Education, Medical Knowledge and the Evolution of Disparities in Health

Anna Aizer*
Brown University and NBER

Laura Stroud
Brown Medical School

September 2009

We show that advances in medical knowledge can increase inequalities in health in the short run, as the more educated are first to act on new knowledge. However, over time, these inequalities decline as knowledge diffuses to the less educated. Our focus is the first Surgeon General's Report on Smoking and Health published in 1964. Using a unique dataset of pregnant women 1959-1966, we find that immediately after publication, the most educated mothers reduced their smoking, in contrast to the least educated who did not. Our findings do not simply reflect differences in reported smoking as we observe similar differential declines in serum cotinine levels. Using FE and IV techniques, we then link declines in smoking among the most educated mothers after the 1964 report with an increase in health disparities among newborns, at least initially. Over time as knowledge "diffuses" to the less educated, they too begin to reduce their smoking. As a result, the education gradient in newborn health, which continued to widen after publication of the 1964 report, peaks in the 1980s and than begins to shrink, returning to initial levels by 2000. These results imply that future advances in medical knowledge and technological innovation are likely to lead to a short run increase in health disparities that decline over the long run.

* Corresponding author: aizer@brown.edu. This work is supported in part by NSF grant # SES 0752755, NIH grant AG023397 and the Robert Wood Johnson Foundation. We thank Tianran Dai for excellent research assistance.

I. Introduction

Health disparities by education in the US are large. Males without a HS degree have a death rate double those with a college degree (Elo and Preston, 1996). Not only are these health disparities large, they are persistent, often originating in childhood and even earlier, in the newborn period (Case, Lubotsky and Paxson, 2002; Currie and Stabile, 2003; Currie and Moretti, 2003). A number of theories have been put forth to explain the observed education disparities in health.[1] We focus on one: that education improves one's ability to obtain, process and act upon medical knowledge, consistent with the theory of the production of health developed by Grossman (1972). If true, one implication is that advances in medical knowledge will lead to improvements in health among the most educated first, followed by eventual improvements among the less educated as medical knowledge diffuses. Thus advances in medical knowledge can lead to an initial *increase* in health inequality that will decline over time as medical knowledge diffuses to the less educated.

To test empirically whether advances in medical knowledge lead to an increase in the education gradient in health in the short run followed by a long run decrease, we focus on how the first major advance in medical knowledge regarding the ill effects of smoking on health affected both the smoking decisions of pregnant women and the health of their newborns. Our analysis is comprised of three parts. First, using a unique, prospective dataset on pregnant women for the period 1959-1966 we estimate the impact of the 1964 publication of the First Surgeon General Report on Smoking and Health on their decision to smoke and the health of

[1] These include the fact that those with more education 1) have more income and greater access to health care, 2) value the future more highly and therefore have greater incentives to invest in health, 3) have jobs with greater access to health insurance and less occupational risk, and 4) face less stress, owing to more supportive social networks and/or higher social rank, which recent work has shown to negatively affect health.

their newborns. We find that the most educated women responded most quickly to advances in medical information and that the education gradient in smoking increased immediately after publication of the report in January 1964. Moreover, we provide evidence that the differential declines in smoking do not simply reflect differences in reporting as serum cotinine levels follow the same pattern. Nor do they appear to reflect differences in income or cognitive ability: when we control for family income and maternal cognitive ability, the education gradient is unchanged. As expected, the education gradient in newborn health, as measured by birth weight, likewise increases immediately after publication of the report.

We follow this with an exploration of how inequalities in smoking and newborn health evolve over time.[2] We find that over time information appears to "diffuse" to the less educated, as evidenced by an eventual convergence in smoking behavior. And when it does, the education gradient in newborn health likewise declines. These two trends in the education gradients in smoking and newborn health, mirror each other almost exactly: increasing until the mid 1980s when they both peak and then declining to 1960s levels by 2000. These results have important implications for our understanding of how future advances in medical knowledge and technological innovation are likely to affect both health and health disparities over time. Moreover, by focusing on prenatal smoking, our results show how advances in medical knowledge can affect inequality in future generations.

Finally, we provide new estimates of the impact of smoking on newborn health using multiple identification strategies and, based on these results, revisit and attempt to reconcile the wide variation in existing estimates of the impact of smoking on birthweight. In so doing, we make two contributions to the existing literature on smoking on birthweight. First, because we

[2] To do so we extend our data from 1959 to the present using additional prospective datasets that include information on maternal education, smoking and birth outcomes: the National Natality Surveys (1969 and 1980) and birth certificate data (1970, 1980, 1990 and 2000).

have siblings for a FE analysis and an independent measure of smoking in the form of serum cotinine levels, we can quantify the extent of bias due to selection and measurement error, separately. Second, we can help to explain why the existing estimates differ so widely based on the estimation technique used. Our results imply that existing fixed effect estimates are too low and IV estimates too high and that the causal effect of smoking on birthweight is moderate in size (and non-linear) and lies between fixed effect and IV estimates.

The rest of the paper is organized as follows: in section II, we discuss the relevant literature; in section III we present results of our analysis of the immediate impact of publication of the 1964 surgeon general's report on the education gradient in prenatal smoking and newborn health; in section IV, we trace the evolution of the education gradient in smoking and birthweight over time to the year 2000. In section V we present the results of our analysis of the impact of prenatal smoking on newborn health based on multiple identification techniques and attempt to reconcile the (highly varying) estimates in the literature. Section VI concludes.

II. Background Literature

A. The Education Gradient in Health

The first major study to document differences in mortality by education in the US was conducted by Kitagwa and Hauser (1973) based on data from 1960. Since then, a number of additional studies have documented significant educational differences in mortality as well as other measures of health (Elo and Preston, 1996; Christenson and Johnson, 1995; Deaton and Paxson, 1999; Cutler and Lleras-Muney, 2006). Even the health of children is highly correlated with the educational attainment of their parents (Case, Lubotsky and Paxson, 2002). Moreover, it appears that these inequalities have been increasing over time (Elo and Preston, 1996; Pappas et al, 1993).

The documented relationship between education and health, however, is not necessarily causal. The positive relationship between education and health could reflect third factors correlated with both or it could reflect the fact that poor health reduces education (reverse causality).[3] More recent work has sought to establish a causal impact of education on health. Using compulsory schooling laws to instrument for education, Lleras-Muney (2003) finds lower mortality rates for the more educated. Currie and Moretti (2003) instrument for female education using the opening and closing of nearby colleges to estimate a significant and positive relationship between maternal education and infant health.

There are multiple potential mechanisms behind the relationship between education and health (Cutler and Lleras-Muney, 2006).[4] We focus on one - that education may improve an individual's ability to learn and make decisions about their health. This is consistent with Grossman's 1972 model of the demand for health in which education improves the efficiency with which individuals produce health. Specifically, education may increase the efficiency of health production by reducing the costs of obtaining medical information and/or using it. There is some empirical evidence to support this. Glied and Lleras-Muney (2008) find that for diseases with more innovation in medical treatment, education gradients in mortality increase, suggesting that the more educated take advantage of new medical innovation more quickly. Rosenzweig

[3] In support of the argument for reverse causality, Case, Fertig and Paxson, 2005 find that sick children are more likely to miss school and obtain less schooling. Bleakley (2002) finds that a de-worming campaign in the American South greatly increased years of schooling.

[4] First, education may lead to better health because it leads to greater income and access to health care. However, the documented relationship between education and health often remains, though somewhat diminished, when controls for income are included, suggesting that income does not explain the entire relationship between education and health (Elo and Preston, 1996; Cutler and Lleras-Muney, 2006). A second possible mechanism may be that the better educated tend to have less risky jobs. Again, Lleras-Muney and Cutler (2006) find that this can explain very little of observed education gradients in health. Third, education may improve health by affecting one's social rank. The Whitehall studies (Marmot, 2002) documented a strong and positive relationship between social rank and health among British civil servants. This is consistent with evidence based on experimental manipulation of social status in animals: Sapolsky (1993) finds that lower ranked animals suffer worse health. A fourth potential mechanism is that the more educated have lower discount rates, thereby increasing their value of the future and increasing their investments in health. However, Fuchs (1982) and Leigh (1990) find little empirical support for this.

and Schultz (1989) show that more educated women have greater success with "complex" contraception methods (eg, the rhythm method). Price and Simon (forthcoming) find that after publication of research on the riskiness of a particular procedure (VBAC), more educated women received differentially fewer of these procedures.[5] Goldman and Smith (2002) find higher rates of compliance with AIDS and diabetes treatment among the more educated.

Another area where education is likely to make a difference is in obtaining and using information on unhealthy behaviors such as smoking. We discuss this in greater detail below.

B. The Education Gradient in Smoking

The less educated are more likely to smoke and this relationship holds regardless of racial background or nativity (Kimbro et al, 2008).[6] Previous work on the relationship between education and smoking falls into two categories. The first is concerned with establishing a causal relationship between education and smoking. Sander (1995), using parental education as an instrument for own education in an IV regression, finds that the highly educated are more likely to quit smoking. De Walque (2004) using exposure to the draft for the Vietnam war as an instrument for college attendance provides suggestive evidence that education reduces smoking. Finally, and most relevant to the present study, Currie and Moretti (2003) use the opening and closings of nearby colleges to instrument for maternal education and find that more educated mothers are less likely to smoke while pregnant.[7]

[5] VBAC refers to "vaginal birth after cesarian section."
[6] Kimbro et al (2008) documents that they are also more likely to drink heavily and less likely to exercise.
[7] Farrell and Fuchs (1986) find that differences in smoking behavior between the more and less educated are present at age 17, before schooling is completed. They argue that this constitutes evidence that the relationship between education and smoking is not causal. However, as de Walque (2004) also argues, these results are insufficient to

A second strand of this literature seeks to link the education gradient in smoking with medical knowledge about the ill effects of smoking on health. Based on retrospective smoking histories collected in 1978-2000, de Walque (2004) finds that the most educated (those with at least a college degree) begin to reduce their smoking as early as the mid 1950s. In contrast, the smoking behavior of high school graduates and those with some college remains roughly steady until 1970 when it too beings to decline. High school drop-outs smoke at increasing rates throughout the 1960s until their smoking begins to decline in the mid 1970s. De Walque (2004) finds no discontinuous break in 1964 with publication of the first SG report on smoking and health.[8]

Meara (2001) and Kenkel (1991) explore whether knowing about the ill effects of smoking can explain the relationship between education and smoking. Meara (2001) uses cross sectional data from 1985 and 1990 that includes information on both maternal prenatal smoking and knowledge about the effects of smoking on health and finds that controlling for knowledge does little to change the estimated education gradient in smoking. Both Meara (2001) and Kenkel (1991) find that smoking knowledge and education have important interactive effects: the smoking behavior of the more educated is more responsive to knowledge than the behavior of the less educated. The authors conclude that the education gradient in smoking may be less attributable to the fact that the more educated have more knowledge about the ill effects of smoking and more to their greater responsiveness to that knowledge.

