Education, Knowledge and the Evolution of Disparities in Health

Anna Aizer[*]
Brown University and NBER

Laura Stroud
Brown Medical School

December 2009

We study how advances in scientific knowledge affect the evolution of disparities in health. Our focus is the 1964 Surgeon General Report on Smoking and Health – the first widely publicized report of the negative effects of smoking on health. Using an historical dataset that includes the smoking habits of pregnant women 1959-1966, we find that immediately after the 1964 Report, more educated mothers immediately reduced their smoking as measured by both self-reports and serum cotinine levels, while the less educated did not. We also find strong peer effects in the response to information: after the 1964 report, educated women surrounded by other educated women were more likely to reduce smoking relative to those surrounded by less educated women. The immediate increase in the education gradient in smoking resulted in a similar increase in the gradient in newborn health as measured by birth weight and fetal death. Over time, the education gradient in both smoking and newborn health continued to increase, peaking in the 1980s and then shrinking, eventually returning to initial levels. These results imply that future advances in medical knowledge may lead to a short run increase in health disparities that declines over the long run.

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 health disparities large, they are persistent, often originating in childhood or even earlier, in the newborn period (Case, Lubotsky and Paxson, 2002; Currie and Stabile, 2003). A number of theories have been put forth to explain the observed education disparities in health.[1] We focus on one: that education is related to one's ability to obtain, process and/or 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 knowledge diffuses. Thus advances in medical knowledge can lead to an initial increase in health inequality that declines over time as the knowledge diffuses.

To examine this empirically, we focus on how the first major advance in medical knowledge regarding the 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 an historical dataset of pregnant women 1959-1966 that, to our knowledge, is the only dataset containing prospective information on smoking habits prior to 1964, we estimate the immediate impact of the first massive dissemination of information on the health effects of smoking (the 1964 Surgeon General Report on Smoking and Health) on the decision to smoke

---

[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.

and newborn health. Previous work analyzing aggregate administrative data on cigarette sales found no decline associated with the 1964 Report (Sloan, Smith and Taylor, 2002). We too find that aggregate cigarette consumption remained constant after the 1964 Report, but that this masked important heterogeneity in the effects of the 1964 Report on smoking. We find that the most educated women reduced their smoking immediately after the 1964 Report in contrast to the least educated who appear to have actually increased their smoking. Thus while aggregate consumption of cigarettes remained constant over this period, the education gradient in smoking increased immediately and substantially. Moreover, the differential declines in smoking by education do not simply reflect differences in reporting, as serum cotinine levels follow the same pattern. Nor do they reflect differences in income or cognitive ability. Consistent with a negative relationship between smoking and newborn health, we find that the education gradient in newborn health, as measured by birth weight and fetal death, also increases immediately after publication of the report.[2]

We also explore the possibility that peers serve as a "social multiplier" - exacerbating the impact of information on the education gradient in smoking.   We hypothesize that the education level of peers affects one's own response to the 1964 Surgeon General's Report via two potential channels. First, more educated peers may serve as an additional, indirect source of information about the health effects of smoking (an information channel). Second, if one's ability to quit is a function, in part, of the smoking behavior of one's peers, as in the cue-theory of consumption, having more educated peers who are themselves less likely to smoke after publication of the report increases one's own probability of quitting (a behavioral channel). We find that the education gradient in smoking increases three times as quickly after the 1964 Surgeon General's

---

[2] There is some debate over whether the relationship between smoking and birth weight is causal which we discuss later (see, for example, Almond, Chay and Lee, 2005).

report among those most segregated by education. In other words, educated women surrounded by other educated women are more likely to reduce their smoking after the report relative to educated women surrounded by less educated women.

We follow the analysis of the short term effect of the 1964 Surgeon General Report on disparities in smoking and newborn health with an exploration of how these disparities evolve over time. 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 closely: increasing until the mid 1980s when they both peak, and then declining to 1960s levels by 2006, the latest year for which data are available. 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, to better understand the connection between the gradients in smoking and newborn health, we provide new estimates of the impact of smoking on newborn health using multiple identification strategies. In so doing, we make two contributions to the existing literature on smoking and birth weight. First, because we have siblings for a fixed effect 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, based on these results we can explain why existing estimates differ based on the estimation technique used. We conclude that the causal effect of smoking on birth weight is moderate in size (and non-linear) and lies between existing 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 and explore the role of peers in increasing the gradient; in section IV, we trace the evolution of the education gradient in smoking and birth weight over time to the year 2006. In section V we present the results of our analysis of the impact of prenatal smoking on newborn health, shedding light on the extent of bias in OLS estimates due to negative selection into smoking and measurement error separately. 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 and 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).

More recent work has sought to establish whether the relationship between education and health is causal. 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. Both studies find a larger effect of education on health in IV estimation relative to OLS, suggesting that any bias in OLS estimates is likely downward due, presumably, to measurement error.

There are multiple potential mechanisms behind the relationship between education and health.[3] We focus on one - that education is related to an individual's ability to learn and make decisions about his or her 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 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.[4] These studies establish a static response to information. They do not exmine how the response may evolve over time. Nor can they necessarily separate an individual's response

---

[3] 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.
[4] VBAC refers to "vaginal birth after cesarian section."

to information from either selective sorting to providers or providers' differential treatment of patients based on their educational status.

Another area where education is likely to matter 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).[5] A number of papers have sought to establish 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 some 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.[6]

A second strand of this literature seeks to link the education gradient in smoking with medical knowledge about the health effects of smoking. Based on retrospective smoking histories collected in 1978-2000, de Walque (2004) finds no discontinuous break in 1964 with publication of the first SG report on smoking and health. Meara (2001) and Kenkel (1991)

---

[5] Kimbro et al (2008) documents that they are also more likely to drink heavily and less likely to exercise.
[6] 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 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.

explore whether knowing about the health 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.

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 health 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. In support of this, 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 percent 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.

