measured between ages 30 and 33.	Difference-in-differences. Exploited a child care reform in 1975 in Norway, which assigned responsibility for child care to local governments, and thus resulted in great variation in child care coverage for children aged 3-6 both between cohorts and across municipalities. T = municipalities where child care expanded a lot; C = municipalities where child care did not change much. Compared changes in outcomes for treatment and control adults who were 3-6 years old before and after the reform. Also investigated heterogeneity of effects. Controlled for various child and family-specific characteristics, as well as municipality fixed effects. Robustness checks: included a time trend, checked for placebo effect by comparing the two groups before the reform, and used sibling fixed effects comparing siblings exposed to the reform to those who were not.	**TT Effects (ITT effects in parentheses)** One more child care place: increases educational attainment by 2.7% (0.4%) decreases probability of dropping out of HS by 22.8% (3.8%) increases probability of college enrollment by 17% (2.5%) decreases probability of having little or no earnings by 23.3% (3.9%) increases probability of having average earnings by 7.4% (1.3%) decreases probability of being on welfare by 31.9% (5.6%) decreases probability of parenthood by 10.4% (1.8%) increases probability of being single with no children by 23.3% (4%) Almost all of the reduction in probability of being a low earner is driven by females. Women more likely to delay child bearing and family formation than men as a result of increased child care. Most benefits of universal child care are for children of low-educated mothers. Subsidized formal child care crowds out informal child care with almost no net increase in total child care use or maternal employment. No impact of child care reform on maternal education. **Cost-Benefit Analysis:** Cost: Annual budgetary cost per child care place = \$5400 Benefit: Increase in 0.35 yrs of education implies an increase in \$27,000 in lifetime earnings.

See Karoly et al. (2005) for more information about some of the studies described in this table. Unless otherwise noted, none of these evaluations was randomized. "T" refers to the treatment, and "C" refers to the control or comparison group. T > C means that the difference was significant at the 5% level.

[a] A very similar study by Gormley et al. (2005) evaluates the effects of Oklahoma's Universal Pre-K program on school readiness using the same regression discontinuity design, but measuring outcomes with the Woodcock-Johnson Achievement Test. They find a 0.79 SD increase in the Letter-Word Identification score, a 0.64 SD increase in the Spelling score, and a 0.38 SD increase in the Applied Problems score.

that projected gains in earnings are enough to offset the cost of the program, so that there is a positive cost/benefit ratio. Carneiro and Ginja (2008) use the same data but a different identification strategy: they focus on families around the cutoff for income eligibility for the program and compare families who are just below (and therefore eligible) to those who are just above (and therefore ineligible). A potential problem with this strategy is that it implicitly assumes that families cannot game the system by reducing their incomes in order to become eligible for the program. Consistent with other studies, they find positive effects of Head Start attendance on adolescents including reductions in behavior problems, grade repetition, depression, and obesity.

Since its inception, Head Start has aimed to improve a broad range of child outcomes (not only test scores). When the program was launched in 1965, the Office of Economic Opportunity assisted the 300 poorest counties in applying for Head Start funds, and these counties were significantly more likely than other counties to receive funds. Using a regression discontinuity design, Ludwig and Miller (2007) show that mortality from causes likely to be affected by Head Start fell among children 5 to 9 in the assisted counties relative to the others. Mortality did not fall in slightly older cohorts who would not have been affected by the introduction of the program.

Frisvold (2006) and Frisvold and Lumeng (2009) also focus on health effects by examining the effect of Head Start on obesity. The former instruments Head Start attendance using the number of Head Start places available in the community, while the later takes advantage of a cut in a Michigan Head Start program which resulted in the conversion of a number of full-day Head Start places to half day places. Both studies find large and significant effects of Head Start on the incidence of obesity. In defense of their estimates, which some might find implausibly large, Frisvold and Lumeng point out that a reduction of only 75 calories per day (i.e. less than a slice of bread or an apple) would be sufficient to yield their results. In small children, even small changes in diet may have large cumulative effects. Anderson et al. (2009) follow Garces et al. and use sibling fixed effects and data from the PSID to estimate the effect of Head Start on smoking as an adult. Again, they find large effects.

Head Start has served as a model for state preschools targeted to low-income children in states such as California, and also for new (non-compulsory) universal preschool programs in Georgia, and Oklahoma. The best available evaluations of universal preschool programs highlight the importance of providing a high quality program that is utilized by the neediest children. Baker et al. (2008) examine the introduction of a universal, $5 per day (later $7), preschool program in the Canadian province of Quebec. The authors find a strong response to the subsidy in terms of maternal labor supply and the likelihood of using care, but they find negative effects on children for a range of outcomes. Lefebvre et al. (2006) focus on the same natural experiment and examine the effects on children's vocabulary scores, which have been shown to be a good predictor of schooling attainment in early grades. They find strong evidence of negative effects.

In interpreting this study, it is important to consider who was affected by the program. Because poor children were already eligible for child care subsidies, the marginal child affected by this program was a child who probably would have stayed home with his or her middle-class, married, mother, and instead was put into child care. Moreover, the marginal child care slot made available by the program was of low quality—the sudden influx of children into care caused the province to place more emphasis on making slots available than on regulating their quality. Hence, the study should be viewed as the consequence of moving middle class children from home care to relatively poor quality care. It is not possible to draw any conclusion from this study about the effect of drawing poor children into care of good quality, which is what model preschool programs and Head Start aim to do.

Gormley and Gayer (2006) examine the effects of Oklahoma's universal pre-K program, which is run through the public schools and is thought to be of high quality. They take advantage of strict age cutoffs for the program and compare children who had just attended for a year to similar children who were ineligible to attend because they were slightly younger. They find a 52% gain in pre-reading skills, a 27% gain in pre-writing skills, and a 21% gain in pre-math skills. These results suggest that a high quality universal pre-K program might well have positive effects, though one would have to track children longer to determine whether these initial gains translate into longer term gains in schooling attainment. Several other recent studies use a similar regression discontinuity design, including Hustedt et al. (2008) and Wong et al. (2008) who examine state pre-K programs in five states. These studies find uniformly positive effects. It has been argued in fact, that the effects of quality state preschool programs are larger than those of Head Start. However, it is difficult to control for pre-existing differences between the Head Start children and children who attend other preschools. For example in Magnuson et al. (2007), the preschool children had systematically higher incomes than those who attended Head Start.

A handful of studies examine the long-term effects of public pre-school or kindergarten programs. Cascio (2009) uses data from four decennial censuses to analyze the impact of introducing kindergarten into public schools in the US, where kindergarten was phased in on a state-by-state basis. Using a cohort-based design, she finds that white children born in adopting states after the reform were less likely to dropout of high school and less likely to be institutionalized as adults. However, she finds no significant effect for blacks, which may be due to significant crowd out of blacks from other programs, such as Head Start. Like Anderson, she finds that the effects were larger for girls. Havnes and Mogstad (2009) study a 1975 policy change in Norway which increased the availability of regulated child care in some areas but not in others. They find that children "exposed" to more child care received more education and were more likely to have earnings as adults. Once again, much of the benefit was concentrated among females, and children of less educated mothers were particularly likely to benefit. In terms of mechanisms, they find

that the increase in formal care largely displaced informal care, without much net effect on the mother's labor force participation.

Finally, it is worth mentioning the "Sure Start" program in England and Wales. This program aimed to provide early intervention services in disadvantaged neighborhoods but allowed a wide variety of program models, which obviously complicates an assessment of the program. An evaluation was conducted by comparing communities that were early adopters to those that adopted later. A second evaluation compared Sure Start children to children from similar neighborhoods who were drawn from the Millennium Cohort study. This second study used propensity scores to balance the samples. The first evaluation found that the most disadvantaged households were actually doing more poorly in intervention areas than in other areas (NESS, 2005), while the second found some evidence of positive effects (NESS, 2008). Following the first evaluation, there has been a move to standardize the intervention and most communities are now offering Sure Start Children's Centers. This latest incarnation of the program remains to be evaluated.

This discussion shows the value of using a framework for the production of child quality as a lens for the interpretation of the program evaluation literature. As discussed above, child human capital is produced using inputs that may come from either the family or from other sources. A program that augments the resources available to the child is likely to have positive effects (subject of course to diminishing returns), while a program that reduces the resources available to the child is likely to have negative effects. Hence, a program that causes poor quality group time to be substituted for relatively high quality maternal time can have a negative effect and vice versa. The important point is that it is possible to intervene effectively and to improve the trajectories of young children.

5.3.4. Health insurance

Health insurance is not an intervention program in the sense of the programs described above. Yet, there is a good deal of evidence that access to health insurance improves children's health at birth and afterwards. Much of the evidence comes from studies of the introduction, or expansion, of health insurance benefits. Some of this literature is summarized in Table 13. For example, Hanratty (1996) examined the introduction of public health insurance in Canada, which was phased in on a province-by-province basis. Using county-level panel data, she finds that the introduction of health insurance was associated with a decline of four percent in the infant mortality rate, and that the incidence of low birth weight also decreased by 1.3% for all parents and by 8.9% for single parents. Currie and Gruber (1996) conduct a similar exercise for the US, focusing on an expansion of public health insurance to pregnant women and infants. They find that the effects vary depending on whether the expansion covered the poorest women, or women somewhat higher in the income distribution. Narrowly targeted expansions that increased the fraction of the poorest women eligible by 30%, reduced low birth weight by 7.8%, and reduced infant mortality by 11.5%. Expansions of eligibility of a similar magnitude to women of higher incomes had very small effects on the incidence of low

Table 13 Effects of Medicaid and other public health insurance on birth and early childhood outcomes.

Study and data	Study design	Results
Effects on birth weight and health at birth		
The efficacy and cost of changes in the Medicaid eligibility of pregnant women (Currie and Gruber, 1996). Note: Authors also conduct an analysis of Medicaid take-up, which is not included in the results here. Data from CPS and Vital Statistics. Data on Medicaid expenditures from the Health Care Financing Administration. Data on the use of medical services by pregnant women from the NLSY. Simulated model for targeted changes estimated for 1979–1992. Simulated model for broad changes estimated for 1987–1992. ($n = 600$).	Exploited variation between states in the timing of expansions of Medicaid eligibility. Use a fixed sample to simulate the fraction eligible under different state rules. Distinguish "targeted" changes affecting very low income women from "broad" changes to women further up the income distribution. Instrument the actual fraction of women eligible in each state and year with the simulated eligibility measure. Controlled for state fixed effects and time varying state characteristics.	The percentage of 15–44 yr old women eligible for Medicaid (had they become pregnant) rose from 12.4% to 43.3% b/n 1979 and 1991. A 30% increase in eligibility leads to a 1.9% decrease in incidence of low birth weight (sig. at 10% level) and a 8.5% decrease in infant mortality rate. For targeted program changes, a 30% increase in eligibility decreases low birth weight (infant mortality) by 7.8% (11.5%). For broad program changes a 30% increase in eligibility decreases low birth weight (infant mortality) by 0.2% (2.9%).

Table 13 (continued)

Study and data	Study design	Results
Canadian National Health Insurance and Infant Health (Hanratty, 1996). County-level panel data on infant mortality from 10 provinces in 1960-1975 from the Census of Canada and from Canada's Division of Vital Statistics. ($n = 204$ counties). Data on birth weight from a sample of all birth records in Canada from 1960 to 1974.	Used variation in timing of implementation of national health insurance across provinces in Canada over 1962-72. Logit of outcomes on a dummy for having national health insurance in a particular county-year, controlling for demographic and socio-economic factors, a time trend, and year fixed effects.	Introduction of national health insurance leads to declines of 4% in infant mortality rates; 1.3% in low birth weight (whole sample); 8.9% in low birth weight (single parents). No impact on birth weight among married women.
Changes in prenatal care timing and low birth weight by race and socioeconomic status: Implications for the Medicaid expansions for pregnant women (Dubay et al., 2001). Data on births from the 1980, 1986, and 1993 Natality Files. $n = 8{,}100{,}000$ births.	Difference-in-difference, subtracting difference in obstetrical outcomes (rates of late initiation of prenatal care and rates of low birth weight) b/n 1980 and 1986 from difference in outcomes b/n 1986 and 1990, within socioeconomic (SES) groups. Also compared changes in obstetrical outcomes in 1986-93 b/n women of low and high SES (since high SES women were not affected by Medicaid expansions). SES defined by marital status and years of schooling. Medicaid expansions occurred in 1986-93. Controlled for year, age of mother, parity, and age-parity interactions.	**Results from diff-in-diff within SES:** Medicaid expansions associated with decreases of: 12%-21% prenatal care initiation after 1st trimester for white women; 3%-5% in low birth weight among white women with <12 years education; 10%-13% in prenatal care initiation after 1st trimester for black women with <12 years education; 13%-27% in prenatal care initiation after 1st trimester for black women with 12-15 years of education; 12%-35% in prenatal care initiation after 1st trimester for black women with >15 years of education. Association with a 3% increase in likelihood of low birth weight for unmarried black women with <12 years education. Similar results using diff-in-diff across SES for 1986-93.

