MOTHER'S EDUCATION AND THE INTERGENERATIONAL TRANSMISSION OF HUMAN CAPITAL: EVIDENCE FROM COLLEGE OPENINGS*

Janet Currie and Enrico Moretti

We examine the effect of maternal education on birth outcomes using Vital Statistics Natality data for 1970 to 1999. We also assess the importance of four channels through which maternal education may improve birth outcomes: use of prenatal care, smoking, marriage, and fertility. In an effort to account for the endogeneity of educational attainment, we use data about the availability of colleges in the woman's county in her seventeenth year as an instrument for maternal education. We find that higher maternal education improves infant health, as measured by birth weight and gestational age. It also increases the probability that a new mother is married, reduces parity, increases use of prenatal care, and reduces smoking, suggesting that these may be important pathways for the ultimate effect on health. Our results add to the growing body of literature which suggests that estimates of the returns to education which focus only on increases in wages understate the total return.

I. Introduction

Over the past century, American women have experienced tremendous increases in average educational attainment. Returns to education are generally quantified in terms of increases in wages. However, many studies report a correlation between maternal education and measures of child health. In fact, this robust relationship was one factor underlying the World Bank's drive over the past decade to promote maternal education in developing countries [World Bank 1993]. But the existing literature is limited in that it focuses on developing countries, emphasizes the effects of improvements in relatively low levels of education, and does not generally attempt to establish whether the

* Claudia Goldin, Donald Kenkel, Thomas Kane, Lance Lochner, Justin McCrary, and seminar participants at University of California Los Angeles, University of California at Berkeley, London School of Economics, IZA, Heidelberg University, National Bureau of Economic Research Summer Institute, Columbia University, City University of New York, University of California, Davis, University of Washington at Seattle, the University of Colorado at Denver, University of California, San Diego, Princeton and Yale Universities, three anonymous referees, and the editors provided very helpful comments. We also thank Michael Greenstone and Kenneth Chay for sharing their Vital Statistics data and David Card for sharing his data on colleges that closed. Rebecca Acosta, Benjamin Bolitzer, Stephanie Riegg, and Anna-Maria Bjornsdotter provided excellent research assistance. We thank NIH (Currie) and the UCLA Senate (Moretti) for financial support.

The Quarterly Journal of Economics, November 2003

effect of maternal education is causal.[1] If higher maternal education does indeed improve child health outcomes, even in a rich country such as the United States, then conventional estimates of the returns to education that focus only on wages may understate the social benefits. Moreover, to the extent that healthier children go on to be more productive and more highly educated adults themselves, there will be an important intergenerational spillover that analyses of wage effects alone will not capture.[2]

In this paper we examine the effect of maternal education on infant health at birth in the United States using data from individual birth certificates from the Vital Statistics Natality files for 1970 to 1999. We also assess the importance of four potential channels through which maternal education may improve child quality. First, we examine the effect of education on the use of prenatal care. Because more educated women earn more, they may be able to afford more health care. Second, more educated women are also likely to marry higher earning men, which will further raise family income. Hence, we examine the extent to which an increase in a woman's education raises the probability of marrying a highly educated man. Third, education may induce women to have healthier behaviors. In particular, we analyze the effect of education on the probability of smoking during pregnancy. Finally, we look at the effect of education on fertility. As Becker's quality/quantity trade-off suggests, education may induce women to have fewer children of higher quality.

The correlation between education, investments in health inputs, and health outcomes might reflect omitted factors such as family background or "forward looking behavior." For example, women with high discount rates tend to have lower educational attainment (because they put less weight on future benefits of schooling) and may be less "careful" mothers. In an effort to isolate the causal effect of education, we use an instrumental variables strategy. We have compiled a new data set on openings of two- and four-year colleges between 1940 and 1996. We use

1. Two papers that do address the question of causality are Desai and Alva [1998] and Thomas, Strauss, and Henriques [1991]. The first suggests that much of the return to female education may be through higher father's earnings, while the second finds that female education is associated with factors such as reading newspapers, and thus seems to have a direct effect on the ability to gather information about health.

2. Lleras-Muney [2001] makes this point in a somewhat different context, showing that there is a return to education in the form of longer life expectancy among adults. Lazear [1983] provides a theoretical analysis of intergenerational externalities in education.

these data about the availability of colleges at the county level as an instrument for maternal education in models of birth outcomes and in the models of potential mediating factors. Our instrumental variables (IV) models control for many potentially unobserved confounding factors by including county-year-of-birth fixed effects, so that our estimates are identified by differences in the availability of educational services among different cohorts of women delivering in the same county and year.

One could argue that colleges tend to open in counties where residents' education is already increasing or is expected to increase, and therefore are not a cause but an effect of increasing education. Although we cannot completely rule out this possibility, we provide several pieces of evidence which suggest that college openings do in fact increase educational attainments. First, we find that maternal education and birth outcomes have no clear trend in the ten years before the opening of a college, but improve afterwards. Second, while the opening of a college in the mother's seventeenth year increases her educational attainment, openings in her twenty-fifth or later years do not. Third, the opening of four-year colleges increases four-year graduation rates, but does not affect the probability of having some college, while the opening of two-year colleges increases the probability of having some college but has a much smaller effect on four-year graduation rates. Fourth, when male-only colleges become coed, maternal education increases, but when female-only colleges become coed, we find no effect on women's education. Fifth, the opening of public colleges has a larger effect than the opening of private colleges, and college openings matter more in relatively "college poor" counties. Finally, we show that the first-stage estimates obtained in the Vital Statistics data (in which we know mother's location only at the time of the birth) are similar to those obtained using data from the Census and from the National Longitudinal Survey of Youth (NLSY), data sets in which we know the actual location of women prior to college age.

We find that higher maternal education improves child quality, as measured by birth weight and gestational age. It also increases the probability that a new mother is married, is associated with higher husband education, reduces parity, increases use of prenatal care, and reduces smoking, suggesting that these are all important pathways for the ultimate effect on health. Hence, our results add to the growing body of literature which suggests that estimates of the returns to education which focus only on increases in wages may significantly understate the total return.[3]

The rest of the paper is laid out as follows. Section II provides essential background, while Section III discusses the data. Section IV describes our empirical strategy. Results appear in Sections V and VI, and are followed by the conclusion.

II. How Mother's Education May Affect Child Quality

Mother's education can affect child quality by changing the household budget constraint or by inducing mothers to have healthier behaviors. Education may affect the budget constraint both directly—because more educated women have higher earnings—and indirectly—by improving a woman's marriage market prospects. For example, Behrman and Rosenzweig [2002] conclude that much of the positive cross-sectional relationship between the schooling of mothers and children is due to assortative mating. Goldin [1992, 1997] examines the implications of positive assortative mating on a woman's "returns" to higher education. She finds that for women, much of the return to college may come in the form of a higher earning spouse. For the cohort that graduated between 1945 and 1960, roughly half of the gain to household income was through this channel.

On the other hand, even holding constant the budget constraint, mother's education may improve child quality if it induces women to have healthier behaviors. For example, education may reduce the probability of smoking, which is the single most important cause of low birth weight in the United States [Lightwood, Phibbs, and Glantz 1999]. Education may also affect child outcomes through its effects on fertility. Becker and Lewis' [1973] discussion of trade-offs between the "quantity" and "quality" of children desired is often given as an explanation of the negative relationship between women's education and their fertility. The model also implies that increases in education will be associated with improvements in child quality.

In what follows, we first attempt to establish whether or not the relationship between maternal education and child outcomes is causal. We then examine some of the pathways through which

3. Education has been shown to have external benefits in terms of increased productivity, reduced crime rates, and improved political behavior. See, for example, Moretti [2002, forthcoming] and Lochner and Moretti [2001].

education may affect child outcomes. As measures of infant health outcomes, we use birth weight and gestational age. Birth weight is the most widely reported single measure of infant health. Infants who are of low birth weight (birth weight less than 2500 grams) or premature (gestation less than 37 weeks) are at substantially higher risk of death than other infants. There are many studies linking low birth weight to future cognitive deficits, for example (see Currie [2000] for a review of some of this literature). Currie and Hyson [1999] show using data from the 1958 British cohort study that low birth weight also has long-term negative impacts on test scores, employment probabilities, and wages among young adults. Thus, if maternal education does affect child health, then this may have significant intergenerational effects.[4]

As pathways through which education may affect child outcomes, we examine whether prenatal care commenced in the first trimester or not, whether the mother smoked during pregnancy, whether she was married, husband's education, and parity (i.e., birth order). While there is controversy in the literature about the effectiveness of prenatal care, timely commencement of prenatal care visits can be regarded as an indicator of a woman's willingness to invest in the pregnancy, and as a marker of other healthy behaviors. At a minimum, it would indicate that the woman had recognized the commencement of the pregnancy and started to take steps to accommodate it. Smoking is a particularly interesting indicator in our context, because it is indicative of a change in behavior rather than a shift in the budget constraint.

III. Data

Our primary source of data is individual-level Vital Statistics Natality records from 1970 to 1999. These data are drawn from birth certificates and are thought to cover virtually all births in the United States. Most states currently report the mother's education, information about when she began prenatal care during pregnancy, whether she smoked during pregnancy, marital status, information about the fathers' education, and information about birth outcomes including birth weight and gestational age.

4. While some recent work has cast doubt on the inference that low birth weight has a causal impact on outcomes [Almond, Chay, and Lee 2002], it remains one of the best markers available for poor infant health.