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

    A. Knowledge about Smoking and Health

Knowledge about the health effects of smoking was accumulating in the scientific literature throughout the 1930s, 1940s, and 1950's. However, it was not until the US Surgeon General issued its first Report on Smoking and Health in 1964, that the information was made widely accessible. The National Library of Medicine characterized the impact of the 1964 Report as follows:

> "[Surgeon General] Terry issued the commission's report on January 11, 1964, choosing a Saturday to minimize the effect on the stock market and to maximize coverage in the Sunday papers. As Terry remembered the event, two decades later, the report "hit the country like a bombshell. It was front page news and a lead story on every radio and television station in the United States and many abroad."[7]

After the first report, additional information "signals" about the effects of smoking on health followed, with the first being warning labels on all cigarette packages starting in 1966.[8] It should be noted that it was not until the SG Report of 1969 that the connection between smoking during pregnancy and birth outcomes was publicized.

    B. Data

---

[7] Profiles in Science- The National Library of Medicine "The Reports of the Surgeon General: The 1964 Report on Smoking and Health" accessed on November 17, 2009 at
http://profiles.nlm.nih.gov/NN/Views/Exhibit/narrative/smoking.html
[8] In 1971, cigarette ads were banned from TV; in 1979 the SG report concluded nicotine was addictive; in 1986 the SG issued a report on second hand smoke and in 1998 there was a settlement reached between 46 state attorney's office and tobacco manufacturers limiting marketing to youths. See Chaloupka and Warner (2000) for a full description.

For the analysis of the immediate impact of the 1964 Surgeon General's (SG) Report on Smoking and Health we use data from the National Collaborative Perinatal Project (NCPP), a prospective survey of 59,391 pregnant women who sought care in one of 12 urban Academic Medical Centers in the years 1959-1966. To our knowledge, this is the only data that contains information on smoking habits prior 1964 that was collected prospectively. The women were randomly recruited to participate in the study 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 either differential sorting to medical providers or providers treating women differently based on their education as all providers in the study were instructed to follow a standard protocol.

Descriptive statistics for the sample are presented in Appendix 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 from age. In column 1 are sample means for the analysis sample (n=50,142); 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

---

[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.

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.

C. Results

Publication of the First 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 these figures. 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 relatively small (about a 5 percentage point difference, or ten percent). Second, smoking is actually increasing over this period among the less educated, so much so that the decline in cigarettes smoked by HS graduates is compensated for by the increase in smoking among HS drop outs. Figure 1C shows that the total number of cigarettes smoked per day over this period remains constant, which is consistent with the constant trend in cigarettes sales over this period documented by Sloan, Smith and Taylor (2002).[10] 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. The fact that the decline is immediate but continues to increase after 1964 must be due to the fact that information diffuses gradually as there was no additional dissemination until warnings on cigarette packaging starting in 1966.

---

[10] Warner (1977) using data on cigarette sales predicts that in the absence of the 1964 Report, cigarette sales would have increased over this period. He interprets the fact that they remained constant as evidence of a negative effect of the 1964 Report on smoking.

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. Moreover, both income and cognitive ability are positively correlated with smoking, reflecting the fact that cigarette smoking in the 1950s and 1960s was more prevalent among those of higher socio-economic status during this time period, though that is no longer the case.[11]

Regression Results – OLS/ Regression Discontinuity Specification

To control for any changes in the composition of pregnant women before and after 1964 that could explain the trends in Figures 1A and 1B, 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 two specifications: a change in intercept (means) and a change in slope (rates). For the former, we estimate the following equation:

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

---

[11] 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.   The trends in smoking by education are not sensitive to the exclusion of the under 19.

In the above specification, Y is either smoking (smoker, cigarettes per day) or a measure of newborn health (birth weight or fetal death). 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) and a vector of personal characteristics including maternal race, age, family income, birth order, offspring gender and city of birth are also included. The main coefficient of interest is β_2 which indicates a change in the education gradient in smoking associated with the 1964 SG report (a change in intercept) that is independent of any underlying quadratic time trend in education and smoking.

The results suggest that the education gradient in smoking and birth weight increase substantially after publication of the First SG Report on Smoking and Health (Table 1). 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 percent 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 birth weight. Before publication of the report, an additional year of schooling is associated with heavier birth weight (23 gram difference). After the report, this increases by 25 percent to 29 grams, though the difference is not significant. Likewise, the gradient in LBW and fetal death increase after 1964, but not significantly. When the sample is

expanded to include the full sample, estimates of the impact of the 1964 Report on the gradient in birth weight and fetal death become significant (columns 6-8). Specifically, the gradient in birth weight increases 32 percent after the 1964 Report and the probability LBW by 25 percent. For fetal death, there is no significant education gradient prior to 1964. After 1964, an additional year of schooling is associated with a .2 percentage point reduction in fetal death (baseline rate of fetal death is 3 percentage points).

Though we include multiple maternal characteristics that might explain the relationship between maternal education and smoking or birth weight in the above analysis, in Table 2 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 Report in the first two columns of Table 2. 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 percentage points and 1 fewer cigarette per day. We observe a very similar increase with respect to education and birth weight (Panel E). 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 in contrast to 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 is not the case, we address this with regression analysis. Specifically, we ran regressions (not presented here) without controlling for family income at birth and found no change in the coefficient on education, suggesting that the relationship between maternal education and smoking is not mitigated through income. In column 3 of Table 2 we control for a measure of maternal cognitive ability and the coefficient on maternal education actually increases to -0.024, not surprisingly since cognitive ability and education are positively correlated and cognitive ability is positively correlated with smoking, at this time.[12] 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 2 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

---

[12] 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 does not appear to be driven by sample selection.

are highly correlated in these data ($\rho=0.72$). Regression results suggest that maternal education

is strongly and negatively associated with serum cotinine levels and that this relationship

increases 40 percent after 1964. After 1964, an additional 2.5 years of schooling is associated

with a 20 percent standard deviation reduction in serum cotinine levels. When we control for

maternal cognitive ability, the coefficient on maternal education increases by one third. The

stronger results for serum cotinine are consistent with smokers smoking fewer cigarettes but

smoking each more intensely after the 1964 report, as was found after tax hikes (Evans and

Farrelly, 1998). They are also consistent with less educated under-reporting after the 1964 SG

Report.