(*continued on next page*)

Table 13 (continued)

Study and data	Study design	Results
Effects of Medicaid expansions and welfare contractions on prenatal care and infant health (Currie and Grogger, 2002). Data on birth outcomes from VSDN files 1990-1996. Data on fetal deaths from vital statistics fetal deaths detailed records. Data on Medicaid administrative reforms from National Governor's Association Maternal and Child Health newsletters. Data on welfare caseloads from the US Department of Health and Human Services. $n = 3{,}985{,}968$ whites; 4,014,935 blacks.	Logit regression that includes dummies for state-level Medicaid administrative reforms; income eligibility cutoffs; rates of welfare participation; unemployment rates; and maternal observable characteristics. Also estimated auxiliary regressions that examine the effect of policy variables on aggregate Medicaid caseloads.	Medicaid caseload increases by 0.233 for each 1% increase in welfare rate. Medicaid caseload increases by 0.664 for each 100% increase in income cut-off (for those not receiving cash benefits). Increase in income cutoff from 100% to 200% of poverty line increases probability of adequate prenatal care by 0.4% for whites. 2 pp increase in welfare rate associated w/ 1.1% increase in probability that prenatal care was initiated in 1st trimester; 0.7% increase in probability of adequate care for whites; 2% increase for both for blacks. Increase in income cutoff from 100% to 200% of poverty line associated w/ a decrease of 1720 fetal deaths per year among blacks. 2% increase in welfare associated w/ 10% decrease in fetal deaths per year among blacks. Most administrative reforms have no effect. But using mail-in forms (instead of in-person interviews) increases probability that prenatal care was initiated in 1st trimester by 0.7% for blacks and shorter forms increase probability that prenatal care was initiated in 1st trimester by 3% for whites. Using mail-in forms decreases probability of low birth weight by 26%; of very low birth weight by 38% for whites.

Table 13 (continued)

Study and data	Study design	Results
Using discontinuous eligibility rules to identify the effects of the federal Medicaid expansions on low-income children (Card and Shore-Sheppard, 2004). Data from SIPP for 1990-93, March CPS for 1990-96, and Health Interview Survey for 1992-1996 on children under 18 years old. $n = 10{,}268$ to 16,196 across the different years in SIPP.	Two sources of identification: "The 133% expansion" (children under age 6 living in families with incomes below 133% of the poverty line became covered in 1989) and "the 100% expansion" (children born after September 30, 1983 in families with incomes below the poverty line became covered). Difference-in-difference design comparing age-6 and age-5 children in families with incomes between 100% and 133% of the poverty line for the 133% expansion, and comparing children born before and after Sep. 30, 1983 for the "100% expansion". Regression of Medicaid enrollment on dummy for being below poverty level, dummy for being born after 9/30/1983, their interaction, dummy for age <6 years old, interaction between dummy for age <6 years old and dummy for being between 100% and 133% of the poverty line, a flexible function of age and family income, and other background characteristics as well as year fixed effects.	The 100% expansion led to 7%-11% take-up rates, while the 133% expansion had <5% take-up rates. No evidence for other insurance crowd-out in SIPP data. Results from CPS data suggest that the 133% expansion led to decline in other health insurance coverage by approximately the same amount as the take-up in Medicaid. Similar results using Health Interview Survey data.

(*continued on next page*)

Table 13 (continued)

Study and data	Study design	Results
Effects of Medicaid managed care on prenatal care and birth outcomes (Aizer et al., 2007). Data on birth outcomes from the California Birth Statistical Master File 1990-2000 and Birth Cohort files for the same time period. Hospital-level data from Vital Statistics records. $n = 55{,}000$ births.	Exploited the county-by-county variation in implementation of Medicaid managed care, which resulted from a phase-in policy in California that required women enrolled in Medicaid to switch to managed care plans. Some counties switched to COHS plan, others switched to 2 or more plan system. Multivariate regression with county fixed effects, and mother fixed effects, as well as other observable characteristics. Robustness checks: Estimating same model for married women only (unlikely to be on welfare, hence unlikely to be subject to MMC); controlling for mobility b/n counties; regression discontinuity design to eliminate time trend effects; adding interaction terms b/n time trend and several dummies; "intent- to-treat" models to control for the MMC adoption being not exogenous to counties.	Probability of starting prenatal care in 1st trimester: decreased by 9%-10% in both COHS and 2-plan counties. Use of induction/stimulation of labor: increased by 43.8% in COHS counties. Use of fetal monitors: increased by 25.9% in COHS counties. Incidence of low birth weight: increased by 15% in both COHS and 2-plan counties. Incidence of short gestation: increased by 15% in both COHS and 2-plan counties. Incidence of neonatal death: increased by 50% in 2-plan counties.

Table 13 (continued)

Study and data	Study design	Results
Effects on later child outcomes		
Medicaid eligibility and the incidence of ambulatory care sensitive hospitalizations for children (Kaestner et al., 2001). Data from the Nationwide Inpatient Sample of the Healthcare Cost and Utilization Project for 1988 and 1992. $n = 36{,}000$.	Difference-in-difference, comparing the change in ambulatory care sensitive (ACS) hospitalizations before and after Medicaid expansions b/n poor and non-poor children. Poverty status determined by median family income in child's zip code of birth. Two treatment groups: income < \$25,000 and \$25,000 < income < \$35,000. Control group: income >\$35,000. Separate estimates for children aged 2-6, 7-9. Controlled for hospital-specific, year, and individual factors. Incidence of ACS hospitalizations calculated using both non-ACS hospitalizations and total births in the denominator.	For children aged 2-6 in families with <\$25,000 income, incidence of ACS hospitalizations due to dehydration, convulsions and non-asthma illnesses declined by 10%-20%. Estimated effect sizes of 40%-80% for those affected by Medicaid expansions. For children aged 2-6 in families with \$25,000 < income < \$35,000, incidence of ACS hospitalizations due to non-asthma illnesses and pneumonia declined by 10%-14% (only when denominator is total births). For children aged 7-9 in families with \$25,000 < income < \$35,000, incidence of ACS hospitalizations due to asthma declined by 22%-30%; hospitalizations due to eye, nose and throat illnesses declined by >100%. No significant effects on children aged 2-6 in \$25,000 < income < \$35,000 group or on children aged 7-9 in <\$ 25,000 income group.

(*continued on next page*)

Table 13 (continued)

Study and data	Study design	Results
Low take-up in Medicaid: does outreach matter and for whom? (Aizer, 2003). Data on monthly Medicaid enrollment for children aged 0-15 in CA from 1996 to 2000 linked with information on the number of community-based application assistants in each zip code in the month of application by age and race. $n = 324{,}331$ for analysis of take-up. n = 121,806 for analysis of hospitalizations.	Examined effects of CA outreach campaign launched in June 1998, which consisted of community-based application assistants and a media campaign to raise awareness about Medicaid, on Medicaid enrollment. Multivariate regression with key explanatory variable being the number of community-based application assistants in each zip code in the month of application by age and race. Included zip code and month fixed effects as well as controls for changes in business cycle and the demographic compositions of the state. Also examined effect of early Medicaid enrollment on ACS hospitalizations by instrumenting Medicaid enrollment with outreach measures.	Access to bilingual application assistants increases new monthly Medicaid enrollment by 4.6% among Hispanic children, and by 6% among Asian children (relative to other children in same neighborhood). A 1000-child increase in Medicaid enrollments decreases ACS hospitalizations by 3.26 hospitalizations (mean not reported, so can't calculate relative effect size).

Table 13 (continued)

Study and data	Study design	Results
Public insurance and child hospitalizations: Access and efficiency effects (Dafny and Gruber, 2005). Data from NHDS on child discharges for 1983 to 1996. Cells defined for 4 age categories (<1, 1-5, 6-10, 11-15) for each state and year. $n = 2308$ cells. Used age-state-year populations from the Census Bureau to calculate hospitalization rates for each cell.	Investigated impact of Medicaid expansions on hospitalizations using the variation in Medicaid eligibility b/n states, over time, and by age. Key explanatory variable is the eligibility rate measured by the fraction of children eligible for Medicaid in each age- state-year cell. Controlled for age, state, and year fixed effects as well as state-year interactions. Used the fraction eligible calculated using a fixed sample to instrument for actual eligibility in each state, year, and age group. Also examined effect of Medicaid eligibility on length of stay in hospital and the number of procedures performed.	A 10 pp increase in Medicaid eligibility leads to a 8.4% increase in total hospitalizations, an 8.1% increase in unavoidable hospitalizations, and no statistically significant impact on avoidable hospitalizations. Assuming access to hospital care increases the likelihood of all kinds of hospitalizations equally, these results imply that increased use of primary care engendered by Medicaid expansions mitigated the increase in total hospitalizations by reducing the increase in avoidable hospitalizations that would have otherwise occurred. A 10 pp increase in eligibility leads to a 3.1% decrease in the length of hospital stay, a 5% increase in the # of procedures performed, and a 6.6% increase in the likelihood of having any procedure performed, i.e. leads to more aggressive care.
Public health insurance, program take-up and child health (Aizer, 2007). See Aizer (2003) entry for description of data.	See Aizer (2003) entry. Additional analysis: Examined heterogeneity in effects on take-up by age. Examined nonlinear effects on take-up. Examined effects of English and Spanish language advertisements on take-up.	Proximity to an additional bilingual application assistant increases new monthly Medicaid enrollment by 7%-9% among Hispanics and by 27%-36% among Asians. Smallest effects for infants who were already largely eligible; largest effects for ages 6-15. Effect is linear for Hispanics; slightly concave for Asians. English language advertisements increase Medicaid enrollment in the following month by 4.7% among all children, and Spanish language advertisements increase Medicaid enrollment in the following month by an additional 2.5% among Hispanic children. Increasing the # of children w/ Medicaid by 10% results in a 2%-3% decline in avoidable hospitalizations.

(*continued on next page*)

Table 13 (continued)

Study and data	Study design	Results
Has public health insurance for older children reduced disparities in access to care and health outcomes? (Currie et al., 2008). Data from the national health interview surveys for 1986-2005. $n = 474{,}164$ children <18 yrs old. Instrument data from the CPS.	Identification due to the fact that expansions in Medicaid/SCHIP eligibility for older children relative to younger children happened at different times in different states. Instrumented for individual Medicaid/SCHIP eligibility using an index of generosity of the state's public health insurance programs calculated using a fixed group of children. Then, estimated impact of public health insurance eligibility on the relationship b/n family income and child health status and on the relationship b/n family income and doctor visits in the previous year. Also tested whether the relationship b/n income and outcomes changed over time within age groups by running separate regressions for 4 age groups: 0-3, 4-8, 9-12, 13-17. Finally, estimated relationship b/n health outcomes and the fraction of children eligible in child's birth cohort in the child's current state of residence for children aged 9-17. Controlled for family background variables and year and state fixed effects.	For children aged 9-12, the coefficient on income in a model of health status declined by 20% over 2000-2005. For children aged 13-17, it declined by 18% over 1996-2000 and by 25% over 2000-05. For children 0-3/4-8/9-12 the coefficient on income in a model of doctor visits declined by 64%/62%/50% between 2000-2005. For children aged 13-17 there was no significant change in the income coefficient. Significant declines in the income coefficient over 1991-1995 and 1996-2000 for children aged 0-3 and 4-8 as well. No statistically significant impact of contemporaneous health insurance eligibility on child health status. A 100 pp increase in the fraction eligible at age 3 would reduce the probability that the average child aged 9-17 is in less than excellent health by 11%. A 100 pp increase in the fraction eligible at ages 1 and 2 would reduce the probability that the average child aged 9-17 had no doctor visit in the past year by 41%.

Table 13 (continued)

Study and data	Study design	Results
The impact of children's public health insurance expansions on educational outcomes (Levine and Schanzenbach, 2009). Data on state-level average scaled test scores from the National Assessment of Educational Progress for 1990 to 2003. ($n = 431$). Data for simulated instruments from March CPS. Data on child health from Vital Statistics for 1984-2003 ($n = 1020$).	Examine the impact of public health insurance at birth on 4th and 8th grade reading and math test scores. Difference-in-difference-in-difference using cross-state variation over time and across ages in children's health insurance eligibility due to Medicaid and SCHIP expansions. Instrumented for public health insurance eligibility using the simulated fraction eligible (as in Currie and Gruber (1996) see above). To test whether changes in educational outcomes are due to improvements in health status or to additional household income generated if the availability of public health insurance crowds out private health insurance, they estimated the direct impact of health at birth on educational outcomes. Controlled for state and year fixed effects and state-specific time trends.	A 50 pp increase in eligibility at birth increased reading test scores by 0.09 SD relative to 4th and 8th grade combined mean scaled score. No effect on math test scores. Expansions to public health insurance eligibility at birth associated w/ 1.6% reduction in low birth weight rate for whole sample and 6.7% reduction in infant mortality rates among women with at least a high school degree. A 50% increase in low birth weight (infant mortality) rate would decrease reading test scores by 0.12 SD (.07SD).