One limitation of the Vital Statistics data is that the number of states reporting some types of information has increased over time, and thus we do not have a balanced panel of counties over time, and the number of counties is not the same for all our outcome measures. In the Data Appendix we describe which of the variables are not available in all years. The effect of this selection is mitigated by the inclusion of county∗year-of-child's-birth dummies in all of our models. Because there are well-known differences in average outcomes by birth order, we focus most of our analyses on first births. The sole exception is when we examine parity, where we include births of all parities in the sample. We also limit the analysis to white mothers, since our identification strategy does not work well for black mothers.[5]

In much of our analysis, we focus on first-time mothers 24 to 45 years old in order to select women whose education is likely to have been completed. We take a 10 percent random sample of the full sample of women in this category for ease of estimation.[6] The first two columns of Table I compare the means for this sample with means obtained for the whole sample of first-time mothers 16–45 years old. Compared with the full sample of first-time mothers, older mothers are more educated, less likely to have negative outcomes, more likely to get prenatal care in the first trimester, much more likely to be married, and less likely to smoke. They also have husbands who are more educated on average. However, the median income in their county in the year when they were seventeen and the number of two- and four-year colleges is comparable across the two samples.

Given these differences between the full sample of first-time mothers, and the sample of older first-time mothers, we also replicate our results using a 10 percent sample of all 16–45 year old first-time mothers. This sample does not suffer from selection on characteristics that may also be associated with education. However, we expect the first stage to be weaker in this sample, since mothers who are sixteen are not yet old enough to have experienced the full potential effect of a college opening on educational attainment.

5. Black women do not appear to be as strongly affected by new college openings as white women.

6. Since our instrument varies at the county-cohort of birth level, the effective sample size for the instrumental variable estimator is the number of cohorts times the number of counties. Using only 10 percent of all births greatly reduces computation time but has no effect on the effective sample size.

TABLE I
SUMMARY STATISTICS

	All 16–45 (1)	All 24–45 (2)	College 24–45 (3)	Some college 24–45 (4)	High school 24–45 (5)	High school dropout 24–45 (6)
Mother's education	12.95	14.21	16.3	14.27	12	9.47
	(2.22)	(2.08)	(0.47)	(0.44)	(0)	(2.15)
Low birth weight	0.054	0.049	0.041	0.047	0.057	0.089
Preterm birth	0.079	0.069	0.062	0.068	0.075	0.100
Prenatal care started 1st trimester	0.827	0.921	0.945	0.925	0.894	0.759
Smoked during pregnancy	.145	0.078	0.023	0.080	0.168	0.340
Married	.772	0.923	0.968	0.930	0.890	0.721
Husband's education	13.1	14.20	15.55	14.39	12.88	10.87
	(2.31)	(2.65)	(1.65)	(1.96)	(1.94)	(2.7)
Mother's age	23.83	28.12	29.02	27.80	27.31	27.30
	(5.14)	(3.56)	(3.6)	(3.1)	(3.3)	(3.6)
Year mother was 17	1980	1974.8	1975.0	1975.6	1974.1	1973.4
	(8.7)	(7.5)	(7.7)	(7.4)	(7.3)	(8.6)
Per capita number of four-year colleges in county, when mother was 17	0.0796 (0.0984)	0.0819 (0.0925)	0.0856 (0.0856)	0.0799 (0.0903)	0.0785 (0.0932)	0.0766 (0.0974)
Per capita number of two-year colleges in county, when mother was 17	0.0548 (0.0850)	0.0502 (0.0696)	0.0492 (0.0492)	0.0519 (0.0672)	0.0497 (0.0739)	0.0514 (0.0850)
Median family income in county when mother was 17	32510 (7150)	33273 (7467)	34082 (7716)	33670 (7120)	32359 (7176)	29942 (7751)
Percentage county that was urban when mother was 17	0.71 (0.27)	0.77 (0.24)	0.80 (0.23)	0.78 (0.23)	0.74 (0.26)	0.70 (0.29)
Sample size	2,042,270	671,468	279,574	169,916	204,394	17,584
Parity (here we include all births rather than 1st births)	2.0 (1.3)	2.4 (1.2)	2.0 (1.1)	2.3 (1.2)	2.5 (1.2)	3.4 (1.2)

Column (1) reports summary statistics for a 10 percent sample of first-time white mothers 16–45. Columns (2) to (6) report summary statistics for a 10 percent sample of first-time white mothers 24 to 45. Averages in the last row include all births, not just first births. Not all variables are available in all years. For smoking, $N = 166{,}183$. For Husband's education, $N = 486{,}255$. Standard deviations are in parentheses.

The last four columns of Table I break out the sample of older mothers by education. The monotonic positive relationship between favorable birth outcomes, input use, and maternal education is very striking. Just to take one example, only 2 percent of college-educated mothers in this age group smoke during pregnancy. The corresponding figures for mothers with some college, high school, and less than high school are 8 percent, 17 percent, and 34 percent. College-educated women are also more likely to come from counties with a higher per capita number of four-year colleges in the year when they were seventeen.

While we have no reason to believe that education is systematically misreported in the Vital Statistics data, we have compared education in the Vital Statistics with that in Census reports. We found that the education distribution in the Vital Statistics data is consistent with the education distribution of first-time mothers in the Census. (See Appendix Table A3 in Currie and Moretti [2002].)

Our IV analysis combines Vital Statistics natality data with a unique data set on the number of two- and four-year colleges by county and year that we compiled for this research. The construction of this data set is described in the Data Appendix. There was a tremendous rise in the number of colleges between 1940 and 1996. In 1940 there were 346 two-year colleges in our sample and 1301 four-year colleges. By 1996 these figures had risen to 1436 and 1808 two- and four-year colleges, respectively. The pattern for two-year colleges is more "S-shaped" than that for four-year colleges reflecting the great expansion of the two-year college system during the 1960s and 1970s. (See Figure 1 in Currie and Moretti [2002].)

We use these data to construct measures of the availability of two- and four-year colleges. Each measure is the number of colleges that existed in the woman's county when she was seventeen years old, divided by the estimated number of 18 to 22 year olds in the county in that year (in thousands). This instrument takes into account the fact that cohort size is likely to have an impact on the availability of college given any fixed number of schools (cf. Card and Lemieux [2001] and Welch [1979]).[7]

The ideal instrumental variable would account not only for the number of colleges, but also for their capacity. Information on the availability of classroom space or capital expenditures is not available for part of the period under consideration. Information on enrollment might be obtainable, but enrollment reflects both the supply of college places and the demand for these places, and thus is not a valid instrument. However, because the first stage

7. The number of individuals 18 to 22 years old is interpolated using data from the Census. See the Data Appendix. Note that measurement error in our denominator should not carry over to our left-hand-side variable since they are measured in two different sources.

using the number of colleges is fairly precisely estimated, our instrument, although imperfect, seems adequate in terms of power.

The most important limitation of the Vital Statistics data is that we observe the mother's residence at the time of the birth of the baby, rather than at the time when she was seventeen. Thus, we are forced to assume that her county of residence is the same at the birth as it was at seventeen. Because young women have high mobility rates, this assumption is potentially problematic. If mothers randomly change location between age seventeen and the time of the birth, then we will tend to understate the extent to which college openings affect women's educational attainment in the first-stage regressions. However, seventeen-year-old women may move in order to attend college and then stay in the new location, or counties that experience college openings may become more attractive to college-educated women for other reasons. We cannot completely rule out these two possibilities, but we discuss evidence below that suggests that this type of mobility is not driving our results.

IV. Identification

Our first-stage model takes the form,

$$\text{EDUC} = b_0 + b_1\text{IV-2} + b_2\text{IV-4} + b_3\text{AGE} + b_4\text{COHORT} + b_5\text{COUNTY}{*}\text{YEARBRTH} + b_6\text{INCOME17} + b_7\text{URBAN17} + u, \tag{1}$$

where IV-2 and IV-4 denote our instrumental variables for two- and four-year colleges; AGE is a vector of single year of maternal age dummies as in (1); COUNTY*YEARBRTH is a vector of indicators for the county and year of the child's birth; INCOME17 and URBAN17 control for the median income and the percent urban in the county when the woman was seventeen, and u is a random error term.

The county/year effects control for many characteristics of the local area that may affect outcomes, such as the availability and quality of medical services, the local business cycle, pollution, etc. Identification in our models comes from the fact that *within each county and year of birth of the baby,* there are mothers who were seventeen before a college opening, and mothers who were seventeen after a college opening.

The income and urbanicity measures control for some of the factors likely to affect the educational attainment of the mother's cohort (such as economic growth and development). In some specifications we test the robustness of our results by including state-mother cohort trends, which allow the cohort effects to vary from state to state. We have also estimated models including state*mother age effects, which yielded results very similar to those reported below.

In the second stage, we regress an outcome variable on the predicted years of education from (1) and all the exogenous covariates:

$$\text{(2)}\quad \text{OUTCOME} = c_0 + c_1\text{PEDUC} + c_2\text{AGE} + c_3\text{COHORT} + c_4\text{COUNTY}{*}\text{YEARBRTH} + c_5\text{INCOME17} + c_6\text{URBAN17} + z,$$

where PEDUC is the predicted years of education from (2), and z is a random error. In all the models, the standard errors allow for potential correlations between the errors within county-year-at-17 clusters.

This specification does not capture the part of the return to schooling that arises through residential location. More educated women can probably afford to live in areas with better hospitals and better environmental quality. By including county*year-of-child's-birth effects, we absorb the effect of these geographical factors on birth outcomes, and in this sense, we underestimate the benefit of schooling.