Next we estimate whether the 1964 report resulted in a change in the <u>rate</u> at which the

education gradient increased over this period. We refer to this as a change in slope model which

is specified as follows:

$$Y= \gamma_1 \text{education} + \gamma_2 \text{ education*year*post SG} + \gamma_3 \text{ education*year}^2\text{*post SG} +$$
$$\gamma_4 \text{ education*year } + \gamma_5 \text{ education*year}^2 + \gamma_6 \text{ year} + \gamma_7 \text{ year}^2 + \gamma_8 X + \varepsilon$$

Y can be an indicator variable for smoking, a count of the number of cigarettes smoked per day

or a serum cotinine level. The results (Table 3) suggest that the time trend in the education

gradient in smoking, defined each of the three ways, does increase significantly after the 1964

report. Prior to the report there is no significant time trend in the education gradient in smoking

(γ_4 small and insignificant in all specifications) but after the report, the gradient appears to

increase in a linear fashion (γ_2=-.004 for smoking, -.064 for cigarettes per day and -3.35 for

cotinine, all significant). Results for newborn health were not significant and not presented here.

Finally, we explore how the education gradient in smoking (defined as cigarettes per day) changes after 1964 using count models (a zero-inflated poisson due to over-dispersion in cigarettes per day). The predictions are nearly identical to those based on linear models.

Based on these results, we conclude that immediately after publication of the first widespread publication about the 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 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 birth weight and fetal death.

Regression Results – Maternal FE Specification

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. This would, for example, control for any differences in female education that may have co-incided with the 1964 Report. This would include passage of the 1964 Higher Education Act which increased financial aid for higher education and Title I (1965) which increased federal funding for schools serving low income children. 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 4A, 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 HS graduates relative to HS drop outs.

We also explore whether other maternal characteristics such as income, maternal health and health of previous children affect the decision to smoke after 1964. 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 low birth weight (LBW). Fixed effect regression estimates of the impact of smoking on birth weight, while suggestive of an increase in the gradient, were not significant and not presented here. We believe this is most likely due to the fact that when we reduce the sample to mothers with a birth right before and right after 1964 we significantly reduce identifying variation in smoking within mother by 80 percent.

---

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

Alternatively, for the sample of women with at least one birth before 1964 and at least one birth after 1964, one can characterize women who quit, women who start and women with no change after the 1964 Report (Table 4B). Women who quit are less likely to be black, more likely to be white, are more educated, and have higher income than those who start smoking over this period. They are very similar, however, to those whose smoking behavior does not change over this period. It is the women who start smoking over this period who are significantly more disadvantaged.

The Role of Peers

In this section we explore whether peers may act as a social multiplier in either the diffusion and/or uptake of new medical knowledge. Specifically, we hypothesize that having more educated peers increases the probability of quitting (or not starting) after publication of the 1964 SG report via two potential channels. First, more educated peers can serve as an additional indirect source of information about the health effects of smoking (an information channel). This assumes that the probability of knowing about the effects of smoking <u>directly</u> from 1964 SG Report increases in education but is less than 1 and that peers can also transmit this information <u>indirectly</u> to each other. As such, the probability that one knows about the effects of smoking on health is higher for the more educated not only because they are more likely to have direct knowledge from the SG Report, but also, if there is sorting on education, because their peers are more likely to have that knowledge and transmit it to them. We refer to this as the information channel. The second mechanism assumes that the cost of quitting, conditional on having information about the ill effects of smoking, is lower among those whose peers do not smoke.[14] As such, having more educated peers who are themselves less likely to smoke after publication

---

[14] This is consistent with the cue theory of consumption (Laibson, 2001).

of the report, increases one's own probability of quitting and/or not starting. We refer to this as the behavioral channel.

If peers do serve as a social multiplier in the diffusion and/or take-up of new medical knowledge, then the education gradient in smoking should increase more after the 1964 Report among those groups characterized by a high degree of residential segregation by education.[15] To examine this empirically, we calculate a dissimilarity index for each of the 11 cities included in the NCPP based on 1960 census data. The dissimilarity index measures the degree of residential segregation in the city and is defined as follows:

$$D = \frac{1}{2} \times \sum_{i=1}^{n} \left| \frac{dropout_i}{DROPOUT} - \frac{grad_i}{GRAD} \right|$$

Where $dropout_i$ refers to the number of HS drop outs in the census tract (i), $DROPOUT$ refers to the number of HS drop outs in the city, $grad_i$ the number of HS graduates in the census tract and $GRAD$ the number of HS graduates in the city. To identify education segregation separate from racial segregation, we calculate a separate index for each racial/ethnic group in each city (white, black, Hispanic). We then stratify the sample based on low vs. high dissimilarity index (cutoff at the median in the sample) and estimate separate regressions of the impact of the 1964 SG Report on the education gradient in smoking for those more vs. less segregated.

We specify both change in intercept (maternal education *post 1964) and change in slope (maternal education*post 1964*year) empirical models. We consider that segregation could be endogenous in this context because mothers who select to reside in more segregated cities may

---

[15] This assumes that the educated residing in segregated cities are more likely to interact with educated peers than those in less segregated cities. Likewise, the less educated residing in segregated cities are less likely to interact with educated peers than those in less segregated cities.

differ in unobserved ways from those choosing less segregated cities. We argue, however, that these regressions are identified for the following reasons. First, to address concerns that more segregated cities differ in important ways that might also affect smoking, we include city fixed effects since there is variation within city in the segregation index, (variation is by race within city). Second, we also allow the smoking behavior of the more (or less) educated to differ by level of segregation. As such, identification in this model comes from the change in smoking behavior by education by level of segregation over time. Any threat to identification would have to come from an omitted variable that varies with the level of education, the level of segregation and the timing of the 1964 Report.

We find that the education gradient in smoking increases much more quickly after the 1964 Surgeon General's report among those most segregated by education (Table 5). For example, those with four more years of schooling reduce their probability of smoking by nearly four percentage points after the 1964 SG Report in segregated areas, while mothers with the same level of education in less segregated areas reduce their smoking by less than one percentage point and the coefficient on the latter is not statistically significant. We take this as evidence that peers serve as a social multiplier in either the diffusion of new medical knowledge and/or the ability to act on such knowledge.