---

draw this conclusion because they do not account for the fact that much of the education gradient in smoking is attributable to differences in quitting behavior which often occurs much after schooling is completed.

[8] Meara (2001) analyzes smoking information collected in the NHIS data for 1966 and 1983 and approximates smoking rates for the pre-1964 period using the share of women who reported that had "ever smoked" in 1966. She finds larger declines for the more educated. However, this is consistent with the fact that the more educated are more likely to quit and cannot necessarily be attributed to changes in information.

A significant difference, however, between this study and that of Kenkel (1991) and Meara (2001) is that the reference period for this study is earlier, during a time when information about the ill effects of smoking was not widespread. Thus, it could still be the case that the increase in the gradient in smoking after 1964 was due to differences in knowledge, and not behavior conditional on knowledge. Kenkel and Liu (2008) find that the education-knowledge gradient increased between the 1950s and 1972 and the 1989 Surgeon General's report states that in 1966 only 34% of all adults thought pregnant women who smoked were more likely to have a premature birth and by 1985, the share had doubled (HHS, 1989). Moreover, the 1989 Report concluded that while (current) general knowledge of the negative consequences of smoking was widespread, it was often superficial, extending only to perceptions of greater risk for lung cancer and heart disease (HHS, 1989). These findings have important implications for the interpretation of our estimates of the impact of the 1964 report on the education gradient in smoking. The differential declines in smoking that we observe with publication of the 1964 report may reflect faster diffusion of knowledge to the more educated or faster modification of behavior based on that knowledge or some combination of the two.

Our analysis differs from existing work in both the quality of the data and focus of the work. The data used by Meara (2001) and De Walque (2004) are retrospective – solicited after publication of the 1964 report and based on recall. This increases the probability of classical measurement error and biased recall. In our analyses based on prospective reports of smoking for the period 1959-1966 which also include arguably more objective measures of smoking (serum cotinine), different trends emerge. Moreover, because of the richness of our data, we can control for multiple factors correlated with education and smoking, including cognitive ability and income, to further isolate the role of education. Most importantly, we trace the evolution in disparities in maternal smoking over a long time horizon and link it to evolving disparities in newborn health.

III. Impact of the 1964 Surgeon General's Report on the Education Gradient in Smoking

A. Data

For the analysis of the impact of the 1964 Surgeon General's Report on Smoking and Health we use data from the National Collaborative Perinatal Project (NCPP), a prospective survey of 59,000 pregnant women who sought care in one of 12 urban Academic Medical Centers in the years 1959-1966. The women were randomly recruited to participate in the survey through public clinics and thus the pregnant women included in the study are characterized by less education and lower income than the general population at the time.[9] This sample selection aids in our ability to compare the behavior and birth outcomes of more and less educated pregnant women because they sought and received the same medical care in terms of both quality and quantity, thereby reducing other potential differences across education groups. This, for example, implicitly allows us to rule out the possibility that the differential declines in smoking are due to differential sorting to medical providers.

Descriptive statistics for the sample are presented in Table 1. Women under the age of 19 or over the age of 35 at delivery were dropped from the analysis sample in order to separate the effects of education and age. In column 1 are sample means for the analysis sample (n=50,142);

[9] The only two things that disqualified women from participating were arriving at the clinic on the day of delivery or expressing a desire or intention to put the child up for adoption.

column 2 contains sample means for a subset of the data that consists of siblings (n=17,530). The average years of schooling is 11 and half of the women in the sample have not completed high school. The sample is racially mixed: 50 percent white, 43 percent black and 7 percent Hispanic (all from Puerto Rico and thus US citizens). Perhaps most remarkable are the rates of smoking among pregnant women as recorded in the third trimester: nearly half of the women smoked and those who did smoked on average half a pack of cigarettes daily.

B. Results

Publication of the First Surgeon General (SG) Report on Smoking and Health on January 11, 1964 resulted in an immediate decline in smoking among women with at least a high school degree relative to those without a high school degree (Figure 1a and 1b). Several observations can be made based on Figure 1a and 1b. First, prior to publication of the report in 1964, overall rates of smoking while pregnant were very high (47 percent) and differences in smoking between HS drop outs and HS graduates were small (5 percentage point difference). Second, smoking is actually increasing slightly over this period among the less educated. Third, among the more educated, smoking shows an immediate decline in 1964 that continues so that by 1966 the difference in smoking rates between HS drop outs and HS graduates is closer to 10 percentage points.

Given that education is correlated with other characteristics such as income and cognitive ability, we explore whether the observed relationship between education and smoking simply reflects trends in these other underlying characteristics by plotting trends in smoking for high and low income mothers and low and high IQ mothers, defined as below and above the median levels in the sample (Figure 2). We observe no trend breaks in smoking around 1964 associated with either income or cognitive ability. In fact, both income and cognitive ability are positively correlated with smoking.[10]

Regression Analyses

To control for any changes in the composition of pregnant women before and after 1964 that could explain the trends visible in Figure 1, we present results of a regression discontinuity analysis of the impact of the 1964 SG report on both prenatal smoking and newborn health. We estimate the following equation:

$$Y = \beta_1 \text{education} + \beta_2 \text{ education*post SG} + \beta_3 \text{ education*year} + \beta_4 \text{ education*year}^2 + \beta_5 \text{ year} + \beta_6 \text{ year}^2 + \beta_7 \text{ post SG} + \beta_8 X + \varepsilon$$

In the above specification, Y is either smoking (smoker, cigarettes per day) or a common measure of newborn health (birthweight). We include as regressors maternal education and its interaction with "post SG" (an indicator equal to 1 for years 1964-1966, after publication of the SG report, and 0 for years 1959-1963, before publication of the report). To allow for a time trend in the education gradient in smoking (or health) we include maternal education interacted with a quadratic time trend (education*year and education* year^2). The main effects (year, year^2, post SG) as well as a vector of personal characteristics including maternal race, age, family income, birth order, offspring gender and city of birth) are also included. The main

[10] In figure 2 with sample is all women because the measure of cognitive ability is missing for more than half the sample so that conditioning on those older than 18 would reduce the sample to very low levels. In contrast, the sample underlying figure 1 is all women at least 19 years old at the time of birth. Because of the difference in sample, in order to compare trends in Figure 2 with those in Figure 1 we present trends in smoking by education for the whole sample in figure 2c.

coefficient of interest is β_2 which indicates a change in the education gradient in smoking associated with the 1964 SG report that is independent of an underlying time trend.

The results, presented in Table 2 suggest that the education gradient in smoking and birthweight increase substantially after publication of the First SG Report on Smoking and Health. Before publication of the report, an additional year of education was associated with a 1.3 percentage point decline in the probability of smoking. After 1964, this increases by 50% from 1.3 to 1.9 and this difference is significant (column 1). If we look at cigarettes per day, we see that before the 1964 report, an additional year of education was associated with smoking .3 fewer cigarettes per day, after the report, it increases by one third to .4 fewer cigarettes per day (column 2). Finally in column 3, we present estimates of the impact of the SG report on the gradient in newborn health as measured by birthweight. Before publication of the report, an additional year of schooling is associated with heavier birthweight (25 gram difference). After the report, this increases by 20 % to 30 grams, though the difference is not significant. When the sample is expanded to include those younger than 19 as well, the estimate increases slightly and becomes significant (columns 4 and 5).

Though we include multiple maternal characteristics that might explain the relationship between maternal education and smoking or birthweight in the above analysis, in Table 3 we present the results of an analysis in which we allow the impact of all maternal characteristics (not just education) to change after the report. Specifically, we present the results of two regressions of smoking/newborn health on maternal education and other characteristics from the period just before (1962-1963) and just after (1964-1965) publication of the 1964 SG Report in the first two columns of Table 3. The negative relationship between maternal education and smoking as measured by whether she smokes at all (Panel A) and cigarettes smoked per day (Panel B) increases by one third after publication of the report. Before publication of the report, a standard deviation increase in years of schooling (2.5 years) was associated with a 3.5 percentage point decrease in the probability of smoking, and with smoking .75 less cigarettes per day. Immediately after the report, this increased to a 4.75 % and 1 fewer cigarette per day. We observe a very similar increase with respect to education and birthweight. However, this trend does not extend to heavy smokers defined as those smoking at least one pack of cigarettes a day (Panel C) whose smoking habits do not change appreciably with the publication of information on the ill-effects of smoking on health.

While the evidence thus far shows that women with more education responded immediately to advances in medical knowledge while the less educated did not, two questions remain. The first concerns the mechanisms that belie the relationship between education and smoking. Education is correlated with other characteristics that might explain the education gradient in smoking, including income and cognitive ability. Understanding whether education is merely a proxy for these other factors has important implications not only for our understanding of the education gradient in health, but the health production process more generally. Though the visual evidence presented in Figure 2 suggests that this not the case, we address this with regression analysis. Specifically, we present the results of regressions of the relationship between smoking and education (post the 1964 SG Report) controlling for income and cognitive ability, respectively. Adding family income at birth to the regression does nothing to change the coefficient estimate on maternal education, suggesting that the relationship between maternal education and smoking is not mitigated through income (column 3). In column 4 we control for a measure of maternal cognitive ability and the coefficient on maternal education actually increases to -0.024, not surprisingly since cognitive abiliaty and education are positively correlated and cognitive ability is positively correlated with smoking, at this time.[11] Results for cigarettes smoked per day are very similar (Panel B).

A second concern is that the observed relationship between maternal education and smoking merely reflects differences in reported smoking. This would be the case if more educated women perceive a stigma associated with smoking after publication of the first Surgeon General's Report that less educated women do not, thereby differentially affecting their reports of smoking. In panel D of Table 3 we present results for a small subset of the sample for which we have third trimester serum cotinine levels. Cotinine and reports of cigarettes smoked per day are highly correlated in these data (ρ=0.72). Regression results suggest that maternal education is strongly and negatively associated with serum cotinine levels and that this relationship increases 40 percent after 1964.[12] After 1964, an additional 2.5 years of schooling is associated with a 20 percent standard deviation reduction in serum cotinine levels. The cotinine results do not diminish when we control for income, but when we control for maternal cognitive ability the coefficient on maternal education increases by one third.

