

Pollution, Ability, and Gender-Specific Investment

Responses to Shocks

Teresa Molina*

June 12, 2018

Abstract

This paper documents that gender-specific labor market conditions and occupation-specific wage parameters can explain why men and women respond differently to early-life shocks. I show that exposure to pollution in the second trimester of gestation leads to significantly lower cognitive ability in adulthood for both genders, but lower high school completion and income for women only. Two labor market features explain this gender difference: (1) women sort into white-collar occupations at higher rates, and (2) schooling and ability are more complementary in white-collar than blue-collar occupations, which I verify by structurally estimating a dynamic model of education and occupational choice.

Keywords: *early life, pollution, gender, occupational choice, education, Mexico*

JEL Codes: *J24, I26, Q53*

*University of Hawaii at Manoa Department of Economics: Saunders Hall 515A, 2424 Maile Way, Honolulu, HI 96822; email: tmolina@hawaii.edu; phone: 808-956-7080. This paper would not exist without the guidance, support, and patience of my advisor, John Strauss, and Achyuta Adhvaryu and Anant Nyshadham. I am also grateful to Dan Bennett, Tushar Bharati, Joe Cummins, Yu-Wei Hsieh, Katie Lim, Matt Kahn, Gaurav Khanna, Mallory Montgomery, Jeff Nugent, Geert Ridder, Neeraj Sood, Jorge Tamayo, and Athleen Venkataramani for helpful conversations and feedback on previous drafts. Thanks to seminar participants at USC, SoCCAM, the WEAI Annual Meeting and Graduate Student Workshop, CSU Long Beach, UH Manoa, World Bank, Erasmus Rotterdam, Zurich, UNSW, Oregon, Melbourne, Monash, PacDev, Hawaii International Conference, PAA, and SEA. I gratefully acknowledge funding from the USC Provost's Ph.D. Fellowship, the USC Dornsife INET graduate student fellowship, the Oakley Endowed Fellowship, and the Gold Family Fellowship. Special thanks to Graciela Teruel and the MxFLS Support Team for providing me with restricted-use data.

1 Introduction

There is substantial evidence documenting that shocks experienced early in life can impact a wide range of outcomes in adulthood,¹ by directly affecting physical and cognitive ability (Almond and Currie, 2011), and by influencing the human capital investment decisions made throughout an individual's life (Almond and Mazumder, 2013). However, much less is known about what drives these investment responses and why they might differ across contexts or individuals. In particular, men and women often respond differently to early life shocks, and several studies have hypothesized that this stems from the different labor market conditions that men and women face (Bhalotra and Venkataramani, 2013; Cutler et al., 2010; Hoddinott et al., 2008). In this paper, I aim to test this hypothesis, for which very little empirical evidence currently exists.

This study investigates how gender-specific occupational choices and occupation-specific wage parameters affect the way that individuals respond to early life shocks, using *in utero* exposure to air pollution as the health shock of interest. Many studies have shown that exposure to air pollution *in utero* can negatively affect birth outcomes (like birth weight or infant mortality),² but this paper contributes to a smaller body of research that looks at the longer-term impact of such exposure (Sanders, 2012; Isen et al., 2017; Bharadwaj et al., 2017; Peet, 2016; Rosales-Rueda and Triyana, 2018).³

Using data from Mexico, I estimate the long-term effects of *in utero* pollution exposure on adult cognitive ability, educational attainment, and income. I use thermal inversions, a meteorological phenomenon that negatively impacts air quality, as an exogenous source of variation in pollution levels. I find that men and women exposed to more thermal inversions (and thus worse pollution) during their second trimester *in utero* score significantly lower on Raven's tests

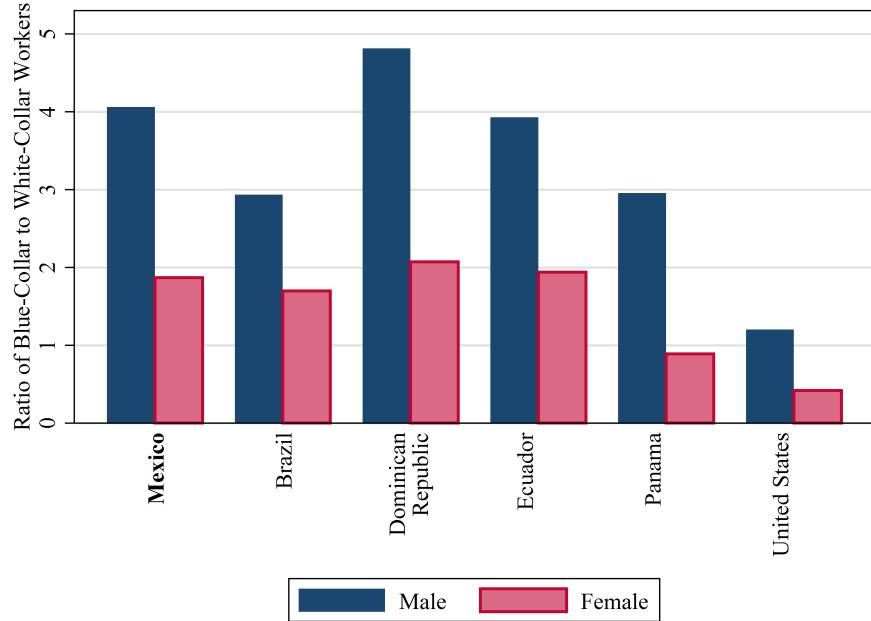
¹In the medical and epidemiological literature, research on the hypothesis (popularized by Barker (1990)) that fetal conditions can have long-term consequences dates back to discoveries of links between birth defects and *in utero* conditions (Jones et al., 1973; McBride, 1961; Von Lenz and Knapp, 1962). Around the same time, a series of studies on the Dutch famine in 1944 linked exposure to the famine with negative outcomes in adulthood (Stein et al., 1975). Almond and Currie (2011) provide a more detailed history of the literature and the economic evidence that followed. See Heckman (2006) and Currie and Vogl (2013) for other reviews of the evidence.

²See Chay and Greenstone (2003), Currie and Neidell (2005), Jayachandran (2009), and others summarized in Currie et al. (2014).

³The bulk of existing research on long-term effects of pollution has focused on exposure to radiation from nuclear accidents (Almond et al., 2009; Black et al., 2014), a much more extreme case of air pollution than what we might be interested in for policy reasons.

of fluid intelligence as young adults. For women only, however, this cognitive shock also leads to lower high school completion and income.

Figure 1. Cross-Country Blue-Collar to White-Collar Ratios, by Gender



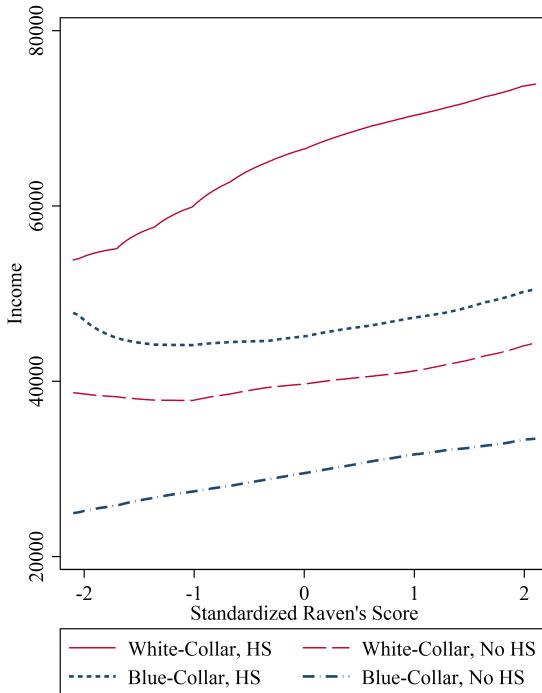
Ratios calculated using weighted counts of adults aged 30 to 50 in the 2010 censuses of the listed countries. Blue-collar and white-collar jobs are identified using the ISCO occupation codes, which are defined as blue-collar or white-collar using the classifications in Vogl (2014), summarized in Table A1.

Larger female schooling responses to early-life health shocks have been commonly documented in existing literature (Bobonis et al., 2006; Maluccio et al., 2009; Field et al., 2009; Maccini and Yang, 2009; Anderson, 2008; Hoynes et al., 2016), but the reasons for these gender differences are less well-explored. The conceptual framework I develop in this paper highlights one potential explanation for the gender difference I find here: in Mexico and across other comparable countries, women are more likely to take up white-collar jobs (as shown in Figure 1). If schooling and ability are more complementary in white-collar jobs than blue-collar jobs, this would imply that individuals expecting to go into a white-collar job (who are more likely to be female) should exhibit larger schooling responses to a cognitive shock.

My empirical results provide support for this hypothesis: once I allow for the effect of pollution to vary across areas with different gender-specific occupational shares, the gender difference in

the schooling response disappears. This result implies that men and women exhibit different schooling responses primarily because of the different occupational choices they make. I am able to rule out other common explanations for gender differences in the effects of shocks, including son preference or gender-specific norms regarding high school completion.

Figure 2. Income-Ability Relationship



Notes: These local linear regressions use individuals aged 30 to 50 in the MxFLS. The blue-collar and white-collar occupation categories are defined using the classifications in Vogl (2014), summarized in Table A1. Income is total earned annual income measured in 2002 Mexican pesos and winsorized at the 99th percentile.

This evidence, however, is just one piece of the puzzle. The second piece involves pinning down whether schooling and ability are more complementary in white-collar compared to blue-collar occupations, which remains an open question. Using data from the Mexican Family Life Survey (MxFLS), Figure 2 illustrates how the income difference between high school graduates and non-graduates varies with cognitive ability for individuals in white-collar jobs and in blue-collar jobs. I plot the relationship between annual income and Raven's test scores, separately for four different schooling-occupation combinations. Interestingly, the income boost enjoyed by high school graduates is *increasing* in ability in white-collar occupations, but *decreasing* in ability in blue-collar occupations. This can be seen in the widening of the gap between the two white-collar

lines and the narrowing of the gap between the two blue-collar lines. Although these figures do not take into account selection into occupation types or schooling, they offer suggestive evidence that the complementarity between schooling and ability may be lower in blue-collar than in white-collar jobs. Taken together, the two features of the labor market illustrated in Figures 1 and 2 have important implications for how men and women respond to their cognitive endowments. A boy and girl of the same age growing up in the same village could respond differently to an early-life cognitive shock because of the different occupations they expect to enter.

Although the occupation-specific complementarity parameters discussed in the previous paragraph are crucial in this context, they are ignored in standard wage function estimations in the existing literature. The major challenge to estimating these parameters is the endogeneity of both schooling and occupational choice, which is ignored by the descriptive exercise in Figure 2. In order to deal with this, I estimate a dynamic discrete choice model, where individuals choose their schooling in the first period and their occupation in the second period, and with the novel feature that the wages determining these decisions exhibit occupation-specific non-separabilities between schooling and ability. My structural estimates reveal that schooling and ability are complements in white-collar but not in blue-collar occupations, which is consistent with my reduced form finding that schooling responses to pollution are smaller for those likely to go into blue-collar jobs.

As mentioned above, it is common for authors to use gender-specific labor market conditions to explain gender differentials in the estimated effects of early-life shocks (Bhalotra and Venkataramani, 2013; Cutler et al., 2010; Hoddinott et al., 2008), but because very few studies directly test for this link, the empirical evidence provided here contributes to this early-life literature. My paper joins two recent exceptions (Pitt et al., 2012; Rosenzweig and Zhang, 2013), which formally model the idea that the occupational sorting tendencies of men and women might explain differential schooling responses across genders. The authors find evidence of gender-specific schooling responses to a physical health shock, similar to what I find in the context of a cognitive ability shock.

Unlike these two papers, I allow for non-separabilities between schooling and cognitive ability

that vary across occupations,⁴ and show that these complementarities help explain why men and women may adjust their schooling differently in response to cognitive shocks. To my knowledge, I am the first to estimate these occupation-specific cross-partial and show that they, together with gender differences in occupational choice, are the primary reason for the gender-specific schooling responses that I find.⁵

The finding that the level of complementarity between schooling and ability differs across occupations offers an important contribution to the literature on human capital formation and the skill production function. In this literature, estimating the complementarities between child endowments and investments is seen as crucial, as these can determine whether educational investments will help remediate or simply exacerbate early-life inequalities across the ability distribution (Cunha and Heckman, 2007; Cunha et al., 2010; Aizer and Cunha, 2012). The estimation of occupation-specific cross-partial in this paper highlights that the effectiveness of schooling in remediating for cognitive ability disadvantage relies heavily on the type of job individuals eventually take up. We can think of individuals “producing” various types of skills, broadly defined – skills that are valued in white-collar jobs and skills that are valued in blue-collar jobs. As a result, although occupational choice is often ignored in estimates of the skill production function, job opportunities and occupational choice matter a great deal for determining when remediation is and is not possible.

One interesting implication of this paper is that interventions aimed at improving cognitive ability could have the added benefit of closing the gender gap in schooling and related outcomes, in contexts where women are more likely to take up white-collar jobs. Relatedly, these results contribute to the literature on gender gaps in labor market outcomes by providing another potentially important explanation for the large relative gains in female education and employment outcomes that have been observed across the globe over the past few decades (Goldin, 1995; Mammen and Paxson, 2000; Rendall, 2010; Olivetti, 2014; Bhalotra et al., 2015; Pitt et al., 2012;

⁴The model underlying both of these papers allows for differential returns to schooling across occupations but does not allow for complementarities between schooling and cognitive ability in these occupation-specific production functions.

⁵The seminal Keane and Wolpin (1997) study and the many that have followed allow the returns to schooling to vary by occupation and sometimes include IQ as an additional covariate, but none allow for occupation-specific cross-partial between schooling and ability.

Rosenzweig and Zhang, 2013).⁶ Because female schooling responds more strongly to cognitive ability shocks than male schooling does, a sustained improvement in population intelligence levels, which has been observed in many contexts, repeatedly documented, and dubbed the “Flynn effect” (Trahan et al., 2014; Flynn, 1984), could have given rise to these larger improvements for women.

In the next section, I outline a conceptual framework that illustrates how local labor market conditions can influence the way schooling decisions respond to an early-life shock. The remainder of the paper is devoted to investigating the implications of this model, using pollution exposure as a shock to cognitive ability. First, I estimate the effects of pollution exposure on the schooling decisions of men and women. Second, I study how gender-specific labor market parameters are driving the gender differences that I find. Finally, I structurally estimate the parameters of white-collar and blue-collar wage functions to determine whether these parameters are consistent with the conclusions of my reduced-form analysis.

2 Model

The purpose of the following model is to identify a few ways in which gender-specific labor market expectations might result in gender differences in the schooling responses to a cognitive endowment shock.

Suppose individuals are born with an ability endowment θ . As adults, individuals can work in one of two occupations: white-collar ($k = w$) or blue-collar ($k = b$). Each job type has a different wage function, where educational attainment E and ability θ are rewarded differently. These functions capture the idea that worker characteristics command different prices in different occupations (Heckman and Scheinkman, 1987) and that schooling and ability may exhibit non-separabilities that vary across sectors, which is reflected in the descriptive evidence in Figure 2.

⁶Previous literature has emphasized other important explanations for this phenomenon. For instance, many of the aforementioned studies argue that economic growth can generate improved economic outcomes for women precisely because it spurs the rise of a less physical sector, while Pitt et al. (2012) and Rosenzweig and Zhang (2013) argue that improvements in physical health and nutrition have played a role in these large relative improvements for women.

I denote the occupation-specific expected wage functions as

$$W_k(E, \theta; \beta_k), \quad (1)$$

where β_k are the parameters that map ability and schooling to wages (i.e., the returns to each of these inputs and the cross-partial between the two). $\frac{\partial W_k}{\partial E} > 0$, $\frac{\partial W_k}{\partial \theta} > 0$, $\frac{\partial^2 W_k}{\partial E \partial E} < 0$, and $\frac{\partial^2 W_k}{\partial \theta \partial \theta} < 0$ for $k = w, b$, but I do not make any assumptions about the signs or relative magnitudes of the occupation-specific cross-partials between schooling and ability.

Individuals can also remain out of the labor market, denoted by $k = n$. The reward function in the home sector takes a similar form ($W_n(E, \theta; \beta_n)$), and should be interpreted as the utility individuals would enjoy if they did not work (which could be a function of their marriage decisions, bargaining power in marriage, productivity in home production, etc.).

The opportunity cost of schooling takes the following form:

$$c(E, \theta; \alpha),$$

where $\frac{\partial c}{\partial E} > 0$ and $\frac{\partial^2 c}{\partial E \partial E} > 0$.