Summary and Interpretation of the Short Run Effects of Information on the Gradient

We conclude based on the results presented thus far that 1) the education gradient in smoking increased immediately after the 1964 SG report and that the gradient continued to increase further still in the two years after for which we have data, 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, 4) peers affect the response to new knowledge about the health effects of smoking, and 5) the education gradient in newborn health, as measured by birth weight and fetal death, likewise increased after the 1964 SG report.

How do we interpret these findings? Should we interpret them as causal? While we do not instrument for education, we could argue that IV estimates would likely be similar to OLS estimates for the following reasons. First, we are able to control for the two largest potential sources of omitted variable bias: income and cognitive ability. Second, existing IV estimates of the impact of education on health using various instruments are very similar, though slightly larger, than OLS estimates.

Alternatively, we argue that a causal interpretation of the relationship between education and the response to information is not crucial. Rather, the objective of this research is to explore how new medical knowledge affects disparities in health – over the short and long term. We focus on education because the more educated are better able to process new information. This may be because education improves this ability or because the more able pursue more education. From the perspective of this research, the distinction is not crucial.

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 rate of 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 birth weight 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. However, we conclude that the omission of single mothers and non-citizens from the NNS data is unlikely to affect our results. [16]

B. Results

---

[16] 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 birth weight 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.

Based on regressions of smoking on maternal education and additional controls (age, race, marital status, birth order and child gender) for the period 1969 - 2006, we find that the education gradient increases until about 1990 when it begins to decline and continuities to do so until 2006, the most recent year for which we have data. Specifically, 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 fewer cigarettes per day), the most recent year for which we have data (columns 2,4,6,8,10 of Table 6). The education gradient in any smoking is similar (columns 1,3,5,7,9 of Table 6). In the second panel we present the same results redefining education as an indicator variable for being in the top 25 percent 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 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.

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 include 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 relationship between education and smoking while pregnant is not operating through income, at least for this period.

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 7 where we present the results of regressions of birth weight 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 7).[17]

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 birth weight and smoking regressions presented in Tables 6 and 7. 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 birth weight. Likewise, as the gradient in smoking declines between 1990 and 2006, so too does the gradient in birth weight. As evident in Figure 3, trends in the relationship between education and smoking and education and birth weight are nearly mirror images of one another.

---

[17] 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.

Yet many other things that likely affect the education gradient in birth weight (such as expansions in Medicaid eligibility for pregnancy) 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 birth weight, 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 birth weight would remain unchanged. In fact, when we control for smoking the education gradient in birth weight is not constant but is much flatter. Controlling for smoking, the education gradient in birth weight 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 birth weight can be explained, in part, by changes in the education gradient in smoking, particular over the period 1969-1990, but less so for the period after 1990.

Thus we conclude 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, peaked around 1990 and then ultimately fell back to levels observed in the 1960s.

For the above analysis, we implicitly assume a negative causal impact of smoking on birth weight, as generally 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 birth weight using multiple identification strategies, we can 1) quantify the bias in OLS estimates due to selection and measurement error, separately, 2) attempt to explain and reconcile the variation in existing estimates, and 3) provide a new estimate of the impact of smoking on birth weight that corrects for bias from both selection and measurement error.

V. Prenatal Smoking and Newborn Health

For this analysis we focus on birth weight as a measure of newborn health, not fetal death which is too infrequent (0.03) to generate stable estimates, particularly with the reduced sample for the fixed effect estimates. However, we acknowledge debate over the appropriateness of birth weight as a measure of newborn health and the presence of evidence both for and against it.[18]

Previous attempts to estimate a causal relationship between smoking and birth weight 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.[19] Studies using fixed effects (Rosenzweig and Wolpin, 1991; 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

---

[18] For example, Almond, Chay and Lee (2005), using variation within twins for identification, find very small negative effects of birth weight on short term health outcomes, with the exception of very low birth weights. Black, Devereaux, and Salvanes (2007) and Oreopolis et al (2008) using the same identification strategy do find moderate long term effects in terms of IQ, income, education and welfare receipt.

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

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.[20]

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.

A. OLS Estimates

OLS estimates suggest that smoking is associated with a decline in birth weight of 187 grams, or 28 percent of a standard deviation (Table 8). Defining smoking as cigarettes smoked per day, we find that each cigarette smoked per day is associated with a reduction in birth weight of 20 grams, and the relationship is decreasing in the number of cigarettes smoked (Column 1,

---

[20] 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 birth weight by 430 grams.

Panel B). In panel 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 percentage point increase in the probability of a LBW birth (column 4) which represents a 50 percent increase and that the relationship is again concave in cigarettes smoked per day: light smokers face a 4 percentage point increase in the probability of a LBW birth, and moderate and heavy smokers face between a 7.4 and 8.8 percentage point increase.

We also explore whether smoking is correlated with prematurity and find that it is, though the relationship is weaker than with birth weight. Smokers deliver on average 2 days earlier (Table 8 column 6) and like birth weight, 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 percentage point increase in the probability of delivering prematurely (the average probability of prematurity is 16 percent in this sample), representing a moderate impact.

B. Fixed Effect 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

(Appendix 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 birth weight based on the sibling subsample are presented in column 2 of Table 8.

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. It should be noted, however, that the

variation in smoking declines considerably when we limit our analysis to that within mothers.

For example, the standard deviation in the number of cigarettes smoked per day is 9.5, overall

(the mean number of cigarettes smoked per day is 6.5). Within mothers, there is still some

variation, but considerably less with the standard deviation falling to 3.0.