Based on these results, we conclude that immediately after publication of the first significant advance in medical information about the ill effects of smoking on health, the negative relationship between maternal education and smoking increased by one third and that this was not driven by differences in income or cognitive ability. Nor does it reflect differences in reporting by education level. Rather, the evidence seems to suggest that schooling itself affects the decision to smoke presumably by reducing the costs of learning and/or acting on new

[11] It should be noted that maternal cognitive ability is missing for many of the women in the sample which is why the sample size decreases in column 4 of Table 3. While the sample for this regression is not random, the result is not driven by sample selection.

[12] This suggests that after the 1964, smokers are smoking fewer cigarettes and smoking each more intensely as was found after tax hikes (Adda and Cornaglia, 2006; Evans and Farrelly, 1998).

knowledge. Moreover, this increase in the gradient in smoking after publication of the report is accompanied by an increase in the gradient in newborn health as measured by birthweight.

While the above analyses of the impact of the First SG Report on Smoking and Health based on cross-sectional comparisons include controls for many potential differences in the characteristics of women before and after publication of the report such as income, marital status, age, number of children, it's possible that important unobserved differences remain. In the next sub-section we address this by limiting the sample to siblings and including maternal fixed effects.

The SG Report and Changes in the Education Gradient in Smoking – Maternal FE Analyses

To control for potential differences in any unobserved characteristics of pregnant women before and after publication of the 1964 SG report, we limit the sample to mothers who had multiple children over this period and include maternal fixed effects (n=17,287), thereby limiting our comparison to the same women before and after publication of the 1964 SG report. For these regressions we include a linear time trend, an indicator for post 1964 and maternal fixed effects.[13] Moreover, we stratify the sample multiple ways to assess whether the smoking decisions of some mothers were more responsive to advances in medical knowledge than others.

On average, we witness a decline in the probability of smoking of 3.5 percent after publication of the report (Table 4, column 1, Panel A) and half a cigarette less per day (Panel B). The decline in heavy smoking is insignificant. In columns 2 and 3 we stratify the sample by maternal education (HS drop out vs. HS graduate). While the change in the probability of smoking is similar for both groups, the decline in cigarettes smoked per day is twice as great for

[13] Because the sample includes only 2 years per mother, on average, it is not feasible to include higher order terms.

HS graduates relative to HS drop outs. Income does not appear to affect decisions to smoke after the 1964 report (columns 4 and 5). There is some suggestive evidence, however, that maternal health and the birth weight of the previous child affect how a mother responds to the 1964 SG report, though the estimates are very imprecise. Mothers defined as “not sick” (having no reported medical condition) are slightly more likely to reduce their smoking relative to those defined as “sick” (at least one medical condition), as are mothers whose previous child was born of normal weight relative to those whose previous child was born LBW. Fixed effect regression estimates of the impact of smoking on birthweight, while suggestive of an increase in the gradient, were not significant and not presented here.

We conclude based on the results presented thus far that 1) the education gradient in smoking increased considerably after the 1964 SG report on smoking in health, 2) this increase did not simply reflect changes in reports of smoking, but actual smoking as measured by serum cotinine levels, 3) the observed negative relationship between education and smoking does not reflect differences in income or cognitive ability, and 4) the education gradient in newborn health, as measured by birthweight, likewise increased after the 1964 SG report. Though we have providence evidence of greater responsiveness of the more educated to information about the ill effects of smoking on health that cannot be explained by differences in income or cognitive ability, it is possible that this reflects differences across schooling categories in individual rates of time preference. We have no evidence to bear directly on this hypothesis. However, this interpretation would not be consistent with the emerging consensus of a causal impact of education on health, health decisions, and smoking decisions, more specifically, based on instrumental variable analyses. Nor would it be consistent with existing literature that shows that differences in rates of time preferences explain little of the disparities in health by socio-economic status (Fuchs, 1982).

IV. The Evolution of the Education Gradient in Smoking and Health

Having shown that publication of the First SG Report on Smoking and Health in 1964 resulted in an immediate increase in the education gradient in smoking and newborn health, we now examine how the gradients in both smoking and health have evolved over time. Specifically, we are interested in both assessing how long it takes for advances in medical knowledge to "diffuse" to the less educated as evidenced by declines in their smoking behavior, and, given the negative relationship between smoking and newborn health, understanding the implications of this "slow diffusion" for the evolution of disparities in newborn health.

A. Data

For this analysis we use multiple sources of data. Vital statistics does not include information on smoking during pregnancy until 1989, so for years prior to 1989 we use data from the National Natality Surveys (NNS) of 1969 and 1980. The NNS were conducted by the National Center for Health Statistics which randomly sampled certificates of live birth and mailed questionnaires to 3,611 new mothers with births in 1969 and 9,941 new mothers with births in 1980. Data on demographic characteristics and birthweight came from birth certificates and the maternal surveys solicited information on income smoking habits while pregnant. The resulting sample, when weighted, is representative of the population of married mothers with US citizenship. Single mothers and non-citizens were excluded from the surveys. In 1970 and 1980, approximately 88 and 73 percent of all births were to married US citizens, respectively, suggesting small potential bias in estimates from 1969 but potentially more bias in 1980 as more women were excluded.

To address any potential bias caused by the omission of these two groups, we compare estimates of the education gradient in birthweight from the NSS with estimates based on vital statistics data which is representative of the entire population of births for 1970 and 1980. We find that they are very similar, suggesting that any bias from the NNS sampling scheme is likely minimal.[14]

B. Results

Based on regressions of smoking on maternal education and additional controls for the period 1969 - 2006, we find that an additional year of education is associated with smoking 0.21 fewer cigarettes per day in 1969, 0.48 fewer cigarettes per day in 1980, 0.45 fewer cigarettes in 1990, 0.23 fewer cigarettes by 2000 and returning to levels below initial 1969 levels by 2006 (-.14), (see columns 2,4,6,8,10 of Table 5). The education gradient in any smoking is similar (columns 1,3,5,7,9 of Table 5). In the second panel we present the same results redefining education as an indicator variable for being in the top 25% of the education distribution in that year. We do this instead of examining categories of education (HS drop out, HS graduate, College) because the content of each of these designations has changed over time. We see the

[14] In a regression with birthweight as the outcome and in which controls for race, maternal age, birth order, and offspring gender are included, the coefficient estimate on a continuous variable in maternal years of schooling in the NNS in 1969 is 7.9 while in the vital statistics data it was 8.9 in 1970and in 1980, the comparable estimates are 12.3 and 11.5. Recall that we cannot repeat this exercise for the education gradient in smoking because the vital statistics data do not include information on smoking until 1989.

same pattern: mothers in the top 25 percent of the education distribution smoked 1.4 fewer cigarettes per day in 1969, 2.5 fewer cigarettes in 1980, 2 fewer cigarettes in 1990, 1 fewer cigarette by 2000 and less than 1 fewer cigarette by 2006. Thus the education gradient in smoking which was relatively small prior to publication of the 1964 SG report, widens over time until sometime between 1980 and 1990 when it begins to fall, declining to levels below 1969 levels by 2006, the most recent year for which data are available.

To isolate the relationship between education and smoking separate from income, we can re-estimate the above education gradient regressions conditioning on income for 1969 and 1980 (because the NNS data also include information on income that the vital statistics data do not). When we do, the education gradient in smoking increases slightly from -.21 to -.28 in 1969 and from -.48 to -.54 in 1980, suggesting that the effect of education on smoking while pregnant is not operating through income.

Finally, we explore the implication of these results for the evolution of health disparities. If advances in medical knowledge benefit the most educated immediately, eventually diffusing to the less educated, then this would suggest that education gradients in health will show a similar pattern: increasing initially and eventually declining. This is what we find in Table 6 where we present the results of regressions of birthweight on maternal education and other demographic characteristics over time. In 1969, an additional year of education is associated with 7.9 more grams at birth, in 1980 it increases to 12.3, declining to 10.9 in 1990, to 8.3 by 2000 and to 6.7 by 2006, which implies a lower education gradient in health in 2006 relative to 1969 (columns 1,3,5,7,9 Table 6).[15]

[15] This is consistent with Racine and Joyce (2004) who document a similar decline in the income gradient in birth outcomes (LBW and mortality) in New York City over the period 1988 -2000.

These findings are consistent with the hypothesis that advances in medical knowledge will initially benefit the most educated, leading to short run increases in health disparities that will, once medical knowledge diffuses to the less educated, eventually decline. Visually, we can see this in Figure 3 where we have plotted regression coefficients on maternal education from the birthweight and smoking regressions presented in Tables 5 and 6. As the education gradient in smoking increases between 1969 and 1980 (that is, as an additional year of schooling is associated with fewer smoked cigarettes), so too does the education gradient in birthweight. Likewise, as the gradient in smoking declines between 1990 and 2006, so too does the gradient in birthweight. As evident in Figure 2, trends in the relationship between education and smoking and education and birthweight are nearly mirror images of one another.

Yet many other things that might affect the education gradient in smoking are also changing over this period, not only changes in smoking behavior. In an attempt to isolate the role of smoking in explaining the education gradient in birthweight, we control for cigarettes smoked in these regressions. If changes in smoking behavior explained all the change in the gradient in health over this period, we would expect that once we control for smoking, the education gradient in birthweight would remain unchanged. In fact, when we control for smoking the education gradient in nirthweight is not constant but is much flatter. Controlling for smoking, the education gradient in birthweight still increases between 1969 and 1980, but only slightly from 6.1 to 8.5 (compared with a rise from 7.9 to 12.3 when we do not control for smoking), then remains unchanged between 1980 and 1990, finally declining to 5.5 by 2000 and 4.2 by 2006. We interpret this as suggestive evidence that changes in the education gradient in birthweight can be explained, in part, by changes in the education gradient in smoking, particular over the period

1969-1990, but less for the period after when changes in other factors also appear to play a strong role.

Thus we conclude based on the evidence presented here that while advances in medical knowledge lead to an immediate increase in the education gradient in smoking and health, over time the behavior of the less educated starts to converge to that of the more educated. As a result, the education gradients in both smoking and health, which initially increased, ultimately fall to levels observed in the 1960s.

For the above analysis, we implicitly assume a negative causal impact of smoking on birthweight, as supported by the existing medical and economic research. However, though previous research has consistently produced estimates that are negative, the estimated size of the effects has varied substantially with the estimation method used. Because of the richness of our data and ability to estimate the impact of smoking on birthweight using multiple identification strategies, we can 1) quantify the bias in OLS estimates due to selection and measurement error, separately, 2) explain and reconcile the variation in existing estimates, and 3) provide a new estimate of the impact of smoking on birthweight.