(*continued on next page*)

Table 13 (continued)

Study and data	Study design	Results
Public insurance, crowd-out and health care utilization (Koch, 2009). Data from the medical expenditure panel survey. Focus on subsample of children <18 yrs whose families have incomes b/n 50% and 400% of poverty line. ($n = 32{,}609$). Data on public insurance reimbursement rates from the American Academy of Pediatrics for 1998-2001.	Regression discontinuity design using family income cut-offs that determine eligibility for Medicaid and SCHIP. To distinguish b/n children who are eligible for public insurance w/out access to private insurance and those whose private insurance may be crowded-out due to expansions in public health insurance, constructed a measure of private-insurance "offer" (=1 if any of the following holds: child is privately insured at the time of the eligibility measurement or a family member is privately employed at the time and either has insurance through the job or turns down the insurance at the job). Estimated impacts of public health insurance eligibility on measures of care utilization (doctor's visits, etc.) as well as on private insurance take-up. Also estimated the differential effect of eligibility on those w/ and w/out private health insurance by interacting the eligibility variable with the "offer" variable. Controlled for background characteristics as well as state and year fixed effects. Also looked at impacts of differential reimbursement rates across states and years on access to care for the publicly insured.	Public insurance eligibility increases public insurance coverage by 21% and decreases private insurance coverage by 24%. So overall decrease in insurance coverage of about 4.8%. Eligibility decreases the number of office visits/self-reported health status/total expenditures by 12%/87%/23%. Increases the number of ER visits by 18%. Eligibility decreases the number of office visits by 18% for those with outside private health insurance. No statistically significant effects on ER visits, prescriptions, or refills. Children just eligible for public health insurance are 4% less likely to have a usual source of care and 4% more likely to go without care. No statistically significant effects on hospital or ER visits. Eligibility increases children's BMI by 2%. Parents of just eligible children are 5% less likely to be given advice about their child's eating healthfully. No significant effects on asthma diagnosis. Conditional on taking medication for asthma, eligibility increases probability of taking inhaled steroids (gold standard treatment) by 32%. A $4 increase in reimbursement for office visits leads to a 5% increase in the number of office visits.

Percent changes (denoted by % instead of "pp") are reported relative to the mean.

birth weight, but reduced infant mortality. This result suggests that among women of somewhat higher income levels, the expansions did not improve health at birth, but may have increased access to life-saving technologies after birth. Currie and Grogger (2002) focus on bureaucratic obstacles to obtaining health insurance by looking at contractions of welfare (women cut from the rolls lost automatic eligibility for Medicaid) as well as outreach measures undertaken by different states. They find that changes that reduced barriers to enrollment increased use of prenatal care and had positive effects on infant health outcomes.

Baldwin et al. (1998) use individual-level data and compare expansions in Washington, which included enhanced prenatal care services, to expansions in Colorado which did not, in a difference-in-differences design. They find reductions in low birth weight among medically high risk infants in Washington. Dubay et al. (2001) conduct a difference-in-difference investigation comparing the outcomes of high and low socioeconomic status women in the 1980-1986 period and in the 1986-1993 period. They find overall improvements in the use of prenatal care for low SES women, but find improvements in birth weight only for some groups of white women. However, this design does not really focus on health insurance *per se*, since the estimates will be affected by any other changes in health care markets between the two periods that had differential effects by SES.

Studies of the effects of health insurance expansions on children often examine preventable hospitalizations (also called ambulatory care sensitive hospitalizations). The idea is that certain conditions, such as childhood asthma, should not result in hospitalizations if they are properly treated on an outpatient basis. Hence, hospitalizations for these conditions are inefficient and indicate that children are receiving inadequate preventive care. Kaestner et al. (2001) use a difference-in-differences design comparing low income and other children before and after Medicaid expansions. They find reductions in preventable hospitalizations of 10% to 20%.

Aizer (2003) examines a California outreach program that increased child enrollments into Medicaid. She finds that an increase in enrollments of 1000 reduces hospitalizations by 3.26. Dafny and Gruber (2005) use a design similar to Currie and Gruber in which actual individual eligibility is instrumented using a "simulated eligibility measure" which is an index of the generosity of the Medicaid program in the state. The reason for adopting instrumental variables estimation is that eligibility for Medicaid is determined by endogenous variables such as parental labor supply. They find that Medicaid eligibility increased hospitalizations overall. However, there was no statistically significant increase in avoidable hospitalizations, suggesting that the increase was mainly due to children with unavoidable conditions gaining greater access to care. They also found increases in the probability of receiving a procedure and reductions in length of stay, suggesting that children who were hospitalized were receiving more aggressive care and that it may have improved their outcomes.

One difficulty with studying child health is that health today is affected by past investments, including health insurance at younger ages. Currie et al. (2008) therefore compare the health effects of contemporaneous eligibility for health insurance among older children to the effect of having been eligible since birth. They find that contemporaneous health insurance coverage has little effect on health status but that eligibility from birth is protective. Levine and Schanzenbach (2009) link health insurance eligibility at birth to 4th and 8th grade scores on the National Assessment of Educational Progress. They find that a 50% point increase in eligibility at birth is associated with a small but significant gain on reading scores at both grades, though there is no effect on math scores. A difficulty with both studies is that neither income at birth nor state of birth are directly observed in the cross sectional data sets that they use, so they must be imputed using current income and state of birth.

Another area of research focuses on the quality of care provided by public health insurance programs. Analysis of this issue is complicated by the possibility that expansions of public insurance cause people to lose private health insurance coverage, a phenomena dubbed "crowdout" (Cutler and Gruber 1996; Dubay and Kenney, 1997; Card and Shore-Shepard, 2004; Gruber and Simon, 2008). If the private insurance that is lost (or dropped) in response to expansions of public insurance is of superior quality to the private insurance, then people's health may suffer. Koch (2009) concludes that recent expansions of public health insurance to children at higher income levels have reduced access to doctor's office visits and increased reliance on emergency rooms. He also shows some evidence consistent with the idea that this is because children are being crowded out of superior (but obviously more expensive) coverage from private heath insurance plans. In fact, it is quite possible that crowding out has increased over time as the public has become familiar with public health insurance plans for children and private health insurance costs have continued to escalate.

Medicaid managed care has also been shown (in at least some cases) to reduce the quality of care. Sloan et al. (2001) conduct a difference-in-difference analysis of Tennessee and North Carolina before and after Tennessee switched its Medicaid patients to managed care. They find that use of prenatal care and birth outcomes deteriorated in Tennessee after the switch. Aizer et al. (2007) examine data from California, where Medicaid managed care was adopted on a county-by-county basis. They also find that compulsory managed care had a negative impact on use of prenatal care and birth outcomes. This may be because the California Medicaid managed care program "carved out" care for sick newborns–that is, they were covered by a state fund rather than by the managed care companies so that companies had little incentive to take actions to improve newborn health.

In summary, health insurance matters for children's outcomes. But quality of care also matters. And it is important to remember that for most children threats to health and

well being come from sources such as injuries, poor nutrition, and toxins rather than only from lack of access to medical care.

6. DISCUSSION AND CONCLUSIONS

There has been an explosion of research into the early determinants of human capital development over the past 10 years. The work surveyed in this chapter conclusively shows that events before five years old can have large long term impacts on adult outcomes. It is striking that child and family characteristics measured at school entry do as much to explain future outcomes as factors that labor economists have more traditionally focused on, such as years of education. Yet evidence for long term effects of early insults should not be a cause of pessimism. While children can be permanently damaged at this age, the damage can be remediated. The picture that emerges is one of vulnerability but also of resilience.

Since early childhood is a new area of research for economists, there remain many unanswered questions. One major question implicit in the structure of this chapter is whether it will ever be possible to estimate human capital production functions. The opening sections of this chapter showed that the production function paradigm provides an extremely useful way of thinking about the problem, and in particular, that it highlights the importance of actions taken by parents and others in exacerbating or mitigating the effects of random shocks. However, to actually estimate the implied production functions would place huge demands on the data, demands that are unlikely to be met in practice.[15] Hence, the evidence we have surveyed is largely reduced form.

A second major question is whether shocks at certain key ages matter more than others? Much has been written about "critical periods". The idea is that certain functionalities must be acquired at a particular point in life, and if they are not acquired at that point, they will either not be acquired at all, or will not be acquired properly. There is to date little hard evidence of critical periods in humans. However, the evidence discussed above certainly suggests that the period while children are *in utero* is one of the most important to their later development. This has important implications for the timing of social interventions designed to mitigate harms—it may be that interventions should be targeted at pregnant women and/or women of child bearing age in addition to young children. But there is insufficient evidence at present to be able to say that insults or interventions at for example zero to one are likely to be more effective than interventions at age three or four. Moreover, the cohort designs used to establish the importance of the fetal period can tell us more about the comparison between the fetal period and the early post natal period than they do about the comparison between the fetal period and say exposures at age five.

[15] See however Appendix D for one thought about how a production function might be estimated.

A related question is whether some types of shocks matter more than others. This chapter surveyed many different types of shocks including exposure to disease, inadequate nutrition, exposure to pollution, injuries, maternal mental health problems, maternal smoking, and maltreatment. However, given that studies of the effects of these shocks rely on different populations, time periods, and methods, it is difficult to get any sense of whether one type of shock is more of a threat to human capital development than any other. Similarly, while it is clear that shocks to health have long-term effects on domains such as education and earnings, it is not clear whether health shocks have direct effects on cognition or learning, or whether they act mainly by affecting future health.

Several studies we reviewed suggested that both shocks and interventions can have different long-term effects on males and females. But these findings are too new for us to be able to predict when this difference will occur, and we have virtually no evidence about why it occurs. One possibility is that gender differences are biological. For example, boys may be less robust than girls so that the same health shock can "cull" boys while girls survive (e.g., see Kraemer (2000); Almond and Mazumder (2008)). In this case, average health of male survivors might be better than that of female survivors. Alternatively, gender differences could reflect differential parental or societal responses to shocks inspired by son preferences or by beliefs about biological gender differences.[16]

Finally, given all of this evidence of long-term effects of early life outcomes, what is the least costly way to intervene to improve outcomes? This is still an open question and our knowledge of the types of programs that are effective (and why) is evolving rapidly. For example, until recently, there was little evidence that income transfers had much effect, so it was easy to surmise that in-kind programs were a more effective way to improve child outcomes. Recent evidence that cash transfers are indeed effective should cause a re-evaluation of the received wisdom on this point, given the inefficiencies involved in providing transfers in-kind. Similarly, the large literature about negative effects of maternal employment in the early years is thrown into question by recent studies showing that large changes in maternity leave policies affected maternal employment without having any detectable impact on long-term child outcomes. This rapid development in our knowledge makes the study of human capital development before five an exciting frontier for research in labor economics.

APPENDIX A

The following acronyms are used in this chapter:

AFDC = Aid to Families with Dependent Children

BCS = British Birth Cohort Study of 1970.

[16] For example, advances in ultrasound technology could have changed the average human capital endowments of boys and girls by allowing parents who prefer sons to invest differentially prenatally (and not only by allowing them to abort female fetuses). See Lhila and Simon (2008) for recent work on this topic.