When we include maternal fixed effects, the coefficient on smoking declines by two

thirds from -175 to -58 (Table 8 column 3), suggesting that once we control for omitted

variables, smoking reduces birth weight by less than 60 grams or 9 percent of a standard

deviation. Smoking an additional cigarette per day reduces birth weight by 12 grams, relative to

the OLS estimate of 20 grams, with the effect decreasing in the number of cigarettes smoked

(Table 8 column 3, 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 birth weight 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 birth weight effects in Table 9. First, we estimate the extent

to which the reduction in birth weight is attributable to less mature birth. For this, we present

results of a fixed effect regression of the impact of smoking on birth weight conditional on

gestation, defined as as birth weight(in grams)/ gestation (in weeks). Conditional on gestation,

smoking is still associated with a decline in birth weight, though the decline is half the

unconditional size (Table 9, column 1).[21] 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. Sources of Selection into Smoking

We attempt to identify the possible omitted variables associated with maternal smoking

and birth weight 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 birth weight and

maternal characteristics include maternal health, maternal size (height and weight), weight gain

during pregnancy and maternal cognitive ability. However, when we include these variables

---

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

(Table 8, columns 5-9) the estimated OLS relationship between smoking and birth weight 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 from the third trimester of 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 10 columns 1 and 2 based on the full and sibling subsample suggest that each additional cigarette smoked per day reduces birth weight by 11.7 grams.[22] The fixed effect estimate in column 3 is much smaller, -3.7. Columns 4 and 5 contain the first and second stages,

---

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

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 2 we provide estimates of the direct (or reduced form) impact of cotinine on birth weight. The estimated impact of cotinine on birth weight 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 upward by 22 percent. Moreover, we can calculate a causal estimate that accounts for both selection and 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, representing a moderate effect.

As a final identification strategy, we instrument for maternal smoking using the publication of the 1964 report interacted with maternal education as an instrument.[23] 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 10, respectively. The first stage is strong (t statistics of 5 and 6). As expected, the point estimate for smoking in 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 and the 95 percent confidence interval includes a wide range.

We draw four main conclusions based on these analyses. First, OLS estimates of the impact of smoking on birth weight are biased upward due to negative selection into smoking and biased downward due to classical measurement error in smoking reports; second, fixed effect estimates which correct for selection bias are significantly attenuated by measurement error; 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 birth weight corrected for selection and measurement error represents a moderate effect: -152 grams for smoking and -9.6 grams per cigarette/day.

Together 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 existing fixed effect estimates, we conclude that they are too small because of exacerbated

---

[23] 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 birth weight directly).

measurement error. This also explains why estimates based on propensity score and selection correction, which are subject to less measurement error (though still some), are larger than the FE estimates and why IV estimates which correct for selection and measurement error are larger still.

VI. Conclusions

The goals of this paper were threefold. First, using the only dataset that, to our knowledge, has information on smoking habits before and after the first wide-spread publication of the effects of smoking on health in 1964, we explored the immediate impact of the 1964 Report on the education gradients in smoking and health. We found that more educated women immediately reduced their smoking in response to the report while the least educated did not. 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. Moreover, we provide evidence of strong peer effects in the response to information. The immediate increase in the gradient in smoking after 1964 was accompanied by an immediate increase in the gradient in newborn health as measured by increases in birth weight and reductions in fetal death.

Second, we examined the long-run effects of an advance in information on the gradient in smoking and health. We showed 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. This 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, 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. 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 advancements in scientific knowledge.

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.

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

Black, S., Devereaux, P and K. Salvanes. 2007. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. Quarterly Journal of Economics.122(1): 409-439.

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.

Chaloupka, Frank and Kenneth Warner. 2000. "The Economics of Smoking" in Handbook of Health Economics, Joseph Newhouse and Anthony Cuyler, eds. Amsterdam: North Holland Press.

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, Farrelly MC (1998) "The Compensating Behavior of Smokers: Taxes, Tar and Nicotine" Rand Journal of Economics. 29(3): 578-595.

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, V. 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.

Goldman, Dana and James Smith (2002) "Can Patient Self-Management Help Explain the SES Health Gradient? Procedings of the National Academy of Sciences. 99(16):10929-10934.

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.

Oreopoulos, Philip, Stabile, Mark, Roos, L., and Walld, R., 2008. "The Short, Medium, and Long Term Effects of Poor Infant Health," Journal of Human Resources, 43(1): 88-138.

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 birth weight. 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.

Sloan, FA, Smith VK and DH Taylor. 2002. "Information, Addiction and Bad Choices: Lessons from a Century of Cigarettes" Economic Letters. 77(2): 147-155.

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.)

Warner, Kenneth (1977) "The Effects of the Anti-Smoking Campaign on Cigarette Consumption" American Journal of Public Health. 67(7): 645-650.

Table 1: Changes in the Education Gradient in Smoking and Birthweight Associated with the 1964 SG Report