Individuals pick the optimal level of education E to maximize their expected future rewards, net of the cost of schooling, as in the maximization problem below. By choosing to model only the schooling decision, which is the focus of this paper, I assume that any major investments parents might make to change θ take place before the crucial schooling decisions are made. This assumption is consistent with the well-documented finding that there are higher returns to investing in a child's skill formation early in life (before primary school) compared to later on (Cunha et al., 2010; Heckman, 2006). Moreover, for children in Mexico, the end of primary school marks the first critical schooling transition period when many drop out (Behrman et al., 2011).

The maximization problem can be written

$$\max_E p_{jg}(E, \theta)q_{jg}(E, \theta)W_w(E, \theta; \beta_w) + (1 - p_{jg}(E, \theta))q_{jg}(E, \theta)W_b(E, \theta; \beta_b)$$

$$+(1 - q_{jg}(E, \theta))W_n(E, \theta; \beta_n) - c(E, \theta; \alpha). \quad (2)$$

$q_{jg}(E, \theta)$ represents the expected probability of an individual entering the labor force. $p_{jg}(E, \theta)$ represents the expected probability of an individual going into a white-collar job, conditional on being in the labor market. The location j subscript captures the idea that these expectations vary across labor markets (over space as well as over time). The g subscript allows for different expectations for each gender.

In order to focus my predictions on a few key features of the model, I simplify the analysis by assuming that p_{jg} and q_{jg} each consist of a gender- and location-specific constant that does not depend on schooling or ability, and a separate term (which does not vary over j or g) that governs how these expectations depend on schooling and ability. That is, $p_{jg} = \bar{p}_{jg} + p(E, \theta)$ and $q_{jg} = \bar{q}_{jg} + q(E, \theta)$.⁷ In the structural model I later use to estimate the wage parameters, however, I explicitly model individuals' labor market decisions and allow these expectations to be endogenously determined.

Using the implicit function theorem, I can show how optimal schooling will respond to a positive shock to θ :

$$\frac{dE^*}{d\theta} = -\frac{A}{B},$$

where the denominator, B , is negative by assumption, and A , which determines the sign of this schooling response, is given by:

$$\begin{aligned} A &= \frac{\partial p}{\partial E} \frac{\partial q}{\partial \theta} (W_w - W_b) + \mathbf{q}_{jg} \frac{\partial p}{\partial E} \left(\frac{\partial W_w}{\partial \theta} - \frac{\partial W_b}{\partial \theta} \right) + \frac{\partial q}{\partial E} \left(\mathbf{p}_{jg} \frac{\partial W_w}{\partial \theta} + (1 - \mathbf{p}_{jg}) \frac{\partial W_b}{\partial \theta} - \frac{\partial W_n}{\partial \theta} \right) \\ &+ \frac{\partial q}{\partial E} \frac{\partial p}{\partial \theta} (W_w - W_b) + \mathbf{q}_{jg} \frac{\partial p}{\partial \theta} \left(\frac{\partial W_w}{\partial E} - \frac{\partial W_b}{\partial E} \right) + \frac{\partial q}{\partial \theta} \left(\mathbf{p}_{jg} \frac{\partial W_w}{\partial E} + (1 - \mathbf{p}_{jg}) \frac{\partial W_b}{\partial E} - \frac{\partial W_n}{\partial E} \right) \\ &+ \mathbf{p}_{jg} \mathbf{q}_{jg} \frac{\partial^2 W_w}{\partial E \partial \theta} + (1 - \mathbf{p}_{jg}) \mathbf{q}_{jg} \frac{\partial^2 W_b}{\partial E \partial \theta} + (1 - \mathbf{q}_{jg}) \frac{\partial^2 W_n}{\partial E \partial \theta} - \frac{\partial^2 c}{\partial E \partial \theta} \end{aligned} \quad (3)$$

⁷ Additionally, I assume that $\frac{\partial^2 p}{\partial E \partial \theta} = \frac{\partial^2 q}{\partial E \partial \theta} = 0$, which means that ability affects labor market expectations in the same way for all education levels.

The boldface text in equation 3 highlights a few potential sources of gender differences in the optimal schooling response to a cognitive shock ($\frac{dE^*}{d\theta}$). First, if men and women have different expectations about what type of job they will end up in or whether they will enter the labor force at all (if $p_{jf} \neq p_{jm}$ or $q_{jf} \neq q_{jm}$), this means that each gender places different weights on the occupation-specific parameters in equation 3, like the returns to schooling or ability and the cross-partial between the two. If these parameters differ substantially across occupations, this will result in men and women responding differently to a cognitive shock.

To highlight one source of gender differentials that my subsequent results reveal to be particularly important, I shut down other channels by assuming that schooling and ability have negligible effects on expectations and that men and women have similar expectations about entering the labor market.⁸ This leaves me with the following equation:

$$A' = \mathbf{p}_{\mathbf{jg}} q_j \frac{\partial^2 W_w}{\partial E \partial \theta} + (1 - \mathbf{p}_{\mathbf{jg}}) q_j \frac{\partial^2 W_b}{\partial E \partial \theta} + (1 - q_j) \frac{\partial^2 W_n}{\partial E \partial \theta} - \frac{\partial^2 c}{\partial E \partial \theta} \quad (4)$$

In equation 4, it is clear that different expectations about future occupational choice will result in different schooling responses. In particular, if schooling and ability are more complementary in white-collar than blue-collar jobs ($\frac{\partial^2 W_w}{\partial E \partial \theta} > \frac{\partial^2 W_b}{\partial E \partial \theta}$), then individuals more likely to get a white-collar job (higher p_{jg}) will increase their schooling more in response to a positive ability shock (higher $\frac{dE^*}{d\theta}$). The intuition is simple: higher p_{jg} places more weight on the cross-partial between schooling and ability in the white-collar wage function ($\frac{\partial^2 W_w}{\partial E \partial \theta}$) and less weight on the cross-partial in the blue-collar wage function ($\frac{\partial^2 W_b}{\partial E \partial \theta}$). As Figure 1 and Table A1 clearly show, working women are on average more likely to take up white-collar jobs than working men, implying that p_{jf} is likely higher than p_{jm} . Given this, women should exhibit larger schooling responses than men if there is a higher degree of complementarity between schooling and ability in white-collar jobs. In

⁸Though men have higher labor force participation than women in my sample population, my sample period coincides with a rapid increase in Mexican female labor force participation (Bhalotra et al., 2015), during which it is likely that a young girl's expected probability of labor force participation exceeded average female labor force participation in the labor market at the time. Though data on these subjective expectations is limited, one study shows that teenage boys and girls in urban Mexico expressed identical expected probabilities of going into the labor force as adults (Attanasio and Kaufmann, 2014).

short, these occupation-specific cross-partialls could be an important source of gender differences that has not been investigated before.

Although this discussion was based on a number of simplifying assumptions, it is important to note that relaxing them would not change the prediction that higher expectations about going into a white-collar job (higher p_{jg}) should lead to larger schooling responses, if schooling and ability are more complementary in white-collar jobs. Without these assumptions, what does become more ambiguous is which gender should exhibit a stronger schooling response to a cognitive shock overall, due to potentially different labor market expectations (q_{jg}) placing different weights on other occupation-specific parameters (the relative returns to ability and the relative returns to education, for example.⁹) Indeed, whether expectations about occupational choice and labor force participation lead to gender differences in the optimal schooling response, and whether these expectations drive women or men to respond more strongly, are empirical questions that I seek to answer in the remainder of this paper.

I investigate the implications of this model in three steps. First, I estimate the effects of a cognitive shock on the schooling decisions of men and women. Second, I study the extent to which gender-specific labor market parameters (related to occupation choice and labor force participation) are driving the gender differences that I find. Finally, I structurally estimate the parameters of each occupation-specific wage function to determine whether these parameters are consistent with the conclusions of my reduced-form analysis. The framework described in this section provides the starting point for the structural model I use to estimate these parameters, which relaxes many of the assumptions made here by fully endogenizing the occupational choice.

3 Background and Data

In this section, I outline the biological reasons for considering pollution exposure as a shock to cognitive ability and describe the data I use to estimate these effects.

⁹It is worth noting, however, that using my estimates of these parameters in section 6.5, lower labor force expectations for women compared to men would still lead to the prediction that $\frac{dE^*}{d\theta}$ should be higher for women than men.

3.1 Pollution

Substantial medical and epidemiological evidence demonstrates that *in utero* exposure to pollution can be harmful to the fetus (Lacasaña et al., 2005; Peterson et al., 2015; Le et al., 2012; Saenen et al., 2015; Backes et al., 2013). Concrete evidence that pins down the biological mechanisms is more limited, but there are a few commonly cited suspected pathways that primarily relate to two types of pollutants: carbon monoxide (CO) and particulate matter (PM-10 or PM-2.5).

As described in detail in Appendix section A.1, both of these pollutants can disrupt the transport of blood, glucose, or oxygen to the fetus, which could in theory have negative impacts on both the physical and cognitive aspects of fetal development. Whether pollution exposure results in primarily physical or cognitive damage likely depends on the timing of exposure (Dobbing and Sands, 1973). For instance, because most neurogenesis takes place in the second trimester of pregnancy, this trimester is seen as a “critical period for the formation of cortical neurons” (Morgan and Gibson, 1991, p.10). In line with this, medical and economic studies on exposure to radiation flag the second trimester as the most sensitive period for brain development (Otake, 1998; Almond et al., 2009; Black et al., 2014).¹⁰ Although day-to-day air pollution and radiation are very different types of pollution, these radiation studies offer some generalizable lessons about the critical periods in brain development: second trimester exposure is also particularly detrimental for other external stressors, like influenza (Schwandt, 2016) and nutritional deficiencies (Morgan and Gibson, 1991).¹¹

As outlined in section 2, individuals make different schooling decisions and earn different wages partially because of heterogeneous levels of ability. Any effect that *in utero* exposure to pollution has on schooling and labor market outcomes is likely working through its biological

¹⁰Otake (1998) document that weeks 8 to 25 (late first and almost entire second trimester) are particularly crucial for brain development. Black et al. (2014) also find that the 3rd, 4th, and 5th months of pregnancy were the critical periods during which exposure to nuclear fallout resulted in lower IQ as adults.

¹¹The critical period highlighted by these studies coincides with crucial processes in the development of the fetal brain. The migration of neurons, from their place of origin to their final location in the brain, peaks in the second trimester and is largely complete by the beginning of the third trimester. Similarly, synaptic connections in the cortex are refined and become more permanent starting in the second trimester; this process peaks by the beginning of the third trimester (Tau and Peterson, 2010).

effect on this unobserved endowment, of which cognitive functioning is an important component.

The ideal data set for an analysis of the long-run effects of *in utero* exposure would consist of pollution data going back to the *in utero* months of my sample individuals, who are adults in the 2002 through 2009 waves of the MxFLS. Currently, pollution measurements for CO, O₃, SO₂, NO₂, PM10, and most recently, PM2.5 are publicly available on Mexico's National Institute of Ecology (INECC) website for a total of 16 cities. However, the majority of this spatially limited data does not go back far enough to study at-birth exposure of adults in the MxFLS. The earliest pollution measurements date back to 1986, but for only CO in Mexico City, for which there are large sections of missing data until about 1993.

The lack of high quality historical data going back far enough to link adults with their *in utero* exposure is a major obstacle to identifying the effects of pollution exposure at birth on later life outcomes, in this context as well as more generally. In order to circumvent this issue, I rely on thermal inversions, a meteorological phenomenon known to worsen air quality, as an exogenous source of variation in pollution levels for which there is data for all of Mexico dating back to 1979.

3.2 Thermal Inversions

Air temperature typically falls with altitude, but when a thermal inversion occurs, this relationship reverses, which results in a warm layer of air sitting above cooler air, trapping pollutants released near the surface. That thermal inversions can negatively impact air quality is well-documented, both in the atmospheric sciences literature (Jacobson, 2002) as well as more recently in the economics literature (Jans et al., 2014; Arceo et al., 2016).¹²

In general, inversions are the result of the combination of various atmospheric forces and geographic conditions. I argue that after controlling for all of the relevant main effects (fixed geographic characteristics, time of year, temperature, humidity, cloud coverage, etc.), the occurrence of a thermal inversion is exogenous: essentially the random interaction of all of the

¹²See Appendix section A.2 and Jacobson (2002) for a more detailed discussion of the different types and causes of inversions.

necessary conditions. For example, out of two separate nights in the same geographic location, during the same time of year, and with same temperature, only one might result in a thermal inversion because of the interaction between the temperature during the day time, humidity, and wind speed of that particular night. Like Jans et al. (2014) and Arceo et al. (2016), I assume that thermal inversions can only affect my outcomes of interest through their effect on pollution levels, once I have controlled for all of the weather controls, geographic fixed effects, and non-linear time trends that I include in my regressions.

I identify thermal inversions in Mexico using the North American Regional Reanalysis (NARR) data, which provides air temperatures just above the surface and at various pressure levels above sea-level on a 0.3 x 0.3 degree grid (roughly 30km by 30 km) across the North American continent.¹³ Using atmospheric modeling techniques, the NARR combines temperature, wind, moisture, and precipitation data from a number of different sources, including weather balloons, commercial aircraft recordings, ground-based rainfall measurements, and satellite data.¹⁴ The resulting data set records, every three hours for each grid point, a wide array of meteorological variables at the surface, a few meters above the surface, and at 29 pressure levels (extending vertically into the atmosphere), from 1000 hPa (roughly equivalent to sea level) to 100 hPa (about 16,000 meters above sea-level).

To identify thermal inversions, I take the air temperature 2 meters above the surface¹⁵ and subtract this from the air temperature recorded at the pressure level 25 hPa lower (roughly 300 meters higher) than the surface pressure at a given location.¹⁶ I identify an inversion episode as any time this difference is greater than zero, consistent with the American Meteorological Society's definition of an inversion (Glickman and Zenk, 2000) and how existing literature identifies them (Beard et al., 2012; Devasthale et al., 2010). I use the 25 hPa increment because this is

¹³NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>.

¹⁴See Mesinger et al. (2006) for more detail about the various data sources and model. See Appendix section A.2 for a discussion of validation checks.

¹⁵2-meter temperature is what is reported by meteorologists in weather reports and is distinct from "skin" surface temperature, which the NARR also records.

¹⁶Because of varying surface altitude across Mexico, I do not take temperature from the same pressure level for all points. For example, for a municipality at sea level, I use the temperature at 975 hPa, whereas for a higher-altitude location in Mexico City, I use the 700 pressure level because the surface pressure is 725.

the smallest increment between pressure levels available in the NARR data. Looking further above the surface (50 hPa or 75 hPa, for example) does not detect many additional inversions and therefore, unsurprisingly, leaves my results virtually unchanged. In general, I am most interested in the inversions close to the surface as they are likely to have the largest effects on air quality.

Like Jans et al. (2014), I focus on nighttime inversions. There is greater variation in the occurrence of nighttime (compared to daytime) inversions over time and across space, which makes nighttime inversions much stronger predictors of pollution in my first-stage checks. Moreover, nighttime inversions are much less visible than daytime inversions and are therefore less likely to generate behavioral responses.¹⁷

In addition to using the NARR to identify thermal inversions, I also utilize this data set's relative humidity, wind speed, and total cloud coverage variables as important controls in all specifications. Although precipitation is also available in the NARR data set, I use ground measurements recorded by Mexico's National Meteorological Service (CONAGUA) to control for rainfall because these should be measured with less error.

As mentioned above, Mexican pollution measures do not date back far enough to enable me to use thermal inversions as an instrument for *in utero* exposure to pollution, as Arceo et al. (2016) do in their study of the contemporaneous effects of pollution. However, using the pollution measures that do exist, I check whether thermal inversions drive pollution levels in the years and cities for which I have pollution data. To establish a link between thermal inversions (I_{jym}) and pollution levels (P_{jym}) in a given municipality j , during the three-month period starting from month m in year y , I run the following regression:

$$P_{jym} = \alpha_1 I_{jym} + \alpha_2' W_{jym} + \mu_j + \delta_y + \alpha_m + v_{jym}. \quad (5)$$

I aggregate to three-month periods here because I eventually analyze the effects of pollution by

¹⁷Daytime inversions are not always visible but are more likely to be seen in warm and humid climates like Mexico's.

trimester. P_{jym} represents CO (8-hour daily maximum) or PM-10 (24-hour mean) averaged over the three month period starting in month m of year y . I_{jym} represents the total number of days (per month) with a nighttime inversion in that same three-month period. I include municipality (μ_j), year (δ_y), and month (α_m) fixed effects. W_{jym} is a vector of flexible weather controls (also averaged across the three month period): linear, quadratic, and cubic terms of minimum, maximum, and mean 2-meter temperature, rainfall, relative humidity, wind speed, and total cloud coverage. In this regression, these weather controls are important because they influence the likelihood of a thermal inversion but also have the potential to directly affect pollution levels. In the later analysis, their inclusion is crucial to ensure that thermal inversions are affecting my outcomes of interest only via the pollution channel, and not through these weather variables.¹⁸ I aggregate to the three-month level because my main analysis studies the effects of pollution by trimester.