BPI = Behavioral Problems Index

BW = birth weight

CESD = Center for Epidemiological Depression scale

CCT = Conditional Cash Transfer

COHS = County Organized Health System

CPS = Current Population Survey

DDST = Denver Developmental Screening Test

ECLS-B = Early Childhood Longitudinal Study—Birth Cohort

ECLS-K = Early Childhood Longitudinal Study—Kindergarten Class of 1998-1999

EITC = Earned Income Tax Credit

EPA = US Environmental Protection Agency

FSP = Food Stamp Program

HAZ = Height for age *z*-score

HOME = Home Observation for Measurement of the Environment Score

IHDP = Infant Health and Development Project

IPUMS = Integrated Public-Use Microdata Samples of the US Census

IV = instrumental variables

LBW = Low Birth Weight (birth weight less than 2500 g)

MMC = Medicaid Managed Care

NBER = National Bureau of Economic Research

NCDS = National Child Development Survey (1958 British Birth Cohort)

NELS = National Education Longitudinal Study

NHANES = National Health and Nutrition Examination Survey

NHDS = National Hospital Discharge Survey

NLSY = National Longitudinal Survey of Youth, 1979 cohort

NLSY-Child = Children of the NLSY 1979 cohort

NLSCY = National Longitudinal Survey of Children and Youth (Canadian)

PIAT = Peabody Individual Achievement Test

pp = percentage points

PPVT = Peabody Picture Vocabulary Test

NSLP = National School Lunch Program

OLS = Ordinary Least Squares

PNM = Post Neonatal Mortality (death after 28 days and before 1 year)

PSID = Panel Study of Income Dynamics

RDA = Recommended Dietary Allowance

REIS = Regional Economic Information System

SCHIP = State Child Health Insurance Program

SD = Standard Deviation

SES = Socio-economic Status

SGA = Small for Gestational Age

SIPP = Survey of Income and Program Participation

SNAP = Supplemental Nutrition Assistance Program (formerly, Food Stamps)

TSIV = Two Sample Instrumental Variables

TVIP = Spanish-speaking version of the Peabody Picture Vocabulary Test

TSP = Total Suspended Particles

USDA = US Department of Agriculture

VSDN = Vital Statistics Detailed Natality files (birth certificate data for US)

WIC = Special Supplemental Nutrition Program for Women, Infants, and Children

WPPSI = Wechsler Preschool and Primary Scale of Intelligence

APPENDIX B

Human capital of a child is produced with a CES technology:

$$h = A\left[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)I_2^{\phi}\right]^{1/\phi}, \tag{9}$$

where μ_g is an exogenous shock to (predetermined) period 1 investments. Parents value their consumption and the human capital of their child:

$$U_p = U(C,h) = B\left[\theta(C)^{\varphi} + (1-\theta)h^{\varphi}\right]^{1/\varphi}, \tag{10}$$

and have the budget constraint:

$$\bar{I}_1 + I_2 + C = \bar{y}.$$

Absent discounting, the marginal utility from consuming equals the marginal utility from investing:

$$\frac{\delta U}{\delta C^*} = \frac{\delta U}{\delta h}\frac{\delta h}{\delta I_2^*}.$$

$$\theta C^{\varphi-1} = (1-\theta)h^{\varphi-1}A[\cdots]^{\frac{1}{\phi}-1}(1-\gamma)I_2^{*\phi-1} \tag{11}$$

$$\theta(\bar{y} - \bar{I}_1 - I_2^*)^{\varphi-1} = (1-\theta)A^{\varphi-1}[\cdots]^{\frac{\varphi-1}{\phi}}A[\cdots]^{\frac{1}{\phi}-1}(1-\gamma)I_2^{*\phi-1} \tag{12}$$

$$\theta(\bar{y} - \bar{I}_1 - I_2^*)^{\varphi-1} = (1-\theta)(1-\gamma)A^{\varphi}[\cdots]^{\frac{\varphi-\phi}{\phi}}I_2^{*\phi-1} \tag{13}$$

$$G(u_g, I_2^*) \equiv \theta(\bar{y} - \bar{I}_1 - I_2^*)^{\varphi-1} - (1-\theta)(1-\gamma)A^{\varphi}[\cdots]^{\frac{\varphi-\phi}{\phi}}I_2^{*\phi-1} = 0. \tag{14}$$

$$\frac{\delta I_2^*}{\delta \mu_g} = -\frac{\frac{\delta G}{\delta \mu_g}}{\frac{\delta G}{\delta I_2^*}}$$

$$= \frac{a(I_2^*)^{\phi-1}[\cdots]^{\frac{\varphi-2\phi}{\phi}}\left(\frac{\varphi-\phi}{\phi}\right)\gamma\phi(\bar{I}_1+\mu_g)^{\phi-1}}{-(\varphi-1)\theta(\bar{y}-\bar{I}_1-I_2^*)^{\varphi-2} - a\left[[\cdots]^{\frac{\varphi-\phi}{\phi}}(\phi-1)I_2^{*\phi-2} + \frac{\varphi-\phi}{\phi}[\cdots]^{\frac{\varphi-2\phi}{\phi}}\phi(1-\gamma)I_2^{*\phi-1}I_2^{*\phi-1}\right]}, \tag{15}$$

using the implicit function theorem and defining a to be $(1-\theta)(1-\gamma)A^{\varphi} \geq 0$.

$$= \frac{(\varphi-\phi)a(I_2^*)^{\phi-1}[\cdots]^{\frac{\varphi-2\phi}{\phi}}\gamma(\bar{I}_1+\mu_g)^{\phi-1}}{(1-\varphi)\theta(\bar{y}-\bar{I}_1-I_2^*)^{\varphi-2} + a[\cdots]^{\frac{\varphi-\phi}{\phi}}I_2^{*\phi-2}\left[(1-\phi)+(\varphi-\phi)(1-\gamma)I_2^{*\phi}/[\cdots]\right]}. \tag{16}$$

For $\varphi > \phi$, (16) is positive, so negative shocks in the first period should be reinforced. Accommodation through preferences (i.e., more consumption and less investment, which lowers h in addition to that caused by μ_g) is optimal.

APPENDIX C

Sibling a has human capital h_a, which is affected by a period 1 investment shock of μ_g:

$$h_a = A\left[\gamma(\bar{I}_{1a} + \mu_g)^{\phi} + (1-\gamma)I_{2a}^{\phi}\right]^{1/\phi}. \tag{17}$$

Sibling b does not experience a shock to first period investments:

$$h_b = B\left[\gamma \bar{I}_{1b}^{\phi} + (1-\gamma)I_{2b}^{\phi}\right]^{1/\phi}. \tag{18}$$

Assume further that first period investments do not distinguish between the two siblings (absent the shock experienced by sibling a):

$$\bar{I}_{1a} = \bar{I}_{1b} = \bar{I}_1.$$

Parents have Cobb-Douglas utility that cares only about the human capital of their two children:

$$U_p = U(h_a, h_b) = (1-\alpha)\log h_a + \alpha \log h_b. \tag{19}$$

The parents exhaust their budget on investments in their children:

$$\bar{y} = 2\bar{I}_1 + I_{2a} + I_{2b}.$$

Denoting $\bar{Y} = \bar{y} - 2\bar{I}_1$ as the budget for second period investments, $I_{2b} = \bar{Y} - I_{2a}$.

To maximize utility, the marginal utilities from investing in siblings a and b should be equal:

$$\frac{\delta U_p}{\delta h_a}\frac{\delta h_a}{\delta I_{2a}} = \frac{\delta U_p}{\delta h_b}\frac{\delta h_b}{\delta I_{2b}}$$

$$\left(\frac{1-\alpha}{h_a}\right)\frac{A}{\phi}\left[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)I_{2a}^{\phi}\right]^{(1/\phi)-1}\phi(1-\gamma)I_{2a}^{\phi-1}$$
$$= \left(\frac{\alpha}{h_b}\right)\frac{B}{\phi}\left[\gamma(\bar{I}_1)^{\phi} + (1-\gamma)I_{2b}^{\phi}\right]^{(1/\phi)-1}\phi(1-\gamma)I_{2b}^{\phi-1} \tag{20}$$

$$(1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)I_{2a}^{\phi}\right]^{-1} I_{2a}^{\phi-1}$$
$$= \alpha\left[\gamma(\bar{I}_1)^{\phi} + (1-\gamma)I_{2b}^{\phi}\right]^{-1} I_{2b}^{\phi-1} \tag{21}$$

$$G(\mu_g, I_{2a}) \equiv (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^{\phi} + (1-\gamma)I_{2a}^{\phi}\right]^{-1} I_{2a}^{\phi-1}$$
$$-\alpha\left[\gamma(\bar{I}_1)^{\phi} + (1-\gamma)(\bar{Y} - I_{2a})^{\phi}\right]^{-1}(\bar{Y} - I_{2a})^{\phi-1} = 0 \tag{22}$$

using budget the constraint: $I_{2b} = \bar{Y} - I_{2a}$. By the implicit function theorem:

$$\frac{\delta I_{2a}^*}{\delta \mu_g} = \frac{-\frac{\delta G}{\delta \mu_g}}{\frac{\delta G}{\delta I_{2a}^*}} \tag{23}$$

$$\frac{\delta G}{\delta \mu_g} = -(1-\alpha) I_{2a}^{\phi-1} \left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-2} \phi\gamma(\bar{I}_1 + \mu_g)^{\phi-1} \tag{24}$$

$$\Rightarrow \texttt{signum}\left[-\frac{\delta G}{\delta \mu_g}\right] = \texttt{signum}[\phi].$$

$$\begin{aligned}
\frac{\delta G}{\delta I_{2a}^*} &= (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-1} (\phi-1) I_{2a}^{\phi-2} \\
&\quad + (-1)(1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-2} \phi(1-\gamma) I_{2a}^{\phi-1} I_{2a}^{\phi-1} \\
&\quad - \left[\alpha\left[\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi\right]^{-1} (\phi-1)(\bar{Y} - I_{2a})^{\phi-2}(-1)\right] \\
&\quad - \left[(-1)\alpha\left[\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi\right]^{-2}\right. \\
&\quad \left. \times \phi(1-\gamma)(\bar{Y} - I_{2a})^{\phi-1}(-1)(\bar{Y} - I_{2a})^{\phi-1}\right] \qquad (25) \\
&= (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-1} (\phi-1) I_{2a}^{\phi-2} \\
&\quad - (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-2} \phi(1-\gamma) I_{2a}^{2\phi-2} \\
&\quad + \left[\alpha\left[\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi\right]^{-1} (\phi-1)(\bar{Y} - I_{2a})^{\phi-2}\right] \\
&\quad - \left[\alpha\left[\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi\right]^{-2} \phi(1-\gamma)(\bar{Y} - I_{2a})^{2\phi-2}\right] \qquad (26) \\
&= (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-1} \\
&\quad \times I_{2a}^{\phi-2}\left[(\phi-1) - \frac{\phi(1-\gamma) I_{2a}^\phi}{\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi}\right] \\
&\quad + \alpha\left[\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi\right]^{-1} (\bar{Y} - I_{2a})^{\phi-2} \\
&\quad \times \left[(\phi-1) - \frac{\phi(1-\gamma)(\bar{Y} - I_{2a})^\phi}{\gamma(\bar{I}_1)^\phi + (1-\gamma)(\bar{Y} - I_{2a})^\phi}\right] \qquad (27) \\
&= (1-\alpha)\left[\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi\right]^{-1} I_{2a}^{\phi-2} \\
&\quad \times \left[\phi\left(1 - \frac{(1-\gamma) I_{2a}^\phi}{\gamma(\bar{I}_1 + \mu_g)^\phi + (1-\gamma) I_{2a}^\phi}\right) - 1\right]
\end{aligned}$$

$$+\alpha\left[\gamma(\bar{I}_1)^{\phi}+(1-\gamma)(\bar{Y}-I_{2a})^{\phi}\right]^{-1}(\bar{Y}-I_{2a})^{\phi-2}$$
$$\times\left[\phi\left(1-\frac{(1-\gamma)(\bar{Y}-I_{2a})^{\phi}}{\gamma(\bar{I}_1)^{\phi}+(1-\gamma)(\bar{Y}-I_{2a})^{\phi}}\right)-1\right]. \tag{28}$$

Because $\phi \leq 1$ and:

$$\frac{(1-\gamma)I_{2a}^{\phi}}{\gamma(\bar{I}_1+\mu_g)^{\phi}+(1-\gamma)I_{2a}^{\phi}}<1 \quad \text{and} \quad \frac{(1-\gamma)(\bar{Y}-I_{2a})^{\phi}}{\gamma(\bar{I}_1)^{\phi}+(1-\gamma)(\bar{Y}-I_{2a})^{\phi}}<1,$$

Eq. (28) is always negative. Therefore:

$$\texttt{signum}\left[\frac{\delta I_{2a}^{*}}{\delta\mu_g}\right]=-\texttt{signum}[\phi].$$

We consider three cases for the substitutability of period 1 and period 2 investments (as captured by ϕ):

1. **Good Substitutability Between Periods 1 and 2** When $\phi > 0$, the optimal I_{2a} moves in the opposite direction from μ_g and parents should *compensate* a negative shock to child a by reducing second period investments in child b. Intuitively, it is easier to substitute though the production function for human capital than it is through the Cobb-Douglas utility function.
2. **Cobb-Douglas Substitutability Between Periods 1 and 2** For $\phi = 0$, the elasticity of substitution between periods is the same as the elasticity of substitution in preferences between the children (both Cobb-Douglas). Here, there is no winning investment response to the shock to child a, i.e. $\frac{\delta I_{2a}^{*}}{\delta\mu_g}=0$, so period 2 investments should be left unchanged.
3. **Poor Substitutability Between Periods 1 and 2** For $\phi < 0$, it is difficult to repair damage from a negative μ_g shock in the second period, so the return to period 2 investments in sibling a is below that for sibling b. Therefore, parents should *reinforce* the first period shock by allocating second period investments away from sibling a.