V. Prenatal Smoking and Newborn Health

Previous attempts to estimate a causal relationship between smoking and birthweight have used multiple techniques to overcome potential bias from omitted variables and measurement error in reports of smoking. These techniques have included maternal fixed effects, propensity score matching, randomized controlled trials and IV, with the results varying with the estimation method used.[16] Studies using fixed effects (Rosenzweig and Wolpin, 1991;

[16] Another method used by Fertig (2009) is to compare OLS estimates of the impact of smoking on birthweight over time to assess the degree of selection into smoking. This method assumes that negative selection into smoking has

Abrevaya, 2006) produce estimates that are smaller than OLS estimates. Those using propensity score techniques produce estimates similar to OLS estimates (Almond, Chay and Lee, 2005) as do those correcting for selection using the Heckman selection technique (Grossman and Joyce, 1990). Sexton and Hebel (1984) conduct a randomized controlled trial of a smoking cessation program for pregnant women and find that women randomly assigned to the smoking cessation program smoked less and delivered babies that weighed significantly (92 grams) more. The final method, IV, produces estimates that are considerably larger than OLS estimates, with the highest estimate nearly double that of OLS.[17]

Our contribution to the existing literature is to assess the bias in OLS estimates due to selection and measurement error, separately. To do so, we exploit 1) the large number of siblings in the data that allows us to include maternal fixed effects to address selection bias and 2) a subset of data with serum cotinine levels which allows us to address the issue of measurement error in reports of smoking in an IV framework. We argue that a careful comparison of the OLS, FE and IV estimates allows us to 1) determine the amount of bias due to selection and measurement error, respectively, 2) produce an estimate that adequately accounts for both and 3) explain the variation in existing estimates. We start with OLS estimates.

A. OLS Estimates

---

grown over time and her finding that OLS estimates based on 1958 data from the UK are considerably smaller than those based on 2000 data suggests that selection can explain a substantial portion of current OLS estimates.

[17] Evans and Ringel (1999) use state cigarette taxes as instruments for smoking while pregnant, yielding IV estimates of 350-600 grams. Lien and Evans (2005) using the same technique but a limited sample yields a smaller IV estimate of 189 grams. Finally, Permutt and Hebel (1989) using data from the 1984 randomized control trial for smoking cessation during pregnancy and instrumenting for smoking status using the randomization, estimate that smoking reduces birthweight by 430 grams.

OLS estimates suggest that smoking is associated with a decline in birthweight of 187 grams, or 28 percent of a standard deviation (Table 7). Defining smoking as cigarettes smoked per day, we find that each cigarette smoked per day reduces birthweight by 20 grams, and the impact is decreasing in the number of cigarettes smoked (Column 1, Panel B). In column C, we explore this non-linearity and categorize women as non-smokers (omitted), light smokers (1-10 cigarettes per day), moderate smokers (11-20 cigarettes per day) and heavy smokers (more than 20 cigarettes per day). Light smokers deliver babies that are on average, 128 grams lighter, while moderate and heavy smokers deliver babies that are nearly 300 grams lighter (43 percent of a standard deviation).

We repeat the analysis defining the birth outcome as a LBW birth and find that smokers face a 6 percent increase in the probability of a LBW birth (column 4) which represents a 50 percent increase and that the impact is again concave in cigarettes smoked per day: light smokers face a 4 percent increase in the probability of a LBW birth, and moderate and heavy smokers face between a 7.4 and 8.8 percent increase.

We also explore whether smoking is correlated with prematurity and find that it is, though the relationship is weaker than with birthweight. Smokers deliver on average 2 days earlier (Table 7 column 6) and like birthweight, there seems to be a concave relationship between number of cigarettes smoked per day and gestation at birth. When we define the outcome as premature (delivery before 36 weeks gestation), we find that smokers face a 2 percent increase in the probability of delivering prematurely (the average probability of prematurity is 16 percent in this sample), so this represents a moderate impact.

To address potential selection bias in OLS estimates, we present fixed effect estimates below.

B. FE Estimates

To assess the extent of selection bias in OLS estimates, we include maternal FE, thereby reducing identifying variation in smoking to that between births to the same mother. The sibling subsample (n=16,483) used for the FE analysis is very similar to the full sample along nearly every dimension including income, education, race, smoking habits and birth outcomes (Table 1). For purposes of comparison and to assess generalizability of the fixed effect estimates that are based on a subset of the full sample, OLS estimates of the impact of smoking on birthweight based on the sibling subsample are presented in column 2 of Table 7. They are very similar to the OLS estimates based on the full sample, suggesting that the FE estimates are likely generalizable to the full sample.

When we include maternal fixed effects, the coefficient on smoking declines by two thirds from -175 to -58 (Table 7 column 3), suggesting that once we control for omitted variables smoking reduces birthweight by less than 60 grams or 9 percent of a standard deviation. Smoking an additional cigarette per day reduces birthweight by 12 grams, relative to the OLS estimate of 20 grams, and the effect is decreasing in the number of cigarettes smoked (Table 7 column3, panel B). Women who smoke more than 10 cigarettes per day on average can expect newborns weighing 125 grams fewer (19 percent of a standard deviation), which is half the OLS estimate, but still represents a moderate effect. The decline in the estimated impact of smoking on the probability of a LBW birth when fixed effects are included is similar to the decline in the birthweight regressions but are not significant due, perhaps, to the low probability of LBW and relatively small sample size.

Interestingly, when fixed effects are included in the regressions estimating the relationship between smoking and gestation/prematurity, the point estimates do not decline at all, and in some cases actually increase, though they become imprecise, suggesting that selection bias may be less of an issue in estimates of the impact of smoking on gestation/prematurity.

We explore the source of the birthweight effects in Table 8. First, we estimate the extent to which the reduction in birth weight is attributable to less mature birth. To do so, we present results of a fixed effect regression in which the outcome is defined as birthweight(in grams)/ gestation (in weeks). Conditional on gestation, smoking is still associated with a decline in birthweight, though the decline is half the unconditional size (Table 8, column 1).[18] Next we explore the extent to which smoking reduces weight as opposed to body size as measured by length and head circumference. To compare estimates across regressions, we log transform the dependent variables: ln(body length), ln(head circumference) and ln(birth weight). We find that smoking has only a small effect on body length and head circumference (columns 2-3, respectively) relative to its impact on weight. This suggests that the newborns of mothers who smoked during pregnancy are considerably thinner but only slightly smaller as measured by length and head circumference.

C. Selection into Smoking

We attempt to identify the possible omitted variables associated with maternal smoking and birthweight that appear to bias upwards OLS estimates. The factors we do observe in these data that are excluded from most datasets and are known to be correlated with birthweight and

[18] Without conditioning on gestation, a standard deviation increase in cigarettes smoked per day, 9, reduces birthweight by 14 percent of a standard deviation, while the same increase in cigarettes smoked per day reduces birthweight conditional on gestation by 8 percent of a standard deviation.

maternal characteristics include maternal health, maternal size (height and weight), weight gain during pregnancy and maternal cognitive ability. However, when we include these variables (Table 8, columns 5-9) the estimated OLS relationship between smoking and birthweight is unchanged, suggesting that these unobservables are not driving the difference between OLS and FE estimates of the relationship between smoking and health at birth.

It may be that despite the richness of the NCPP we still do not observe the omitted variables that explain the considerably smaller maternal FE estimates. Alternatively, the much smaller FE estimates might reflect the fact that classical measurement error in the smoking variable which leads to attenuation bias is magnified in a FE stetting (Grilliches, 1979). We explore this next.

D. Measurement Error in Smoking Reports

To assess potential measurement error in these data we instrument for maternal smoking with a measure of serum cotinine measured in the third trimester of their pregnancy. This measure is highly correlated with reports of maternal smoking but it is also correlated with any unobservables that may be correlated with both smoking and newborn health. As such, this instrument only addresses bias due to measurement error in reports of smoking, it does not address bias due to selection into smoking.

Comparison of results from OLS, FE and IV regressions enables us to assess the extent of bias in the OLS estimate due to omitted variables (selection) and measurement error, respectively and to adjust (upwards) our FE estimate to account for measurement error in smoking. OLS results in Table 9 columns 1 and 2 based on the full and sibling subsample suggest that each additional cigarette smoked per day reduces birthweight by 11.7 grams. [19] The fixed effect estimate in column 3 is much smaller, -3.7. Column 4 and 5 contain the first and second stages, respectively, of an IV regression. Cotinine is a very strong predictor of reported smoking (t statistic = 31). However, we do not interpret the cotinine measure as the "true" measure as it is also subject to measurement error (recall that an instrument need not be a "better" measure of the endogenous variable, only an alternative one). Specifically, it is a single spot measure that reflects only relatively recent smoking which may or may not be the average level. Moreover, it can also reflect second hand smoke.

When we do instrument for reports of smoking with cotinine, the resulting IV estimate is much larger than the OLS estimate: -18.6 vs -11.7, yielding a reliability ratio of 0.63 in the measure of smoking and suggesting that measurement error in smoking leads to a considerable downward bias in the OLS estimate. As previously noted, attenuation bias due to classical measurement error is exacerbated in a fixed effect setting (Grilliches, 1979) with this "exacerbation" increasing in the degree of correlation in the smoking reports of siblings which in this case is high (ρ=0.77). It should be noted that the greater impact of smoking instrumented with cotinine does not reflect the fact that conditional on smoking, those with a higher cotinine levels suffer worse birth outcomes. In appendix Table 1 we provide estimates of the direct (or reduced form) impact of cotinine on birthweight. The estimated impact of cotinine on birthweight is in fact slightly smaller than the estimated impact of reported smoking.

Based on this, we conclude that OLS estimates are biased upward by 59 percent due to negative selection into smoking and biased downward by 37 percent due to classical measurement error. On net, the results suggest that OLS estimates are biased downward by 22 percent. Moreover, we can calculate a causal estimate that accounts for both selection and

[19] We exclude the quadratic term because we have only one instrument.

measurement error by adjusting the fixed effect estimates for classical measurement error in smoking. When we do, we find that smoking results in 9.6 fewer grams at birth per cigarette smoked per day and -152 grams for smoking.

As a final identification strategy, we instrument for maternal smoking using the publication of the 1964 report interacted with maternal education as an instrument.[20] This instrument is designed to mitigate bias from both measurement error and selection. The results of the first and second stages are presented in the last two columns of table 7, respectively. The first stage is strong (t statistics of 5 and 6). As expected, the point estimate for the second stage (-15.6 for cigarettes smoked per day and -311 for smoking) falls between the fixed effect and IV estimates based on cotinine, reflecting the fact that this second IV estimate corrects for both measurement error and selection. However, the coefficient is imprecisely estimated.