In contrast to most studies using policy changes as an instrument, the possible correlation between changes in number of colleges and *contemporaneous* changes in other factors affecting infant outcomes is not problematic. For example, if counties that experience increases in the number of colleges also experience increases in the number of hospitals, our IV estimates are still valid, because we control for county*year-of-birth effects. Even if a new hospital took some time to serve significant numbers of pregnant women, to the extent that the delay affected all women at time t similarly, the effect would be captured by these interactions.

The two main concerns regarding the validity of our instrumental variables are the fact that the geographical location of new colleges may not be random, and the endogenous mobility of women who move to attend college. We now discuss these two concerns in turn.

It is possible that college openings reflect expected increases in demand for education. If demand for education were correlated with omitted factors that also increased investments in health, then changes in this omitted factor rather than college openings would drive both increases in educational attainments and improvements in health, not vice versa, and our estimates could overestimate the effect of schooling on birth outcomes. On the other hand, it is also possible that state legislatures place colleges in a compensatory way, i.e., in underserved areas that have low college attendance. For example, legislation establishing the new University of California Merced campus begins: "The San Joaquin Valley is the most populous region of the state without a University of California campus, and has one of the lowest rates of college participation of all regions in California. Access to postsecondary education is determined, in significant measure, by a student's proximity to college campuses" (California State Education Code, Section 92160). Similarly, other state plans for higher education emphasize increasing access to education as a main goal of public investments.[8] These policy statements suggest that to the extent that there is any correlation, the location of new public colleges may be negatively rather than positively correlated with average academic attainments in the area, which would lead us to underestimate the effect of education on birth outcomes.

The location of private colleges may be governed by different considerations. On the one hand, it is possible that private colleges are located where demand is anticipated to grow. On the other hand, private college location could be affected by land prices, which would dictate location in less developed areas. Ultimately, whether new colleges locate in areas with larger expected increases in schooling is an empirical question.

The second problem is that women may move to places with college openings in order to attend college, or, more educated women may find these counties attractive places to live for other reasons (such as jobs, or the availability of college-educated men). While highly selective institutions recruit students from across the country, most of the four-year and virtually all of the two-year colleges in our sample are not of this type. A typical new institution is more likely to resemble Coleman College (LaMesa, CA) or Aims Community College (Greeley, CO) than Harvard or Prince-

8. See examples on www.sheeo.org/govern/gov-panrep.htm.

ton. Although most of the new schools in our sample are nonselective institutions that one would expect to enroll mainly local women, it is still possible that at least some of our results are explained by the endogenous mobility of women.

We address these concerns in several ways. First, we look at whether education appears to be increasing in counties which open colleges in the years prior to college openings. If openings determine education rather than vice versa, we should find little evidence of a pretrend in schooling prior to the college opening. We also look for the presence of pretrends in birth outcomes.

Second, we ask whether new four-year colleges have a stronger effect on the completion of four-year college degrees than on the completion of two-year degrees, and whether new two-year colleges have a stronger effect on the completion of two-year degrees than on the completion of four-year degrees, as argued by Rouse [1995]. For example, a finding that new four-year colleges have the same effect on the completion of four-year degrees and two-year degrees would be a sign of spurious correlation, because it would suggest that new colleges open in counties where education is increasing at every level.

A third test of the validity of our instruments is to estimate first-stage models that include measures of college availability when the mother was 25 or older in addition to measures of college availability when the woman was 17. If our identification strategy is valid, then measures of opportunities at age 17 should have a larger effect on the woman's educational attainment than measures of college availability taken when she is past the usual age of college attendance. Conversely, if college-educated women 25 or older move to counties that experience college openings, then measures of college availability taken at 25 or older should be stronger predictors of a woman's educational attainment than our measure of college availability measured at age 17.

Fourth, we ask whether the opening of new public colleges has a different impact than the opening of new private colleges. If our story is correct, and our instrument captures reductions in the cost of attendance for the marginal woman, then we expect the opening of a public college to have a larger impact than the opening of a private college. Women who decide to attend college because a new college opens in their county are probably from disadvantaged backgrounds, and therefore are more likely to attend public colleges. Presumably, most women who migrate to attend private colleges would have gone to college anyway, even if there were no college in their county of residence. Finding larger effects for private colleges than for public colleges would cast doubt on the validity of our instrument.

Fifth, we examine conversions of single-sex colleges to coed colleges.[9] Male-only colleges that become coed increase women's access to college and hence should have a positive effect on female education; but female-only colleges that become coed could even reduce access to college (by crowding women out) and so should not have any positive effect on women's education. Similarly, we ask whether the opening of a new college has a greater effect on educational attainments in regions that are "college poor." That is, we expect that the opening of a new college should have a smaller effect if there are many colleges close by than if there are few.[10]

Finally, we try to assess the magnitude of the bias introduced by endogenous mobility (which is probably the main limitation of our instrument) using data from the 1980 and 1990 Census Public Use Micro Samples (PUMS) and from the NLSY. The Census allows us to identify the location of respondents five years prior to the survey date and has the advantage of very large sample sizes. A limitation of the Census data is that it reports PUMAs (Public Use Microdata Areas) rather than counties. PUMA are smaller than counties in urban areas, and larger than counties in rural areas. For example, for the nine-county San Francisco Bay Area, there are 48 PUMAs. Hence, in urban areas, we use MSAs (Metropolitan Statistical Areas), while in rural areas we use PUMAS to define location of residence. The NLSY identifies counties, and has more complete histories of location, but a relatively small sample size (especially given our focus on white mothers).

We estimate the analogue to our first-stage regression in the Census sample using information on the educational attainments of women 19–22 and the availability of colleges in the location of

9. Alternatively, we could examine the effects of opening single-sex colleges as a specification check. But there are very few openings of single-sex colleges in the period under consideration.

10. In previous work we also estimated models that control for the average educational attainment of fathers (of all ages) in the county when the woman was seventeen years old. If new colleges were more likely to be built in areas with high demand for college education, then part of this latent demand for college education by women will be captured by measures of the educational attainments of men. We found that inclusion of this variable reduced the first-stage effect of IV-4 somewhat, but had little effect on our TSLS results. (See Currie and Moretti [2002].)

residence five years prior to the Census, that is at the time the women were 14–17 years old. The location of residence at age 14–17 is less likely to be contaminated by endogenous mobility than current location of residence. We conduct a similar experiment in the sample of women who had had at least one child by 1996 from the NLSY. The youngest women in this sample were 32. Specifically, we regress the mother's highest grade completed as of 1996 on the availability of colleges in the county and year that she was fourteen. As we show below, we find evidence that the availability of colleges at the time the woman is 14–17 has a significant effect on her educational attainment in both the Census and the NLSY. Further, the estimates are quite similar to those obtained in the Vital Statistics data, suggesting that our first-stage result is not driven by endogenous mobility.

If education is endogenous, and mothers of better "quality" tend to have higher education, OLS will tend to overestimate the true effect of schooling on birth outcomes. Hence, OLS estimates will be larger than valid IV estimates. On the other hand, there are at least two reasons why IV may exceed OLS.[11]

First, random measurement error may cause OLS estimates to be biased toward zero, though for reasons discussed below, we do not feel that measurement error of this type is driving our results. Second, the marginal benefit of schooling for individuals whose education has been affected by college openings may be larger than the average benefit for the population. Our IV estimates reflect the effect of education for women who would not have gone to college if it had not been for the fact that a college opened in their county of residence. These women are likely to come from disadvantaged backgrounds relative to those who were not constrained by college location. Card [2000, 2001] reviews a series of recent studies that measure the effects of education on earnings using instrumental variables such as compulsory schooling laws, differences in the accessibility of schools, and

11. Lochner and Moretti [2001] show that spillovers can make IV estimates high relative to OLS, when the instrumental variable varies at some aggregate level. In our case, if an individual's decision to engage in healthy behaviors or to buy health inputs depends on average education levels or the average behavior of other individuals in their cohort and county, IV (using county-cohort level instruments as we do) will estimate the combined effect of own education on outcomes as well as the effect of average cohort education on outcomes. That is, IV will estimate the sum of the individual effect and the spillover effect. OLS will only estimate the individual effect of education. In a previous version of this paper, we argued that the results presented below may indicate spillover effects, especially for prenatal care (see Currie and Moretti [2002] for details).

other institutional features of the education market. These studies typically find that instrumental variables estimates *exceed* OLS estimates of the returns to education, suggesting that the marginal returns to education among the groups affected by changes in the instrument are higher than the average returns in the population.

This line of argument is relevant when estimating the effect of schooling on child quality. The women affected by college openings may have high marginal returns to schooling, both in terms of earnings and the adoption of healthier behaviors, so that we may find IV estimates that are larger than OLS.[12] Some of the specification tests described in the previous section will shed light on whether women's educational choices appeared to be constrained by lack of college availability.

V. Results

We begin our analysis of college openings by graphically showing how maternal education in the sample of first-time mothers 24 and older changed in the ten years before and after the opening of a four-year college. The first quadrant of Figure I shows the average years of schooling before and after a four-year college opening, after controlling for county dummies and dummy variables for the mother's cohort of birth.[13] For purposes of illustration, we select openings where there were no other openings in the same county either ten years before or ten years after the year of the index college opening. (In the formal analysis below we use all the openings.) The figure shows that, on average, maternal education levels are higher in the ten years following the opening than they were in the ten years preceding. Moreover, there is no

12. For the sake of comparison with other studies, we have estimated the returns to education using all white individuals 24–60 in the 1980 and 1990 Censuses and using our measure of college availability as an instrument. The OLS estimate of the return (and its standard error) was .068 (.0007) while the IV estimate was .126 (.041), where the regressions also controlled for gender, age, decade of birth, Census year, and location fixed effects. As discussed below, counties are not reported in the Census, so we use MSAs to define urban areas and PUMAs to define nonurban areas. The estimated effects are somewhat sensitive to the definition of local areas, which is not surprising since the measure of college availability depends on the definition of the area.