Importantly, the direction of these investment responses did **not** depend on α, the relative weight parents place in their utility function on the human capital of child a versus child b. Favoring the human capital formation of a particular child—even the child that experiences the negative endowment shock—does not affect the direction of the optimal investment response. Nor do differences in the "overall" productivity of the child, i.e. efficiency parameter $A \neq B$ in Eqs (17) and (18), alter the direction of the optimal investment response to μ_g. Thus, empirical evidence suggesting either reinforcing or compensating investments within the family does not reveal information

on parental preferences absent additional information on the production function for human capital.

APPENDIX D

In general, we need to observe the baseline investments $\bar{I}_1$ and $\bar{I}_2$ to estimate parameters of the production function ϕ and γ. However, nearly all datasets with measures of human capital h and an observable investment shock μ_g lack measures of human capital investments $\bar{I}_1$ and $\bar{I}_2$. We can still make progress in estimating parameters of the production function despite not observing $\bar{I}_1$ and $\bar{I}_2$, so long as we expect baseline investment levels to be similar: $\bar{I}_1 \cong \bar{I}_2$. For $\mu_g = \mu_g'$,[17] Eq. (4) reduces to:

$$\frac{1-\gamma}{\gamma}. \tag{29}$$

That is, we can observe damage to h from the shock μ_g' in second period investments relative to the damage from the first period shock μ_g, which isolates γ (while remaining silent on the magnitude of ϕ).

With an estimate of γ in hand, we can then estimate ϕ by using the total derivative in investment shocks, i.e. how human capital changes as we change *both* first and second period investments (by an equal amount). In an overlapping generations framework, this would require a shock lasting two childhood periods (or longer), and "half-exposed" cohorts on either end of the shock. Damage to the fully exposed cohort relative to the cohort exposed in period 1 alone is:

$$= \frac{1}{\gamma}\left[\frac{\gamma(\bar{I}_1+\mu_g)^{\phi}+(1-\gamma)(\bar{I}_1+\mu_g)^{\phi}}{\gamma(\bar{I}_1+\mu_g)^{\phi}+(1-\gamma)\bar{I}_1^{\phi}}\right]^{\frac{1-\phi}{\phi}}, \tag{30}$$

using the assumption that $\bar{I}_1 = \bar{I}_2$ and $\mu_g = \mu_g'$,

$$= \frac{1}{\gamma}\left[\frac{(\bar{I}_1+\mu_g)^{\phi}}{\gamma(\bar{I}_1+\mu_g)^{\phi}+(1-\gamma)\bar{I}_1^{\phi}}\right]^{\frac{1-\phi}{\phi}}, \tag{31}$$

$$= \frac{1}{\gamma}\left[\frac{1}{\gamma+(1-\gamma)\frac{\bar{I}_1}{\bar{I}_1+\mu_g}^{\phi}}\right]^{\frac{1-\phi}{\phi}}, \tag{32}$$

$$= \frac{1}{\gamma}\left[\gamma+(1-\gamma)\left(1+\frac{\mu_g}{\bar{I}_1}\right)^{-\phi}\right]^{\frac{\phi-1}{\phi}}. \tag{33}$$

[17] The assumption that $\mu_g = \mu_g'$ simplifies the algebra, but ϕ and γ can still be estimated for $\mu_g \neq \mu_g'$ so long as μ_g and μ_g' are observed. We thank Christine Pal for pointing this out.

Similarly, damage to the "doubly exposed" cohort relative to that experiencing a shock in just the *second* childhood period is:

$$= \frac{1}{1-\gamma}\left[(1-\gamma)+\gamma\left(1+\frac{\mu_g}{\bar{I}_1}\right)^{-\phi}\right]^{\frac{\phi-1}{\phi}}, \tag{34}$$

Eqs (33) and (34) constitute two equations in the two unknowns $\bar{I}_1$ and ϕ, with γ known from Eq. (29).

REFERENCES

2001. An overstretched hypothesis? The Lancet 357 (9254), 405.

Agüero, Jorge M., Carter, Michael R., Woolard, Ingrd, 2009. The impact of unconditional cash transfers on nutrition: the South African child support grant. Manuscript, February.

Aizer, Anna, 2003. Low take-up in Medicaid: does outreach matter and for whom? American Economic Review Papers and Proceedings 93 (2), 238–241.

Aizer, Anna, 2007. Public health insurance, program take-up and child health. The Review of Economics and Statistics 89 (3), 400–415.

Aizer, Anna, Currie, Janet, Moretti, Enrico, 2007. Does managed care hurt health? Evidence from Medicaid mothers. The Review of Economics and Statistics 89 (3), 385–399.

Aizer, Anna, Stroud, Laura, 2009. Education, medical knowledge and the evolution of disparities in health. manuscript, Brown University, September.

Aizer, Anna, Stroud, Laura, Buka, Stephen, 2009. Maternal stress and child well-being: evidence from siblings. manuscript. Brown University, February.

Almond, Douglas, 2006. Is the 1918 Influenza Pandemic over? Long-term effects of *in utero* influenza exposure in the post-1940 US population. Journal of Political Economy 114 (4), 672–712.

Almond, Douglas, Chay, Kenneth Y., Lee, David S., 2005. The costs of low birth weight. The Quarterly Journal of Economics 120 (3), 1031–1084.

Almond, Douglas, Edlund, Lena, Palme, Mårten, 2009. Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in sweden. Quarterly Journal of Economics 124 (4), 1729–1772. URL http://www.mitpressjournals.org/doi/abs/10.1162/qjec.2009.124.4.1729.

Almond, Douglas, Hoynes, Hilary W., Schanzenbach, Diane Whitmore, 2010. Inside the war on poverty: the impact of food stamps on birth outcomes. The Review of Economics and Statistics (forthcoming).

Almond, Douglas, Mazumder, Bhashkar, 2005. The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data. American Ecnomic Review, Papers and Proceedings 95 (2).

Almond, Douglas, Mazumder, Bhashkar, 2008. Health capital and the prenatal environment: the effect of maternal fasting during pregnancy. Working Paper 14428, National Bureau of Economic Research, October, URL http://www.nber.org/papers/w14428.

Anderson, Kathryn H., Foster, James E., Frisvold, David E., 2009. Investing in health: the long-term impact of Head Start on smoking. Working Paper 1465-7295, Western Economic Association International.

Anderson, Michael, 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. Journal of the American Statistical Association 103 (484), 1481–1495.

Angrist, Joshua D., Pischke, Jörn-Steffen, 2009. Mostly Harmless Econometrics: an Empiricist's Companion. Princeton University Press, Princeton, New Jersey.

Aos, Steve, Lieb, Roxanne, Mayfield, Jim, Miller, Marna, Pennucci, Annie, 2004. Benefits and costs of prevention and early intervention programs for youth. Technical appendix, Washington State Institute for Public Policy, September.

Ashenfelter, Orley, Rouse, Cecilia, 1998. Income, schooling, and ability: evidence from a new sample of identical twins. Quarterly Journal of Economics 113 (1), 253–284.
URL http://www.mitpressjournals.org/doi/abs/10.1162/003355398555577.

Baker, Kevin, 2008. Do cash transfer programs improve infant health: evidence from the 1993 expansion of the earned income tax credit. Manuscript, University of Notre Dame.

Baker, Michael, Gruber, Jonathan, Milligan, Kevin, 2008. Universal child care, maternal labor supply, and family well-being. Journal of Political Economy 116 (4), 709–745.
URL http://www.journals.uchicago.edu/doi/abs/10.1086/591908.

Baker, Michael, Milligan, Kevin, 2010. Evidence from maternity leave expansions of the impact of maternal care on early child development. Journal of Human Resources 45 (1), 1–32.

Baldwin, Laura-Mae, Larson, Eric H., Connell, Frederick A., Nordlund, Daniel, Cain, Kevin C., Cawthon, Mary L., Byrns, Patricia, Rosenblatt, Roger A., 1998. The effect of expanding Medicaid prenatal services on birth outcomes. American Journal of Public Health 88 (11), 1623–1629.

Banerjee, Abhijit, Duflo, Esther, Postel-Vinay, Gilles, Watts, Timothy, 2010. Long run health impacts of income shocks: wine and phylloxera in 19th century France. The Review of Economics and Statistics (forthcoming).

Banzhaf, Spencer, Walsh, Randall P., 2008. Do people vote with their feet? an empirical test of tiebout. American Economic Review 98 (3), 843–863.

Barker, D.J.P. (Ed.), 1992. Fetal and infant origins of adult disease. British Medical Journal.

Barnett, W.S., Masse, Leonard N., 2007. Comparative cost-benefit analysis of the Abecedarian program and its policy implications. Economics of Education Review 26, 113–125.

Barnett, Steven, Belfield, Clive R., Nores, Milagros, Schweinhart, Lawrence, 2006. The High/Scope Perry Preschool program: cost-benefit analysis using data from the age-40 follow-up. The Journal of Human Resources 41 (1), 162–190.

Barreca, Alan, 2010. The long-term economic impact of *in utero* and postnatal exposure to malaria. Journal of Human Resources 45 (4), 865–892.

Basiotis, Peter, Kramer-LeBlanc, Carol S., Kennedy, Eileen T., 1998. Maintaining nutrition security and diet quality: the role of the Food Stamp program and WIC. Family Economics 11 (1-2), 4–16.

Baten, Jörg, Crayen, Dorothee, Voth, Joachim, 2007. Poor, hungry and stupid: numeracy and the impact of high food prices in industrializing Britain, 1780-1850. Economics Working Papers 1120, Department of Economics and Business, Universitat Pompeu Fabra, October.
URL http://ideas.repec.org/p/upf/upfgen/1120.html.

Baum, Charles L., 2008. The effects of food stamp benefits on weight gained by expectant mothers. Discussion Paper 1343-08, Institute for Research on Poverty, September.

Becker, Gary S., Tomes, Nigel, 1976. Child endowments and the quantity and quality of children. The Journal of Political Economy 84 (4), S143–S162.

Behrman, Richard E., Butler, Adrienne S. (Eds.), 2007. Preterm Birth: Causes, Consequences, and Prevention. Institute of Medicine Report. National Academies Press.

Behrman, Jere R., Pollak, Robert A., Taubman, Paul, 1982. Parental preferences and provision for progeny. Journal of Political Economy 90 (1), 52–73.

Behrman, Jere R., Rosenzweig, Mark R., 2004. Returns to birthweight. The Review of Economics and Statistics 86 (2), 586–601.

Behrman, Jere R., Rosenzweig, Mark R., Taubman, Paul, 1994. Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. Journal of Political Economy 102 (6), 1131–1174.

Van Den Berg, Gerard J., Lindeboom, Maarten, Portrait, France, 2006. Economic conditions early in life and individual mortality. American Economic Review 96 (1), 290–302.

Van Den Berg, Gerard J., Lindeboom, Maarten, Portrait, France, 2010. Long-run longevity effects of a nutritional shock early in life: the Dutch potato famine of 1846-1847. Journal of Health Economics (in press).

Berger, Lawrence M., Hill, Jennifer, Waldfogel, Jane, 2005. Maternity leave, early maternal employment, and child health and development in the us. The Economic Journal 115, F29–F47.

Berger, Lawrence M., Paxson, Christina, Waldfogel, Jane, 2009. Income and child development. Children and Youth Services Review 31 (9), 978–989.

Bingley, Paul, Walker, Ian, 2007. There's no such thing as a free lunch: altruistic parents and the response of household food expenditures to nutrition program reforms. Discussion Paper 2007-19, University College Dublin Geary Institute, June.

Bitler, Marianne, Currie, Janet, 2005. Does WIC work? the effect of WIC on pregnancy and birth outcomes. Journal of Policy Analysis and Management 24 (1), 73–91.

Black, Maureen M., Cutts, Diana B., Frank, Deborah A., Geppert, Joni, Skalicky, Anne, Levenson, Suzette, Casey, Patrick H., Berkowitz, Carol, Zaldivar, Nievves, Cook, John T., Meyers, Alan F., Herren, Tim, 2004. Special supplemental nutrition program for women, infants, and children participation and infants' growth and health: a multi-site surveillance study. Pediatrics 114 (1), 169–176.

Black, Sandra E., Devereux, Paul J., Salvanes, Kjell G., 2005. The more the merrier? The effect of family size and birth order on children's education. Quarterly Journal of Economics 120 (2), 669–700.