Table 1 reports the results of this regression, using data from 1994, when more complete data was being recorded, to 2009, the last year of available pollution data. Even after controlling for a complete set of fixed effects and weather controls, inversions are positively and significantly related to both CO and PM-10 levels. The F-statistics in this “quasi-first-stage” exceed conventional thresholds for strong instruments.

3.3 Mexican Family Life Survey

All outcome variables come from the Mexican Family Life Survey (MxFLS), a nationally representative longitudinal household survey that began in 2002 and conducted follow-ups in 2005 and 2009. In addition to collecting standard demographic, schooling, and employment information, this survey also measured several physical biomarkers (like height) and administered Raven’s tests of fluid intelligence. I use these measures of cognitive ability and height, along with educational attainment and earned annual income as my main outcomes of interest. I include

¹⁸It is worth noting that this is only necessary because I am specifically interested in estimating the long-run effects of pollution exposure. The weather fluctuations I am controlling for could also be thought of as exogenous, and could in theory be considered a shock to the early-life endowment, but given our limited knowledge of the long-run effects of pollution exposure (especially in developing countries), I try to isolate the pollution shock from other weather shocks.

Table 1. Relationship between Thermal Inversions and Pollution, 3-Month Periods

	CO	PM-10
Average Monthly Inversions During 3-Month Period	0.0140*** (0.00254)	0.474*** (0.0828)
N	23821	21292
Mean of DV	2.306	55.17
Fstat	30.449	32.835

Notes:

* p< 0.1 ** p< 0.05 *** p< 0.01

Standard errors (clustered at municipality level) in parentheses.

CO and PM-10 represent three-month averages of the 8-hour daily maximum (in ppm) and 24-hour daily mean (in $\mu\text{g}/\text{m}^3$), respectively. All regressions control for month, year, and municipality fixed effects, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, average monthly relative humidity, average monthly precipitation, and average monthly cloud coverage during each relevant 3-month period.

individuals found in any wave of the survey in order to obtain as large of a sample as possible. For all outcomes except for Raven's scores, I take the outcome from the most recent wave in which the individual was interviewed. For Raven's tests, I use each individual's first test score in order to minimize the effect that test-taking experience (either from the survey or elsewhere) may have on their scores.¹⁹

Another key variable obtained from the MxFLS is municipality of birth, a restricted-use variable that enables me to link adults and adolescents (including those who have migrated) with thermal inversion exposure specific to their birthplace at their time of birth. More details about the construction of individual-level variables are provided in Appendix section A.3.

Table 2 reports summary statistics, by gender, for the outcomes and main regressors for all individuals with non-missing thermal inversion data (implying a non-missing birth month, birth municipality, and birth year after 1978) and who were at least 15 years of age in the last MxFLS wave in which they appeared. These are the individuals old enough to have been included in the migration module of the survey, which obtains information about place of birth. I report raw means for Raven's test scores and height in this table but use standardized variables in

¹⁹It should be noted that Raven's tests were identical across waves.

Table 2. Summary Statistics for Reduced Form Sample

Variable Name	Female			Male		
	Mean	S.D.	N	Mean	S.D.	N
Outcome Variables						
Raven's test score (% correct)	0.55	0.229	5455	0.56	0.226	4865
Height (cm)	155.20	7.722	5506	166.06	10.27	4892
Years of schooling	9.52	3.075	5634	9.20	3.074	5081
Annual income	23473.04	22428.0	998	29745.26	72818.3	2157
Control Variables						
Mother's Education	6.00	3.853	5204	6.36	3.804	4566
Father's Education	6.34	4.279	4832	6.69	4.226	4258
Age for Raven's Test variable	17.24	3.319	5455	17.12	3.398	4865
Age for height variable	20.38	4.491	5506	19.86	4.507	4892
Age for schooling variables	20.45	4.424	5634	19.94	4.463	5081
Age for income variable	22.61	3.656	998	22.37	3.715	2157
Full Sample						
Dependent Variables		Mean	SD	10th pctile	Median	90th pctile
Average monthly inversions during trimester 1		18.09	8.206	5.93	19.23	28
Average monthly inversions during trimester 2		17.93	8.235	5.69	18.94	28
Average monthly inversions during trimester 3		17.80	8.288	5.54	18.94	28
						N
						10848

Notes: Sample includes individuals with non-missing thermal inversion data who were at least 15 years of age in the last MxFLS wave in which they appeared.

the regressions.²⁰ The sample size for annual income is much smaller compared to the other variables, primarily because I restrict to those who report work as their primary activity in the week prior to the survey.²¹ I do this in order to exclude those still in school but working part-time, whose income is likely a poor representation of their labor market productivity or lifetime earning potential. Note that this restriction does not apply to the rest of the outcome variables. On average, individuals in this sample are exposed to approximately 18 inversion nights per month during any given trimester.

²⁰I standardize test scores using the full sample mean and standard deviation. For height, I use WHO standards for everyone under 20 and for the remainder of the sample simply standardize using the gender-specific mean and standard deviation of the sample population 20 and older. I identify and drop gross outliers.

²¹They make up about 40% of this relatively young sample. About 1,000 more are dropped due to missing income data.

3.4 Mexican Labor Market Data

In order to investigate the interaction between labor market conditions and pollution exposure, I use occupation information (specifically, white-collar shares) from the 1990, 2000, and 2010 Mexican censuses (Minnesota Population Center, 2015). Following Atkin (2016), I collapse to the commuting zone level and link this labor market information to individuals using their commuting zone of residence during their school-aged years. Commuting zones, which I discuss in more detail in Appendix section A.4, are municipalities or groups of municipalities that better represent local labor markets: for instance, large metropolitan areas or neighboring municipalities.

In some specifications, I include the gender-specific share of white-collar workers directly calculated from the census, where I categorize white-collar and blue-collar workers using the occupation categorizations in Vogl (2014), described in Table A1. I assign these values to individuals based on the census conducted closest to the year in which they turned 12. In other specifications, I predict values for years in between censuses and assign individuals to the predicted values from the exact year they turned 12. To calculate these predicted values, I use a shift-share strategy, similar to Bartik (1991) and others, which involves predicting economic variables for geographic regions (like states or, in this case, commuting zones) by combining national industry-level growth rates and baseline industry compositions for these regions of interest (see section A.4 for more details). I calculate national industry-level growth rates from Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH), a nationally representative household survey that was first conducted by Mexico's National Institute of Statistics and Geography (INEGI) in 1982, and every two years since 1992. I cannot obtain municipality-level data directly from this data set because it is only representative at the national level (and at the state level for a limited number of states and years).

4 Empirical Strategy

To estimate the effects of pollution, I regress my outcomes of interest on thermal inversion counts over several three month periods prior to and after a child's birth. In addition to helping to

overcome the pollution data limitations described above, using thermal inversions also addresses the endogeneity of pollution. Pollution is not randomly assigned: individuals born in highly polluted areas are different from those born in less polluted areas. While location fixed effects can be used to alleviate these residential sorting concerns, they do not control for location-specific trends in pollution that may coincide with trends in the outcomes of interest.

In this framework, thermal inversions can be thought of as an “instrument” that generates exogenous variation in an endogenous variable that I do not observe. This endogenous variable is not a particular pollutant but rather, air quality in general. The approach of this paper is not designed to estimate the dose-response function of specific pollutants: rather, it offers a well-identified way to learn whether being exposed to higher pollution while *in utero* has discernible effects in the long term.

For individual i , born in municipality j , in year y and month m , whose outcome Y_{ijymw} comes from survey wave w , I estimate the following specification:

$$Y_{ijymw} = \alpha_0 + \sum_{k=-7}^4 \beta_k I_{jym}^{3k} + \sum_{k=-7}^4 \alpha'_k W_{jym}^{3k} + \gamma' X_i + \mu_j + (\delta_y \times \nu_w) + \eta_m + \epsilon_{ijymw}. \quad (6)$$

I_{jym}^a represents the average number of monthly thermal inversions that took place in individual i 's municipality of birth during the three month period starting a months after the individual's birth month (where negative values indicate months before birth). I include all three month periods starting a year before conception (21 months before birth) until a year after birth in order to identify critical periods and ensure that any effects I find in the *in utero* period are not being driven by serial correlation in the thermal inversion variable year to year. Omitting the thermal inversion variables from before and after pregnancy could result in their effects loading onto the trimester coefficients. The coefficients on inversions prior to conception also serve as a falsification check, as pollution exposure before a child is conceived should not have direct effects on that child's outcomes.

W_{jym}^a is a vector of weather controls (minimum, maximum, and mean temperatures, rain,

total cloud coverage, relative humidity, and wind speed), averaged over each three-month period, along with their squares and cubes. In this specification, municipality fixed effects (μ_j) address cross-sectional pollution endogeneity concerns, including residential sorting issues, by ensuring that identification comes from within-municipality variation over time. Year (δ_y) and month fixed effects (η_m) control flexibly for long-term and seasonal time trends. The interaction of year and wave dummies ($\delta_y \times \nu_w$) capture both wave and age effects. Controls X_i include gender, mother's education, and father's education, for which I set missing values to zero and include dummies for missing values. In more rigorous specifications, I add various combinations of location fixed effects and location-specific trends in order to allow for differential long-term and seasonal trends across geographic areas (including state-by-season fixed effects, state-by-quadratic year trends, municipality-by-season fixed effects, and year-month fixed effects). In choosing these fixed effects and trends, I am somewhat restricted by sample size – municipality-month fixed effects, for example, cannot be estimated with much precision given that the median municipality-month combination has only 9 individuals.

I run these regressions for the full sample and then separately for men and women. In order to explore whether gender-specific labor market conditions play a role in determining schooling responses to shocks, I also estimate a specification that interacts various labor market variables with the trimester coefficients of interest.

I cluster the standard errors at the municipality level.²² As stated above, I am restricted to individuals born in 1979 or later due to the availability of the NARR data, and those who are at least 15 years of age in their most recent MxFLS interview. Because I am identifying off of variation within municipalities over time (controlling non-linearly for year and month effects), I also drop individuals in municipalities with very small numbers of individuals (less than 30), which make up less than 5% of the full sample.

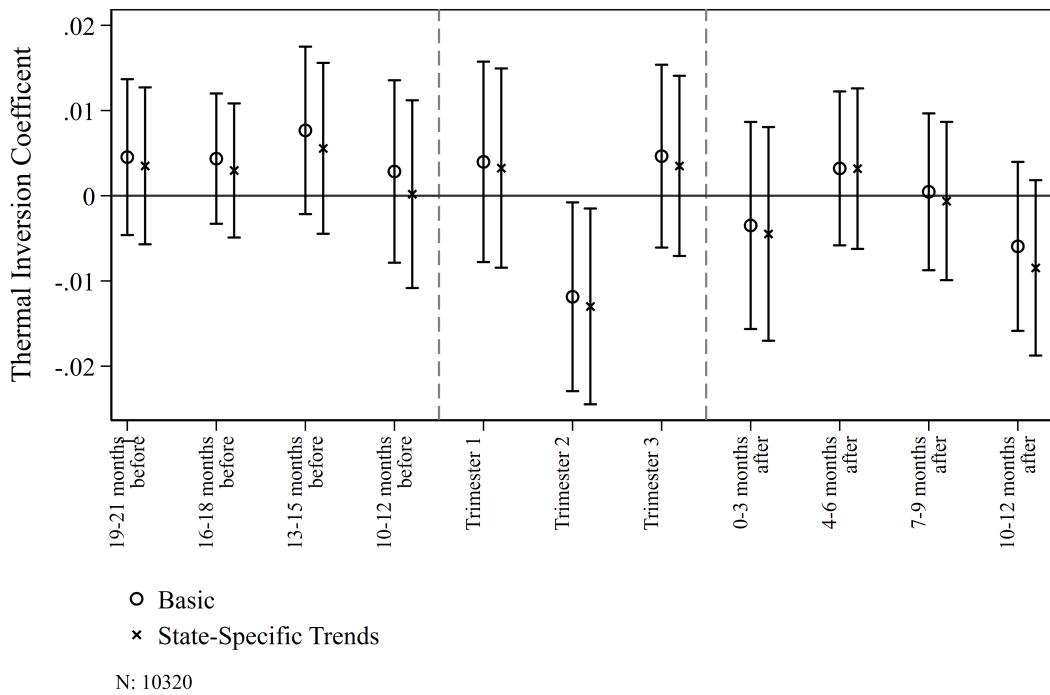
²²There are 150 municipalities in the final sample.

5 Reduced-Form Results

In this section, I begin by documenting the overall and gender-specific effects of pollution. I then investigate the labor market mechanisms driving the gender differences that I find. Finally, I address potential threats to identification.

5.1 Main Reduced Form Results

Figure 3. Effects of Pollution on Raven's Test Z-Scores



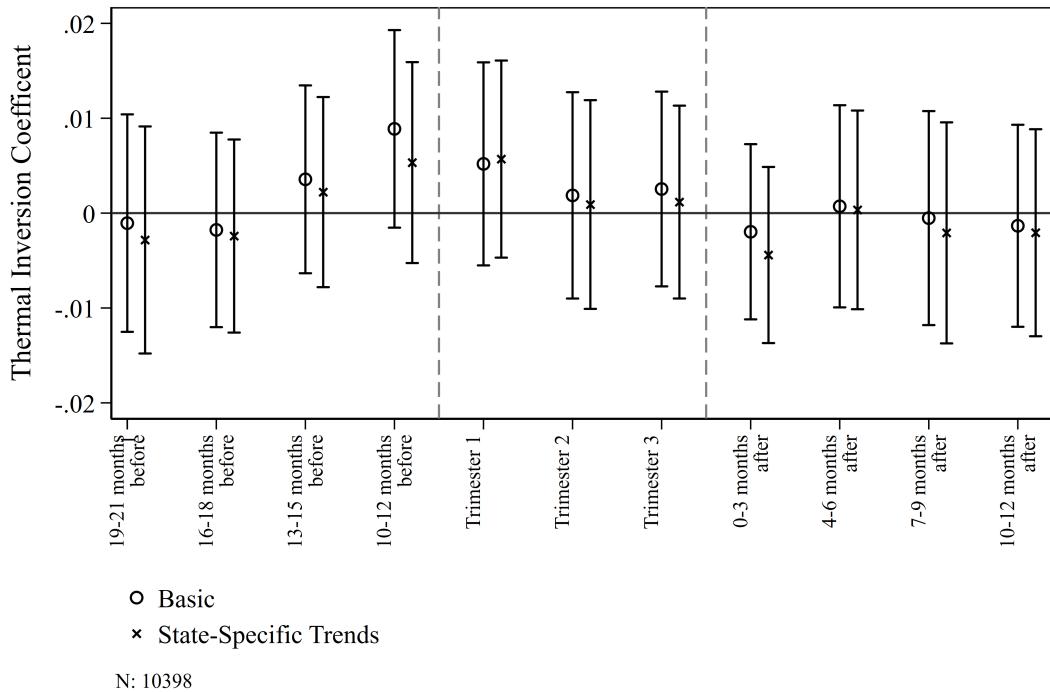
N: 10320

Notes: Intervals represent 95% confidence intervals. “Basic” coefficients are from regressions that control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, gender, mother’s education, father’s education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. “State-Specific Trends” include all basic controls, state-by-season fixed effects and state-by-quadratic year trends. See Table A3, columns 1 and 2, for corresponding estimates.