13. In particular, we regressed years of schooling on dummies for time equal to $t - 10$, $t - 8$, $t - 6$, $t - 4$, $t - 2$, t, $t + 2$, $t + 4$, $t + 6$, $t + 8$, $t + 10$, county dummies and cohort dummies. The figure shows the coefficients on dummies for time equal to $t - 10$, $t - 8$, $t - 6$, $t - 4$, $t - 2$, t, $t + 2$, $t + 4$, $t + 6$, $t + 8$, $t + 10$, where t is the time of college opening. Figures II and III are obtained in a similar way.

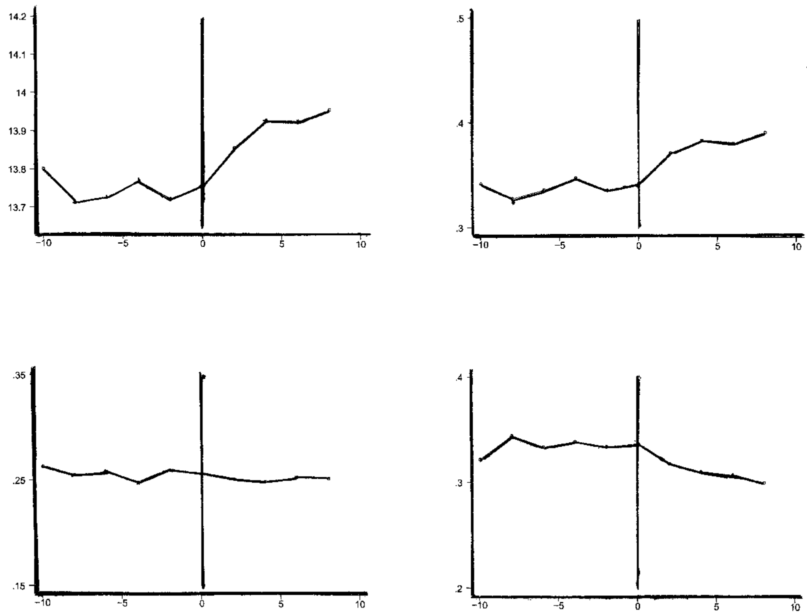

FIGURE I
Schooling Achievement of Women before and after the Opening of a Four-Year College

The top-left panel shows the average years of schooling in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The top-right panel shows the conditional probability of college education in the years preceding and the years following the opening of a four-year college, controlling for county dummies and mother cohort dummies. The bottom left panel is for community college; the bottom-right panel is for high school. Time 0 is when the college opens and the mother is seventeen years old. The figures include only college openings that occurred at least ten years apart from other openings in the same county.

particular trend in maternal education in the county prior to the opening of a new four-year college.

The remaining quadrants of Figure I show the same story by level of education completed, rather than years of education. The pattern for four-year college attendance is quite similar to the pattern for average years of education shown in the first panel. This is consistent with the hypothesis that most of the increase in education that arises from the opening of a four-year college comes from increases in the fraction of women with a four-year college education. Furthermore, most of the increase in the fraction of women with a four-year college degree appears to be coming from a reduction in the number of women who have only a high school degree. This lends some credibility to the hypothesis that the college opening induced some women who would have otherwise stopped studying after graduating from high school to go to college. There is little trend in either the number of women with some college (but no four-year degree) or in the number of high school dropouts (though the latter is not shown).

Detecting pretrends in outcomes is even more important than detecting pretrends in education. Figure II shows the impact of college openings on the incidence of low birth weight and premature birth. Figure II shows that while there were no clear trends in the incidence of low birth weight or prematurity in the ten years prior to the opening of a new college in the county, there is a drop in the incidence of these negative outcomes following the opening.

Figure III is like Figure II, but considers college openings that occur when the mother is 25, rather than college openings in the year that she was seventeen. Most women should be unaffected by a college opening that occurs past their college age. As expected, we see no clear effect of these later openings on outcomes.

We now turn to a more formal analysis of the effect of college openings on schooling. Our first-stage results are shown in Table II. Column (1) of panel 1 shows that an increase of one four-year college per 1000 persons 18 to 22 would result in almost a full year more of maternal education among women 24 and over at the time of first birth. The mean of this variable in our sample is .082 indicating that at the mean, new college openings increased maternal education by approximately .08 years. The effect of a new two-year college is substantially smaller: an increase of one college per 1000 young adults would increase schooling by about a fifth of a year. Evaluated at the sample mean of .05 two-year colleges per 1000, this coefficient implies that increases in the availability of two-year colleges increased maternal schooling by about .01 years.[14]

The model in column (1) accounts for differences across cohorts by including decade of birth dummies. One concern is that

14. A second way to interpret these estimates is to ask how many more college-educated women resulted from the opening of, for example, a new four-year college? Model 2 of Table II indicates that such an opening increased the probability that a first-time mother 24 or older had four years of college by about 20 percentage points. Since the average county has approximately 10,000 18 to 22 year olds, this implies that one new four-year college would increase the probability that these mothers had a college education by 2 percentage points. Relative to the baseline of 42 percent, this is an increase of about 5 percent in the probability of college. Since the average county in our sample has 165 first-time moms 24 to 45, a 5 percent increase would imply that approximately eight more of women per county per year received a college education as a result of the college opening.

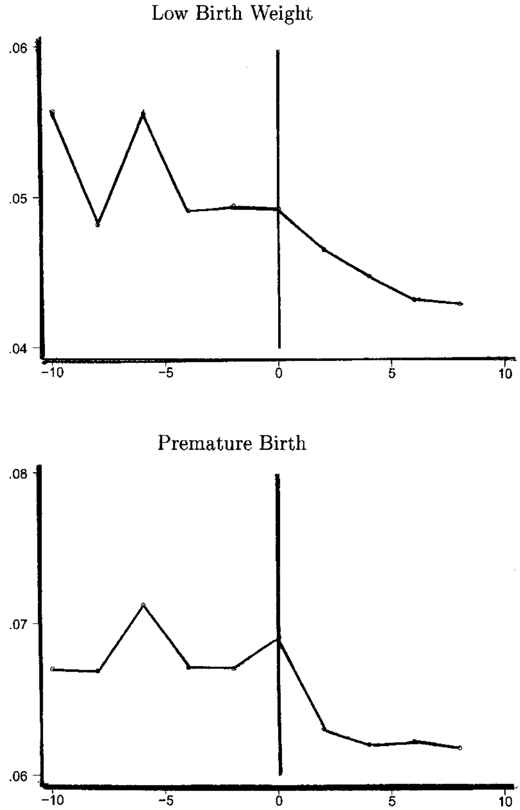

FIGURE II
Average Birth Outcomes, before and after the Opening of a Four-Year College—Mother Is 17 When the College Opens

The top panel shows the conditional probability of low birth weight in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The bottom panel shows the conditional probability of premature birth in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. In both figures, time 0 is when the college opens and the mother is seventeen years old. The figures include only college openings that occurred at least ten years apart from other openings in the same county.

cohort effects were stronger in states where more schools were built, possibly because those states had higher economic growth or experienced in-migration of higher human capital individuals. Column (2) shows that adding state-specific cohort trends to the model has little effect on the estimates. We have also estimated models that include state-specific quadratic trends in the mother's

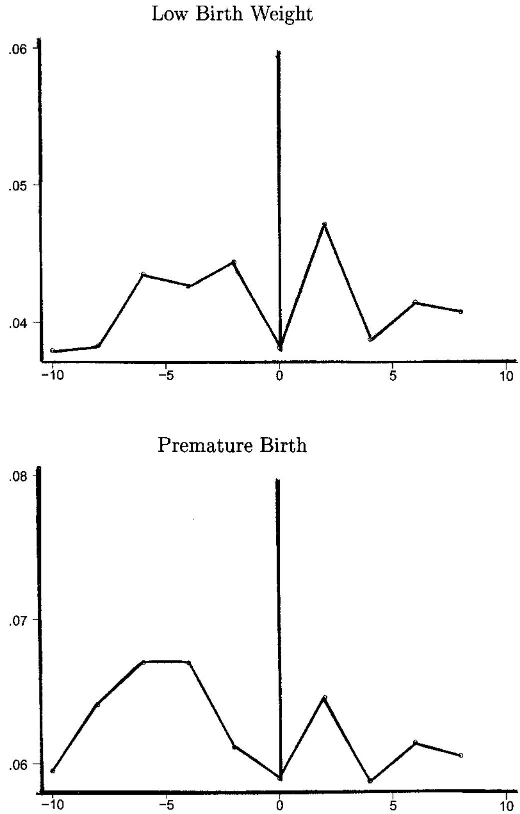

FIGURE III

Average Birth Outcomes, before and after the Opening of a Four-Year College—Mother Is 25 When the College Opens

The top panel shows the conditional probability of low birth weight in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. The bottom panel shows the conditional probability of premature birth in the years preceding and the years following the college opening, controlling for county dummies, and cohort dummies. In both figures, time 0 is when the college opens and the mother is 25 years old. The figures include only college openings that occurred at least ten years apart from other openings in the same county.

cohort, or state-specific mother cohort effects (with cohorts grouped in ten-year intervals), or state-mother-age effects (with ages grouped in five-year intervals) with very similar results. (Second-stage results are also insensitive to these additional controls.)