Black, Sandra E., Devereux, Paul J., Salvanes, Kjell G., 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. Quarterly Journal of Economics 122 (1), 409–439.

Blanden, Jo, Gregg, Paul, MacMillan, Lindsey, 2006. Explaining intergenerational income persistence: non-cognitive skills, ability and education. CMPO Working Paper Series 06/146, University of Bristol, The Centre for Market and Public Organisation, April. URL www.bris.ac.uk/Depts/CMPO/.

Blau, David, Currie, Janet, 2006. Pre-school, day care, and after-school care: Who's minding the kids? 2: 1163–1278.

Bleakley, Hoyt, 2007. Disease amd development: evidence from hookworm eradication in the south. The Quarterly Journal of Economics 122 (1), 73.

Blow, Laura, Walker, Ian, Zhu, Yu, 2005. Who benefits from child benefit. manuscript, February.

Blundell, Richard, 2006. Earned income tax credit policies: impact and optimality: the Adam Smith lecture, 2005. Labour Economics (ISSN: 0927-5371) 13 (4), 423–443. European Association of Labour Economists, 2nd World Conference SOLE/EALE, Fairmont Hotel San Francisco, USA, 2-5 June 2005.

Boyer, George R., 1990. An Economic History of the English Poor Law. CUP, Cambridge, 1750–1850.

Bozzoli, Carlos, Deaton, Angus S., Quintana-Domeque, Climent, 2010. Child mortality, income and adult height. NBER Working Paper No. 12966, March. Demography (forthcoming).

Brooks-Gunn, Jeanne, Markham, Lisa B., 2005. The contribution of parenting to ethnic and racial gaps in school readiness. The Future of Children 15 (1), 139–168.

Butler, J.S., Raymond, Jennie E., 1996. The effect of the Food Stamp program on nutrient intake. Economic Inquiry 34, 781–798.

Card, David, Shore-Sheppard, Laura, 2004. Using discontinuous eligibility rules to identify the effects of the federal Medicaid expansions on low-income children. Review of Economics and Statistics 86 (3), 752–766.

Carneiro, Pedro, Ginja, Rita, 2008. Preventing behavior problems in childhood and adolescence: evidence from Head Start. manuscript, March.

Cascio, Elizabeth U., 2009. Do investments in universal early education pay off? long-term effects of introducing kindergartens into public schools. Working Paper 14951, National Bureau of Economic Research, May.

Case, Anne, Fertig, Angela, Paxson, Christina, 2005. The lasting impact of childhood health and circumstance. Journal of Health Economics 24, 365–389.

Case, Anne, Lee, Diana, Paxson, Christina, 2008. The income gradient in children's health: a comment on Currie, Shields, and Wheatley price. Journal of Health Economics 27 (3), 801–807.

Case, Anne, Lubotsky, Darren, Paxson, Christina, 2002. Economic status and health in childhood: the origins of the gradient. American Economic Review 92 (5), 1308–1334.

Case, Anne, Paxson, Christina, 2008a. Height, health, and cognitive function in old age. American Economic Review: Papers & Proceedings 98 (2), 463–467.

Case, Anne, Paxson, Christina, 2008b. Stature and status: height, ability, and labor market outcomes. Journal of Political Economy 116 (3), 499–532.

Case, Anne, Paxson, Christina, 2009. Early life health and cognitive function in old age. American Economic Review Papers and Proceedings 99 (2), 104–109.

Case, Anne, Paxson, Christina, 2010a. Causes and consequences of early life health. Working Paper 15637, National Bureau of Economic Research, January. URL http://www.nber.org/papers/w15637.

Case, Anne, Paxson, Christina, 2010b. The long reach of childhood health and circumstance: evidence from the Whitehall II study. Working Paper 15640, National Bureau of Economic Research, January. URL http://www.nber.org/papers/w15640.

Chatterji, Pinka, Brooks-Gunn, Jeanne, 2004. Wic participation, breastfeeding practices, and well-child care among unmarried, low-income mothers. American Journal of Public Health 94 (8), 1324–1327.

Chatterji, Pinka, Markowitz, Sarah, 2000. The impact of maternal alcohol and illicit drug use on children's behavior problems: evidence from the children of the national longitudinal survey of youth. Working Paper 7692, National Bureau of Economic Research.

Chatterji, Pinka, Bonuck, Karen, Dhawan, Simi, Deb, Nandini, 2002. WIC participation and the initiation and duration of breastfeeding. Discussion Paper No. 1246-02. Madison, WI: Institute for Research on Poverty.

Chay, Kenneth Y., Greenstone, Michael, 2003a. Air quality, infant mortality, and the Clean Air Act of 1970. Working Paper 10053, National Bureau of Economic Research, October. URL http://www.nber.org/papers/w10053.

Chay, Kenneth Y., Greenstone, Michael, 2003b. The impact of air pollution on infant mortality: evidence from the geographic variation in pollution shocks induced by a recession. Quarterly Journal of Economics 118 (3), 1121–1167.

Chay, Kenneth Y., Guryan, Jonathan, Mazumder, Bhashkar, 2009. Birth cohort and the black-white achievement gap: the roles of access and health soon after birth. Working Paper 15078, National Bureau of Economic Research, June. URL http://www.nber.org/papers/w15078.

Chiang, Alpha C., 1984. Fundamental Methods of Mathematical Economics, third edition, McGraw-Hill, New York.

Condliffe, Simon, Link, Charles R., 2008. The relationship between economic status and child health: evidence from the United States. American Economic Review 98 (4), 1605–1618.

Conley, Dalton, Bennett, Neil G., 2001. Birth weight and income: interactions across generations. Journal of Health and Social Behavior (ISSN: 00221465) 42 (4), 450–465. URL http://www.jstor.org/stable/3090189.

Conley, Dalton, Pfeiffer, Katherine M., Velez, Melissa, 2007. Explaining sibling differences in achievement and behavioral outcomes: the importance of within- and between-family factors. Social Science Research 36 (3), 1087–1104.

Conley, Dalton, Strully, Kate W., Bennett, Neil G., 2006. Twin differences in birth weight: the effects of genotype and prenatal environment on neonatal and post-natal mortality. Economics and Human Biology 4 (2), 151–183.

Costa, Dora L., Lahey, Joanna N., 2005. Predicting older age mortality trends. Journal of the European Economic Association 3 (2–3), 487–493.

Crimmins, Eileen M., Finch, Caleb E., 2006. Infection, inflammation, height, and longevity. Proceedings of the National Academy of Sciences 103 (2), 498–503.

Cunha, Flavio, Heckman, James, 2007. The technolgy of skill formation. American Economic Review 97 (2), 31–47.

Cunha, Flavio, Heckman, James J., 2008. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. Journal of Human Resources 43 (4), 738–782. URL http://jhr.uwpress.org/cgi/content/abstract/43/4/738.

Cunha, Flavio, Heckman, James, Schennach, Susanne, 2010. Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78 (3), 883–931.

Currie, Alison, Shields, Michael, Price, Stephen W., 2007. Is the child health/family income gradient universal? Evidence from England. Journal of Health Economics 26 (2), 213–232.

Currie, Janet, 2003. US food and nutrition programs. In: Moffitt, Robert (Ed.), Means-tested Transfer Programs in the United States. University of Chicago Press for NBER, Chicago, IL, pp. 199–290.

Currie, Janet, Cole, Nancy, 1993. Welfare and child health: the link between AFDC participation and birth weight. The American Economic Review 83 (4), 971–985.

Currie, Janet, Gahvari, Firouz, 2008. Transfers in cash and in-kind; theory meets the data. Journal of Economic Literature 46 (2), 333–383.

Currie, Janet, Grogger, Jeffrey, 2002. Medicaid expansions and welfare contractions: offsetting effects on prenatal care and infant health. Journal of Health Economics 21 (2), 313–335.

Currie, Janet, Gruber, Jonathan, 1996a. Health insurance eligibility, utilization of medical care, and child health. The Quarterly Journal of Economics 111 (2), 431–466.

Currie, Janet, Gruber, Jonathan, 1996. Saving babies: the efficacy and cost of recent changes in the Medicaid eligibility of pregnant women. Journal of Political Economy 104 (6), 1263. URL http://www.journals.uchicago.edu/doi/abs/10.1086/262059.

Currie, Janet, Hyson, Rosemary, 1999. Is the impact of shocks cusioned by socioeconomic status? The case of low birth weight. American Economic Review 89 (2), 245–250.

Currie, Janet, Moretti, Enrico, 2007. Biology as destiny? short- and long-run determinants of intergenerational transmission of birth weight. Journal of Labor Economics 25 (2), 231–264. URL http://www.journals.uchicago.edu/doi/abs/10.1086/511377.

Currie, Janet, Moretti, Enrico, 2008. Short and long-run effects of the introduction of food stamps on birth outcomes in California. In: Schoeni, R.F., House, J., Kaplan, G., Pollack, H. (Eds.), Making Americans Healthier: Social and Economic Policy as Health Policy. Russell Sage, New York.

Currie, Janet, Neidell, Matthew, 2005. Air pollution and infant health: what can we learn from California's recent experience? Quarterly Journal of Economics 120 (3), 1003–1030. URL http://www.mitpressjournals.org/doi/abs/10.1162/003355305774268219.

Currie, Janet, Stabile, Mark, 2003. Socioeconomic status and child health: why is the relationship stronger for older children? American Economic Review 93 (5), 1813–1823.

Currie, Janet, Stabile, Mark, 2006. Child mental health and human capital accumulation: the case of ADHD. Journal of Health Economics 25 (6), 1094–1118.

Currie, Janet, Tekin, Erdal, 2006. Does child abuse cause crime? Working Paper 12171, National Bureau of Economic Research, April. URL http://www.nber.org/papers/w12171.

Currie, Janet, Thomas, Duncan, 1995. Does Head Start make a difference? American Economic Review 85 (3), 341–364.

Currie, Janet, Thomas, Duncan, 1999a. Does Head Start help Hispanic children? Journal of Public Economics (ISSN: 0047-2727) 74 (2), 235–262.

Currie, Janet, Thomas, Duncan, 1999b. Early test scores, socioeconomic status and future outcomes. Working Paper 6943, National Bureau of Economic Research, February.
URL http://www.nber.org/papers/w6943.

Currie, Janet, Walker, Reed, 2009. Traffic congestion and infant health: evidence from E-Zpass. Working Paper 15413, National Bureau of Economic Research, October.
URL http://www.nber.org/papers/w15413.

Currie, Janet, Widom, Cathy S., 2009. Long-term consequences of child abuse and neglect on adult economic well-being. Manuscript, September.

Currie, Janet, Yelowitz, Aaron, 2000. Are public housing projects good for kids? Journal of Public Economics (ISSN: 0047-2727) 75 (1), 99–124.

Currie, Janet, Decker, Sandra, Lin, Wanchuan, 2008. Has public health insurance for older children reduced disparities in access to care and health outcomes? Journal of Health Economics 27 (6), 1567–1581.

Currie, Janet, Neidell, Matthew, Schmieder, Johannes F., 2009. Air pollution and infant health: lessons from New Jersey. Journal of Health Economics 28 (3), 688–703.

Currie, Janet, Stabile, Mark, Manivong, Phongsack, Roos, Leslie L., 2010. Child health and young adult outcomes. The Journal of Human Resources 45 (3), 517–548.

Cutler, David M., Miller, Grant, Norton, Douglas M., 2007. Evidence on early-life income and late-life health from America's Dust Bowl era. Proceedings of the National Academy of Sciences 104 (33), 13244–13249. URL http://www.pnas.org/content/104/33/13244.abstract.

Dafny, Leemore, Gruber, Jonathan, 2005. Public insurance and child hospitalizations: access and efficiency effects. Journal of Public Economics 89 (1), 109–129.
URL http://ideas.repec.org/a/eee/pubeco/v89y2005i1p109-129.html.

Dahl, Gordon, Lochner, Lance, 2008. The impact of family income on child achievement: evidence from the earned income tax credit. Working Paper 14599. National Bureau of Economic Research, December. URL http://www.nber.org/papers/w14599.

Dahl, Gordon B., Lochner, Lance, 2005. The impact of family income on child achievement. Working Paper 11279, National Bureau of Economic Research, April. URL http://www.nber.org/papers/w11279.