To display my reduced form results, I graphically illustrate the estimated coefficients from equation 6. All corresponding tables are available in Appendix section A. Figure 3 reports the estimated biological effects of pollution on standardized Raven’s test scores, a measure of cognitive ability. In addition to the coefficients from the baseline specification, I plot the coefficients estimated from a specification that adds state-specific quadratic year trends and state-specific

quarter of the year dummies, hereafter referred to as season dummies. Across both specifications, thermal inversions in the second trimester have a significant negative impact on Raven's test scores. In the specification with state-specific trends, I estimate a coefficient of -0.013, which implies that a standard deviation increase in average monthly thermal inversions per trimester (8.2) leads to a 0.106 standard deviation decline in Raven's test scores. I do not find any significant effects associated with any of the other three-month periods. This is consistent with the medical and economic literature discussed in section 3.1, which flags the second trimester as a crucial period for brain development (Otake, 1998; Almond et al., 2009; Black et al., 2014; Schwandt, 2016; Morgan and Gibson, 1991).²³

Figure 4. Effects of Pollution on Height Z-Scores



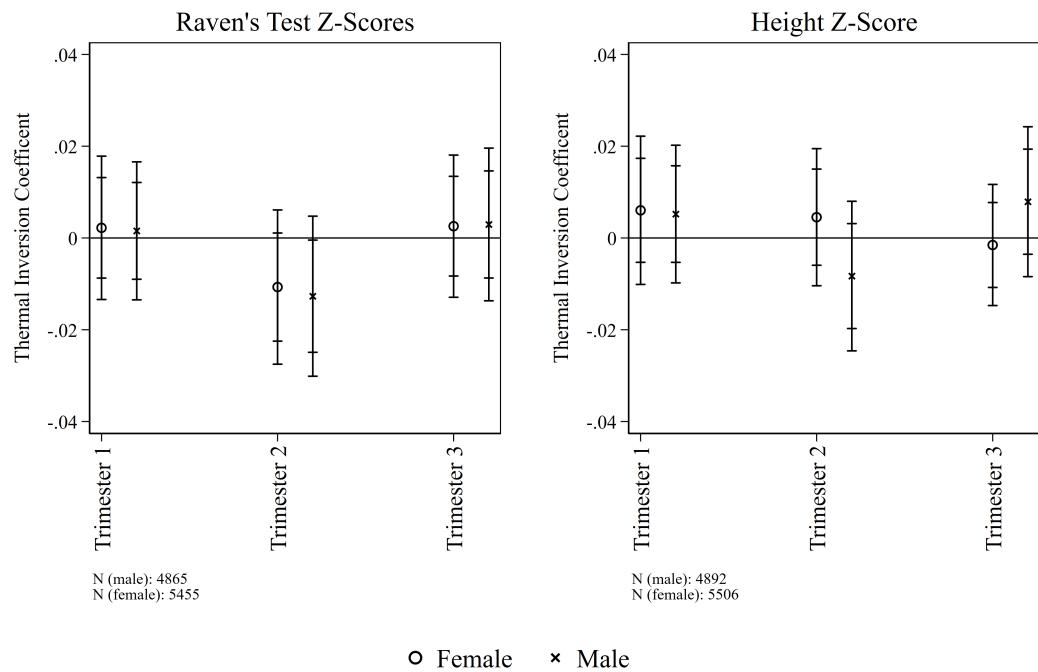
Notes: Intervals represent 95% confidence intervals. "Basic" coefficients are from regressions that control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. "State-Specific Trends" include all basic controls, state-by-season fixed effects and state-by-quadratic year trends. See Table A3, columns 3 and 4, for corresponding estimates.

In contrast, Figure 4 shows no evidence of a robust relationship between pollution exposure (in

²³In the following set of results, I show that it is also the second trimester coefficient, specifically, that has a significant impact on schooling and income outcomes, alleviating concerns that its statistical significance in the cognitive ability regression is simply a result of multiple hypothesis tests and Type 1 error.

any period) and height, which is often used as a cumulative measure of the quality of health and nutritional inputs early in life (Thomas and Strauss, 1997; Maccini and Yang, 2009; Vogl, 2014) and has been shown to be causally linked to fetal health measures like birth weight (Behrman and Rosenzweig, 2004; Black et al., 2007). These results suggest that pollution did not substantially hinder the *physical* development of fetuses and therefore that the negative impact of *in utero* pollution exposure was primarily cognitive.

Figure 5. Effects of Pollution on Cognitive and Physical Health, by Gender

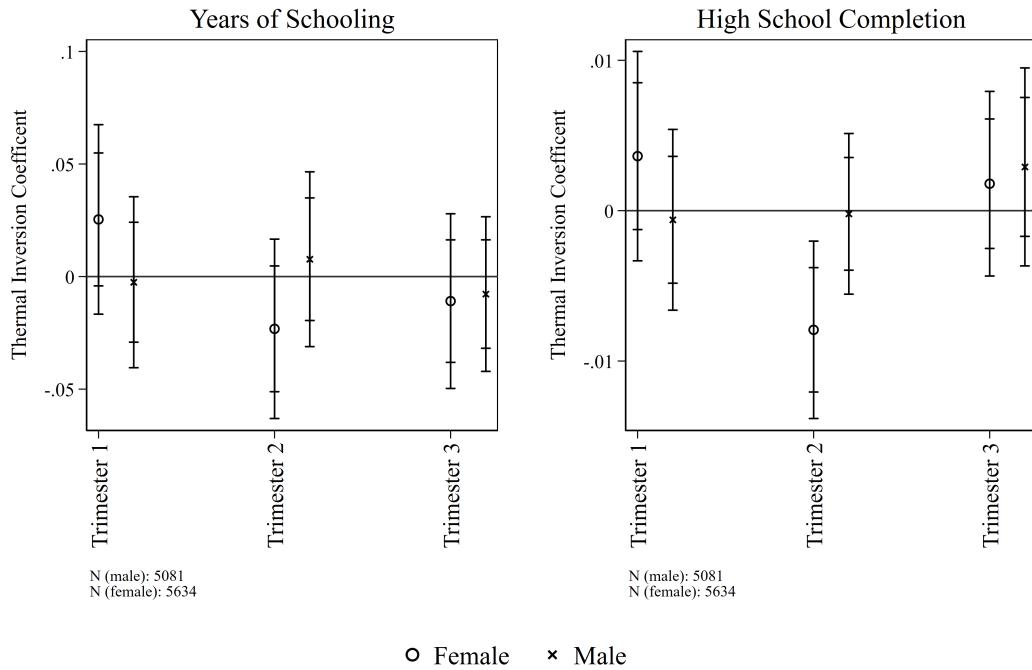


Notes: Separate regressions are conducted for men and women. Intervals represent 95% and 83% confidence intervals, useful for a rough visual detection of significant differences across groups (at the 5% level). Controls include birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, state-by-season fixed effects, state-by-quadratic year trends, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. See Table A4, columns 2 and 4, for corresponding estimates. Although not plotted here, inversions in all other three-month periods are included in these regressions.

In order to study differences across gender, I run these regressions separately for men and women. In the following figures, I plot the coefficients from the state-specific trend specification for males and females on the same graph, reporting only the three trimester coefficients (even though all regressions control for the remaining three-month periods). In the Appendix, I report these trimester coefficients, along with their differences and associated standard errors. In addition to 95% confidence intervals, I also plot 83% confidence intervals, which can be used

for a rough visual detection of differences across groups (significant at the 5% level). The first panel of Figure 5 shows that the second trimester estimates for the effect of pollution on Raven's scores are very similar in magnitude for males and females: -0.0107 for females compared to -0.0127 for males, which are not significantly different from each other. Neither coefficient is significant individually, likely due to the smaller sample sizes, but given the significance of the negative effect in the full sample, the main takeaway from this table is that cognition appears to be affected by pollution in similar ways for men and women. For height, in the second panel of Figure 5, there are no significant gender differences.

Figure 6. Effects of Pollution on Educational Attainment, by Gender



Notes: Separate regressions are conducted for men and women. Intervals represent 95% and 83% confidence intervals, useful for a rough visual detection of significant differences across groups (at the 5% level). Controls include birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, state-by-season fixed effects, state-by-quadratic year trends, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. See Table A5, columns 2 and 4, for corresponding estimates. Although not plotted here, inversions in all other three-month periods are included in these regressions.

It is important to note that the effects being estimated here are reduced form effects: they are the result of the biological effects of pollution as well as a series of investments made by parents up until the age at which the Raven's tests are administered and height is measured.²⁴ The

²⁴See Cunha and Heckman (2007), Cunha and Heckman (2008), and Cunha et al. (2010) for a commonly used

purpose of this analysis is not to tease out the biological effect from the investment responses, as the data is not well-suited for this question: for the sample that I am using, information on early parental investments is not available. What is important for the goals of this paper is the fact that thermal inversions provide exogenous variation in cognitive ability, which allows me to study how schooling decisions respond to exogenously determined cognitive endowments.

Having established that *in utero* exposure to pollution acted as a negative and primarily cognitive endowment shock that did not affect men and women differentially, I next ask whether there were any differences in male and female schooling responses to this shock. Clear gender differences are apparent in Figure 6. Though both panels depict a similar pattern, the result is more pronounced in the regression on high school completion: thermal inversions had a significant negative impact on high school completion for women only. The male coefficient, on the other hand, is positive, statistically indistinguishable from zero, and significantly different from the female coefficient at the 10% level.

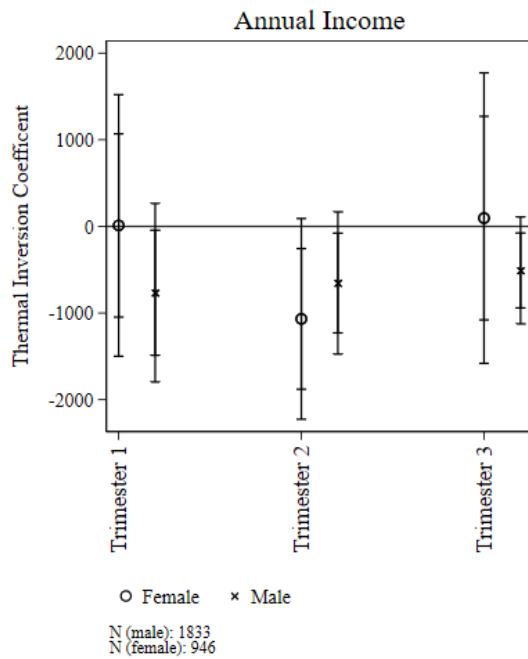
High school graduation appears to be the only milestone affected by pollution: Appendix Table A7 shows that *in utero* thermal inversions had no significant impact on elementary school or junior high school completion for either gender. This suggests that this cognitive shock primarily affected later-life schooling decisions and had little effect on early parental education decisions.

Figure 7 reports the effects of thermal inversions on income, again by gender, among those that report work as their primary activity in the previous week. This deliberately excludes individuals who may be working part time while still in school and whose annual income would not be an appropriate measure of their labor market productivity. Once again, I find that thermal inversions in the second trimester have a significant negative effect on female income. As stated above, the fact that it is the second trimester coefficient that is significant in this regression (as well as in the cognitive ability and female high school completion regressions) alleviates concerns that these statistically significant coefficients are just a result of Type 1 error from multiple hypothesis tests.

dynamic framework for the production function of skill).

The effect of second trimester pollution on men is smaller in magnitude and not significantly different from zero, but still negative, sizable, and not significantly different from the female coefficient. Unlike the high school completion results, Figure 7 does not offer clear-cut evidence for stark gender differences. Although it appears that pollution affected incomes primarily for women, there are also some non-negligible effects on men, which would be consistent with existing examples of early-life circumstances that significantly affected male labor market outcomes despite having very little effect on their schooling decisions (Hoddinott et al., 2008; Rosenzweig and Zhang, 2013; Politi, 2015).

Figure 7. Effects of Pollution on Income, by Gender



Notes: Separate regressions are conducted for men and women. Intervals represent 95% and 83% confidence intervals, useful for a rough visual detection of significant differences across groups (at the 5% level). Controls include birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, state-by-season fixed effects, state-by-quadratic year trends, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. See Table A6, column 2, for corresponding estimates. Although not plotted here, inversions in all other three-month periods are included in these regressions.

These results should be interpreted with caution. Because this is a young sample (aged 15 to 34), the estimated coefficients represent the effect of pollution on early career outcomes, which might be very different from the effects on lifetime income. In particular, the career wage trajectories of men and women likely differ drastically; the direction and magnitude of the gender

differences found here may not be the same as those in lifetime income effects. These regressions also ignore selection into the sample of working individuals: I do not, however find evidence that thermal inversions affected the working decision for either gender (results available upon request). Because of the limitations associated with the income analysis, the remainder of the paper focuses on explaining the well-identified gender difference in the schooling response, about which the framework in section 2 had clear predictions.

5.2 Labor Market Mechanisms

According to the model in section 2, gender differences in schooling responses to shocks can arise from gender differences in expectations about occupational sorting or labor force participation. I begin by testing the extent to which occupational sorting tendencies might be explaining the gender differences I find. To do this, I take advantage of differences across space and over time in the local share of the male and female workforce in white collar jobs, which I argue are reasonable proxies for p_{jg} , the expected probability of taking up a white-collar job.²⁵ In particular, the gender-specific proportion of white collar employment in the local labor market during a child's critical schooling transition period should be positively related to the expectation that she will end up in a white collar job (p_{jg}). Like Rosenzweig and Zhang (2013), I focus on the local labor market in which a child is residing at age 12. In Mexico, the end of elementary school is a critical transition period during which a large proportion of children drop out (Behrman et al., 2011). Moreover, for the majority of individuals in my sample, I have data on their municipality of residence at age 12 specifically.

For the following analysis, I create an indicator equal to 1 if the predicted white-collar share among men (for males) or women (for females), in the commuting zone in which the individual was residing at age 12, falls in the top quartile of the overall predicted white-collar proportion

²⁵Although it is difficult to capture expectations without subjective expectations data, the existing literature suggests current labor market conditions can serve as a reasonable proxy for p_{jg} . For example, Jensen (2010) finds that 70% of survey respondents in the Dominican Republic report that people in their community were their main source of information about expected earnings. Similarly, Nguyen (2008) shows that information about current labor market conditions can affect parental and child expectations about future returns. In a slightly different context, Attanasio and Kaufmann (2012) use current conditions in the marriage market – gender ratios for various education categories – to proxy for marriage market expectations.

distribution. The results in Table 3 use a discrete transformation of proportions predicted by combining municipality-level occupation distributions from the census with national-level industry growth rates from ENIGH, using an industry shift-share strategy similar to Bartik (1991) and others (See Appendix section A.4 for more details). However, the pattern of results is robust to the use of a continuous instead of a discrete measure, as well as simply assigning individuals the relevant value from the census decade in which they turned 12 (Table A8).

I begin this exercise by reporting, in columns 1 and 3 of Table 3, the trimester coefficients from the fully-interacted specification used to generate the second panel of Figure 6, which demonstrates the significant gender difference in the effect of thermal inversions on high school completion. In columns 2 and 4, I add inversion-by- p_{jg} interactions to investigate the extent to which this gender difference is being driven by gender-specific labor market opportunities. The negative effect of second trimester inversions on high school completion is concentrated among individuals more likely to go into a white-collar job: in both specifications, the coefficient on the second trimester interaction is negative and significant at the 5% level, while the main effect is much smaller and insignificant. This establishes a clear link between labor market conditions and investment responses to shocks, likely operating through the effect the current labor market has on expectations.²⁶ More importantly, the gender difference that appears in columns 1 and 3 completely disappears when the labor market interactions are included: the male interaction is much smaller in magnitude than the p_{jg} interaction and insignificant. The drastic decrease in the second trimester male interaction with the inclusion of the p_{jg} interactions demonstrates that the gender difference in this context is driven by the different occupational choices made by men and women. Importantly, this helps me rule out other common explanations for gender differences in the effects of shocks, including gender discrimination or gender-specific norms regarding high school completion. If the larger negative effect of pollution on female schooling was driven by son preference or parental beliefs that men should complete high school no matter their ability level, the gender difference should persist even after controlling for gender-specific white-collar

²⁶The finding that parental and child expectations can influence child schooling decisions is consistent with evidence from subjective expectations data from urban Mexico (Kaufmann, 2014; Attanasio and Kaufmann, 2014).