The remaining panels of Table II show how the change in educational attainment breaks down. A new four-year college (per

TABLE II
THE EFFECT OF COLLEGE OPENINGS ON MATERNAL EDUCATION

	(1)	(2)
Model 1: Years of schooling		
Four-year colleges	0.950	0.952
	(0.046)	(0.046)
Two-year colleges	0.176	0.158
	(0.046)	(0.046)
R^2	0.11	0.12
Model 2: 4 years or more of college		
Four-year colleges	0.198	0.199
	(0.010)	(0.010)
Two-year colleges	0.025	0.021
	(0.010)	(0.010)
R^2	0.10	0.10
Model 3: Some college		
Four-year colleges	0.000	0.046
	(0.006)	(0.006)
Two-year colleges	0.032	0.032
	(0.007)	(0.007)
R^2	0.02	0.03
Model 4: High school only		
Four-year colleges	−0.156	−0.156
	(0.009)	(0.009)
Two-year colleges	−0.053	−0.051
	(0.010)	(0.010)
R^2	0.07	0.07
Model 5: Less than high school		
Four-year colleges	−0.026	−0.026
	(0.003)	(0.003)
Two-year colleges	−0.003	−0.003
	(0.004)	(0.004)
R^2	0.03	0.03
State-mother cohort trends	No	Yes

Standard errors are in parentheses and are clustered at the county-year of mother's birth level. All models include mother age effects, mother ten-year cohort of birth effects, county*year-of-child's-birth effects, median county income, and percent urban in year when mother was seventeen. "Four-year" and "Two-year" are the number of four-year colleges and two-year colleges existing in a county in the year when the mother was 17 normalized by the number of 18–22 year olds in the county in that year. The sample includes all first births to white mothers 24 to 45.

1000) increases the probability of obtaining a four-year degree by 19 percentage points, but has little effect on the probability that only "some college" is completed. A new two-year college (per 1000) increases the probability of obtaining a four-year degree by 2.5 percentage points, but has a larger effect on the probability of some college, increasing it by 3.2 percentage points. Like the figures, model 4 indicates that most of the increases in schooling following a college opening are coming from a decrease in the number of women obtaining only a high school degree.

Estimates in panels 2 to 5 of Table II, and in particular the comparison of model 2 with model 3, are inconsistent with the hypothesis that colleges systematically open in areas where education is increasing (or is expected to increase) at all levels. Models 2 and 3 suggest that the opening of a four- or two-year college affects the education distribution at the "right" level, and has little spillover onto other levels of schooling.[15]

Table III shows alternative specifications of our first-stage equation (2), in which measures of college availability when the mother was 25 or older are added to the base specification shown in Table II. Finding that college availability when the mother was 25 or older had *a positive* impact on schooling would cast doubt on our instrumental variables strategy. The coefficients on college availability at age seventeen remain generally similar to those in Table II, confirming that this measure of college availability has a positive effect on maternal education. Surprisingly, the coefficients on college at later ages (at age 25 and 35) in column (1) are actually negative, though they are also small relative to the effects of college availability at age 17.[16] We perform an F-test to test whether the coefficients on college availability at seventeen are equal to the coefficients on college availability at later ages, and we always reject this restriction. Thus, Table III provides evidence that it is college availability at seventeen rather than at older ages that matters. This suggests that our results regarding the effects of college openings on maternal education are not driven by underlying trends in college enrollments or by the migration of women older than 24

15. We have also estimated first-stage models separately for three different cohorts of mothers; those who were seventeen between 1940 and 1960, those who were seventeen between 1960 and 1980, and those who were seventeen between 1980 and 2000 (results are not reported in the table). In Table II the coefficient on four-year colleges was .95. The comparable coefficients for the effect of four-year colleges on the three cohorts are 1.19, .903, and 1.53 (all statistically significant). The comparable coefficients for the effects of two-year colleges are .007, .173, and .214 (but only the middle coefficient is statistically significant). Thus, the effect of having a four-year college open in one's own county is not noticeably weaker for the younger cohorts than for the older ones, whereas the effects of two-year colleges are greatest for those cohorts who experienced the large boom in two-year college construction. We speculate that part of the reason for this pattern may be the increase in the average size of colleges over the period under consideration.

16. If we include the additional controls for average father's education discussed in footnote 10, then only the college availability measures at age seventeen are statistically significant.

TABLE III
THE EFFECT OF COLLEGE OPENINGS AT AGES LATER THAN 17 ON MATERNAL YEARS OF SCHOOLING

	(1)	(2)
Model 1: Add colleges at age 25		
Four-year colleges at age 17	1.259	1.330
	(0.161)	(0.161)
Two-year colleges at age 17	0.317	0.299
	(0.104)	(0.103)
Four-year colleges at age 25	−0.371	−0.454
	(0.180)	(0.178)
Two-year colleges at age 25	−0.180	−0.177
	(0.105)	(0.105)
Effect at age 17 = Effect at age 25 (p-value)	0.000	0.000
Model 2: Add colleges at age 35		
Four-year colleges at age 17	1.221	1.234
	(0.151)	(0.152)
Two-year colleges at age 17	0.250	0.236
	(0.106)	(0.106)
Four-year colleges at age 35	−0.325	−0.344
	(0.179)	(0.180)
Two-year colleges at age 35	−0.184	−0.174
	(0.105)	(0.105)
Effect at age 17 = Effect at age 35 (p-value)	0.000	0.00
State∗mother cohort trend	No	Yes

Standard errors are in parentheses and are clustered at the county-year of mother's birth level. All models include mother age effects, mother ten-year cohort of birth effects, county∗year-of-child's-birth effects, median county income, and percent urban in year when mother was seventeen. "Four-year" and "Two-year" are the number of four-year colleges and two-year colleges existing in a county in the year when the mother was 17 normalized by the number of 18–22 year olds in the county in that year. Similar definitions apply to colleges at 25 and 35. The sample includes all first births to white mothers 24 to 45.

seeking jobs (or even husbands) in the county of the new college. We present further specification checks below.

OLS estimates of the effects of maternal education on infant health, marriage, husband's education, probability of smoking, and parity are shown in column (1) of Table IV for the sample of older mothers. These estimates suggest that an additional year of maternal education reduces the probability of low birth weight by 0.5 percentage points (or 10 percent), and lowers the risk of prematurity by 0.44 percentage points (or 6 percent) relative to the means in Table I. Table IV also suggests that much of the positive impact of higher education on birth outcomes may be coming through reductions in smoking: an additional year of education reduces the probability of smoking by more than 30 percent. The increases in four-

TABLE IV
THE EFFECT OF MATERNAL EDUCATION ON INFANT HEALTH, AND HEALTH INPUTS

	OLS 24–45 (1)	IV 24–45 (2)	OLS 24–45 (3)	IV 24–45 (4)	OLS 16–45 (5)	IV 16–45 (6)
1. Low birth weight	−0.0050	−0.0098	−0.0050	−0.0099	−0.0053	−0.0096
	(0.0001)	(0.0038)	(0.0001)	(0.0038)	(0.0001)	(0.0036)
2. Preterm birth	−0.0044	−0.010	−0.0044	−0.010	−0.0038	−0.098
	(0.0001)	(0.0044)	(0.0001)	(0.0044)	(0.0001)	(0.0044)
3. Prenatal care	0.0114	0.0234	0.0116	0.0241	0.0111	0.0466
	(0.0001)	(0.0055)	(0.0001)	(0.0054)	(0.0001)	(0.0074)
4. Smoked during pregnancy	−0.0305	−0.0583	−0.0305	−0.0623	−0.0336	−0.0364
	(0.0004)	(0.0118)	(0.0004)	(0.0118)	(0.0004)	(0.0159)
5. Married	0.0206	0.0128	0.0207	0.0129	0.0219	0.0107
	(0.0002)	(0.0040)	(0.0002)	(0.0040)	(0.0002)	(0.0060)
6. Husband's education	0.607	0.988	0.604	0.992	0.595	0.876
	(0.0019)	(0.040)	(0.0019)	(0.040)	(0.0019)	(0.045)
7. Parity	−0.121	−0.092	−0.121	−0.088	−0.124	−0.103
	(0.000)	(0.010)	(0.000)	(0.010)	(0.000)	(0.018)
State-mother cohort trends	No	No	Yes	Yes	No	No

Standard errors are in parentheses and are clustered at the county-year of mother's birth level. All models include mother age effects, mother ten-year cohort of birth effects, county*year-of-child's-birth effects, median county income, and percent urban in year when mother was seventeen. The instrumental variables in columns (2), (4), and (6) are the number of four-year colleges and two-year colleges in a county in the year when the mother was 17 normalized by the number of 18–22 year olds in the county and year. Models in rows 1 to 6 include all white mothers with parity equal to 1. Models in row 7 include all white mothers of any parity.

and two-year colleges are estimated to have reduced smoking by 3 percent and 2 percent, respectively.

Effects on the use of prenatal care may be interpreted either as the result of higher income, or as a change in behavior. In any case, an additional year of education increases the probability that prenatal care began in the first trimester by one percentage point. However, marriage and prenatal care utilization rates are high for all groups in our sample, which means that there is less scope for education to affect these outcomes than smoking. We find that an additional year of education increases the probability of marriage by one percentage point and increases husband's education by .6 of a year.