Dalenius, Tore, Reiss, Steven P., 1982. Data swapping: a technique for disclosure control. Journal of Statistical Planning and Inference 6 (1), 73–85.
Datar, Ashlesha, Kilburn, Rebecca, Loughran, David S., 2010. Endowments and parental investments in infancy and early childhood. Demography 47 (1), 145–162.
Dearden, Lorraine, 1998. Ability, family, education, and earnings in Britain. Working Paper 98/114, Institute for Fiscal Studies, July.
Del Bono, Emilia, Ermisch, John, Francesconi, Marco, 2008. Intrafamily resource allocations: a dynamic model of birth weight. IZA Discussion Papers 3704, Institute for the Study of Labor (IZA), September. URL http://ideas.repec.org/p/iza/izadps/dp3704.html.
Deming, David, 2009. Early childhood intervention and life-cycle skill development. American Economic Journal: Applied Economics 1 (3), 111–134.
Deschnes, Olivier, Greenstone, Michael, 2007. Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the US. NBER Working Paper No. 13178, June.
Doblhammer, Gabriele, Vaupel, James W., 2001. Lifespan depends on month of birth. Proceedings of the National Academy of Sciences 98 (5), 2934–2939.
Doyle Jr., Joseph J., 2008. Child protection and adult crime: using investigator assignment to estimate causal effects of foster care. Journal of Political Economy 116 (4), 746–770.
URL http://www.journals.uchicago.edu/doi/abs/10.1086/590216.
Doyle, Orla, Walker, Ian, Harmon, Colm P., 2005. The impact of parental income and education on the health of their children. Discussion Paper 5359, Centre for Economic Policy Research, December.
Dubay, Lisa, Joyce, Ted, Kaestner, Robert, Kenney, Genevieve M., 2001. Changes in prenatal care timing and low birth weight by race and socioeconomic status: implications for the medicaid expansions for pregnant women. Health Services Research 36 (2), 373–398.
Duggan, Anne K., McFarlane, Elizabeth, Fuddy, Loretta, Burrell, Lori, Higman, Susan M., Windham, Amy, Sia, Calvin, 2004. Randomized trial of a statewide home visiting program: impact on preventing child abuse and neglect. Child Abuse and Neglect 28 (6), 597–622.
Duggan, A.K., McFarlane, E.C., Windham, A.M., Rohde, C., Salkever, D.S., Fuddy, L., Rosenberg, L.A., Buchbinder, S.B., Sia, C.C.J., 1999. Evaluation of Hawaii's Healthy Start Program. The Future of Children 9 (1), 66–90.
DuMont, Kimberly, Mitchell-Herzfeld, Susan, Greene, Rose, Lee, Eunju, Lowenfels, Ann, Rodriguez, Monica, 2006. Healthy families New York (hfny) randomized trial: impacts on parenting after the first two years. Working Paper Series: evaluating Healthy Families New York 1, New York State Office of Children & Family Services, June.
Dustmann, Christian, Schönberg, Uta, 2009. The effect of expansions in maternity leave coverage on children's long-term outcomes. SSRN eLibrary.
Eissa, Nada, Liebman, Jeffrey B., 1996. Labor supply response to the earned income tax credit. The Quarterly Journal of Economics 113 (3), 653–691.
Farrelly, Matthew C., Nimsch, Christian T., James, Joshua, 2003. State cigarette excise taxes: Implications for revenue and tax evasion. Report 08742.000, RTI International, Research Triangle Park, NC, May.
Fergusson, David M., Grant, Hildegard, Horwood, L. John, Ridder, Elizabeth M., 2006. Randomized trial of Early Start program of home visitation. Pediatrics 117 (3), 803–809.
Fertig, Angela R., Reingold, David A., 2007. Public housing, health, and health behaviors: Is there a connection? Journal of Policy Analysis and Management 26 (4), 831–859.
Field, Erica, Robles, Omar, Torero, Maximo, 2009. Iodine deficiency and schooling attainment in Tanzania. American Economic Journal: Applied Economics 1 (4), 140–169.
Figlio, David, Hamersma, Sarah, Roth, Jeffrey, 2009. Does prenatal WIC participation improve birth outcomes? new evidence from Florida. Journal of Public Economics 93 (1–2), 235–245.
Fletcher, Jason M., Lehrer, Steven F., 2009. Using genetic lotteries within families to examine the causal impact of poor health on academic achievement. Working Paper 15148, National Bureau of Economic Research, July. URL http://www.nber.org/papers/w15148.
Frank, Richard G., Meara, Ellen, 2009. The effect of maternal depression and substance abuse on child human capital development. Working Paper 15314, National Bureau of Economic Research, September.
Frisvold, David E., 2006. Head Start participation and childhood obesity. Working Paper 06-WG01, Vanderbilt University, February.

Frisvold, David E., Lumeng, Julie C., 2009. Expanding exposure: can increasing the daily duration of Head Start reduce childhood obesity? manuscript, April.
Garces, Eliana, Thomas, Duncan, Currie, Janet, 2002. Longer-term effects of Head Start. American Economic Review 92 (4), 999–1012.
Gertler, Paul J., 2004. Do conditional cash transfers improve child health? Evidence from Progresa's control randomized experiment. Health, Health Care, and Economic Development 94 (2), 336–341.
Gluckman, Peter D., Hanson, Mark A., 2006. Adult disease: echoes of the past. European Journal of Endocrinology 155 (Suppl. 1), S47–S50.
Gormley Jr., William, 2008. The effects of Oklahoma's pre-K program on Hispanic children. Social Science Quarterly 89 (4), 916–936.
Gormley Jr., William, Gayer, Ted, 2006. Promoting school readiness in Oklahoma: an evaluation of Tulsa's pre-K program. The Journal of Human Resources 40 (3), 533–558.
Gormley Jr., William, Gayer, Ted, Phillips, Deborah, Dawson, Brittany, 2005. The effects of universal pre-K on cognitive development. Developmental Psychology 41 (6), 872–884.
Goux, Dominique, Maurin, Eric, 2005. The effects of overcrowded housing on children's performance at school. Journal of Public Economics 89 (5-6), 797–819.
Griliches, Zvi, 1979. Sibling models and data in economics: beginnings of a survey. Journal of Political Economy 87 (s5), S37. URL http://www.journals.uchicago.edu/doi/abs/10.1086/260822.
Grossman, Michael, 1972. On the concept of health capital and the demand for health. Journal of Political Economy 223–255.
Gueorguieva, Ralitza, Morse, Steven B., Roth, Jeffrey, 2009. Length of prenatal participation in WIC and risk of delivering a small for gestational age infant: Flordida, 1996-2004. Journal of Maternal and Child Health 13 (4), 479–488.
Gundersen, Craig, Ziliak, James, 2004. Poverty and macroeconomic performance across space, race, and family structure. Demography 41 (1), 61–86.
Hanratty, Maria J., 1996. Canadian national health insurance and infant health. American Economic Review (ISSN: 00028282) 86 (1), 276–284. URL http://www.jstor.org/stable/2118267.
Harding, Kathryn, Galano, Joseph, Martin, Joanne, Huntington, Lee, Schellenbach, Cynthia J., 2007. Healthy Families America effectiveness: a comprehensive review of outcomes. Journal of Prevention and Intervention in the Community 34 (1-2), 149–179.
Havnes, Tarjei, Mogstad, Magne, 2009. No child left behind: universal child care and children's long-run outcomes. Discussion Paper 582, Statistics Norway, Research Department, May.
Heckman, James, 2007. The economics, technology, and neuroscience of human capability formation. PNAS 104 (33), 13250–13255.
Heckman, James, Moon, Seong H., Pinto, Rodrigo, Savelyev, Peter A., Yavitz, Adam, 2009. The rate of return to the High/Scope Perry preschool program. Working Paper 15471, National Bureau of Economic Research, November.
Heckman, James J., Rubinstein, Yona, 2001. The importance of noncognitive skills: lessons from the GED testing program. American Economic Review: Papers & Proceedings 91 (2), 145–149.
Heider, Fritz, 1934. The influence of the epidemic of 1918 on deafness: a study of birth dates of pupils registered in schools for the deaf. American Journal of Epidemiology 19, 756–765.
Howard, Kimberly S., Brooks-Gunn, Jeanne, 2009. The role of home visiting programs in preventing child abuse and neglect. The Future of Children 19 (2), 119–146.
Hoynes, Hilary W., Page, Marianne, Stevens, Ann H., 2009. Is a WIC start a better start? evaluating WIC's impact on infant health using program introduction. Working Paper 15589, National Bureau of Economic Research, December.
Hoynes, Hilary W., Schanzenbach, Diane, 2009. Consumption responses to in-kind transfers: evidence from the introduction of the Food Stamp program. American Economic Journal: Applied Economics 1 (4), 109–139.
Hsin, Amy, 2009. Is biology destiny? Birth weight and differential parental treatment. Manuscript, Population Studies Center, University of Michigan.
Hustedt, Jason T., Barnett, W. Steven, Jung, Kwanghee, Figueras, Alexandra, 2008. Impacts of new Mexico pre-K on children's school readiness at kindergarten entry: Results from the second year of a growing initiative. Research report, National Institute for Early Education, Rutgers University, June.

Jacob, Brian, 2004. Public housing, housing vouchers, and student achievement: evidence from public housing demolitions in Chicago. American Economic Review 94 (1), 233–258.
Johnson, Rucker C., Schoeni, Robert F., 2007. The influence of early-life events on human capital, health status, and labor market outcomes over the life course. Working Paper 07-05, National Poverty Center, February.
Joyce, Theodore, Gibson, Diane, Colman, Sylvie, 2004. The changing association between prenatal participation in WIC and birth outcomes in New York city. Working Paper 10796, National Bureau of Economic Research, September.
Joyce, Theodore, Racine, Andrew D., Yunzal-Butler, Cristina, 2007. Reassessing the WIC effect: evidence from the pregnancy nutrition surveillance system. Journal of Policy Analysis and Management 87 (2), 277–303.
Kaestner, Robert, Joyce, Ted, Racine, Andrew, 2001. Medicaid eligibility and the incidence of ambulatory care sensitive hospitalizations for children. Social Science and Medicine 52 (2), 305–313.
Karoly, Lynn A., Kilburn, Rebecca M., Cannon, Jill S., 2005. Early Childhood Interventions: Proven Results, Future Promise. RAND Corporation.
Karoly, Lynn A., Greenwood, Peter W., Everingham, Susan S., Hoube, Jill, Kilburn, M. Rebecca, Rydell, C. Peter, Sanders, Matthew, Chiesa, James, 1998. Investing in our children: what we know and don't know about the costs and benefits of early childhood interventions. Consensus report, Rand, July.
Kehrer, Barbara H., Wolin, Charles M., 1979. Impact of income maintenance on low birth weight: evidence from the Gary experiment. The Journal of Human Resources 14 (4), 434–462. autumn.
Kelly, Elaine, 2009. The scourge of Asian flu: *in utero* exposure to pandemic influenza and the development of a cohort of British children. Institute for Fiscal Studies, Working Paper 09/17, September. University College London.
Kermack, W.O., McKendrick, A.G, McKinlay, P.L., 1934. Death-Rates in Great Britain and Sweden. The Lancet 698–703.
Khanam, Rasheda, Nghiem, Hong S., Connelly, Luke B., 2009. Child health and the income gradient: evidence from Australia. Journal of Health Economics 28 (4), 805–817.
Kling, Jeffrey R., Liebman, Jeffrey B., Katz, Lawrence F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.
Kling, Jeffrey R., Ludwig, Jeffrey, Katz, Lawrence F., 2005. Neighborhood effects on crime for female and male youth: evidence from a randomized housing voucher experiment. The Quarterly Journal of Economics 120 (1), 87–130.
Knittel, Christopher R., Miller, Douglas L., Sanders, Nicholas J., 2009. Caution drivers! children present: Traffic, pollution and infant health. Manuscript, UC Davis.
Koch, Thomas G., 2009. Public insurance, crowd-out, and health care utilization. Manuscript, August.
Kowaleski-Jones, Lori, Duncan, Greg J., 2002. Effects of participation in the WIC program on birthweight: evidence from the National Longitudinal Survey of Youth. Am J Public Health 92 (5), 799–804. URL http://ajph.aphapublications.org/cgi/content/abstract/92/5/799.
Kraemer, Sebastian, 2000. The fragile male. BMJ 321 (7276), 1609–1612. URL http://www.bmj.com.
Lee, Bong J., Mackey-Bilaver, Lee, 2007. Effects of WIC and Food Stamp program participation on child outcomes. Children and Youth Services Review 29 (4), 501–517.
Lefebvre, Pierre, Merrigan, Philip, Verstraete, Matthieu, 2006. Impact of early childhood care and education on children's preschool cognitive development: Canadian results from a large quasi-experiment. Working Paper 06-36, Centre interuniversitaire sur le risque, les politiques conomiques et lemploi (CIRPEE), October.
Levine, Philip B., Schanzenbach, Diane W., 2009. The impact of children's public health insurance expansions on educational outcomes. Forum for Health Economics and Policy 12 (1), Article 1.
Levine, Phillip B., Zimmerman, David J., 2000. Children's welfare exposure and subsequent development. Working Paper 7522, National Bureau of Economic Research, February.
URL http://www.nber.org/papers/w7522.
Lhila, Aparna, Simon, Kosali I., 2008. Prenatal health investment decisions: does the childs sex matter? Demography 45 (4), 885–905.
Lien, Diana S., Evans, William N., 2005. Estimating the Impact of large cigarette tax hikes: the case of maternal smoking and infant birth weight. J. Human Resources XL (2), 373–392.
URL http://jhr.uwpress.org/cgi/content/abstract/XL/2/373.