We draw four main conclusions based on these analyses. First, OLS estimates of the impact of smoking on birthweight are biased upward due to negative selection into smoking and biased downward due to measurement error in smoking reports; second, fixed effect estimates which correct for selection bias are significantly attenuated by measurement errors; third, a comparison of OLS and IV estimates based on cotinine suggest a reliability ratio of .63 in reports of prenatal smoking (cigarettes/day); fourth, the true causal estimate of smoking on birthweight corrected for selection and measurement error represents a moderate effect only slightly smaller than OLS estimates (-152 grams for smoking and -9.6 grams per cigarette/day). These estimates (from OLS, FE and IV) can help to explain the substantial variation in existing estimates that are based on different estimation methods. With respect to the fixed effect estimates, we conclude that they are too small because of exacerbated measurement error. This also explains why

[20] The two main effects, publication of the report and maternal education, are included in the first and second stages of the IV regression (eg, they are not excluded as instruments but rather are allowed to affect birthweight directly).

estimates based on propensity score and selection correction, which are subject to less measurement error, are larger than the FE estimates. With respect to IV estimates, our results generally support those of Brachet (2005) who presents evidence that the IV estimates based on cigarette taxes are inflated by as much as 43 percent due to misclassification of the endogenous binary smoking variable. This finding would imply that IV estimates in the presence of less misclassification of the endogenous binary variable would yield lower IV estimates. Our second set of IV results is consistent with this: less misclassification in reports of smoking in the 1960s yield IV estimates roughly one quarter lower than those based on data from the 1980s and 1990s which are characterized by greater misclassification.

VI. Conclusions

The goals of this paper were threefold. First, using data that represented an important improvement over previous datasets, we showed that after publication of the first SG report on smoking and health in 1964, more educated women immediately reduced their smoking while the least educated did not and that these differences cannot be explained by differences in income or cognitive ability, nor do they simply reflect differences in reported smoking as differential declines in serum cotinine levels were also observed. Second, we provided evidence that the education gradient in smoking which increased immediately after 1964, continued to increase until the mid 1980s when it began to decline as the smoking behavior of the less educated began to converge to that of the more educated. We followed this with evidence that the trend in the education gradient in smoking witnessed over the past half century was mirrored by trends in the education gradient in birth weight which initially increased after the 1964 report, and likewise, beginning in the mid 1980s, began to decline. Finally, we provided new evidence on the impact of smoking on newborn health using FE and IV techniques that allowed us to assess the degree of measurement error and selection in smoking reports separately and provide a causal estimate that is moderate in size and adjusts for both sources of bias.

We conclude based on these findings that increasing health disparities is a likely byproduct of advances in medical knowledge or technologies, which the more educated are quicker to adopt. Moreover, as we showed here, the increase in inequality can persist to the next generation. However, over time, the disparities decline as the less educated eventually adopt the new information or technology. Obviously, this does not imply that scientific progress should be eliminated even though it will, at least initially, increase health disparities. It can, however, explain why despite efforts to reduce inequalities in health, they continue to persist in an era characterized by continuous medical advancements.

References – INCOMPLETE

Abrevaya J. 2006. Estimating the effect of smoking on birth outcomes using a matched panel data approach. Journal of Applied Econometrics 21: 489–519.

Adda J, Cornaglia F. 2006. Taxes, cigarette consumption, and smoking intensity. American Economic Review 96(4):1013–1028.

Almond D, Chay KY, Lee DS. 2005. The costs of low birth weight. Quarterly Journal of Economics 120(3):1031–1083.

Bleakley, H. 2007. Disease and development: Evidence from hookworm eradication in the American South. Quarterly Journal of Economics 122(1): 73-117.

Brachet T. 2005. Maternal smoking, misclassication, and infant health. Unpublished manuscript available at http://works.bepress.com/tbrachet/1

Case, A., A. Fertig and C. Paxson. 2005. The lasting impact of childhood health and circumstance. Journal of Health Economics 24: 365-389.

Case, A., D. Lubotsky, and C. Paxson. 2002. Economic status and health in childhood: the origins of the gradient. American Economic Review 92(5): 1308-1334.

Cutler, D.M., and A. Lleras-Muney. 2007. Education and health: evaluating theories and evidence, in J.S. House, R.F. Schoeni, G.A. Kaplan, and H. Pollack (eds.), The Health Effects of Social and Economic Policy, New York: Russell Sage Foundation.

Currie, J., and M. Stabile. 2003. Socioeconomic status and health: Why is the relationship stronger for older children? American Economic Review 93(5): 1813-1823.

Deaton, A., D. Cutler, and A. Lleras-Muney. 2006. The determinants of mortality. Journal of Economic Perspectives 20(3): 97-120.

Deaton, A. 2003. Health, inequality, and economic development. Journal of Economic Literature 41(1): 113-158.

Deaton, A. and C. Paxson. 2004. Mortality, income, and income inequality over time in Britain and the United States, in D.A. Wise (ed.), Perspectives on the Economics of Aging. Chicago: University of Chicago Press, 247-279.

De Walque, D. 2004. Education, Information and Smoking Decisions: Evidence from Smoking Histories, 1940-2000. World Bank Policy Research Working Paper 3362.

Elo, I. T. and S.H. Preston. 1996. Educational differentials in mortality: United States, 1979-85. Social Science and Medicine 42(1): 47-57.

Evans WN, Ringel JS. 1999. Can higher cigarette taxes improve birth outcomes? Journal of Public Economics 72:135–154.

Fertig, A. 2009. Selection and the Effect of Prenatal Smoking. Health Economics.

Fuchs, Victor. 1982. "Time Preferences and Health: An Exploratory Study" in Victor Fuchs (ed) Economic Aspects of Health. Chicago: the University of Chicago Press.

Glied, S. and A. Lleras-Muney. 2008. 'Health inequality, education, and medical innovation'. Demography 45(3): 741-761.

Grossman, M. 1972. The Demand for Health—A Theoretical and Empirical Investigation. New York: National Bureau of Economic Research.

Grossman M, Joyce TJ. 1990. Unobservables, pregnancy resolutions, and birth weight production functions in NewYork City. Journal of Political Economy 98(5): 983–1007.

Kenkel, D. 1991. Health Behavior, Health Knowledge and Schooling. Journal of Political Economy. 99(2):287-305.

Kitagawa, E. M. and P. M. Hauser. 1973. Differential Mortality in the United States: A Study in Socioeconomic Epidemiology, Cambridge, MA: Harvard University Press.

Lien DS, Evans WN. 2005. Estimating the impact of large cigarette tax hikes: the case of maternal smoking and infant birth weight. Journal of Human Resources 40(2): 373–392.

Lleras-Muney, A. 2005. The relationship between education and adult mortality in the United States. Review of Economic Studies 72: 189-221

Marmot, M.G., G.D. Smith, S. Stansfield, et al. 1991. Health inequalities among British civil servants: The Whitehall II study. Lancet 337: 1387-1393.

Meara E. 2001. Why is health related to socioeconomic status? The case of pregnancy and low birth weight, April.Unpublished manuscript available as NBER Working Paper 8231.

Pappas, Gregory, Susan Queen, Wolber Hadden and Grail Fisher (1993), "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986," The New England Journal of Medicine.

Permutt T, Hebel R. 1989. Simultaneous-equation estimation in a clinical trial of the effect of smoking on birthweight. Biometrics 45(2): 619–622.

Price, Joseph and Kosali Simon (forthcoming) "Patient Information and the Impact of New Medical Research" Journal of Health Economics.

Racine AD, Joyce, TJ. 2004. The Disappearing Income Gradient in New York City Birth Outcomes: Thirteen Years of Convergence from 1988 to 2001. 114: e51-e57.

Sapolsky, R.M. 1993. Endocrinology alfresco: Psychoendocrine studies of wild baboons. Recent Progress in Hormone Research 48: 437-468.

Sexton, M, Hebel R. 1984. A Clinical Trial of Change in Maternal Smoking and Its Effect on Birth Weight. Journal of the American Medical Association 25(1):911-915.

US Office of the Surgeon General. 1964. Smoking and health: report of the advisory committee of the surgeongeneral of the public health service. Technical Report, Public Health Service, Office of the Surgeon General.(Public Health Service Publication No. 1103.)

Table 1: Descriptive Statistics: National Collaborative Perinatal Project 1959-1966

	Full Sample		Sibling Sample	
	mean	Std dev	mean	Std dev
Maternal Characteristics				
Maternal education (years)	10.86	2.55	10.82	2.39
HS Drop out	0.52		0.53	
HS graduate	0.33		0.35	
Some college	0.09		0.07	
College +	0.06		0.05	
SRA Rank Quotient (IQ)	89.77	19.86	90.28	19.62
Family income (in 2007 $)	$ 27,200	$ 15,050	$ 27,570	$ 14,270
Married	0.80		0.84	
White	0.49		0.52	
Black	0.43		0.44	
Hispanic	0.07		0.03	
Asian	0.00		0.00	
Maternal age	24.20	4.56	24.23	4.47
male	0.51		0.51	0.50
Birth order	2.72	2.21	3.16	2.12
Smoking Variables				
Smoker	0.48		0.49	
Cigarettes per Day	6.07	9.31	6.44	9.46
Cigarettes per Day conditional on Smoking	12.70	9.80	13.10	9.70
Birth Outcomes				
Birth weight	3108	660	3080	706
Gestation (weeks)	38.68	4.83	38.38	5.18
Low Birth Weight	0.12		0.14	
Observations	50142		17530	

note: Sample excludes all women less than 19 at time of birth

Table 2 Changes in the Education Gradient in Smoking and Birthweight Associated with the 1964 SG Report on Smoking and Healtl

				Full Sample	
	Smoker	Cigarettes per day	Birthweight	Birthweight	LBW
Maternal education*post 1964	-0.006	-0.092	5.54	7.363	-0.004
	[0.003]	[0.053]	[3.739]	[3.524]	[0.002]
Maternal education	-0.013	-0.305	24.754	25.78	-0.011
	[0.006]	[0.078]	[5.367]	[5.173]	[0.003]
Maternal education*year	-0.001	-0.017	-5.158	-5.195	0.002
	[0.002]	[0.029]	[2.246]	[2.267]	[0.001]
Maternal education*year^2	0	0.001	0.339	0.284	0
	[0.000]	[0.003]	[0.251]	[0.265]	[0.000]
Black	0.101	0.273	-83.004	-85.777	0.072
	[0.038]	[0.590]	[49.058]	[46.794]	[0.020]
White	0.244	5.08	100.044	103.786	0.017
	[0.037]	[0.564]	[47.909]	[45.802]	[0.019]
Hispanic	-0.081	-1.787	13.566	14.115	0.041
	[0.042]	[0.646]	[54.919]	[51.641]	[0.025]
Birth Order	0.015	0.389	14.004	13.715	-0.001
	[0.002]	[0.035]	[2.345]	[2.233]	[0.001]
Married	-0.13	-2.609	34.005	25.204	-0.009
	[0.009]	[0.174]	[10.625]	[8.861]	[0.005]
Ln(family income in $1000)	0.047	0.674	57.149	20.249	-0.004
	[0.018]	[0.327]	[20.939]	[6.676]	[0.004]
Male	0.005	0.098	111.571	110.299	-0.025
	[0.006]	[0.107]	[7.045]	[6.340]	[0.004]
Maternal age	-0.007	-0.037	2.422	3.43	0
	[0.001]	[0.018]	[1.104]	[0.967]	[0.001]
Year	7.347	93.225	-12,526.04	-11,876.55	2.892
	[7.529]	[80.317]	[7,579.202]	[8,122.320]	[4.250]
Year squared	-0.002	-0.024	3.198	3.034	-0.001
	[0.002]	[0.020]	[1.929]	[2.067]	[0.001]
Post 1964	0.048	0.673	-98.002	-113.066	0.055
	[0.035]	[0.591]	[40.599]	[37.205]	[0.022]
Observations	40663	40663	39958	48393	49436
R-squared	0.06	0.1	0.05	0.06	0.02

Robust standard errors in brackets
Also included are indicators for city of birth.