Instrumental variables estimates are shown in column (2) of Table IV for the sample of 24–45 year old women. All IV estimates have the same sign as the OLS estimates, and are statistically different from zero. To help in interpreting the magnitude of the estimated effects, consider that the increase in maternal education between the cohort of women who went to college in the 1940s and the 1950s and the cohort of women who went to college in the 1980s is about 1.6 years. The probability of low birth weight and preterm birth decreased by 6 percentage points and 3 percentage points, respectively. Our estimates suggest that 12 percent of the decrease in the probability of low birth weight and 20 percent of the decrease in the probability of preterm birth can be attributed to increased maternal education. Although we present IV estimates for husband's education, it may not be plausible to assume that the college availability instruments affect the education of the husband only through the education of the wife (that is, only through assortative mating—there may also be a general effect on the distribution of available husband's education). The next two columns of Table IV present estimates from models that also include state-mother cohort effects. These estimates are very similar to those discussed above.

It is striking that the IV estimates are larger than the OLS estimates for all outcomes except marriage and parity, where the IV estimates are lower than the OLS. One interpretation is that educated women have unobserved characteristics that also make them more marriageable and likely to have fewer children. Controlling for these characteristics using instrumental variables reduces the estimated effect of education on marriage and parity but does not eliminate it. On the other hand, IV estimates of the effects of education on prenatal care utilization and birth outcomes exceed OLS estimates. If these results were driven only by measurement error in the mother's education, then one might expect IV to be systematically greater than OLS for all choices of the dependent variable.

The last two columns of Table IV present estimates using the full sample of 16–45 year old first-time mothers. As discussed above, the advantage of using this sample is that it is representative of all first-time mothers. A disadvantage is that we cannot expect the instruments to have a good deal of explanatory power for the youngest mothers. This problem is explored further in Table VI below. However, Table IV indicates that the instrumental variables estimates we obtain (that is, the effect of the "treatment" on those who are actually "treated") is remarkably similar in this sample.

A possible interpretation of these results is that there is a higher than average return to educating the marginal woman in terms of infant health. If the only people affected by changes in college availability were those who altered their years of schooling, then our estimates would measure the effects of additional years of college education at the quality of the new local schools. However, some inframarginal individuals who would have left the county to attend college, may now attend locally instead. To the extent that the quality of the new school differs from the one that would have been attended, our IV results will also pick up this quality effect. Hence, we also show reduced-form estimates of the effects of college availability on all our outcomes in Appendix 2. These estimates also indicate that increases in college availability are associated with both positive behaviors and better infant health outcomes. Four-year college availability has consistently larger and more significant effects, though in the full sample of 16–45 year olds, both two- and four-year college availability are statistically significant.

VI. Additional Specification Tests

Table V presents evidence about four of the additional specification checks described in Section IV above. The first panel shows separate estimates for four-year public and private colleges. Virtually all of the two-year colleges in our sample are public institutions, while roughly half of four-year colleges are public. As discussed above, if college openings in a woman's county affect maternal education by reducing the costs of attendance, then we might expect to see a larger effect for public colleges than for private colleges, given the usual differences in tuition. Table V shows that this is indeed the case. The effect of a new four-year public college is over three times greater than the effect of a new four-year private college. Furthermore, this finding, together with the fact that the location of public colleges is often compensatory, is inconsistent with the hypothesis that private colleges tend to locate in areas with higher expected increases in schooling. There are not enough private two-year colleges to perform the same test on two-year colleges.[17]

Panel 2 examines transformations of colleges from single-sex to coed. If our identification strategy is valid, then the transformation of a male-only college into a coed college should increase maternal education, while the transformation of a female-only

17. We have examined the sensitivity of our IV results to the exclusion of private four-year colleges (most two-year colleges are public). We found that we could not reject the hypothesis that the estimates were the same whether or not these colleges were excluded.

TABLE V
FURTHER EVIDENCE ON THE EFFECT OF COLLEGE OPENINGS ON MATERNAL EDUCATION

	Coefficient on four-year colleges (1)	Coefficient on two-year colleges (2)
1. Separate estimates for public and private four-year colleges		
Public colleges	1.814	0.181
	(0.075)	(0.046)
Private colleges	0.503	—
	(0.053)	
2. Transformation of single-sex colleges to coed colleges		
Male-only to coed	0.976	—
	(0.318)	
Female-only to coed	0.305	—
	(0.411)	
3. Number of colleges available in contiguous counties—dividing in two groups		
Contiguous counties have less than median number of colleges	1.074 (0.086)	0.299 (0.069)
Contiguous counties have more than median number of colleges	0.884 (0.051)	0.027 (0.060)
4. Number of colleges available in contiguous counties—linear interaction		
Number of colleges in own county	1.078	0.288
	(0.076)	(0.062)
Number of colleges in own county* Number of colleges in own contiguous counties	−1.236 (0.465)	−1.154 (0.498)

Standard errors are in parentheses, and are clustered at the county-year of mother's birth level. The sample includes mothers 24–45 at time of first birth. All coefficients in the first panel come from one regression which includes unrestricted mother age effects, unrestricted mother ten-year cohort of birth effects, county*year-of-child's-birth effects, median county income in year when mother was seventeen, and percent of the county that is urban in year when the mother was seventeen. In panel 1 there are not enough private two-year colleges to perform the same test for two-year colleges. In panels 3 and 4 the four coefficients in each panel come from a single regression which includes main effects for the Number of colleges in own contiguous counties as well as the other variables included in our baseline regressions.

college into a coed college should have no effect (or a negative effect if women are crowded out) on maternal education. In our sample, there are 104 cases of four-year colleges switching from male-only to coed, and 85 cases of four-year colleges switching from female-only to coed. Virtually all two-year colleges began as coed institutions. Table V shows that having a college change from being male-only to being coed, has an effect on maternal education that is similar to the effect of a new four-year college opening shown in Table II. The effect of having a college change from being female-only to coed is smaller, and not statistically significant.[18]

In the third and fourth panels of Table V, we ask whether the effects of new college openings are greater in locations far from other colleges. In panel 3 we divide counties into two groups according to whether the number of colleges (normalized by cohort size) in adjacent counties is above or below the median. These variables are then interacted with our measures of two- and four-year college availability. The results indicate that new four-year colleges do have a greater effect in college-poor areas, but the difference between college-rich and college-poor areas is not statistically significant. On the other hand, new two-year colleges are found to have a much greater impact on maternal education in college-poor areas.

In panel 4 we adopt a continuous measure of the availability of colleges in adjacent counties, defined as the average number of colleges in adjacent counties (again normalized by the relevant cohort sizes). This specification suggests that the effect of a new college falls off rapidly with the number of colleges in contiguous areas: an increase of one four-year college per 1000 residents 18 to 22 is estimated to increase maternal education by almost a year in areas with no colleges in contiguous counties. But the effect falls to zero as the number of colleges in contiguous counties rises to one per 1000 residents in the relevant age cohort.

Finally, we turn to the issue of endogenous mobility. There are two types of endogenous mobility that are potentially problematic for our instrumental variables strategy, given that in the Vital Statistics data we observe women only at the time of the first birth, and not at the time they were actually seventeen (or younger). First, women may move to counties that experience college openings to attend college. Second, college-educated women may move to counties that experience college openings after they graduate from college but before they have their first child. This second type of mobility may arise if labor market or marriage market opportunities for college graduates improve with the opening of a college. In Table III we presented evidence

18. Because the standard errors are large, one needs to be careful about the interpretation of these estimates. For example, one cannot rule out the possibility that the male-coefficient is as small as 0.40 or that the female-coefficient is as large as 1.00.

that indicates that the second type of mobility does not appear to be very significant.

Mobility of the first type is more worrisome. It is undoubtably the case that some women move to areas with higher college availability in order to attend college. The question is whether such mobility is important enough to account for a significant portion of the documented correlation between the number of colleges and maternal education (i.e., our first-stage coefficient). We want to make clear that, because of data limitations, we cannot completely rule out this possibility. We do have some indirect evidence that we hope may shed some light on the magnitude of the problem.

First, consider that new colleges in our sample appear to be, for the most part, nonselective institutions that are unlikely to attract a significant portion of the student body from far away. This is, of course, just a conjecture, since data on the county of residence of students prior to their enrollment do not exist. The only available data (the Integrated Postsecondary Education Data System) report the state of residence prior to enrollment. An examination of these data for 2303 colleges in 1998 showed that, while old, more established institutions attract a significant number of out-of-state students, new, less established colleges attract mostly in-state students. For example, while more than a third of the students enrolled in four-year colleges that are 100 years old or older are from out of state, only 9 percent of students enrolled in colleges ten years or younger are from out of state. The corresponding figures for two-year colleges are even lower. Unsurprisingly, community college students are overwhelmingly local, irrespective of the age of the college. See Table A1 in Currie and Moretti [2002] for details.

Although suggestive, these numbers are less than ideal. On one hand, since student mobility has steadily increased since 1940 [Hoxby 1996], mobility in most of the years we consider is likely to be lower than figures for 1998 would suggest. On the other hand, student mobility across states is likely to underestimate mobility across counties.

We explore the issue of mobility more directly using data from the NLSY and the 1980 and 1990 5 percent Public Use Micro Samples of the Census. We make use of the fact that, unlike the Vital Statistics, both the Census and the NLSY have information on the residential history of respondents. As explained in Section IV, we use this information to estimate different versions of our first-stage regression and compare them in order to assess the extent to which our results are driven by the fact that we observe location only at the time of the birth.