Table 3. Effects of Pollution on High School Graduation, by White Collar Opportunities

	(1)	(2)	(3)	(4)
Average monthly inversions...	HS Completion	HS Completion	HS Completion	HS Completion
Trimester 1	0.00375 (0.00353)	0.00462 (0.00534)	0.00363 (0.00353)	0.00447 (0.00553)
Trimester 2	-0.00773** (0.00306)	-0.00172 (0.00415)	-0.00792*** (0.00299)	-0.000896 (0.00435)
Trimester 3	0.000748 (0.00308)	0.000168 (0.00571)	0.00179 (0.00311)	0.000764 (0.00579)
Trimester 1 x 1(Male)	-0.00460 (0.00476)	-0.00437 (0.00596)	-0.00423 (0.00478)	-0.00429 (0.00621)
Trimester 2 x 1(Male)	0.00702* (0.00393)	0.00258 (0.00448)	0.00771* (0.00405)	0.00221 (0.00483)
Trimester 3 x 1(Male)	0.00240 (0.00429)	0.00185 (0.00593)	0.00112 (0.00431)	0.000707 (0.00584)
Trimester 1 x 1(Predicted white collar proportion in top quartile)		-0.000873 (0.00456)		-0.000643 (0.00476)
Trimester 2 x 1(Predicted white collar proportion in top quartile)		-0.00758** (0.00359)		-0.00891** (0.00370)
Trimester 3 x 1(Predicted white collar proportion in top quartile)		0.000228 (0.00454)		0.000691 (0.00455)
N	10715	10572	10715	10572
Dependent variable mean	0.266	0.264	0.266	0.264
Additional Fixed Effects	None			state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for the following variables and their interactions with a male indicator (as well as the main effect of gender): birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period, as well as inversions in all other three-month periods. In columns 2 and 4, the main effect of the white collar variable and the interactions with inversions in all other three month periods are also included. Predicted white collar proportions calculated using census data and annual industry growth rates from ENIGH. See Data Appendix for details on the construction of predicted white collar proportions.

opportunities. In my discussion of threats to identification in section 5.3.3, I provide a more detailed discussion about why these results help us rule out son preference as an explanation.

In addition to gender-specific occupational choices, the framework in section 2 highlighted gender-specific labor force participation as a potential source of gender differences in schooling responses. Given that the white-collar proportions already explain such a large portion of the gender difference, there should be little scope for labor force participation to play a role, unless gender-specific white-collar proportions are correlated with gender-specific labor force participation rates (and are simply picking up the effects of this variable). In order to investigate this possibility, I repeat the same analysis conducted in Table 3, but include interactions between the inversion variables of interest and a proxy for q_{jg} , the gender-specific expected likelihood of entering the labor market. Similar to the above analysis, I use gender-specific labor force participation rates from the census and assign them to individuals based on their municipality of residence at age 12 and the census year closest to their 12th birthday. In Table 4, I report the results from a regression that includes two sets of interactions: inversions interacted with a high labor force participation indicator and inversions interacted with a high white-collar proportion indicator. These results make it clear that labor force participation plays no role in explaining why men and women respond differently to these inversion shocks. All of the labor force participation interactions are small in magnitude and statistically insignificant, while the coefficient estimates for the white-collar proportion interactions are slightly larger in magnitude compared to the previous results. The fact that labor force participation plays such a limited role could be an indication that adult female labor force participation rates do not adequately represent the labor force participation expectations of young girls, especially during a period of rapidly increasing female employment.²⁷ It could also be an indication that the gap between white-collar and blue-collar complementarities, whose importance in determining the schooling response (in equation 3) depends on white-collar expectations, is larger than the gaps in the occupation-specific returns to schooling or ability, the importance of which depends on labor force expectations.²⁸

²⁷One survey that collected subjective expectations data shows that teenage boys and girls in Mexico expressed identical expected probabilities of going into the labor force as adults (Attanasio and Kaufmann, 2014).

²⁸This, indeed, turns out to be true – see Table 7.

In short, differential occupational sorting, rather than labor force participation, appears to be the main reason why female schooling responds more to a cognitive shock.

5.3 Threats to Identification

5.3.1 Fertility Timing

The validity of the above analysis relies on the assumption that mothers in a given municipality who experience many thermal inversions during their second trimester are not systematically different from mothers in that same municipality who experience fewer thermal inversions in that same period. One way of testing this is to regress observable maternal characteristics on the thermal inversion variables of interest. Columns 1 and 3 of Table 5 report the results of regressions of maternal years of schooling and an indicator for whether an individual's mother ever worked on thermal inversions in the second trimester.²⁹ In both columns, there is no systematic relationship between inversion exposure and these two maternal characteristics. In Columns 2 and 4, I report the regression results from running the entire specification used for the above analysis (excluding the maternal and paternal schooling controls), with these two maternal characteristics as my dependent variables. None of the trimester coefficients are significantly different from zero (and all are small in magnitude), suggesting that, conditional on all of the fixed effects and weather controls, thermal inversion exposure is truly exogenous to these maternal characteristics.

Of course, these two characteristics may not represent all of the observed or unobserved dimensions that could be systematically correlated with thermal inversion exposure. Perhaps the more relevant variables are those related to maternal characteristics in the year before birth, which are not available in this data set. For example, thermal inversions are more common in winter, and pregnant mothers who are in their second trimester during winter give birth in the spring. In areas where the maize harvest is in the spring, mothers who can afford to give birth in the spring might be less likely to be working in agriculture, for example, than mothers who choose instead to give birth in the fall. In the current specification, month fixed effects help

²⁹These are the only two maternal characteristics which are recorded in a comparable way for individuals with parents living in the household and individuals whose parents do not live in the same household.

Table 4. Effects of Pollution on High School Graduation, by White Collar Opportunities and Labor Force Participation

	(1)	(2)
Average monthly inversions...	HS Completion	HS Completion
Trimester 1	0.00455 (0.00550)	0.00186 (0.00494)
Trimester 2	-0.000664 (0.00430)	0.000352 (0.00434)
Trimester 3	0.000439 (0.00585)	0.00204 (0.00555)
Trimester 1 x 1(Male)	-0.00328 (0.00662)	-0.00253 (0.00627)
Trimester 2 x 1(Male)	0.000382 (0.00527)	-0.000531 (0.00512)
Trimester 3 x 1(Male)	0.00338 (0.00652)	0.00315 (0.00588)
Trimester 1 x 1(White collar variable in top quartile)	-0.000716 (0.00475)	0.00204 (0.00384)
Trimester 2 x 1(White collar variable in top quartile)	-0.00917** (0.00367)	-0.00951** (0.00372)
Trimester 3 x 1(White collar variable in top quartile)	0.00108 (0.00467)	-0.000482 (0.00448)
Trimester 1 x 1(Labor force participation variable in top quartile)	-0.00196 (0.00412)	-0.00189 (0.00419)
Trimester 2 x 1(Labor force participation variable in top quartile)	0.00283 (0.00403)	0.00305 (0.00402)
Trimester 3 x 1(Labor force participation variable in top quartile)	-0.00421 (0.00438)	-0.00544 (0.00439)
N	10572	10677
Dependent variable mean	0.264	0.265
Additional Fixed Effects	state-by-season, state-by-quadratic-year	
White Collar Variable	Predicted	Assigned by census
Agricultural Share Variable	Assigned by census	Assigned by census

Notes:

Standard errors (clustered at municipality level) in parentheses. * p<0.1 ** p<0.05 *** p<0.01

All regressions control for the following variables and their interactions with a male indicator (as well as the main effect of gender): birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period, as well as inversions in all other three-month periods. The main effect of the white collar variable, the labor force participation variable, and each one's interactions with inversions in all other three month periods are also included. See Data Appendix for details on the construction of predicted white collar proportions.

Table 5. Maternal Characteristics and Thermal Inversions

	(1)	(1)	(2)	(2)
Average monthly inversions...	Mother's Education	Mother's Education	1(Mother Worked)	1(Mother Worked)
BEFORE CONCEPTION				
19-21 months before birth		-0.00180 (0.0177)		0.00134 (0.00238)
16-18 months before birth		-0.0229 (0.0194)		-0.00329 (0.00219)
13-15 months before birth		-0.0295 (0.0184)		-0.00150 (0.00266)
10-12 months before birth		-0.0137 (0.0183)		0.00132 (0.00256)
DURING PREGNANCY				
Trimester 1		0.00429 (0.0233)		-0.00149 (0.00269)
Trimester 2	-0.000340 (0.0175)	-0.00930 (0.0215)	-0.0000635 (0.00132)	0.00196 (0.00248)
Trimester 3		0.0125 (0.0202)		0.00237 (0.00309)
AFTER BIRTH				
0-2 months after birth		-0.00557 (0.0178)		-0.00119 (0.00262)
3-5 months after birth		-0.00262 (0.0199)		0.000120 (0.00281)
6-8 months after birth		0.0121 (0.0170)		0.00153 (0.00257)
9-11 months after birth		0.0280 (0.0202)		-0.00234 (0.00256)
N	10322	9770	11104	10496
Mean of dependent variable	6.105	6.170	0.462	0.466
Basic Controls	No	Yes	No	Yes
		state-by-season, state-		state-by-season, state-
Additional Fixed Effects	None	by-quadratic-year	None	by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

The “Basic Controls” included in columns 2 and 4 include: birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, gender, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period.

account for this, but are an incomplete solution if these seasonal effects vary over time or space. In order to better control for time-varying or municipality-specific seasonal effects, I run two additional specifications. In the first specification, I replace the state-season fixed effects with municipality-season fixed effects. In the second specification, I keep these municipality-season fixed effects and replace the year and month fixed effects with interacted year-month dummies. The latter allows for monthly trends to differ non-linearly over time, which would be important if the incentives to time births have changed over the two decade period spanning the birth years in my sample. As Appendix Figures A1 and A2 show, my main results are robust to these specification changes: pollution significantly reduces Raven's test scores for the whole sample and high school completion for women only.

5.3.2 Mortality Selection

Given that *in utero* exposure to pollution is known to affect infant mortality, one important concern is whether my results are being driven by selective mortality. First, it is worthwhile to note that if the infants that do not survive as a result of pollution exposure are mostly from the left tail of the ability distribution, my estimated effects should be an underestimate of pollution's true impact. However, in order to verify whether selective mortality is an issue in my setting, I check whether thermal inversions before birth have any effect on cohort size or cohort gender composition. Using all individuals in the MxFLS born after 1979 and old enough in at least one survey wave to have been asked about their place of birth, I first calculate the total number of individuals and fraction that is male for each birth municipality, birth month, and birth year combination. With each observation representing a year-month-municipality, I regress these aggregate values on thermal inversions during pregnancy and in the year before and after. My results, reported in Table 6, show no evidence for selective mortality in this sample.

While the absence of any pollution-driven changes in cohort size may seem inconsistent with previous studies documenting a positive link between pollution and infant mortality (Arceo et al., 2016; Jayachandran, 2009; Currie and Neidell, 2005; Chay and Greenstone, 2003), it does not necessarily rule out the possibility that thermal inversions led to higher infant mortality in this

sample as well. These null effects are consistent with a situation in which thermal inversions increased infant mortality by accelerating the deaths of infants who would have died before reaching adolescence or adulthood in the absence of pollution. By the time I observe my sample, pollution-driven changes in its composition do not appear to be a substantial concern.

5.3.3 Correlates of White Collar Proportions

In the investigation of labor market mechanisms summarized in Table 3, I have interpreted the gender-specific white-collar shares as representing individuals' expected probabilities of going into a white-collar job. This interpretation may be flawed, however, if these white-collar shares are simply capturing the effects of omitted variables that are correlated with these shares.

For example, if white-collar proportions are correlated with higher pollution levels (due to greater economic activity and urbanization, for example), it could be the case that the stronger negative effect I find in high white-collar areas is due to a non-linear relationship between thermal inversions and pollution. That is, if thermal inversions exacerbate pollution more in highly polluted areas (compared to less polluted areas), this would lead to larger reduced form effects in highly polluted areas. It is important to note, however, that while this would explain the significant negative coefficient on the interaction between white-collar proportions and thermal inversions, this would not be able to explain why the gender difference disappears after controlling for these variables. Another piece of evidence that rules out this alternative explanation can be found in Table A9. Here, I repeat the analysis conducted in Table 3, instead using cognitive ability as the dependent variable. If the high white-collar proportions were simply capturing larger effects in more polluted areas, there should also be a similar pattern with cognitive ability: stronger negative effects in high white-collar areas. On the contrary, in this cognitive ability regression, I find no differential effect across the different labor market conditions, and the effect sizes (for males and females) are not affected by the inclusion of these labor market controls.

In order for an omitted variable to be driving both the significant negative white-collar interaction and the disappearance of the gender difference in Table 3, it has to be gender-specific (like the white-collar proportion variables I construct) and on average different for men and

Table 6. Effects of Pollution on Cohort Size and Gender Composition

Average monthly inversions...	(1) Cohort size	(2) Cohort size	(3) Fraction male	(4) Fraction male
BEFORE CONCEPTION				
19-21 months before birth	-0.00351 (0.00329)	-0.00482 (0.00322)	-0.000354 (0.00271)	0.000394 (0.00269)
16-18 months before birth	-0.00127 (0.00320)	-0.00123 (0.00324)	0.000777 (0.00263)	0.00122 (0.00275)
13-15 months before birth	0.00514 (0.00323)	0.00462 (0.00327)	-0.00109 (0.00278)	-0.000581 (0.00285)
10-12 months before birth	0.00369 (0.00361)	0.00411 (0.00378)	0.00254 (0.00259)	0.00368 (0.00259)
DURING PREGNANCY				
Trimester 1	0.00161 (0.00331)	0.00178 (0.00344)	-0.00134 (0.00246)	-0.000932 (0.00256)
Trimester 2	0.00216 (0.00357)	0.00295 (0.00348)	-0.000612 (0.00284)	-0.000372 (0.00277)
Trimester 3	-0.00557 (0.00467)	-0.00593 (0.00483)	-0.000615 (0.00240)	-0.000327 (0.00233)
AFTER BIRTH				
0-2 months after birth	-0.000247 (0.00397)	0.000248 (0.00403)	0.000183 (0.00244)	0.00142 (0.00257)
3-5 months after birth	0.000784 (0.00379)	0.000110 (0.00408)	-0.00167 (0.00267)	-0.000926 (0.00276)
6-8 months after birth	-0.00101 (0.00345)	-0.00100 (0.00354)	-0.000345 (0.00248)	-0.000430 (0.00261)
9-11 months after birth	-0.00439 (0.00361)	-0.00472 (0.00365)	-0.00131 (0.00307)	-0.00119 (0.00316)
N (municipality-year-months)	10108	10108	10098	10098
Mean of dependent variable	1.352	1.352	0.482	0.482
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year	None	state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

In these regressions, each observation represents a unique municipality-month-year combination. All regressions control for month, year, and municipality fixed effects, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period.

women. Specifically, I need only to be concerned about variables that are correlated with white-collar proportions, within each gender, and which have different means for men and women. Variables that are positively (negatively) correlated with white-collar proportions *and* which are on average higher (lower) for women than for men are a concern. Fortunately, most correlates of white-collar shares do not fulfill this criteria.

For example, one might be concerned that high white-collar shares for a particular gender represent higher income or better job opportunities, specifically for that gender. Unlike white-collar shares, however, incomes are on average higher for men than women, which means that this variable would not be able to generate the pattern of results in Table 3. Another concern is that high white-collar shares could be related to features of the marriage market. Specifically, high white-collar shares for a particular gender may be associated with later marriage or later parenthood for that gender. However, in order for this to produce the pattern of results in Table 3, men would have to marry and have children at a younger age than women on average, which is not the case.

Another explanation that may come to mind is son preference. As mentioned above, it is difficult to reconcile a gender discrimination explanation with the finding that gender differences become insignificant with the inclusion of white-collar controls. However, if preferences for particular genders are correlated with white-collar shares, this could still be an issue. Suppose p_{jg} is simply capturing the effects of a “preference” or “cultural norms” variable: either the preference for gender g in location j , or the cultural expectation that gender g will complete high school in location j . A typical son preference or cultural norms argument might argue that sons are expected to complete high school no matter what, because of discrimination or different expectations about male and female educational trajectories, which would explain why girls are the only ones who reduce their schooling in response to a cognitive shock. For this story to be true, we would have to see a weaker negative effect of the shock among individuals for whom this “preference” variable is high, which would mean a positive coefficient on the interaction between “preference” and thermal inversions. In order to reconcile this with my results, this requires p_{jg} and “preferences” to be *negatively* correlated, which is somewhat counter-intuitive. If a particular

gender is “preferred” and expected to complete high school, the natural assumption would be that they would also be expected to go into higher-paying white-collar jobs.

Though I can rule out gender-specific incomes, gender-specific marital and fertility behavior, and son preference as alternative explanations for my results in Table 3, gender-specific agricultural industry shares do satisfy the criteria of being negatively correlated with white-collar proportions and on average lower for men than women. In other words, areas with high white-collar occupation shares also tend to have low shares of workers in the agricultural industry, while in addition, men on average tend to work in the agricultural industry at higher rates than women. If agricultural shares (which are defined by industry) rather than white-collar shares (which are defined by occupation type) are responsible for the results in Table 3, this would rule out white-collar expectations as the main mechanism and imply a different underlying model from the one outlined in section 2. In order to check whether this is the case, I repeat the analysis conducted in Table 3, this time including interactions between thermal inversions and agricultural industry shares. The inclusion of these interactions does not affect any coefficients from the previous regressions. In Appendix Table A10, I still find that there are significantly larger negative effects for individuals living in high white-collar areas, and no significant differences across individuals facing different agricultural industry shares.