Table 3: Changes in The Gradient in Smoking and Birthweight

Panel A: Smoker	1962/1963	1964/1965	1964/1965	1964/1965
Maternal education	-0.014	-0.019	-0.019	-0.024
	[0.002]	[0.002]	[0.002]	[0.003]
Married	-0.141	-0.116	-0.115	-0.119
	[0.013]	[0.011]	[0.011]	[0.016]
White	0.218	0.065	0.053	0.049
	[0.120]	[0.091]	[0.095]	[0.124]
Black	0.08	-0.073	-0.085	-0.099
	[0.121]	[0.091]	[0.095]	[0.125]
Hispanic	-0.115	-0.183	-0.195	-0.189
	[0.125]	[0.094]	[0.098]	[0.130]
Asian	-0.183	-0.273	-0.284	-0.141
	[0.121]	[0.095]	[0.099]	[0.196]
Maternal age	-0.01	-0.006	-0.006	-0.006
	[0.001]	[0.001]	[0.001]	[0.001]
Male	-0.017	-0.008	-0.008	0.003
	[0.009]	[0.008]	[0.008]	[0.010]
Birth Order	0.019	0.012	0.012	0.01
	[0.003]	[0.002]	[0.002]	[0.003]
Ln(family income in $1000)			-0.002	0.001
			[0.009]	[0.011]
SRA rank quotient				0
				[0.000]
Observations	14184	16305	16294	9691
R-squared	0.06	0.05	0.05	0.06
Panel B: Cigarettes Smoked per Day				
Maternal education	-0.298	-0.404	-0.408	-0.565
	[0.040]	[0.036]	[0.037]	[0.052]
Ln(family income in $1000)			0.07	0.217
			[0.169]	[0.217]
SRA rank quotient				0.02
				[0.007]
Observations	14184	16305	16294	9691
R-squared	0.1	0.1	0.1	0.11
Panel C: Heavy Smoker (at least 1 pack per day)				
Maternal education	-0.005	-0.005	-0.005	-0.006
	[0.001]	[0.001]	[0.001]	[0.001]
Ln(family income in $1000)			0	0.004
			[0.004]	[0.005]
SRA rank quotient				0
				[0.000]
Observations	14184	16305	16294	9691
R-squared	0.04	0.04	0.04	0.05
Panel D: Cotinine				
Maternal education	-3.844	-10.268	-10.247	-14.975
	[1.573]	[2.433]	[2.452]	[2.403]
Ln(family income in $1000)			-1.292	3.547
			[10.210]	[10.802]
SRA rank quotient				0.589
				[0.283]
Observations	375	318	318	279
R-squared	0.05	0.08	0.08	0.14
Panel E: Birthweight				
Maternal education	10.517	13.015	12.417	15.7
	[2.933]	[2.383]	[2.406]	[3.119]
Ln(family income in $1000)			21.666	23.531
			[10.618]	[13.059]
SRA rank quotient				-0.236
				[0.399]
Observations	12662	14666	14657	8842
R-squared	0.06	0.05	0.05	0.06

Robust standard errors in brackets

Table 4: Changes in Smoking Behavior Across Birth Stratified by Maternal Characteristic, Includes FE

Panel A: Smoker	All	<HS	>=HS	Poor	Non Poor	Sick	Non Sick	LBW	Not LBW	White	Black
post 1964	-0.035	-0.031	-0.038	-0.04	-0.03	-0.034	-0.052	0.015	-0.041	-0.023	-0.044
	[0.013]	[0.017]	[0.021]	[0.021]	[0.017]	[0.016]	[0.057]	[0.037]	[0.014]	[0.018]	[0.021]
year	0.007	0.011	0.001	0.015	0	0.006	0.016	-0.007	0.008	-0.002	0.016
	[0.004]	[0.006]	[0.007]	[0.007]	[0.006]	[0.005]	[0.016]	[0.014]	[0.005]	[0.006]	[0.007]
Observations	17287	9429	7858	6596	10691	14086	3201	2399	14502	9041	7564
R-squared	0.89	0.89	0.89	0.9	0.89	0.9	0.95	0.94	0.89	0.9	0.89

Panel B: Cigarettes per Day	All	<HS	>=HS	Poor	Not Poor	Sick	Non Sick	LBW	Not LBW	White	Black
post 1964	-0.507	-0.344	-0.709	-0.425	-0.547	-0.426	-0.634	-0.401	-0.512	-0.545	-0.421
	[0.231]	[0.341]	[0.303]	[0.416]	[0.265]	[0.285]	[0.887]	[0.665]	[0.249]	[0.334]	[0.332]
year	0.458	0.53	0.365	0.51	0.408	0.449	0.491	0.389	0.462	0.465	0.445
	[0.074]	[0.110]	[0.093]	[0.128]	[0.085]	[0.093]	[0.243]	[0.223]	[0.079]	[0.105]	[0.108]
Observations	17287	9429	7858	6596	10691	14086	3201	2399	14502	9041	7564
R-squared	0.9	0.9	0.92	0.89	0.92	0.91	0.96	0.93	0.9	0.91	0.87

Panel C: >20 Cigarettes per Day	All	<HS	>=HS	Poor	Non Poor	Sick	Non Sick	LBW	Not LBW	White	Black
post 1964	-0.009	-0.009	-0.009	0.005	-0.018	-0.01	-0.01	-0.025	-0.007	-0.013	-0.005
	[0.008]	[0.012]	[0.011]	[0.014]	[0.010]	[0.010]	[0.035]	[0.027]	[0.009]	[0.013]	[0.010]
year	0.01	0.012	0.007	0.007	0.011	0.011	0.005	0.011	0.01	0.013	0.006
	[0.003]	[0.004]	[0.003]	[0.004]	[0.003]	[0.003]	[0.011]	[0.008]	[0.003]	[0.004]	[0.003]
Observations	17287	9429	7858	6596	10691	14086	3201	2399	14502	9041	7564
R-squared	0.77	0.77	0.76	0.73	0.79	0.79	0.84	0.8	0.76	0.78	0.7

Robust standard errors clustered on mother in brackets

Table 5: The Education Gradient in Smoking Over Time

	1969		1980		1990		2000		2006	
Panel A: Education in years	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day
Maternal education in years	-0.0121	-0.213	-0.0202	-0.478	-0.0295	-0.452	-0.0193	-0.223	-0.0125	-0.1377
	[0.00451]	[0.0858]	[0.00249]	[0.0534]	[0.00013]	[0.0022]	[9.36e-05]	[0.0013]	[127.00]	[114.49]
White	0.157	3.802	-0.141	-3.328	0.0711	1.368	0.0273	0.518	0.0566	0.6375
	[0.0518]	[0.748]	[0.0146]	[0.261]	[0.0015]	[0.0219]	[0.0009]	[0.0099]	[82.26]	[85.33]
Black	0.108	0.449	-0.0889	-2.07	-0.0103	-0.587	-0.0801	-0.919	-0.0413	-0.5465
	[0.0567]	[0.794]	[0.0649]	[1.442]	[0.0017]	[0.0244]	[0.0010]	[0.0115]	[44.37]	[56.42]
Maternal age	0.00118	0.076	-0.00902	-0.104	-0.0022	-0.0162	-0.00187	-0.0164	-0.097	-0.9874
	[0.00326]	[0.068]	[0.00171]	[0.0352]	[0.0001]	[0.0017]	[7.52e-05]	[0.0010]	[113.23]	[107.76]
Male	-0.00863	0.478	0.00325	0.238	-0.00021	0.00144	0.000416	-0.00229	-0.001	-0.0514
	[0.0204]	[0.408]	[0.0110]	[0.230]	[0.00057]	[0.0093]	[0.0004]	[0.0054]	[1.98]	[8.58]
birth order	0.00163	0.0879	-0.00291	-0.0824	0.0216	0.374	0.0208	0.243	-0.0008	-0.0065
	[0.00741]	[0.155]	[0.00421]	[0.0845]	[0.00023]	[0.00408]	[0.000178]	[0.00249]	[1.65]	[1.20]
Married					-0.207	-2.883	-0.145	-1.493	-0.004	-0.0315
					[0.0012]	[0.0203]	[0.000799]	[0.0112]	[49.73]	[33.02]
Hispanic					-0.1	-1.517	-0.108	-1.184	-0.1164	-1.1098
					[0.0015]	[0.0216]	[0.000821]	[0.00941]	[142.49]	[112.66]
Native born					0.0873	1.091	0.0688	0.552	0.0195	0.2069
					[0.0010]	[0.0141]	[0.000562]	[0.00661]	[96.85]	[83.01]
Observations	2417	2417	6734	6734	1517622	1502984	1741927	1722596	1107945	1107945
R-squared	0.006	0.017	0.029	0.03	0.11	0.094	0.115	0.084	0.1	0.08

	1969		1980		1990		2000		2006	
Panel B: HS Graduate	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day
Mother HS graduate	-0.132	-2.382	-0.0895	-2.018	-0.175	-2.973	-0.113	-1.474	-0.0786	-0.9125
	[0.0259]	[0.534]	[0.0181]	[0.400]	[0.00128]	[0.0242]	[0.00102]	[0.0156]	[70.00]	[63.08]
Observations	2417	2417	6734	6734	1517622	1502984	1741927	1722596	1105797	1105797
R-squared	0.015	0.024	0.023	0.023	0.099	0.088	0.105	0.078	0.09	0.07

	1969		1980		1990		2000		2006	
Panel C: Top 25% Education Distribution	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day	Smoker	Cigs/Day
Mother in top 25% education distribution	-0.0668	-1.426	-0.119	-2.537	-0.136	-2.008	-0.0943	-1.014	-0.0731	-0.7448
	[0.0228]	[0.442]	[0.0119]	[0.243]	[0.000594]	[0.00938]	[0.000418]	[0.00529]	[156.78]	[138.88]
Observations	2417	2417	6734	6734	1517622	1502984	1741927	1722596	1107945	1107945
R-squared	0.007	0.018	0.034	0.032	0.11	0.093	0.114	0.081	0.1	0.08

Robust standard errors in brackets

Notes: Regressions for 1969 and 1980 based on National Natality Surveys with population weights. No state FE included.