We begin by comparing the first-stage estimates for the Vital Statistics sample to estimates for the Census and the NLSY. As in Table II, in the top panel of Table VI we regress years of schooling on the (normalized) number of four- and two-year colleges that existed when the respondent was 14 to 17 in the location where she currently lives. Column (1) reports estimates for mothers 16 to 45 in our Vital Statistics sample. Column (2) reproduces estimates from Table II for mothers 24 to 45. Column (3) reports estimates for mothers 19 to 22 in our Vital Statistics sample. Column (4) reports equivalent estimates for all white women 19 to 22 in the 1980 and 1990 Censuses who report information on the location of residence five years prior to the Census (that is, when the respondent was 14 to 17). Column (5) reports estimates for all white mothers in the NLSY who had had a first birth by 1996 (these women are 20 to 41 at the time of their first birth).[19]

The comparison of column (2) with column (3) indicates that the first-stage coefficients are lower in younger samples. This is expected, because very young women have not had time for the availability of college to have its full effects. However, the availability of two- and especially of four-year colleges has a positive effect on educational attainment in all of the Vital Statistics samples. More importantly, estimates from the Census and the NLSY are qualitatively similar to estimates obtained from Vital Statistics data (when women in similar age ranges are compared) for four-year colleges, though they are not statistically significant for two-year colleges.

We now turn to the question of how different our first-stage estimates would have been if we had observed location of residence at age seventeen. The second panel of Table VI shows models estimated using data from the Census and NLSY which correspond to what we would estimate in the Vital Statistics, if data on the location of the mother's residence at age seventeen (or younger) were available in that data set. In the Census we re-

19. Models in column (4) include age, decade of birth, year, and location fixed effects. Models in column (5) include age, year, the AFQT score, mother's own mother's highest grade completed, and whether there was a father in the mother's household when she was fourteen as well as current state of residence fixed effects.

TABLE VI
THE EFFECT OF COLLEGE OPENINGS ON YEARS OF SCHOOLING, IN THE VITAL STATISTICS, CENSUS, AND NLSY

	Vital statistics age 16–45 (1)	Vital statistics age 24–45 (2)	Vital statistics age 19–22 (3)	Census age 19–22 (4)	NLSY age 20–39 (5)
Model 1					
Four-year colleges in current location of residence when respondent was 14–17	.674 (.037)	.950 (.046)	0.514 (0.030)	0.712 (0.174)	1.188 (.598)
Two-year colleges in current location of residence when respondent was 14–17	.059 (.029)	.176 (.046)	0.075 (0.025)	0.268 (0.177)	−.226 (.598)
Model 2					
Four-year colleges in location of residence at age 14–17	—	—	—	0.671 (0.175)	1.133 (.632)
Two-year colleges in location of residence at age 14–17	—	—	—	0.259 (0.175)	−.118 (.551)

Standard errors are in parentheses and are clustered at the county-year of mother's birth level. Sample in column (4) includes white women 19–22 from the 1980 and 1990 5 percent Census of Population. The dependent variable is years of schooling (highest grade completed as of 1996 in the NLSY). The models in the first two columns are the first stages from the Vital statistics regressions discussed above. The column (3) and (4) models include age, decade of birth, year, and location fixed effects. Models in column (5) are estimated using all white women who had had a child by 1996 and include the AFQT score, mother's own mother's highest grade completed, and whether there was a father in the mother's household when she was fourteen as well as current state of residence fixed effects. In column (3) $N = 375{,}066$; in column (4) $N = 305{,}225$; and in column (5), $N = 1{,}473$ in Model 1; and $N = 1{,}423$ in Model 2. In the NLSY, we use location as of the year the woman was fourteen. If highest grade completed or the current location in 1996 was missing in the NLSY, we imputed it using grades reported in interviews going back to 1990.

gress years of education as of age 19 to 22 on the availability of colleges in the location of residence and year when the woman was 14 to 17. In the NLSY we regress the highest grade completed as of 1996 on the availability of colleges in the county of residence and year that the woman was age 14.[20]

A comparison of the coefficients on the availability of four-year colleges in panels 1 and 2 indicates that the first stage based on current location of residence is remarkably similar to what would be obtained using the location of residence at

20. Although it is not reported in the table, we also estimated a model using NLSY data in which the education of the mother at the time of the first birth was regressed on college availability at age seventeen. This model produced a coefficient (standard error) of 1.397 (.580) on the four-year college variable and of −.001 (.593) on two-year colleges.

14–17. For example, the coefficients for four-year colleges for the Census are 0.712 (0.174) and 0.671 (0.175) in panel 1 and 2, respectively. The coefficients for two-year colleges for the Census are 0.268 (0.177) and 0.259 (0.175). Since the location of residence *before* the respondent reaches college age is less likely to be contaminated by endogenous mobility than the current location of residence, this similarity of coefficients lends some credibility to the assumption that endogenous mobility is not driving our results.

VII. Conclusions

We provide new evidence regarding the effect of maternal education on infant health, and on a series of other factors that are likely to influence infant health including smoking, use of prenatal care, marriage, and fertility. Our estimates suggest that increases in maternal education over the past 30 years have had large positive effects on birth outcomes. We estimate that an additional year of education reduces the incidence of low birth weight by approximately 10 percent, and reduces the incidence of preterm birth by 6 percent, on average. These effects arise because education affects maternal behavior (by reducing smoking by more than 30 percent, for example). Consistent with prior research, our estimates also suggest that education improves a woman's marriage market and reduces expected fertility.

In addition, our first-stage estimates are of interest in their own right, since they demonstrate that the boom in the construction of new colleges during the 1960s and 1970s had a significant impact on the education of mothers. These improvements in the education infrastructure increased schooling of women in the average county by .08 years and .01 years for four- and two-year colleges, respectively. The finding that college availability affects educational attainment is consistent with results in Card [1995], who uses college proximity as an instrumental variable to estimate returns to schooling, and in Goldin and Katz [1999], who show that the growth in the availability of public universities in the period 1890 to 1940 (which immediately precedes the 1940 to 1999 period that we focus on in this paper) significantly increased access to higher education and college enrollment.

On average, a change of .09 years in education reduced low birth weight by about 1 percent, and reduced the incidence of preterm birth by about half of 1 percent, according to our OLS estimates. If we use the IV estimates to calculate the impact on those most likely to have been affected, then the increase in education induced by the college openings is estimated to have reduced the incidence of low birth weight and preterm delivery by closer to 2 percent and 1 percent, respectively. While these may seem like small improvements, the costs of low birth weight (or conditions for which low birth weight is a marker) and prematurity are large. For example, it is estimated that between birth and age fifteen, low birth weight children incur an additional $5.5 to $6 billion more in health, education, and other costs than children of normal birth weight [March of Dimes 2002].

Educational attainment in the United States has increased dramatically in the past century. After decades of debate, economists now have a good idea of the benefits that education generates in terms of increased wages. Yet, there is increasing evidence that education has other benefits that are not reflected in the wage of educated individuals. Evidence in this paper indicates that improved infant health is an important benefit of education that may not be fully reflected in the wages of educated mothers. Furthermore, if educating mothers improves the health, educational attainment, and labor market outcomes of children, increases in educational attainment today may benefit multiple future generations.

Data Appendix

A. Reporting of Key Variables in the Vital Statistics Data

States began reporting maternal education at different points, so that we do not have a balanced panel of states. Appendix 1 shows which states did not report maternal education over time. Vital Statistics did not begin reporting smoking until 1989. In 1989, 43 states report. States that do not report include CA, IN, LA, NB, NY, OK, and SD. By 1992 the states that were not reporting were down to CA, IN, NY, and SD. In the most recent years, only CA and NY are missing. Finally, husband's education is not reported in detail after 1991. Models examining smoking or husband's education are estimated using a smaller data set than those examining other outcomes.

B. Construction of the College Openings Data Set

We began with a listing of accredited two- and four-year institutions that existed in fall 1996 from the National Center for Education Statistic's Integrated Postsecondary Education Data System (IPEDS) and excluded very small schools (those with less than 200 students in 1996). For four-year colleges we then searched the *Peterson's Guide to Four-Year Colleges* (1999) and the *Barron's Profiles of American Colleges* (1996) for information about the starting date of each college. If the college was listed in both sources and the founding date was the same in the two guides, then we accepted that date. If these conditions were not met, then we searched the internet for information about the founding date of the school.

In reading through college histories, we attempted to choose a date that was as close as possible to the date when actual undergraduate academic instruction began. For example, if the university began as a vocational school and later added academic instruction, we chose the later date. However if the school began as a "Normal school" or a teacher's college we did use that date. Similarly, if the founding date of the school was listed as the date when land was purchased, we used the date when instruction actually began. If the school began as a graduate school or divinity school and later added undergraduate instruction, we used the later date. If a university was formed through the merger of two schools, we used the date at which it opened its doors to the public.

We excluded the following categories of schools: schools of psychology, law schools, seminaries, Bible colleges and other mainly religious schools, schools that offer only distance or only internet learning, medical schools and medical centers, graduate schools and schools that were purely research facilities, and foreign universities offering degrees in the United States in their native languages.

Information on two-year colleges was collected in a similar fashion. There were essentially two types of two-year colleges that existed in 1996: those that offered one- or two-year degree programs as well as transfer programs to four-year schools, and those that offered only very specific vocational programs. However, we found that most two-year colleges offering broader programs now began as vocational institutions. Thus, for the sake of consistency, we decided to keep all two-year colleges, excluding only hairdressing and beauty schools. Excluding institutions that did not offer an Associates Degree had very little effect on our results, however.