6 Structural Model and Estimation

The above analysis has shown that female schooling decisions are more strongly affected by a cognitive endowment shock than male schooling decisions. This, in theory, could be due to gender-specific expectations about entering the labor force or about working a white-collar job. However, my empirical results show that occupations seem to matter most: gender differences are primarily driven by the lower representation of women in blue-collar jobs. Recall that the framework in section 2 predicted a larger schooling response among individuals likely to go into white-collar jobs *if* schooling and ability are more complementary in white-collar than in blue-collar occupations. To verify whether this last statement is indeed true, I estimate the parameters

of occupation-specific wage functions using a model of schooling and occupational choice. This section describes the model, estimation methods, and sample used to obtain these parameters, and concludes by presenting these estimates.

Schooling and occupational choice are endogenous decisions that likely depend on each other. As a result, separate white-collar and blue-collar Mincer-style regressions may not yield consistent estimates of the parameters of interest. The endogeneity of schooling in a wage regression has long been acknowledged as an important issue to consider when attempting to obtain causal estimates of the return to schooling (Griliches and Mason, 1972). Studies have used both structural approaches and instrumental variables strategies to deal with this issue (Belzil, 2007). Similarly, self-selection into occupations is also recognized as a potential source of bias in the estimation of occupation-specific wage parameters (Roy, 1951; Heckman and Sedlacek, 1985), for which structural approaches are a common solution.

In order to deal with both sources of endogeneity and take into account that schooling decisions may also depend on expected occupational choice, I model the schooling decision with this consideration in mind. A natural starting point is the maximization problem in section 2, where individuals choose the optimal education level to maximize expected future rewards net the cost of schooling, given their expectations about their future occupation. In order to use this model to estimate my wage parameters of interest, I make two changes. First, I discretize the education decision to make the model more tractable. Specifically, individuals choose between completing high school ($E=1$) and not completing high school ($E=0$), which was the only schooling milestone affected by the cognitive shock in the reduced form analysis. In this framework, the maximization problem is now a decision rule: individuals choose to go to high school if

$$p_{jg}(1, \theta)q_{jg}(1, \theta)W_w(1, \theta; \beta_w) + (1 - p_{jg}(1, \theta))q_{jg}(1, \theta)W_b(1, \theta; \beta_b) + (1 - q_{jg}(1, \theta))W_n(1, \theta; \beta_n) - c(1, \theta; \alpha) > p_{jg}(0, \theta)q_{jg}(0, \theta)W_w(0, \theta; \beta_w) + (1 - p_{jg}(0, \theta))q_{jg}(0, \theta)W_b(0, \theta; \beta_b) + (1 - q_{jg}(0, \theta))W_n(0, \theta; \beta_n) - c(0, \theta; \alpha).$$

The second change I make is to explicitly model the occupation and labor force decisions that determine the probabilities of going into the white-collar sector ($p(E, \theta)$) and going into the

labor force at all ($q(E\theta)$), which allows this decision process to now be represented as a two-period discrete choice model.

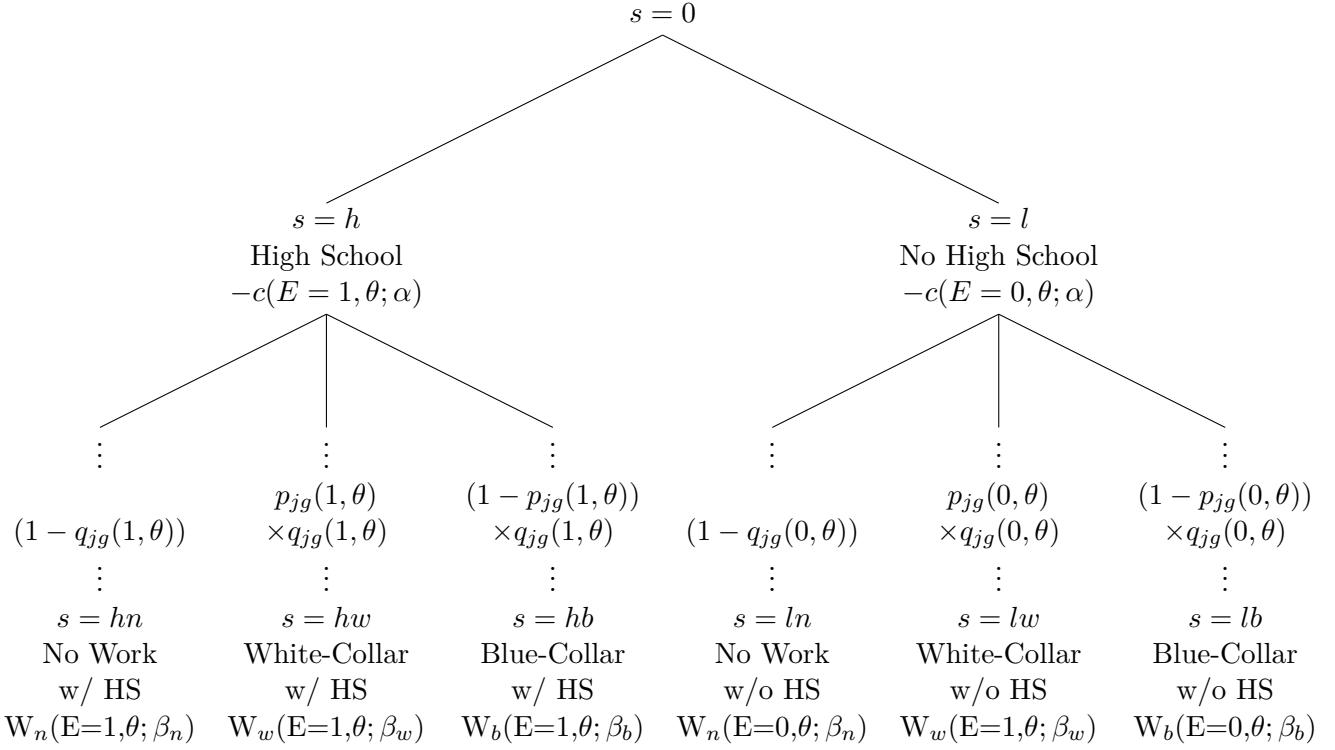
6.1 Decision Tree

This dynamic discrete choice model is summarized in Figure 8, which illustrates the agent's decision tree and all possible states $s \in S = \{0, h, l, hn, hw, hb, ln, lw, lb\}$. This set-up is similar to that of Eisenhauer et al. (2015), which models several discrete schooling decisions from high school enrollment to college completion. I focus on a single schooling decision (high school completion) but expand the model to include a labor market decision after graduation or dropout.

As described above, individuals first decide whether to obtain a high school degree in the first period, during which they incur costs $c(E, \theta; \alpha)$. In the second period, individuals decide whether to work in a white-collar job, work in a blue-collar job, or remain out of the labor force. As described in section 2, their rewards in each of these respective choices are given by $W_w(E, \theta; \beta_w)$, $W_b(E, \theta; \beta_b)$, and $W_n(E, \theta; \beta_n)$.

Figure 8 highlights that using a two-period model requires some drastic simplifications. For example, once individuals have chosen their employment status and occupation, they do not change it. While this is certainly a strong assumption, the majority of individuals in my sample do not switch occupation types across waves, as I discuss in more detail in section 6.4. Like the model in Eisenhauer et al. (2015), this framework can be viewed as a deliberately simplified version of the dynamic discrete choice model of education and occupational choice in Keane and Wolpin (1997). These simplifications make it possible to estimate a joint schooling and occupational choice model when I do not have the long-run annual panel data that is typically required when modeling and predicting the evolution of wages over the life cycle. Because these life cycle predictions are not the focus of this paper, I use a simpler model which allows me to estimate occupation-specific wage parameters while accounting for the endogeneity of schooling and occupational choice, using only one cross-section of individuals.

Figure 8. Decision Tree



6.2 Wage Equations

In this model, the value of each state depends on the immediate net rewards, as well as the expected future value of all feasible states made available by entering that state. I denote the current state $s \in S = \{0, h, l, hn, hw, hb, ln, lw, lb\}$. When an individual picks s' , they earn net rewards

$$R(s') = Y(s') - C(s').^{30}$$

In the four states in which the individual is working, $Y(s')$ is simply equal to the discounted sum of annual income earned during their working lives (which I assume to be from age 30 to 50) given that they have chosen s' . In other words, if they choose a white-collar job,

$$Y(s') = W_w(E(s'), \theta; \beta_w) = \sum_{t=0}^{20} \delta^t w_w(E(s'), \theta, t; \beta_w) \quad \forall s' \in \{hw, lw\}, \quad (7)$$

³⁰The use of net rewards in each state is another way in which the structural model departs from the maximization problem in section 2. Here, the costs in each state are important for identification.

and if they choose a blue-collar job,

$$Y(s') = W_b(E(s'), \theta; \beta_b) = \sum_{t=0}^{20} \delta^t w_b(E(s'), \theta, t; \beta_b) \quad \forall s' \in \{hb, lb\}, \quad (8)$$

where $w_k(E(s), \theta, t; \beta_k)$ is the income earned in state s (which determines the occupation type k) at time t ,³¹ expressed as a linear function of observed characteristics and an unobserved component. Like in equation 1, wages depend on schooling and ability θ . Unobserved by the researcher but known to the agent, ability is captured by an individual's Raven's test score plus a standard normal error,

$$\theta = \text{Raven's Score} + \nu,$$

in order to allow for other dimensions of skill to be included in this measure of labor market ability. Because θ enters all relevant equations of this model, ν allows for correlation in the unobservables that govern the schooling choice, occupational choice, and wages. This is assumed to be the only source of dependency among the unobservables in the model.

Let A_1 to A_3 represent three indicator variables for each of the three 5-year age categories spanning ages 36 to 50, which leaves the age group 30 to 35 as the omitted category. As in the Eisenhauer et al. (2015) framework, which allows for a different set of coefficients for each state, I allow for the effects of ability and experience to vary not only by job type but also by schooling level. Stochastic shocks $\epsilon(s, t)$ are independently normally distributed with mean zero (and variance normalized to 1). These components form the per-period wage functions for each occupation type:

$$\begin{aligned} w_w(E(s), \theta, t; \beta_w) = & \beta_{w0} + \beta_{w1}E(s) + \beta_{w2}\theta + \beta_{w3}E(s)\theta + \\ & \sum_{j=4}^6 \beta_{wj}A_{j-3}(t) + \sum_{j=7}^9 \beta_{wj}A_{j-6}(t)E(s) + \sum_{j=10}^{k_w} \beta_{wj}X_{wj} + \epsilon(s, t) \\ & \forall s \in \{hw, lw\} \end{aligned} \quad (9)$$

³¹Time, measured in years from the beginning of an individual's working life, is equal to age minus 30.

$$\begin{aligned}
w_b(S(s), \theta, t; \beta_b) = & \beta_{b0} + \beta_{b1}E(s) + \beta_{b2}\theta + \beta_{b3}E(s)\theta + \\
& \sum_{j=4}^6 \beta_{bj}A_{j-3}(t) + \sum_{j=7}^9 \beta_{bj}A_{j-6}(t)E(s) + \sum_{j=10}^{k_b} \beta_{bj}X_{bj} + \epsilon(s, t) \\
\forall s \in \{hb, lb\}.
\end{aligned} \tag{10}$$

The coefficients that map schooling and ability to wages are occupation-specific, which implies that a fixed set of characteristics will map to a different level of wages in white-collar and blue-collar jobs. This is consistent with the existence of two types of skill: one that is rewarded in white-collar occupations and one that is rewarded in blue-collar occupations, each formed by different functions of individual characteristics. Individuals are therefore choosing their job type in a generalized version of the Roy (1951) economy.

β_{w3} and β_{b3} are important parameters that are typically left out of wage equations in similar dynamic discrete choice models. These represent the non-separability between schooling and ability in each occupational sector, or $\frac{\partial^2 W_w}{\partial E \partial \theta}$ and $\frac{\partial^2 W_b}{\partial E \partial \theta}$ using the previous notation. Values greater than zero indicate the existence of complementarities between schooling and ability in the wage function. Most importantly, however, the difference between β_{w3} and β_{b3} will indicate whether schooling and ability are more complementary in one occupation type than the other.

Individuals do not earn “adult” income in period 1, while high-school-aged. Even if they do earn “adolescent” wages during this period, I do not observe this for the vast majority of individuals in the data. However, I do allow for opportunity cost of schooling – which includes foregone wages – to vary across individuals, as shown in the cost functions below.

6.3 Relative Cost Equations

At each node, I normalize the non-stochastic portion of the cost of one state (states l , hb , and lb for nodes $s = 0$, $s = h$, and $s = l$, respectively) to equal zero because only relative costs can be identified. Relative costs depend on θ , a vector of observed characteristics Q_s , and stochastic shocks $\eta(s)$. $c(h)$ represents the total cost of obtaining a high school degree. $c(hw)$ and $c(lw)$ represent the costs of going into a white-collar job, relative to a blue-collar job, for individuals

with and without high school degree.

$$c(h) = c_h + \alpha_h \theta + \sum_{j=1}^{q_h} \delta_{hj} Q_{hj} + \eta(h) - \eta(l) \quad (11)$$

$$c(hw) = c_{hw} + \alpha_{hw} \theta + \sum_{j=1}^{q_{hw}} \delta_{hwj} Q_{hwj} + \eta(hw) - \eta(hb) \quad (12)$$

$$c(lw) = c_{lw} + \alpha_{lw} \theta + \sum_{j=1}^{q_{lw}} \delta_{lwj} Q_{lwj} + \eta(lw) - \eta(lb) \quad (13)$$

Net rewards for states hn and ln are defined below. Although individuals do receive non-monetary rewards in these two states, I do not observe these rewards and cannot separately identify them from costs. Instead, I allow for net rewards $(Y(s') - C(s'))$ to be a function of E , θ , and observable characteristics:³²

$$Y(hn) - c(hn) = W_n(1, \theta, \beta_n) - c(hn) = -(c_{hn} + \beta_{hn} \theta + \sum_{j=1}^{q_{hn}} \delta_{hnj} Q_{hnj} + \eta(hn) - \eta(hb)) \quad (14)$$

$$Y(ln) - c(ln) = W_n(0, \theta, \beta_n) - c(ln) = -(c_{ln} + \beta_{ln} \theta + \sum_{j=1}^{q_{ln}} \delta_{lnj} Q_{lnj} + \eta(ln) - \eta(lb)). \quad (15)$$

The shocks in the cost function, $\eta(s)$, are assumed to be independent across all s . I assume that they are drawn from a Type 1 extreme value distribution, with scale factors specific to each node: ρ_{η_n} for the initial schooling decision, ρ_{η_h} for the occupation decision among high school graduates, and ρ_{η_l} for the occupation decision among non-graduates. Although accompanied by a number of strong assumptions,³³ this error structure greatly reduces the computational burden of estimating the model as it allows for an analytic expression of the likelihood function and the calculation of standard errors using the information matrix.

In period 1, individuals choose whether or not to go to high school based on current rewards and the continuation value of each choice. In period 2, they choose whether to work and what type of job to take, by comparing the expected net benefits of each choice. Individuals observe the cost shocks $\eta(s)$ before they decide on their next state, but only observe the reward shocks

³²Here, as in the white-collar costs, costs are defined relative to the cost of taking a blue-collar jobs.

³³This implies, for instance, a constant error variance across occupational choices within each schooling branch and the independence of irrelevant alternatives.

$\epsilon(s)$ after making their choice.

This set-up generates decision rules and transition probabilities (outlined in Appendix section B), which I use to construct a likelihood function. I estimate the model using maximum likelihood, integrating over the ability error term ν using 100 simulations, and setting the discount rate to 0.04.

6.4 Structural Estimation Sample and Variables

To estimate this model, I use the MxFLS and generate additional labor market controls from the 1970 to 2000 decennial censuses. I restrict to those 30 and older to exclude late high school graduates and those who have not entered their main career occupation. I also restrict to those 50 and younger to exclude early retirees and to avoid classifying individuals into an occupation they may have switched into later in their careers (in preparation for retirement, for example). It should be noted that this sample is distinct from the one used in the reduced form analysis because thermal inversion data is only available for a relatively young sample (born after 1979), while this model requires information from adults later in life.