	Sample Restricted to 19-34 Year Olds					Full Sample		
	Smoker	Cigarettes per day	Birthweight	LBW	Fetal Death	Birthweight	LBW	Fetal Death
Maternal education*post 1964	-0.006	-0.094	6.01	-0.002	-0.001	7.997	-0.003	-0.002
	[0.003]	[0.054]	[4.054]	[0.002]	[0.001]	[3.496]	[0.002]	[0.001]
Maternal education	-0.013	-0.305	23.205	-0.008	0.001	24.681	-0.012	0.001
	[0.006]	[0.078]	[5.934]	[0.003]	[0.001]	[4.949]	[0.003]	[0.001]
Maternal education*year	-0.001	-0.018	-4.193	0.001	-0.001	-4.037	0.002	-0.001
	[0.002]	[0.029]	[2.574]	[0.001]	[0.001]	[2.069]	[0.001]	[0.001]
Maternal education*year^2	0	0.001	0.255	0	0	0.122	0	0
	[0.000]	[0.003]	[0.297]	[0.000]	[0.000]	[0.236]	[0.000]	[0.000]
Black	0.101	0.27	-125.867	0.076	0.032	-124.143	0.07	0.025
	[0.038]	[0.589]	[49.510]	[0.021]	[0.004]	[45.344]	[0.020]	[0.005]
White	0.243	5.075	81.665	0.024	0.018	83.693	0.015	0.015
	[0.037]	[0.563]	[47.872]	[0.019]	[0.002]	[44.130]	[0.019]	[0.004]
Hispanic	-0.082	-1.792	-30.568	0.049	0.035	-29.52	0.035	0.031
	[0.042]	[0.645]	[55.226]	[0.026]	[0.006]	[50.065]	[0.024]	[0.006]
Birth Order	0.015	0.391	15.928	0	-0.001	16.695	-0.002	-0.001
	[0.002]	[0.035]	[2.532]	[0.001]	[0.001]	[2.186]	[0.001]	[0.001]
Married	-0.131	-2.611	32.608	-0.015	-0.004	27.807	-0.01	-0.003
	[0.009]	[0.175]	[11.333]	[0.006]	[0.002]	[9.273]	[0.005]	[0.002]
Ln(family income in $1000)	0.02	0.27	25.177	-0.004	0	20.904	-0.004	-0.001
	[0.007]	[0.126]	[8.399]	[0.005]	[0.002]	[6.809]	[0.004]	[0.002]
Male	0.005	0.098	111.052	-0.025	0.002	110.454	-0.026	0.003
	[0.006]	[0.107]	[7.460]	[0.004]	[0.002]	[6.471]	[0.004]	[0.002]
Maternal age	-0.007	-0.038	0.987	0.001	0.002	1.611	0.001	0.002
	[0.001]	[0.018]	[1.181]	[0.001]	[0.000]	[0.805]	[0.000]	[0.000]
Year	7.363	94.205	-13.59	3.505	0.153	-17.36	3.956	1.745
	[6.879]	[83.732]	[9.243]	[4.948]	[2.278]	[6.834]	[2.840]	[2.031]
Year squared	-0.002	-0.024	3.47	-0.001	0	4.431	-0.001	0
	[0.002]	[0.021]	[2.356]	[0.001]	[0.001]	[1.735]	[0.001]	[0.001]
Post 1964	0.049	0.688	-102.28	0.036	0.01	-122.471	0.044	0.016
	[0.034]	[0.592]	[43.918]	[0.025]	[0.012]	[36.574]	[0.020]	[0.010]
Observations	40668	40668	40601	40863	40863	53120	53459	53459
R-squared	0.06	0.1	0.05	0.01	0	0.06	0.02	0

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

Table 2: Changes in The Gradient in Smoking and Birthweight - Stratified Specifications

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]
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 3: Changes in the Education Gradient in Smoking - Spline Specifications

	Smoker	Cigarettes per day	cotinine(ng/ml)
Maternal education*year*post 1964	-0.004	-0.064	-3.353
	[0.002]	[0.030]	[1.755]
Maternal education*year^2*post 1964	0	0.008	0.318
	[0.000]	[0.004]	[0.231]
Maternal education*year	-0.007	-0.013	-1.548
	[0.007]	[0.082]	[5.804]
Maternal education*year^2	0.001	-0.001	0.345
	[0.001]	[0.010]	[0.693]
Maternal education	0	-0.234	-5.425
	[0.015]	[0.162]	[12.383]
Year	0.104	0.862	16.457
	[0.065]	[0.848]	[68.253]
Year^2	-0.012	-0.063	-4.032
	[0.007]	[0.103]	[7.959]
Post 1964	0.086	0.746	107.171
	[0.047]	[0.855]	[54.370]
Observations	44695	44695	923
R-squared	0.05	0.09	0.06

Robust standard errors in brackets

Table 4A: Changes in Smoking Behavior Across Birth Stratified by Maternal Characteristic - FE Specification

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 4B: Characteristics of Quitters and Starters

Maternal Characteristic	Quitter	Starter	No Change in Smoking
White	0.46	0.6	0.46
Black	0.51	0.35	0.5
Maternal Education	10.7	10.2	10.6
Married	0.82	0.68	0.8
Family income (in $1000)	30.7	23.99	27.6
HS Drop-Out	0.56	0.68	0.58
Obs	653	653	8369

Sample consists of 9675 births to 4035 mothers with at least one birth before and one birth after 1964

Table 5: Changes in the Education Gradient in Smoking and Residential Segregation

Change in Intercept Specification	Low	High	Drop Hispanic	
			Low	High
Maternal Education*post 1964	-0.001	-0.009	-0.001	-0.009
	[0.003]	[0.004]	[0.003]	[0.005]
Maternal education	-0.024	-0.007	-0.024	-0.012
	[0.002]	[0.004]	[0.002]	[0.005]
Observations	24522	21296	24522	18626
R-squared	0.06	0.04	0.06	0.04

Change in Slope Specification (Spline)	Low	High	Drop Hispanic	
			Low	High
Maternal education*year*post 1964	0	-0.003	0	-0.002
	[0.001]	[0.001]	[0.001]	[0.001]
Maternal education*year	-0.001	-0.002	-0.001	-0.003
	[0.001]	[0.002]	[0.001]	[0.002]
Observations	24522	21296	24522	18626
R-squared	0.06	0.04	0.06	0.04

Robust standard errors in brackets

Also included as covariates: maternal indicators for race, marital status, maternal age,
 family income, birth order, child gender, city FE and a linear time trend

Note: Hispanics have very high segregation values and are dropped from the analysis sample in columns 3 and 4 to segregation"
test whether the results for "high" segregation are driven by their inclusion. They are not .

Table 6: The Education Gradient in Smoking Over Time

Panel A: Education in years	1969		1980		1990		2000		2006	
	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

Panel B: HS Graduate	1969		1980		1990		2000		2006	
	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

Panel C: Top 25% Education Distribution	1969		1980		1990		2000		2006	
	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 7: The Education Gradient in Birthweight Over Time

	1969		1980		1990		2000		2006	
Panel A: Education in years	(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 8: Impact of Smoking on Birthweight and Gestation

Panel A: Smoker	Birthweight (grams)			LBW		Gestation (weeks)		Premature	
	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	Birthweight (grams)			LBW		Gestation (weeks)		Premature	
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	Birthweight (grams)			LBW		Gestation (weeks)		Premature	
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

	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 deviation	-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 dev	-8%	-7%	-4%	-21%	-23%	-23%	-22%	-23%	-23%

Table 9: Impact of Smoking on Birthweight - Extensions

Table 10: 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: 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