An important potential problem is that some schools that existed over part of our sample period may have exited the sample by 1996. In order to assess the severity of this problem, we used data generously supplied by David Card about colleges that were in existence at five-year intervals over our sample period. This database was compiled using the Department of Education's CASPAR data base. We estimate that roughly 11 percent of schools that were in existence at some point had exited by 1996. Adding these schools to our database did not affect the estimates reported below.

Our instrumental variable is the number of colleges that existed in the woman's county when she was 17 years old divided by the estimated number of 18 to 22 year olds in the county in that year. Since population numbers are not available by age group, county, and year, we impute the number of 18 to 22 year olds in the county using the county population in each year (interpolating between Census years) and information about the number of 18 to 22 year olds in the state in each Census year. That is, we assume that 18 to 22 year olds are distributed across each state roughly in proportion to county populations. We experimented with two alternative imputation strategies, and found similar results.

In order to identify conversions of single-sex to coed colleges, we started with the 1968 to 1983 waves of the HEGIS (Higher Education General Information Survey) data available from the National Center for Education Statistics. These data identify each school as all male, all female, coed, or "coordinate" (which indicates that the institution had integrated classes even if it did not formally admit both men and women). We matched the data to the list of four-year colleges that we had already constructed from the 1996 IPEDS data. This enabled us to identify which of the colleges began as single-sex institutions, and the year that they went coed, if the year was between 1969 and 1983. For colleges that were still single sex in 1983, but had become coed by 1996, we checked college web sites in order to determine the year of conversion. We found that the typical school that changed from single-sex to coed was a small religious institution (rather than a Princeton or a Dartmouth). Also, we did not find any instances of two-year colleges that began as single-sex institutions. Colleges that were listed as "coordinate" at some point, generally held this status for one or two years before becoming coed. We used the first date that they were listed as coed as the date of conversion.

Reliance on the HEGIS data means that we may have missed conversions of single-sex colleges that occurred before 1968. Harworth, Maline, and DeBra [2002] cite data from the Education Directories produced by the U.S. Department of Education which show that there were 276 women's colleges in 1945, 248 women's colleges in 1955, and 252 women's colleges in 1960. They also cite problems with these data which suggest that one should not place a great deal of weight on small variations in the number of colleges reporting. The HEGIS data indicate that there were 258 women's colleges in 1969, but that the number declined rapidly after that (to 194 in 1970, and to 103 by 1983). Thus, our procedures capture the era of greatest change in the number of female colleges after 1955, though we miss one earlier period with many conversions between 1945 and 1955.

APPENDIX 1: STATES MISSING MATERNAL EDUCATION IN THE VITAL STATISTICS DATA, BY YEAR

1969	AL	AR	CA	CT	DE	DC	FL	GA	ID	MD	NM	PA	TX	WA
1970	AL	AR	CA	CT	DE	DC		GA	ID	MD	NM	PA	TX	WA
1971	AL	AR	CA	CT	DE	DC		GA	ID	MD	NM	PA	TX	WA
1972	AL	AR	CA	CT		DC		GA	ID	MD	NM	PA	TX	WA
1973	AL	AR	CA						ID	MD	NM	PA	TX	WA
1974	AL	AR	CA						ID		NM	PA	TX	WA
1975	AL	AR	CA						ID		NM	PA	TX	WA
1976		AR	CA						ID		NM		TX	WA
1977		AR	CA						ID		NM		TX	WA
1978			CA								NM		TX	WA
1979			CA								NM		TX	WA
1980–1988			CA										TX	WA
1989–1991														WA
1992+	All states report.													

APPENDIX 2: REDUCED-FORM ESTIMATES

	24–45 (1)	24–45 (2)	16–45 (3)
1. Low birth weight			
Four-year colleges	−.0096	−.0097	−.0055
	(.0036)	(.0036)	(.0023)
Two-year colleges	.0008	.0011	−.0058
	(.0043)	(.0043)	(.0026)
2. Preterm birth			
Four-year colleges	−.0105	−.0108	−.0063
	(.0041)	(.0041)	(.0047)
Two-year colleges	.0041	.0052	−.0004
	(.0052)	(.0052)	(.0031)
3. Prenatal care			
Four-year colleges	.0234	.0241	.0287
	(.0052)	(.0052)	(.0047)
Two-year colleges	−.0084	−.0090	.0121
	(.0066)	(.0066)	(.0053)
4. Smoked during pregnancy			
Four-year colleges	−.1148	−.1239	−.0627
	(.0024)	(.0024)	(.0024)
Two-year colleges	−.0066	−.0002	−.0004
	(.0033)	(.0033)	(.0034)
5. Married			
Four-year colleges	.0123	.0125	.0162
	(.0037)	(.0037)	(.0007)
Two-year colleges	−.0002	−.0002	.0003
	(.0004)	(.0004)	(.0007)
6. Husband's education			
Four-year colleges	.8890	.8941	.7041
	(.0491)	(.0492)	(.0301)
Two-year colleges	.1340	.1142	.0920
	(.0531)	(.0532)	(.0283)
7. Parity			
Four-year colleges	−.0809	−.0833	−.0877
	(.0091)	(.0091)	(.0092)
Two-year colleges	−.0076	−.0052	−.0093
	(.0010)	(.0010)	(.0011)
State-mother cohort trends	No	Yes	No

See notes to Table IV.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, AND THE NATIONAL BUREAU OF ECONOMIC RESEARCH

REFERENCES

Almond, Douglas, Kenneth Chay, and David Lee, "Does Low Birth Weight Matter? Evidence from the U. S. Population of Twin Births," Xerox, Department of Economics, University of California, Berkeley, August 2002.

Becker, Gary S., and H. Gregg Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy,* LXXXI (1973), S279–288.

Behrman, Jere R., and Mark R. Rosenzweig, "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review,* XCII (March 2002), 323–334.

Card, David, "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp,* L. N. Christophides, E. K. Grant, and R. Swidinsky, eds. (Toronto: University of Toronto Press, 1995), pp. 201–222.

——, "The Causal Effect of Education on Earnings," in *The Handbook of Labor Economics,* Vol. 3, Orley Ashenfelter and David Card, eds. (New York: North-Holland, 2000).

——, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica,* LXIX (2001), 1127–1160.

Card, David, and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics,* CXVI (2001), 705–746.

Currie, Janet, "Child Health in Developed Countries," in *The Handbook of Health Economics,* Anthony Culyer and Joseph Newhouse, eds. (New York: North-Holland, 2000).

Currie, Janet, and Rosemary Hyson, "Is the Impact of Health Shocks Cushioned by Socioeconomic Status?: The Case of Low Birth Weight," *American Economic Review,* LXXXIX (1999), 245–250.

Currie, Janet, and Enrico Moretti, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data," National Bureau of Economic Research Working Paper No. 9360, 2002.

Desai, Sonalde, and Soumya Alva, "Maternal Education and Child Health: Is There a Strong Causal Relationship?" *Demography,* XXXV (1998), 71–81.

Goldin, Claudia, "The Meaning of College in the Lives of American Women: The Past One Hundred Years," National Bureau of Economic Research Working Paper No. 4099, 1992.

——, "Career and Family: College Women Look to the Past," in Francine Blau and Ronald Ehrenberg, eds., *Gender and Family Issues in the Workplace* (New York: Russell Sage Press, 1997).

Goldin, Claudia, and Lawrence Katz, "The Shaping of Higher Education: The Formative Years in the United States, 1840 and 1940," *Journal of Economic Perspectives,* XIII (1999), 37–62.

Harworth, Irene, Mindi Maline, and Elizabeth DeBra, "Women's Colleges in the United States: History, Issues and Challenges," http://www.ed.gov/offices/OERI/PLLI/webreprt.html, August 19, 2002.

Hoxby, Caroline, "The Effects of Geographic Integration and Increasing Competition in the Market for College Education," National Bureau of Economic Research Working Paper No. 6323, 1997.

Lazear, Edward, "Intergenerational Externalities," *Canadian Journal of Economics,* XVI (1983), 212–228.

Lightwood, James, Cairan Phibbs, and Stanton Glantz, "Short-Term Health and Economic Benefits of Smoking Cessation: Low Birth Weight," *Pediatrics,* CIV (1999), 1312–1320.

Lleras-Muney, Adriana, "The Relationship between Education and Adult Mortality in the U. S.," Princeton University, Department of Economics, xerox, 2001.

Lochner, Lance, and Enrico Moretti, "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," National Bureau of Economic Research Working Paper No. 8606, 2001.

March of Dimes. Health Library: "Low Birthweight/Prematurity," www.modimes.org/HealthLibrary/355-1477.htm, March 2002.

Moretti, Enrico, "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions," 2002.

——, "Estimating the Social Return to Education: Evidence from Longitudinal and Repeated Cross-Sectional Data," *Journal of Econometrics,* forthcoming.

Rouse, Cecilia, "Democratization or Diversion? The Effect of Community College on Educational Attainment," *Journal of Business and Economic Statistics,* XIII (1995), 217–224.

Thomas, Duncan, John Strauss, and Maria-Helena Henriques, "How Does Mother's Education Affect Child Height?" *Journal of Human Resources,* XXVI (Spring 1991), 183–211.

Welch, Finis, "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," *Journal of Political Economy,* LXXXVII (1979), s65–97.

The World Bank, *World Development Report 1993: Investing in Health* (New York: Oxford University Press, 1993).