I take the schooling, occupation, and total annual income for each individual from the most recent survey wave in which they were aged between 30 and 50 and living in Mexico.³⁴ Other variables I obtain from the MxFLS include age, gender, parental schooling (the average of maternal and paternal years of schooling), and an urban indicator for the individual's place of residence. To represent the non-stochastic portion of ability θ , I use the individual's standardized Raven's test score from the first test they took.

As mentioned earlier, switching across occupation types is ignored by the model. Fortunately, 85% of the individuals whom I observe more than once between the ages 30 and 50 never switch between white-collar and blue-collar jobs. More importantly, however, only 5% of those surveyed in all three waves switch more than once.³⁵ Most individuals appear to be picking a job type and

³⁴Although the MxFLS tracks migrants, even those that move to the U.S., data from the detailed interviews of these U.S. migrants are not publicly available. As a result, I include migrants in my analysis, but I use their income and occupation information from the most recent wave in which they were still in Mexico. Doing this alleviates concerns about comparing income earned in the vastly different labor markets of the U.S. and Mexico.

³⁵While the existence of switchers does suggest that an individual's occupational choice is at least partially

staying in it, or else choosing an occupation and eventually ending up there (potentially after dabbling in the other occupation type first).

Like the occupational decision, the “no work” decision is also a permanent one in this framework. In order to avoid erroneously placing individuals who are only temporarily out of work in this category, I only include individuals who report having never worked before in this group. Individuals who are currently out of work and therefore missing occupation and wage information, but who report having worked before, are dropped from the analysis, though the wage parameter estimates are robust to the inclusion of these individuals in the “no work” category.

Zone-level labor market variables that serve as exclusion restrictions are calculated from the 1970 to 2000 decennial censuses. For school-aged variables, I assign individuals, by commuting zone, to the value from the census at the beginning of the decade in which they turned 12. For early working age variables, I assign individuals to the census at the beginning of the decade in which they turned 22.³⁶ As a cost shifter in the white-collar cost equations (Q_{hw} and Q_{lw}), I use the gender-specific share of men or women working in white-collar jobs in an individual’s municipality of residence in their early working years. As a cost shifter in the cost equation for not working (Q_{hn} and Q_{ln}), I use the adult unemployment rate while the individual was working-aged. As cost shifters in the schooling opportunity cost equation (Q_h), I use gender-specific youth employment rates during an individual’s school-aged years. Specifically, for each commuting zone, I calculate the proportion of boys (for males) or girls (for females) aged 16 to 18 who report being currently employed. Appendix Table A2 reports summary statistics for all of the relevant variables described in this section.

6.5 Estimates of Wage Function

Table 7 reports the estimates for the wage parameters from the model described above. The parameter estimates for the cost functions can be found in Appendix Table A12. Tests of model

determined by time-varying shocks or learning, it does not invalidate the important assumption underlying this framework: that individuals can calculate, with reasonable accuracy, the probability that they end up in a particular job type for the majority of their career.

³⁶Because there is no 1980 census, for individuals whose school aged or working age census was the 1980 census, I use the 1990 census instead.

fit are discussed in Appendix section B.3.

In Table 7, the first two columns report the coefficient estimates and standard errors for the white-collar wage parameters, the second two for the blue-collar wage parameters, and the last two report the differences between the two. Average wages are higher in white-collar jobs than in blue-collar jobs. The age patterns differ across occupation types as well; white collar occupations offer greater rewards for experience in the older age categories, but the standard errors of these differences are large. Individuals in urban areas earn higher income on average. Conditioning on schooling and ability, men earn significantly higher wages than women in both occupational sectors. Notably, the male advantage is significantly larger in blue-collar jobs than in white-collar jobs, which is consistent with men having a comparative advantage in more physical blue-collar occupations, offering a possible explanation for why men and women sort into occupations differently in the first place. Of course, this pattern of coefficients is also consistent with other explanations – like occupation-specific gender discrimination – and pinning this down, though an important question, is outside the scope of this paper.

What is most relevant to the model predictions, however, and what has yet to be estimated by existing work, is the relative magnitude of the occupation-specific coefficients on the interaction between high school completion and ability. This term is positive and significant in white-collar jobs, offering evidence for complementarities between ability and educational investments in these occupations. In blue-collar occupations, on the other hand, this term is negative but not statistically significant. The difference between the two coefficients is significant at the 1% level.

These estimates, together with the coefficients on the main effects of high school graduation and ability, provide some interesting insights. There is a positive and significant return to ability in blue collar jobs, with or without a high school degree. In white-collar jobs, however, ability does not seem to improve wages unless the individual has graduated from high school, in which case the returns to ability are quite large (larger in magnitude than in blue collar occupations). If we think of ability as a measure of how quickly one can learn new skills, then it is not surprising that ability is important even in blue-collar jobs, where more productive workers (who have acquired more skills or learned better ways to complete tasks) can earn higher wages. To

understand why the importance of ability is more heterogeneous in white-collar jobs, we can think about a high school degree in two ways. First, completing high school might provide individuals with knowledge and skills that are useful primarily in the white-collar sector and difficult to learn elsewhere. In this case, a high-ability individual who has not had the opportunity to learn skills that are important for white-collar jobs will not earn higher wages than a lower-ability individual (who has also not had this opportunity). For high school graduates, however, ability will increase the amount of skills acquired and therefore eventual wages. Alternatively, high school graduation can be thought of as a signal, one that enables high school graduates to work at “higher-tier” white-collar jobs, where tasks are more complex and career wage trajectories are steeper. In this scenario, high ability individuals will earn more than low ability individuals in these “higher-tier” jobs because they will acquire new skills and earn quicker promotions. In “lower-tier” white-collar jobs, however, where learning new skills is less important and salaries rise much more slowly, the importance of ability might be severely diminished.

Because occupation-specific wage functions separately capture two different types of skill (one that is rewarded in white-collar jobs and one that is rewarded in blue-collar jobs), this result offers a nuanced contribution to the discussion about the production function of skill and whether there are complementarities between cognitive ability accumulated during childhood and schooling investments in adolescence (Cunha and Heckman, 2007; Cunha et al., 2010; Aizer and Cunha, 2012). In short, the existence or strength of complementarities can be heterogeneous across different occupation-specific skills. In the blue-collar context, a low ability individual can “make up” for their low ability by obtaining a high school degree. On the other hand, the strong complementarity between schooling and ability in white-collar jobs means that education would serve to exacerbate income inequalities across the ability distribution. It is important to note that this could be due to the real productivity effects of schooling varying across ability levels and occupations, or due to the signaling value of schooling varying across these dimensions – pinning down the primary reason for these heterogeneous complementarities would be an interesting question for future research.

These estimates confirm that the reduced form results discussed earlier do indeed support

Table 7. Wage Function Parameter Estimates

	White Collar		Blue Collar		WC -- BC Difference	
	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Constant	10.340	(0.115)***	9.832	(0.053)***	0.509	(0.126)***
HS	0.237	(0.119)**	0.198	(0.128)	0.039	(0.167)
Ability	-0.015	(0.077)	0.164	(0.021)***	-0.179	(0.080)**
HS x Ability	0.214	(0.089)**	-0.070	(0.086)	0.284	(0.124)**
HS: Age 35-40	0.115	(0.099)	0.188	(0.153)	-0.072	(0.183)
HS: Age 41-45	0.141	(0.092)	0.246	(0.160)	-0.105	(0.184)
HS: Age 46-50	0.676	(0.130)***	0.346	(0.213)	0.330	(0.250)
No HS: Age 35-40	0.288	(0.144)**	-0.005	(0.053)	0.293	(0.154)*
No HS: Age 41-45	0.243	(0.140)*	-0.016	(0.054)	0.259	(0.150)*
No HS: Age 46-50	0.315	(0.216)	0.117	(0.067)*	0.198	(0.226)
Male	0.372	(0.065)***	0.684	(0.040)***	-0.312	(0.076)***
Urban	0.199	(0.085)**	0.418	(0.040)***	-0.219	(0.093)**
Parental Education	0.036	(0.008)***	0.013	(0.004)***	0.024	(0.006)***

Notes:

Standard errors are calculated analytically using the information matrix.

* p< 0.1 ** p< 0.05 *** p< 0.01

the model predictions from section 2. Because women tend to sort disproportionately into jobs where schooling and ability are more complementary, they exhibit larger schooling responses to a cognitive endowment shock.

7 Conclusion

This study offers evidence that gender differences in investment responses can arise from gender-specific labor market expectations. Long-term cognitive damage caused by pollution exposure *in utero* affects the schooling decisions and income of women, but not of men. The gender difference in the schooling response is largely driven by the different labor market opportunities faced by men and women. In particular, men are more likely to end up in the blue-collar sector, where I show that schooling and ability exhibit a lower degree of complementarity than in white-collar jobs.

The finding that schooling and cognitive ability are more complementary in white-collar than in blue-collar jobs has important implications for our understanding of whether education can

help remediate for early-life inequalities across the ability distribution. Because a higher degree of complementarity offers less scope for remediation, complementarities that are much stronger in white-collar jobs (and in fact non-existent in blue-collar jobs) suggest that remediation is less likely for individuals who eventually end up working in these white-collar occupations. In short, occupation types can determine whether education exacerbates or remediates early-life ability inequalities, which could shed some light on social mobility in different contexts – why this varies across space and might change over time.

By emphasizing the interaction between early-life shocks and labor market opportunities, this paper speaks to two important questions in the early life literature. First, how do early-life shocks interact with policy interventions or economic conditions later in life? Whether these events are policy interventions (Adhvaryu et al., 2015; Rossin-Slater and Wüst, 2015; Gunnsteinsson et al., 2014), economic shocks (Bharadwaj et al., 2017), or the labor market conditions studied in this paper, the fact that they interact with early-life conditions in ways we may not yet fully understand has implications for future policy and the interpretation of existing results. Second, this paper sheds light on the substantial heterogeneity – both across and within studies – in the estimated schooling responses reported in the existing literature.³⁷ Given that labor market conditions vary over time, across space, and across groups, this heterogeneity can be explained by the main result of this paper: individuals facing different labor market opportunities respond differently to early-life shocks.

³⁷For example, many studies find that health conditions early in life have a substantial impact on educational attainment (Almond, 2006; Bleakley, 2007), while others find no effect (Venkataramani, 2012; Cutler et al., 2010), or much smaller effects for certain groups (Maluccio et al., 2009; Maccini and Yang, 2009; Field et al., 2009; Bleakley, 2010).

References

Adhvaryu, A., Molina, T., Nyshadham, A., and Tamayo, J. (2015). Helping Children Catch Up: Early Life Shocks and the Progresa Experiment.

Aizer, A. and Cunha, F. (2012). The production of human capital: Endowments, investments and fertility. Technical report, National Bureau of Economic Research.

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

Almond, D. and Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives*, pages 153–173.

Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, (November).

Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annu. Rev. Econ.*, (January 2013).

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Arceo, E., Hanna, R., and Oliva, P. (2016). Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City. *The Economic Journal*, 126(591):257–280.

Atkin, D. (2016). Endogenous Skill Acquisition and Export Manufacturing in Mexico. *American Economic Review*, 106(8):2046–85.

Attanasio, O. and Kaufmann, K. (2012). Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations.

Attanasio, O. P. and Kaufmann, K. M. (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, 109:203–216.

Aubard, Y. and Magne, I. (2000). Carbon monoxide poisoning in pregnancy. *BJOG: An International Journal of Obstetrics & Gynaecology*, 107(7):833–838.

Backes, C. H., Nelin, T., Gorr, M. W., and Wold, L. E. (2013). Early life exposure to air pollution: how bad is it? *Toxicology letters*, 216(1):47–53.

Barker, D. J. (1990). The fetal and infant origins of adult disease. *BMJ: British Medical Journal*, 301(6761):1111.

Bartik, T. J. (1991). *Who benefits from state and local economic development policies*. W. E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Beard, J. D., Beck, C., Graham, R., Packham, S. C., Traphagan, M., Giles, R. T., and Morgan, J. G. (2012). Winter temperature inversions and emergency department visits for asthma in Salt Lake County, Utah, 2003-2008. *Environmental health perspectives*, 120(10):1385–1390.

Behrman, J. R., Parker, S. W., and Todd, P. E. (2011). Do conditional cash transfers for schooling generate lasting benefits? A five-year followup of PROGRESA/Oportunidades. *Journal of Human Resources*, 46(1):93–122.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to Birthweight. *The Review of Economics and Statistics*, 86(2):586–601.

Belzil, C. (2007). The return to schooling in structural dynamic models: a survey. *European Economic Review*, 51(5):1059–1105.

Bhalotra, S. and Venkataramani, A. (2013). Cognitive Development and Infectious Disease: Gender Differences in Investments and Outcomes.

Bhalotra, S. R., Fernández, M., and Venkataramani, A. S. (2015). The Right Tail and the Right Tale: The Gender Wage Gap in Mexico. *Gender Convergence in the Labor Market (Research in Labor Economics, Volume 41)* Emerald Group Publishing Limited, 41:299–341.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. (2017). Gray matters: fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists*, 4(2):505–542.

Black, S., Bütikofer, A., Devereux, P., and Salvanes, K. (2014). This Is Only a Test? Long-Run and Intergenerational Impacts of Prenatal Exposure to Radioactive Fallout. *Scientific American*, pages 1–50.

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

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

Bleakley, H. (2010). Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American economic journal. Applied economics*, 2(2).

Bobonis, G. J., Miguel, E., and Puri-Sharma, C. (2006). Anemia and school participation. *Journal of Human resources*, 41(4):692–721.

Chay, K. Y. and Greenstone, M. (2003). The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, 118(3):1121–1167.

Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. *The American Economic Review*, 97(2):31–47.

Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, 43(4):738–782.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Currie, J. and Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn From California’s Recent Experience? *Quarterly Journal of Economics*, 120(3):1003–1030.

Currie, J. and Vogl, T. (2013). Early-Life Health and Adult Circumstance in Developing Countries. *Annu. Rev. Econ.*, 5:1–36.

Currie, J., Zivin, J. G., Mullins, J., and Neidell, M. (2014). What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution? *Annual Review of Resource Economics*, 6(March):217–247.

Cutler, D., Fung, W., Kremer, M., Singhal, M., and Vogl, T. (2010). Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics*, pages 72–94.

Devasthale, A., Willén, U., Karlsson, K.-G., and Jones, C. (2010). Quantifying the clear-sky temperature inversion frequency and strength over the arctic ocean during summer and winter seasons from airs profiles. *Atmospheric Chemistry and Physics*, 10(12):5565–5572.

Dobbing, J. and Sands, J. (1973). Quantitative growth and development of human brain. *Archives of disease in childhood*, 48(10):757–767.

Eisenhauer, P., Heckman, J. J., and Mosso, S. (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International economic review*, 56(2):331–357.

Field, E., Robles, O., and Torero, M. (2009). Iodine Deficiency and Schooling Attainment in Tanzania. *American Economic Journal: Applied Economics*, pages 140–169.

Flynn, J. R. (1984). The mean iq of americans: Massive gains 1932 to 1978. *Psychological bulletin*, 95(1):29.

Glickman, T. S. and Zenk, W. (2000). *Glossary of meteorology*. American Meteorological Society.

Goldin, C. (1995). The U-Shaped Female Labor Force Function in Economic Development and Economic History. *Investment in Women’s Human Capital*, page 61.

Griliches, Z. and Mason, W. M. (1972). Education, Income, and Ability. *Journal of Political Economy*, 80(3):S74–S103.

Gunnsteinsson, S., Adhvaryu, A., Christian, P., Labrique, A., Sugimoto, J., Shamim, A. A., and West Jr., K. P. (2014). Vitamin A and Resilience to Early Life Shocks. *Working Paper*.

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 1688(June):2005–2007.

Heckman, J. and Scheinkman, J. (1987). The importance of bundling in a Gorman-Lancaster model of earnings. *The Review of Economic Studies*, 54(2):243–255.

Heckman, J. J. and Sedlacek, G. (1985). Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market. *Journal of political Economy*, 93(6):1077–1125.

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., and Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The lancet*, 371(9610):411–416.

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-Run Impacts of Childhood Access to the Safety Net. *The American Economic Review*, 106(4):903–934.

Isen, A., Rossin-Slater, M., and Walker, W. R. (2017). Every breath you take—every dollar you’ll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy*, 125(3):848–902.

Jacobson, M. Z. (2002). *Atmospheric pollution: history, science, and regulation*. Cambridge University Press.