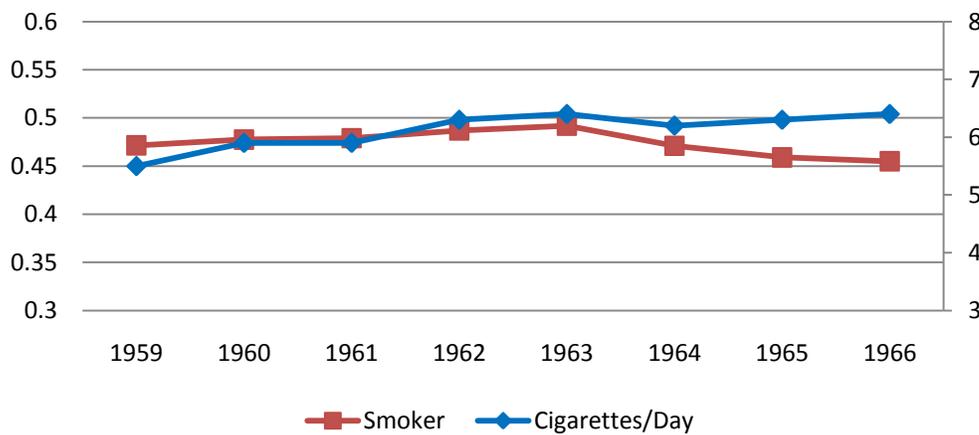|	Birthweight				Cigarettes per day
	(1)	(2)	(3)	(4)	(5)
Positive cotinine	-178.006				
	[53.507]				
Smoker		-211.383			
		[40.992]			
Ln(cotinine)			-97.373		
			[18.207]		
Ln(cigarettes per day)				-115.568	
				[28.115]	
cotinine(ng/ml)					0.093
					[0.007]
Cotinine*LBW					-0.006
					[0.016]
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 by Education
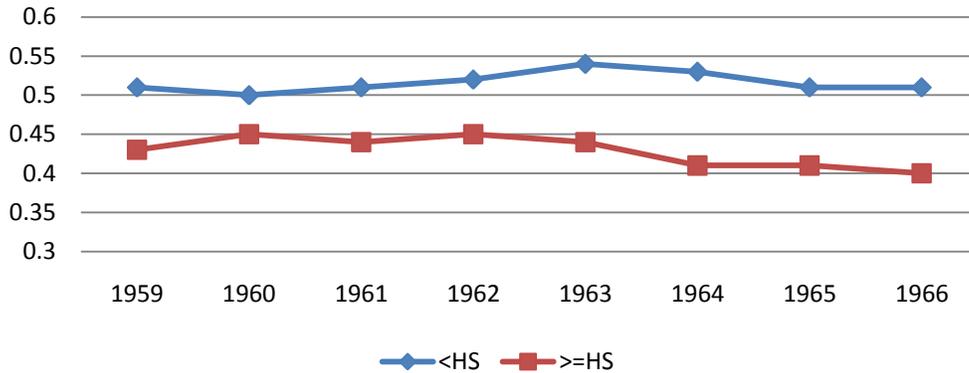
Figure 1B: Trends in Cigarettes per Day by Education
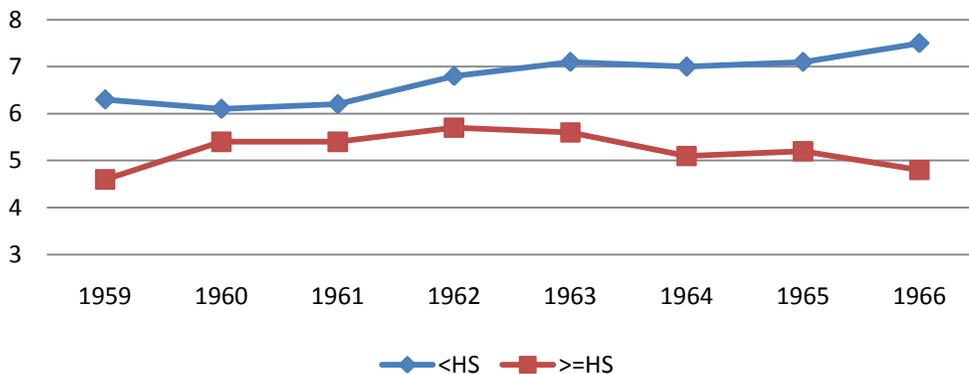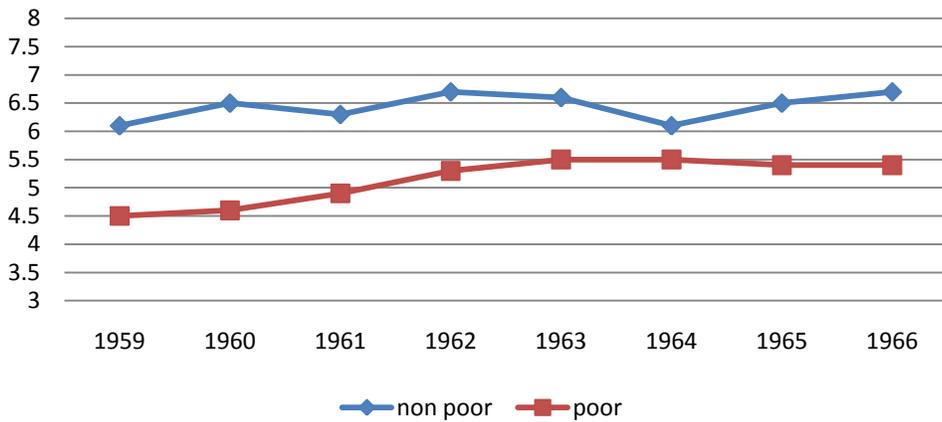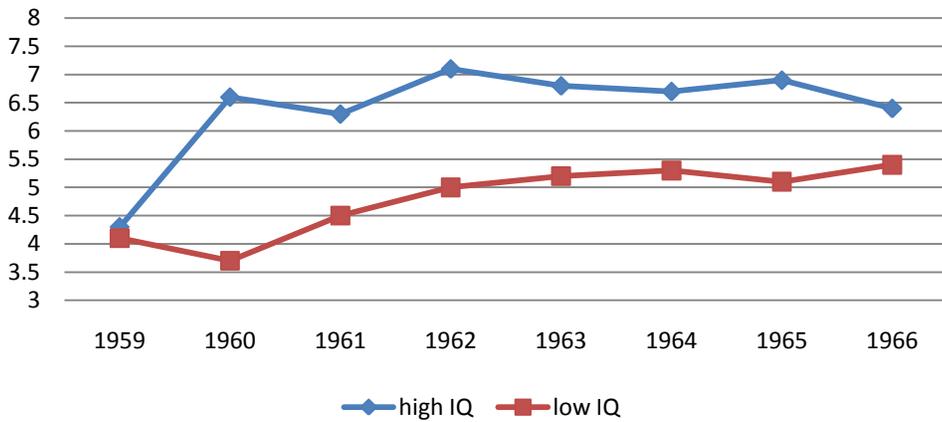