Lin, Ming-Jen, Liu, Jin-Tan, 2009. Do lower birth weight babies have lower grades? twin fixed effect and instrumental variable evidence from Taiwan. Social Science & Medicine 68 (10), 1780–1787.

Logan, W.P.D., Glasg, M.D., 1953. Mortality in London fog incident, 1952. Lancet 1 (1), 336–338.

Love, John M., Kisker, Ellen E., Ross, Christine, Constantine, Jill, Boller, Kimberly, Chazan-Cohen, Rachel, Brady-Smith, Christy, Fuligni, Allison S., Raikes, Helen, Brooks-Gunn, Jeanne, Tarullo, Louisa B., Schochet, Peter Z., Paulsell, Diane, Vogel, Cheri, 2005. The effectiveness of Early Head Start for 3-year-old children and their parents: lessons for policy and programs. Developmental Psychology 41 (6), 885–901.

Ludwig, Jens, Miller, Douglas L., 2005. Does Head Start improve children's life chances? evidence from a regression discontinuity design. Discussion Paper 1311-05, Institute for Research on Poverty, November.

Ludwig, Jens, Miller, Douglas L., 2007. Does Head Start improve children's life chances? evidence from a regression discontinuity design. Quarterly Journal of Economics 488 (1), 159–208.

MacDorman, Marian F., Martin, Joyce A., Mathews, T.J., Hoyert, Donna L., Ventura, Stephanie J., 2005. Explaining the 2001-2002 infant mortality increase in the United States: data from the linked birth/infant death data set. International Journal of Health Services 35 (3), 415–442.

MacMillan, Harriet L., Thomas, B. Helen, Jamieson, Ellen, Walsh, Christine A., Boyle, Michael H., Shannon, Harry S., Gafni, Amiram, 2005. Effectiveness of home visitation by public health nurses in prevention of the recurrence of child physical abuse: a randomized controlled trial. Lancet 365 (9473), 1786–1793.

Macours, Karen, Schady, Norbert, Vakis, Renos, 2008. Cash transfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment. Working Paper 4759, The World Bank: Development Research Group, Human Development and Public Services Team, October.

Magnuson, Katherine A., Ruhm, Christopher, Waldfogel, Jane, 2007. Does pre-kindergarten improve school preparation and performance? Economics of Education Review 26 (1), 33–51.

Masse, Leonard, Barnett, W. Steven, 2002. A Benefit-cost Analysis of the Abecedarian Early Childhood Intervention. National Institute for Early Education Research, Rutgers, NJ.

McCormick, Marie C., Brooks-Gunn, Jeanne, Buka, Stephen L., Goldman, Julie, Yu, Jennifer, Salganik, Mikhail, Scott, David T., Bennett, Forrest C., Kay, Libby L., Bernbaum, Judy C., Bauer, Charles R., Martin, Camilia, Woods, Elizabeth R., Martin, Anne, Casey, Patrick H., 2006. Early intervention in low birth weight premature infants: results at 18 years of age for the infant health and development. Pediatrics 117 (3), 771–780.

McIntosh, Emma, Barlow, Jane, Davis, Hilton, Stewart-Brown, Sarah, 2009. Economic evaluation of an intensive home visiting programme for vulnerable families: a cost-effectiveness analysis of public health intervention. Journal of Public Health 31 (3), 423–433.

McLeod, Jane, Kaiser, Karen, 2004. Childhood emotional and behavioral problems and educational attainment. American Sociological Review 69 (5), 636–658.

Meyer, Bruce D., Rosenbaum, Dan T., 2001. Welfare, the earned income tax credit, and the labor supply of single mothers. Quarterly Journal of Economics 116 (3), 1063–1114.
URL http://www.mitpressjournals.org/doi/abs/10.1162/00335530152466313.

Milligan, Kevin, Stabile, Mark, 2008. Do child tax benefits affect the wellbeing of children? evidence from Canadian child benefit expansions. Working Paper 14624, National Bureau of Economic Research, December. URL http://www.nber.org/papers/w14624.

Morris, Pamela, Duncan, Greg J., Rodrigues, Christopher, 2004. Does money really matter? estimating impacts of family income on children's achievement with data from random-assignment experiments. Manuscript.

Newman, Sandra J., Harkness, Joseph M., 2002. The long-term effects of public housing on self-sufficiency. Journal of Policy Analysis and Management 21 (1), 21–43.

Nilsson, Peter, 2008. Does a pint a day affect your child's pay? the effect of prenatal alcohol exposure on adult outcomes. Centre for Microdata Methods and Practice Working Paper CWP22/08, The Institute for Fiscal Studies, August.

Nilsson, Peter, 2009. e long-term effects of early childhood lead expsoure: evidence from sharp changes in local air lead levels induced by the phase-out of leaded gasoline. manuscript, Uppsala University, September.

Oaxaca, Ronald L., 1973. Male-female wage differentials in urban labor markets. International Economic Review 14, 693–709.

Olds, David L., Henderson, Charles R., Kitzman, Harriet J., Eckenrode, John J., Cole, Robert E., Tatelbaum, Robert C., 1999. Prenatal and infancy home visitation by nurses: recent findings. The Future of Children 9 (1), 44–65. Spring-Summer.

Olds, David L., Kitzman, Harriet, Hanks, Carole, Cole, Robert, Anson, Elizabeth, Sidora-Arcoleo, Kimberly, Luckey, Dennis W., Henderson Jr., Charles R., Holmberg, John, Tutt, Robin A., Stevenson, Amanda J., Bondy, Jessica, 2007. Effects of nurse home visiting on maternal and child functioning: age-9 follow-up of a randomized trial. Pediatrics 120 (4), 832–845.

Olds, David L., Robinson, JoAnn, O'Brien, Ruth, Luckey, Dennis W., Pettitt, Lisa M., Henderson Jr., Charles R., Ng, Rosanna K., Sheff, Karen L., Korfmacher, Jon, Hiatt, Susan, Talmi, Ayelet, 2002. Home visiting by paraprofessionals and by nurses: a randomized controlled trial. Pediatrics 110 (3), 486–496. URL http://pediatrics.aappublications.org/cgi/content/abstract/110/3/486.

Oreopoulos, Phil, Stabile, Mark, Walld, Randy, Roos, Leslie, 2008. Short, medium, and long-term consequences of poor infant health: an anlysis using siblings and twins. Journal of Human Resources 43 (1), Winter.

Oreopoulos, Philip, 2003. The long-run consequences of living in a poor neighborhood. Quarterly Journal of Economics 118 (4), 1533–1575.
URL http://www.mitpressjournals.org/doi/abs/10.1162/003355303322552865.

Orr, Larry, Feins, Judith D., Jacob, Robin, Beecroft, Erik, Sanbonmatsu, Lisa, Katz, Lawrence F., Liebman, Jeffrey B., Kling, Jeffrey R., 2003. Moving to opportunity: interim impacts evaluation. Final report, US Department of Housing and Urban Development, Office of Policy Development and Research, June.

Pitt, Mark, Rosenzweig, Mark, Hassan, Nazmul, 1990. Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries. American Economic Review 80 (5), 1139–1156.

Propper, Carol, Rigg, John, Burgess, Simon, 2007. Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort. Health Economics 16 (11), 1245–1269.

Puma, Michael, Bell, Stephen, Cook, Ronna, Heid, Camilla, Shapiro, Gary, Broene, Pam, Jenkins, Frank, Fletcher, Philip, Quinn, Liz, Friedman, Janet, Ciarico, Janet, Rohacek, Monica, Adams, Gina, Spier, Elizabeth, 2010. Head start impact study: Final report; executive summary. Technical report, Administration for Children and Families. US Department of Health and Human Services, January.

Rasmussen, Kathleen Maher, 2001. The "fetal origins" hypothesis: challenges and opportunities for maternal and child nutrition. Annual Review of Nutrition 21, 73–95.

Reyes, Jessica Wolpaw, 2007. Environmental policy as social policy? the impact of childhood lead exposure on crime. Working Paper 13097, National Bureau of Economic Research, May. URL http://www.nber.org/papers/w13097.

Reznek, Arnold, 2007. Methods of confidential data protection. VirtualRDC powerpoint presentation, Cornell University, April. URL http://www.vrdc.cornell.edu/news/.

Rosenzweig, Mark R., Paul Schultz, T., 1982. Market opportunities, genetic endowments, and intrafamily resource distribution: child survival in rural India. American Economic Review 72 (4), 803–815.

Rosenzweig, Mark R., Wolpin, Kenneth I., 1988. Heterogeneity, intrafamily distribution, and child health. Journal of Human Resources 23 (4), 437–461.

Royer, Heather, 2009. Separated at girth: US twin estimates of the effects of birth weight. American Economic Journal: Applied Economics 1 (1), 49–85.

Ruhm, Christopher J., 2004. Parental employment and child cognitive development. The Journal of Human Resources 39 (1), 155–192.

Rush, David, Stein, Zena, Susser, Mervyn, 1980. Diet in Pregnancy: a Randomized Controlled Trial of Nutritional Supplements. Alan R. Liss, Inc., New York.

Salm, Martin, Schunk, Daniel, 2008. The role of childhood health for the intergenerational transmission of human capital: evidence from administrative data. Discussion Paper 3646, The Institute for the Study of Labor (IZA), August.

Sanbonmatsu, Lisa, Kling, Jeffrey R., Duncan, Greg J., Brooks-Gunn, Jeanne, 2006. Neighborhoods and academic achievement. Journal of Human Resources 41 (4), 649–690.

Schanzenbach, Diane W., What are food stamps worth. American Economic Journal - Economic Policy (forthcoming).
Scherjon, Sicco A., Oosting, Hans, Ongerboer de Visser, Bram W., deWilded, Ton, Zondervan, Hans A., Kok, Joke H., 1996. Fetal brain sparing is associated with accelerated shortening of visual evoked potential latencies during early infancy. American Journal of Obstetrics and Gynecology 175 (6), 1569–1575.
Sloan, Frank A., Conover, Christopher J., Rankin, Peter J., 2001. Effects of Tennessee Medicaid managed care on obstetrical care and birth outcomes. Journal of Health Politics, Policy, and Law 26 (6), 1291–1324.
Smith, James P., 2009. The impact of childhood health on adult labor market outcomes. The Review of Economics and Statistics 91 (3), 478–489.
Smith, James P., Smith, Gillian C., 2008. Long-term economic costs of psychological problems during childhood. Manuscript, October.
Smolensky, E., Gootman, J. Appleton, 2003. Working Families and Growing Kids: Caring for Children and Adolescents. The National Academies Press, Washington, DC.
Solon, Gary, 1999. Chapter Intergenerational mobility and the labor market. In: Ashenfelter, Orley, Card, David (Eds.), Handbook of Labor Economics, vol. 3. Elsevier, pp. 1761–1800.
St. Pierre, Robert G., Layzer, Jean I., 1999. Using home visits for multiple purposes: the comprehensive child development program. The Future of Children 9 (1), 134–151.
Stein, Zena, Susser, Mervyn, Saenger, Gerhart, Marolla, Francis, 1975. Famine and Human Development: the Dutch Hunger Winter of 1944-1945. Oxford University Press, New York.
Tominey, Emma, 2007. Maternal smoking during pregnancy and early child outcomes. CEP Discussion Papers dp0828, Centre for Economic Performance, LSE, October.
URL http://ideas.repec.org/p/cep/cepdps/dp0828.html.
Vujic, Suncica, Konig, Pierre, Webbinnk, Dinand, Martin, Nick, 2008. The effect of childhood conduct disorder on human capital. Discussion Paper 3646, CPB Netherlands Bureau for Economic Policy Analysis, November.
Wong, Vivian C., Cook, Thomas D., Barnett, W. Steven, Jung, Kwanghee, 2008. An effectiveness-based evaluation of five state pre-kindergarten programs. Journal of Policy Analysis and Management 27 (1), 122–154.
Zweifel, Peter, Breyer, Friedrich H.J., Kifmann, Mathias, 2009. Health Economics, second edition, Oxford University Press.