Table 6: The Education Gradient in Birthweight Over Time

Panel A: Education in years	1969		1980		1990		2000		2006	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Maternal education in years	7.921	6.079	12.26	8.537	10.86	8.17	8.27	5.51	6.7238	4.1928
	[5.075]	[5.014]	[3.132]	[3.115]	[0.1674]	[0.2129]	[0.1699]	[0.1910]	[29.62]	[18.48]
Cigarettes per day		-8.648		-7.797		-14.88		-17.32		-18.382
		[1.225]		[0.734]		[0.0868]		[0.1285]		[92.79]
White	112.3	145.2	-269.5	-295.5	166.35	170.26	180.72	187.66	146.678	158.3968
	[71.95]	[72.13]	[20.73]	[20.98]	[2.0228]	[2.7197]	[1.7517]	[2.0462]	[77.30]	[83.48]
Black	-92.39	-88.5	-39.26	-55.4	-84.16	-100.66	-40.98	-58.31	-55.6605	-65.707
	[78.51]	[78.51]	[64.63]	[62.65]	[2.3577]	[3.0356]	[2.1446]	[2.4013]	[22.67]	[26.80]
Maternal age	6.835	7.492	4.396	3.583	-0.23	-1.09	2.02	1.52	31.1836	114.3356
	[3.816]	[3.778]	[2.125]	[2.118]	[0.1422]	[0.1641]	[0.1377]	[0.1494]	[27.16]	[48.52]
Male	106	110.1	132.2	134.1	126.50	127.89	114.80	115.68	114.6417	30.2381
	[23.61]	[23.37]	[13.01]	[12.93]	[0.7849]	[0.8975]	[0.7792]	[0.8425]	[110.41]	[26.46]
birth order	13	13.76	7.313	6.67	15.65	22.69	13.55	18.54	2.8838	114.5214
	[8.934]	[8.854]	[5.321]	[5.350]	[0.3088]	[0.3582]	[0.3029]	[0.3268]	[15.98]	[110.81]
Hispanic					140.56	83.94	155.68	114.40	7.5195	2.3044
					[2.0608]	[2.9101]	[1.7638]	[2.1190]	[19.04]	[12.82]
Native born					-30.98	-0.53	-22.74	-15.13	132.4854	59.2672
					[1.4218]	[1.8189]	[1.2557]	[1.4116]	[56.25]	[40.68]
Married					134.32	99.62	89.10	65.69	79.6678	11.3234
					[1.2954]	[1.5988]	[1.1814]	[1.3222]	[54.96]	[28.74]
Observations	2417	2417	6734	6734	2024664	1501708	2032666	1721462	1107945	1107945
R-squared	0.025	0.046	0.04	0.054	0.053	0.078	0.04	0.053	0.04	0.05

Panel B: HS Graduate	1969		1980		1990		2000		2006	
Mother HS graduate	71.44	51.16	110.6	94.92	86.07	63.14	55.81	38.73	40.5553	23.6345
	[30.77]	[30.50]	[20.66]	[20.55]	[1.4000]	[1.7274]	[1.4296]	[1.6288]	[21.24]	[12.46]
Cigarettes per day		-8.514		-7.777		-14.95		-17.47		-18.5426
		[1.231]		[0.730]		[0.0864]		[0.1282]		[93.75]
Observations	2417	2417	6734	6734	2024664	1501708	2032666	1721462	1105797	1105797
R-squared	0.026	0.046	0.041	0.056	0.053	0.078	0.039	0.053	0.04	0.05

Panel C: Top 25% Education Distribution	1969		1980		1990		2000		2006	
Mother in top 25% education distribution	21.78	9.391	38.44	18.36	61.02	34.55	40.61	24.46	42.2488	28.6327
	[26.84]	[26.73]	[14.68]	[14.85]	[0.8509]	[0.9788]	[0.8737]	[0.9413]	[35.19]	[23.77]
Cigarettes per day		-8.693		-7.912		-14.94		-17.41		-18.2815
		[1.226]		[0.736]		[0.0868]		[0.1284]		[92.26]
Observations	2417	2417	6734	6734	2024664	1501708	2032666	1721462	1107945	1107945
R-squared	0.024	0.045	0.039	0.053	0.054	0.078	0.039	0.053	0.04	0.05

Robust standard errors in brackets

Notes: Regressions for 1969 and 1980 based on National Natality Surveys with population weights. No state FE included.

Regressions for 1990 and 2000 based on vital statistics data. State FE included but results very similar if excluded

Table 7: Impact of Smoking on Birthweight and Gestation

	Birthweight (grams)			LBW		Gestation (weeks)		Premature	
Panel A: Smoker	OLS	OLS	FE	OLS	FE	OLS	FE	OLS	FE
Smoker	-187.105	-174.905	-58.171	0.059	0.015	-0.254	-0.247	0.021	0.033
	[7.052]	[13.944]	[32.936]	[0.008]	[0.020]	[0.083]	[0.256]	[0.008]	[0.023]
Maternal education	9.293	10.082		-0.003		-0.024		-0.001	
	[1.584]	[3.287]		[0.002]		[0.020]		[0.002]	
Ln(family income in $1000)	26.994	34.696	15.897	-0.004	0.001	0.155	0.187	-0.017	-0.016
	[7.137]	[15.046]	[20.178]	[0.008]	[0.014]	[0.097]	[0.176]	[0.010]	[0.016]
Married	8.592	28.08	-23.474	-0.028	0.026	-0.125	-0.508	-0.002	0.05
	[9.665]	[21.399]	[37.516]	[0.012]	[0.023]	[0.138]	[0.276]	[0.013]	[0.028]
White	9.706	52.445		0.003		-0.451		0.052	
	[69.293]	[173.420]		[0.054]		[0.509]		[0.034]	
Black	-204.32	-168.69		0.061		-1.605		0.143	
	[69.876]	[173.929]		[0.055]		[0.521]		[0.035]	
Hispanic	-137.997	-158.692		0.108		-1.926		0.198	
	[73.472]	[194.657]		[0.081]		[0.638]		[0.065]	
Asian	-279.99	-267.392		-0.033		-0.939		-0.005	
	[87.758]	[200.010]		[0.064]		[0.652]		[0.042]	
Maternal age	1.19	-2.472	-51.507	0.003	0.035	-0.015	-0.437	-0.001	0.029
	[0.797]	[1.802]	[15.989]	[0.001]	[0.010]	[0.010]	[0.114]	[0.001]	[0.010]
Male	116.485	120.436	137.983	-0.034	-0.037	-0.238	-0.088	0.021	0.016
	[6.458]	[11.847]	[15.004]	[0.007]	[0.010]	[0.075]	[0.114]	[0.008]	[0.011]
Birth Order	18.721	22.717	41.068	0	-0.008	-0.036	-0.059	0.006	0.003
	[2.081]	[4.502]	[15.264]	[0.003]	[0.009]	[0.026]	[0.117]	[0.003]	[0.010]
Observations	46907	16483	16493	16901	16912	16840	16851	16840	16851
R-squared	0.08	0.08	0.81	0.03	0.71	0.04	0.69	0.04	0.7
Panel B: Cigarettes per Day									
Cigarettes per day	-20.615	-20.371	-12.398	0.007	0.003	-0.031	-0.028	0.002	0.003
	[0.870]	[1.660]	[3.996]	[0.001]	[0.002]	[0.010]	[0.029]	[0.001]	[0.003]
Cigarettes per day squared	0.324	0.331	0.213	0	0	0.001	0.001	0	0
	[0.027]	[0.049]	[0.095]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
Observations	46907	16483	16493	16901	16912	16840	16851	16840	16851
R-squared	0.09	0.09	0.81	0.04	0.71	0.04	0.69	0.04	0.7
Panel C: Heavy Smoking									
1-10 Cigarettes per Day	-127.738	-104.842	-43.618	0.04	0.01	-0.122	-0.22	0.014	0.034
	[8.135]	[15.905]	[32.570]	[0.009]	[0.020]	[0.100]	[0.257]	[0.010]	[0.023]
11-20 Cigarettes per Day	-268.549	-269.061	-126.574	0.088	0.037	-0.46	-0.357	0.034	0.03
	[9.978]	[19.513]	[46.332]	[0.011]	[0.029]	[0.111]	[0.363]	[0.011]	[0.033]
>20 Cigarettes per Day	-272.429	-257.716	-121.178	0.074	0.013	-0.335	-0.413	0.022	0.009
	[17.551]	[31.006]	[70.163]	[0.017]	[0.048]	[0.180]	[0.514]	[0.017]	[0.049]
Observations	46907	16483	16493	16901	16912	16840	16851	16840	16851
R-squared	0.09	0.09	0.81	0.03	0.71	0.04	0.69	0.04	0.7

Robust standard errors in brackets

Table 8: Impact of Smoking on Birthweight - Extensions

	Weight for gest (grams/weeks)	Ln(body length)	Ln(head circ)	Ln(weight)	Birthweight - OLS Regressions				
Cigarettes per day	-0.2219	-0.0006	-0.0003	-0.0066	-20.615	-20.559	-19.773	-20.47	-20.775
	[0.1388]	[0.0004]	[0.0003]	[0.0025]	[0.870]	[0.870]	[0.903]	[0.869]	[1.136]
Cigarettes per day squared	0.0041	0	0	0.0001	0.324	0.323	0.293	0.318	0.336
	[0.0027]	[0.0000]	[0.0000]	[0.0001]	[0.027]	[0.027]	[0.028]	[0.027]	[0.036]
Maternal education					7.968	7.871	8.728	7.424	2.41
					[1.588]	[1.590]	[1.756]	[1.566]	[2.251]
Ln(family income in $1000)	0.7214	0.0001	0.0022	0.0116	27.298	27.081	22.406	27.381	29.931
	[1.0436]	[0.0023]	[0.0018]	[0.0180]	[7.121]	[7.115]	[7.387]	[7.095]	[9.380]
Mother has 1 health condition						9.298			
						[8.981]			
Mother has 2-3 health conditions						11.024			
						[9.001]			
Mother has >3 health conditions						-41.51			
						[12.456]			
Maternal weight gain, net							7.486		
							[0.450]		
Maternal height(inches)								1.763	
								[0.388]	
Maternal weight pre-pregnancy (lbs)								0.009	
								[0.029]	
Maternal Cognitive Ability (SRA - Rank Quotient)									0.556
									[0.294]
Observations	16435	15985	16055	16840	46907	46907	38327	46611	21952
R-squared	0.68	0.75	0.77	0.69	0.09	0.09	0.1	0.09	0.1
Robust standard errors in brackets									
FE Included	Y	Y	Y	Y	N	N	N	N	N
mean of dependent variable	79	3.9	3.5	8	3149	3149	3149	3149	3149
std deviation of dependent variable	21.57	0.08	0.07	0.26	718	718	718	718	718
std deviation of cigarettes per day	9	9	9	9	9	9	9	9	9
Increase of 1 cigarette per day on outcome as % std deviatio	-1%	-1%	0%	-3%	-3%	-3%	-3%	-3%	-3%
Impact of std dev increase in smoking on outcome	-1.7	0.0	0.0	-0.1	-165.2	-164.7	-160.0	-164.3	-165.6
Impact of std dev increase in smoking on outcome as % std d	-8%	-7%	-4%	-21%	-23%	-23%	-22%	-23%	-23%