Figure 1C: Trends in Maternal Smoking

Sample: women ages 19-34

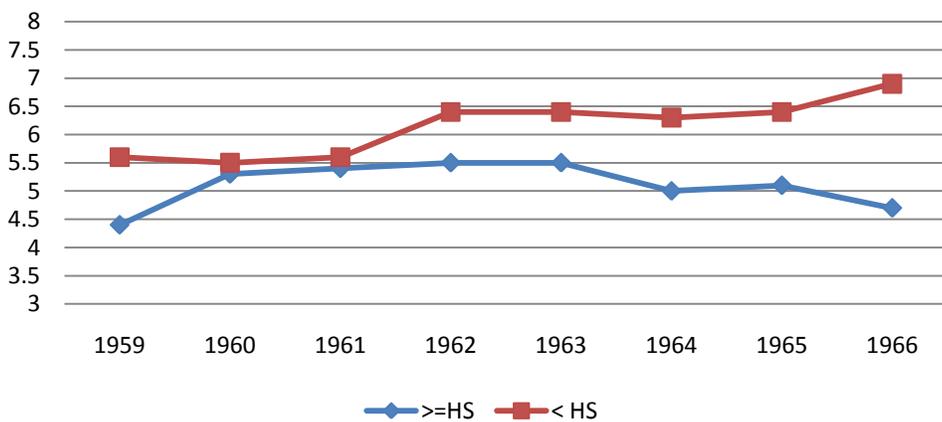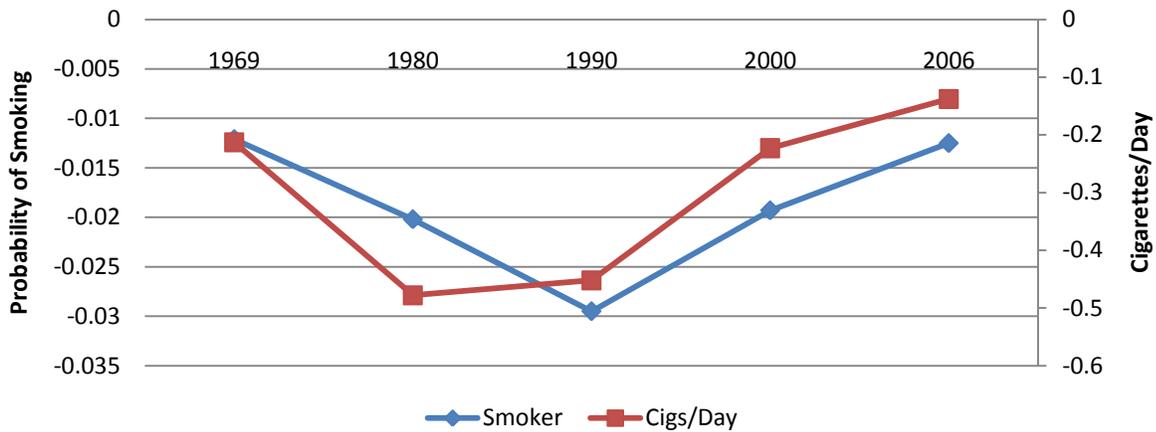Fig 2A: Cigs/Day by Income

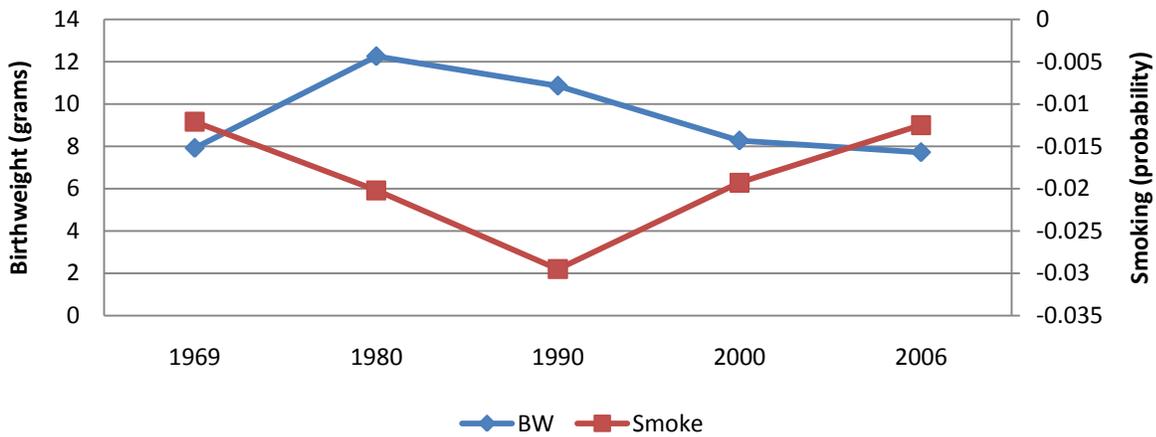Fig 2B: Cigs/Day by Cognitive Ability

Fig 2C: Cigarettes/Day by Education

Sample: all women

Fig 3A: Education Gradient in Smoking Over Time

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

Note: Each point represents a coefficient estimate on education in a regression in which the outcome is either an indicator for smoking, the number of cigarettes smoked per day or birth weight. Other covariates listed in Table 7 also included.