Table 9: Impact of Smoking on Birthweight: Comparison of OLS, FE and IV Estimates

Panel A: Cigarettes per Day	OLS	OLS	FE	First Stage	IV	First Stage	IV
Cigarettes per day	-11.64	-11.74	-3.661		-18.563		-15.634
	[0.340]	[0.593]	[1.338]		[2.345]		[11.522]
Maternal education	7.399	6.732		-0.302	6.847	-0.295	5.917
	[1.420]	[2.667]		[0.120]	[8.403]	[0.024]	[4.505]
Ln(family income)	24.143	29.692	4.295	-0.551	-38.424	0.26	25.148
	[6.298]	[12.152]	[15.315]	[0.618]	[42.238]	[0.090]	[6.942]
Maternal age	-0.273	-3.661	-72.927	0.054	1.036	-0.033	-0.412
	[0.863]	[1.588]	[12.160]	[0.069]	[4.747]	[0.012]	[0.952]
Birth Order	21.177	26.573	46.648	0.026	14.783	0.39	22.74
	[1.837]	[3.457]	[9.683]	[0.161]	[11.001]	[0.026]	[4.869]
Male	110.77	101.047	118.141	-0.592	138.511	0.042	110.937
	[5.970]	[10.690]	[11.316]	[0.494]	[33.856]	[0.085]	[5.999]
Married	5.248	13.395		-2.089	-18.625	-2.462	-4.593
	[8.384]	[16.696]		[1.073]	[73.971]	[0.119]	[29.598]
White	132.277	103.258		0.333	6.098	4.513	150.463
	[30.906]	[62.132]		[3.227]	[220.824]	[0.442]	[60.904]
Black	-122.461	-166.249		-3.63	-287.494	-0.315	-123.592
	[31.850]	[63.557]		[3.291]	[225.217]	[0.456]	[32.068]
Hispanic	-15.952	-20.682				-1.773	-23.156
	[35.036]	[74.448]				[0.501]	[40.782]
Post 1964	-30.818	-20.965	8.9	-0.033	-147.534	1.961	-32.367
	[12.450]	[21.929]	[21.869]	[0.982]	[67.181]	[0.424]	[13.248]
Birth year	-35.019	-26.287	50.705	1.694	-44.01	0.351	-33.514
	[7.944]	[14.656]	[19.626]	[0.775]	[53.087]	[0.114]	[9.064]
Birth year squared	4.252	3.188	0.588	-0.167	6.403	-0.025	4.142
	[0.986]	[1.828]	[1.875]	[0.091]	[6.197]	[0.014]	[1.037]
Cotinine(ng/ml)				0.093			
				[0.003]			
Education*post 1964						-0.216	
						[0.036]	
Observations	42238	15126	15160	859	859	42397	42238
R-squared	0.07	0.07	0.76	0.54	0.11	0.1	0.07

Panel B: Smoker	OLS	OLS	FE	First Stage	IV	First Stage	IV
Smoker	-187.105	-174.905	-58.171		-462.589		-311.829
	[7.052]	[13.944]	[32.936]		[60.473]		[231.128]
cotinine(ng/ml)				0.004			
				[0.000]			
Education*post 1964						-0.011	
						[0.002]	
Observations	46907	16483	16493	859	859	42397	42238
R-squared	0.08	0.08	0.81	0.38	0.04	0.06	0.06
Standard errors in brackets							
Sample	Full	Sibling	Sibling	Cotinine	Cotinine	Full	Full

Appendix Table 1 Impact of Smoking and Cotinine on Brithweight

	Birthweight				Cigarettes per day
	(1)	(2)	(3)	(4)	
Positive cotinine	-178.006				
	[53.507]				
Smoker		-211.383			
		[40.992]			
Ln(cotinine)			-97.373		
			[18.207]		
Ln(cigarettes per day)				-115.568	
				[28.115]	
cotinine(ng/ml)					0.093
					[0.007]
Cotinine*LBW					-0.006
					[0.016]
LBW					1.562
					[1.643]
Maternal education	12.398	7.955	1.513	5.404	-0.265
	[11.663]	[11.619]	[12.586]	[12.806]	[0.120]
Ln(family income in $1000)	-14.647	-13.661	-42.003	-41.594	0.135
	[55.616]	[55.372]	[68.498]	[71.546]	[0.576]
Married	126.66	70.784	113.25	88.639	-1.721
	[91.348]	[91.144]	[109.853]	[116.674]	[1.312]
White	-293.67	-179.417	-388.826	-389.064	3.634
	[115.081]	[118.151]	[266.624]	[259.377]	[0.972]
Black	-463.391	-376.189	-558.995	-621.378	-0.355
	[131.078]	[132.644]	[280.186]	[275.053]	[1.138]
Maternal age	-3.662	-5.11	-6.455	-8.147	-0.034
	[4.982]	[4.902]	[6.136]	[6.003]	[0.059]
Male	113.855	105.798	83.968	99.329	-0.035
	[39.336]	[38.722]	[49.099]	[49.698]	[0.520]
Birth Order	4.637	4.045	1.641	6.69	0.135
	[12.569]	[12.227]	[14.846]	[14.894]	[0.156]
Observations	970	968	526	524	968
R-squared	0.08	0.1	0.13	0.11	0.53
Robust standard errors in brackets					
Sample	Full	Full	Smokers	Smokers	Full

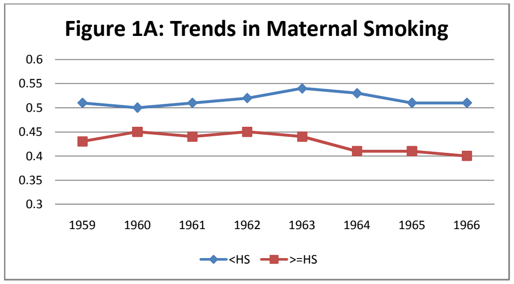

Figure 1A: Trends in Maternal Smoking

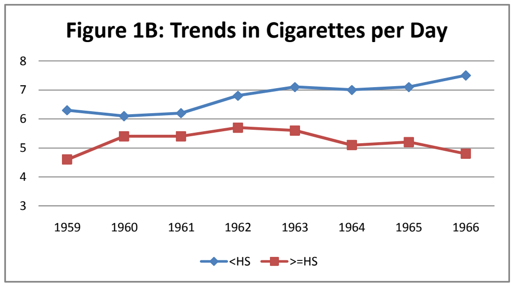

Figure 1B: Trends in Cigarettes per Day

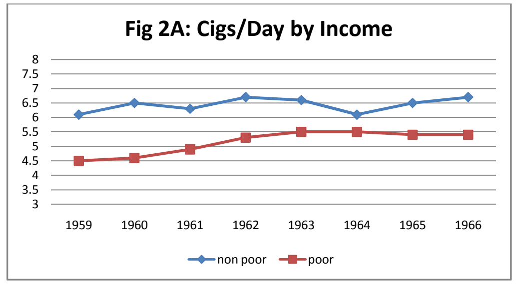

Fig 2A: Cigs/Day by Income

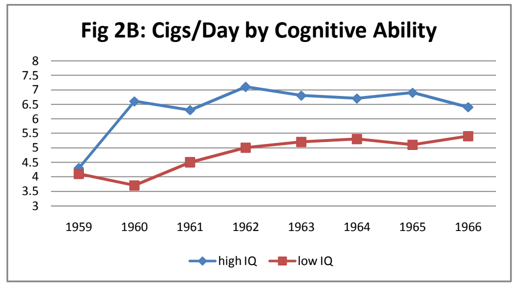

Fig 2B: Cigs/Day by Cognitive Ability

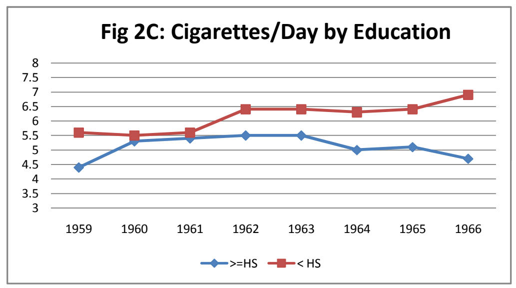

Fig 2C: Cigarettes/Day by Education

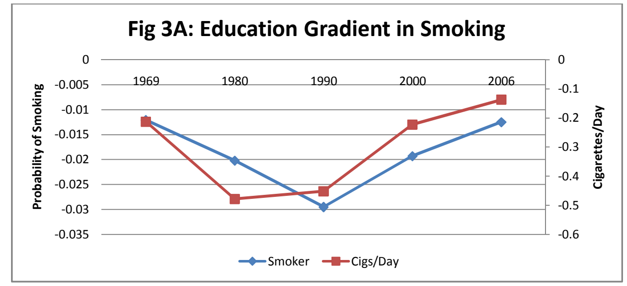

Fig 3A: Education Gradient in Smoking

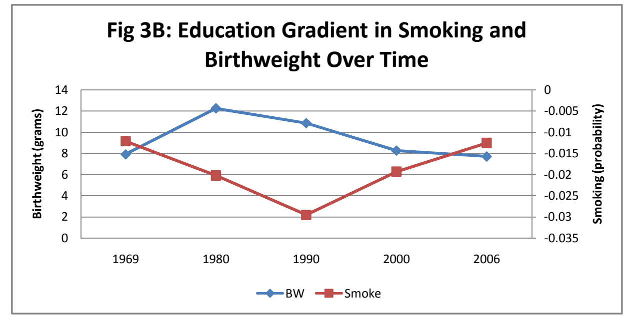

Fig 3B: Education Gradient in Smoking and Birthweight Over Time