Jans, J., Johansson, P., and Nilsson, J. P. (2014). Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes. (7929).

Jayachandran, S. (2009). Air Quality and Early-Life Mortality: Evidence from Indonesia’s Wildfires. *Journal of Human Resources*, 44(4):916–954.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2).

Jones, K., Smith, D., Ulleland, C., and Streissguth, A. (1973). Pattern of malformation in offspring of chronic alcoholic mothers. *The Lancet*, 301(7815):1267–1271.

Kaufmann, K. M. (2014). Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns. *Quantitative Economics*, 5(3):583–630.

Keane, M. P. and Wolpin, K. I. (1997). The Career Decisions of Young Men. *Journal of political Economy*, 105(3):473–522.

Lacasaña, M., Esplugues, A., and Ballester, F. (2005). Exposure to ambient air pollution and prenatal and early childhood health effects. *European journal of epidemiology*, 20(2):183–199.

Le, H. Q., Batterman, S. A., Wirth, J. J., Wahl, R. L., Hoggatt, K. J., Sadeghnejad, A., Hultin, M. L., and Depa, M. (2012). Air pollutant exposure and preterm and term small-for-gestational-age births in Detroit, Michigan: long-term trends and associations. *Environment international*, 44:7–17.

Maccini, S. and Yang, D. (2009). Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, 99(3):1006–1026.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *The Economic Journal*, 119(537):734–763.

Mammen, K. and Paxson, C. (2000). Women's work and economic development. *The Journal of Economic Perspectives*, 14(4):141–164.

McBride, W. (1961). Thalidomide and Congenital Abnormalities. *The Lancet*, 278(7216):1358.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E. H., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3):343–360.

Minnesota Population Center (2015). *Integrated Public Use Microdata Series, International: Version 6.4 [Machine-readable database]*. University of Minnesota, Minneapolis.

Morgan, B. and Gibson, K. R. (1991). Nutritional and environmental interactions in brain development. *Brain maturation and cognitive development: Comparative and cross-cultural perspectives*, pages 91–106.

Nguyen, T. (2008). Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar.

Olivetti, C. (2014). The Female Labor Force and Long-Run Development. *Human Capital in History: The American Record*, page 161.

Otake, M. (1998). Review: Radiation-related brain damage and growth retardation among the prenatally exposed atomic bomb survivors. *International journal of radiation biology*, 74(2):159–171.

Peet, E. D. (2016). Environment and Human Capital: The Effects of Early-Life Exposure to Pollutants in the Philippines.

Perera, F. P., Whyatt, R. M., Jedrychowski, W., Rauh, V., Manchester, D., Santella, R. M., and Ottman, R. (1998). Recent developments in molecular epidemiology: a study of the effects of environmental polycyclic aromatic hydrocarbons on birth outcomes in Poland. *American Journal of Epidemiology*, 147(3):309–314.

Peterson, B. S., Rauh, V. A., Bansal, R., Hao, X., Toth, Z., Nati, G., Walsh, K., Miller, R. L., Arias, F., Semanek, D., et al. (2015). Effects of Prenatal Exposure to Air Pollutants (Polycyclic Aromatic Hydrocarbons) on the Development of Brain White Matter, Cognition, and Behavior in Later Childhood. *JAMA psychiatry*.

Pitt, M. M., Rosenzweig, M. R., and Hassan, N. (2012). Human capital investment and the gender division of labor in a brawn-based economy. *The American economic review*, 102(7):3531.

Politi, D. (2015). The effects of the generalized use of iodized salt on occupational patterns in Switzerland. Working Paper.

Rendall, M. (2010). Brain versus brawn: the realization of women's comparative advantage. Available at SSRN 1635251.

Rosales-Rueda, M. and Triyana, M. (2018). The persistent effects of early-life exposure to air pollution: Evidence from the indonesian forest fires. *Journal of Human Resources*, pages 0117–8497R1.

Rosenzweig, M. R. and Zhang, J. (2013). Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins. *Journal of Development Economics*, 104:245–260.

Rossin-Slater, M. and Wüst, M. (2015). Are Different Early Investments Complements or Substitutes? Long-Run and Intergenerational Evidence from Denmark.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146.

Rubalcava, L. and Teruel, G. (2006). Mexican Family Life Survey, Second Round. www.ennvih-mxfls.org.

Rubalcava, L. and Teruel, G. (2007). The Mexican Family Life Survey: First Wave. www.mxfls.uia.mx.

Rubalcava, L. and Teruel, G. (2013). Mexican Family Life Survey, Third Round. www.ennvih-mxfls.org.

Saenen, N., Plusquin, M., Bijnens, E., Janssen, B., Gyselaers, W., Cox, B., Fierens, F., Molenberghs, G., Penders, J., Vrijens, K., et al. (2015). In Utero Fine Particle Air Pollution and Placental Expression of Genes in the Brain-Derived Neurotrophic Factor Signaling Pathway: An ENVIRONAGE Birth Cohort Study.

Sanders, N. J. (2012). What Doesn't Kill you Makes you Weaker: Prenatal Pollution Exposure and Educational Outcomes. *Journal of Human Resources*, 47(3):826–850.

Schwandt, H. (2016). The Lasting Legacy of Seasonal Influenza: In-utero Exposure and Human Capital Development. Technical report.

Stein, Z., Susser, M., Saenger, G., and Marolla, F. (1975). Famine and human development: The Dutch hunger winter of 1944–1945.

Tau, G. Z. and Peterson, B. S. (2010). Normal development of brain circuits. *Neuropsychopharmacology*, 35(1):147–168.

Thomas, D. and Strauss, J. (1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77(1):159–185.

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., and Hiscock, M. (2014). The flynn effect: A meta-analysis. *Psychological bulletin*, 140(5):1332.

Venkataramani, A. S. (2012). Early life exposure to malaria and cognition in adulthood: evidence from Mexico. *Journal of health economics*, 31(5):767–80.

Vogl, T. S. (2014). Height, skills, and labor market outcomes in Mexico. *Journal of Development Economics*, 107:84–96.

Von Lenz, W. and Knapp, K. (1962). Die thalidomid-embryopathie. *Deutsche Medizinische Wochenschrift*, 87(24):1232–1242.

Appendix

A Additional Tables, Background, and Data Details

- Table A1 reports ISCO occupation shares, by gender, from the 2010 Mexican census. I group the occupation categories into two groups using the brain-intensive and brawn-intensive classification used by Vogl (2014), who calculates average skill and strength intensities of each occupational category using job requirement scores in the Dictionary of Occupational Titles.
- Table A2 reports summary statistics for all of the relevant variables for the structural estimation described in section 6.
- Tables A3 to A6 provide the coefficient estimates, standard errors, and observation counts for the graphs in section 5.1. For each variable, the first column includes the basic fixed effects (municipality, month, and year) and the second adds state-specific season fixed effects and state-specific quadratic trends. Table A7 provides the coefficient estimates from identical regressions that use elementary and junior high school completion as dependent variables.
- Table A8 demonstrates the robustness of the labor market mechanism results to the use of other proxies of p_{jg} . The regression in column 2 assigns individuals with the relevant white-collar proportion from the census decade in which they turned 12. Results are consistent with those in Table 3. Column 3 reports the results from using a continuous version of the discrete measure used in Table 3, demeaned so that the main effects can be interpreted as the effects for the average individual. Though more imprecise, these results are consistent with the previous finding that gender-specific labor market opportunities appear to be playing a more important role than gender itself – the male interaction with second trimester inversions is much smaller in magnitude than in column 1.
- Figures A1 and A2 show the robustness of my main results to the inclusion of additional

fixed effects.

- Table A9 repeats the analysis conducted in Table 3, except using Raven's test scores as the dependent variable. Unlike the effect of second trimester pollution on high school completion, the effect of second trimester pollution on cognitive ability is not heterogeneous across white-collar proportions.
- Table A10 repeats the analysis conducted in Tables 3 and A8, but adds interactions between thermal inversions and agricultural shares. Zone-specific agricultural shares are calculated using industry codes in the census and matched to individuals using the commuting zone in which they lived at age 12 and the census decade during which they turned 12. The main results from Table 3 are robust to the inclusion of these additional interactions.

Table A1. Occupation Distributions by Gender

ISCO Occupation Code & Description	Male	Female
White-Collar ("Brains")	19.76	34.85
1 Legislators, senior officials and managers	5	4.47
2 Professionals	7.39	7.88
3 Technicians and associate professionals	4.08	12.04
4 Clerks	3.29	10.46
Blue-Collar ("Brawn")	80.25	65.15
5 Service workers and shop and market sales	17.62	29.63
6 Skilled agricultural and fishery workers	12.25	1.7
7 Crafts and related trades workers	24.23	8.26
8 Plant and machine operators and assemblers	14.19	5.44
9 Elementary occupations (domestic workers, laborers, etc)	11.96	20.12

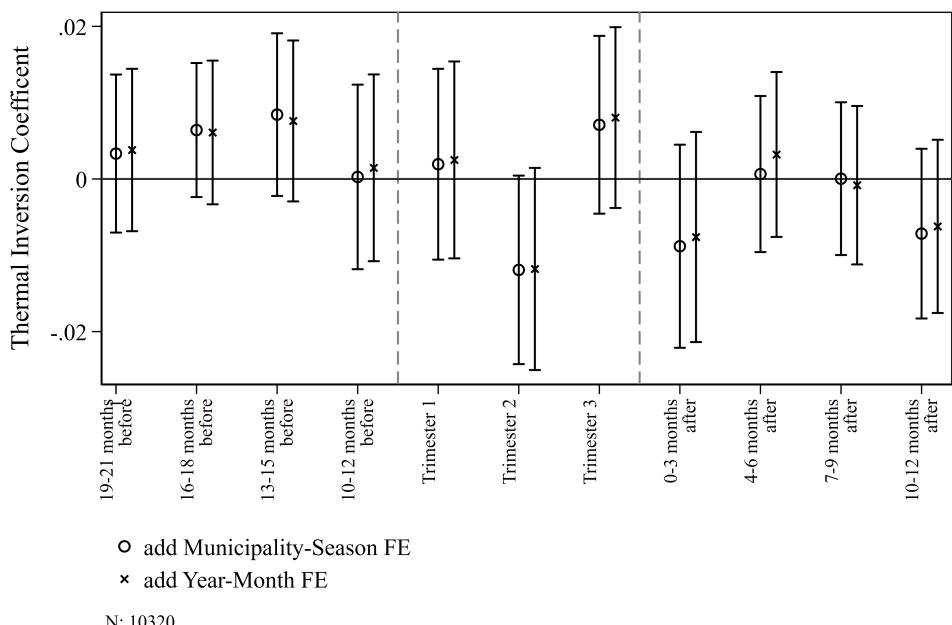
Notes: Brain and brawn categorizations from Vogl (2014). Weighted percentages calculated from working adults aged 30 to 50 in the 2010 Mexican census.

Table A2. Summary Statistics for Structural Estimation Sample

Variable Name	Female			Male		
	Mean	S.D.	N	Mean	S.D.	N
Individual-Level Variables						
Total annual income (inverse hyperbolic sine)	4.64	5.336	3129	10.77	1.942	2136
1(Completed high school)	0.23	0.419	3129	0.28	0.450	2136
1(White collar occupation)	0.19	0.391	3129	0.24	0.428	2136
1(Blue collar occupation)	0.25	0.433	3129	0.74	0.440	2136
1(Not employed)	0.56	0.496	3129	0.02	0.144	2136
Raven's test score (% correct)	0.50	0.240	3129	0.54	0.238	2136
Age	39.01	5.844	3129	38.30	5.699	2136
1(Urban)	0.60	0.489	3129	0.64	0.481	2136
Parental education	3.42	3.323	3129	3.65	3.310	2136
Labor Market Variables						
White collar proportion	0.44	0.119	3129	0.17	0.0781	2136
Unemployment rate	0.02	0.0137	3129	0.02	0.0111	2136
Youth employment rate, per 10 children	1.99	1.079	3129	4.99	1.589	2136

Notes: Sample includes individuals aged 30 to 50 who are either currently employed (with non-missing occupation and income information) or reported never having worked before. Individual-level variables are from the Mexican Family Life Survey. Labor market variables are from the 1970 to 2000 censuses and matched to individuals by their commuting zone of residence. Parental education is the average of maternal and paternal years of schooling. White collar proportion is the fraction of adult males (for men) and adult females (for women) working in white collar jobs during individual's early working years. Unemployment rate is the adult unemployment rate during an individual's early working years. Youth employment rate is the proportion of boys (for men) or girls (for women) aged 16 to 18 who report being employed during an individual's school-aged years.

Figure A1. Effects of Pollution on Cognitive Ability, with Additional Fixed Effects



Notes: Intervals represent 95% confidence intervals. All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, state-by-quadratic year trends, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period.

Table A3. Effects of Pollution on Cognitive and Physical Health

Average monthly inversions...	(1) Raven's test z-score	(2) Raven's test z-score	(3) Height z-score	(4) Height z-score
BEFORE CONCEPTION				
19-21 months before birth	0.00453 (0.00463)	0.00351 (0.00466)	-0.00105 (0.00580)	-0.00283 (0.00605)
16-18 months before birth	0.00435 (0.00387)	0.00297 (0.00398)	-0.00177 (0.00519)	-0.00242 (0.00515)
13-15 months before birth	0.00767 (0.00497)	0.00556 (0.00507)	0.00356 (0.00501)	0.00221 (0.00507)
10-12 months before birth	0.00284 (0.00542)	0.000185 (0.00557)	0.00888* (0.00527)	0.00533 (0.00536)
DURING PREGNANCY				
Trimester 1	0.00398 (0.00595)	0.00324 (0.00592)	0.00519 (0.00542)	0.00570 (0.00526)
Trimester 2	-0.0119** (0.00561)	-0.0130** (0.00581)	0.00187 (0.00550)	0.000907 (0.00557)
Trimester 3	0.00465 (0.00543)	0.00350 (0.00535)	0.00254 (0.00519)	0.00116 (0.00514)
AFTER BIRTH				
0-2 months after birth	-0.00349 (0.00615)	-0.00448 (0.00634)	-0.00196 (0.00467)	-0.00442 (0.00470)
3-5 months after birth	0.00321 (0.00457)	0.00318 (0.00477)	0.000718 (0.00539)	0.000341 (0.00530)
6-8 months after birth	0.000462 (0.00466)	-0.000623 (0.00470)	-0.000524 (0.00571)	-0.00208 (0.00589)
9-11 months after birth	-0.00594 (0.00502)	-0.00846 (0.00521)	-0.00133 (0.00539)	-0.00207 (0.00552)
N	10320	10320	10398	10398
Mean of dependent variable	0.0164	0.0164	-1.008	-1.008
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year	None	state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, gender, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period.

Table A4. Effects of Pollution on Cognitive and Physical Health by Gender

Average monthly inversions...	(1) Raven's test z-score	(2) Raven's test z-score	(3) Height z-score	(4) Height z-score
FEMALE				
Trimester 1	0.00464 (0.00774)	0.00221 (0.00790)	0.00182 (0.00810)	0.00603 (0.00818)
Trimester 2	-0.00971 (0.00832)	-0.0107 (0.00850)	0.00565 (0.00729)	0.00453 (0.00756)
Trimester 3	0.00392 (0.00770)	0.00257 (0.00784)	-0.00178 (0.00668)	-0.00153 (0.00668)
N	5455	5455	5506	5506
Dependent variable mean	-0.00429	-0.00429	-1.043	-1.043
MALE				
Trimester 1	0.00193 (0.00754)	0.00155 (0.00761)	0.00746 (0.00754)	0.00521 (0.00759)
Trimester 2	-0.0139* (0.00814)	-0.0127 (0.00883)	-0.00539 (0.00805)	-0.00830 (0.00825)
Trimester 3	0.00438 (0.00862)	0.00294 (0.00842)	0.0107 (0.00831)	0.00790 (0.00826)
N	4865	4865	4892	4892
Dependent variable mean	0.0397	0.0397	-0.970	-0.970
MALE-FEMALE DIFFERENCE				
Trimester 1	-0.00272 (0.0101)	-0.000663 (0.0109)	0.00564 (0.0111)	-0.000822 (0.0115)
Trimester 2	-0.00422 (0.0117)	-0.00199 (0.0122)	-0.0110 (0.0108)	-0.0128 (0.0110)
Trimester 3	0.000465 (0.0120)	0.000376 (0.0120)	0.0124 (0.0105)	0.00943 (0.0106)
N	10320	10320	10398	10398
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year	None	state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. Separate regressions are conducted for men and women.

Table A5. Effects of Pollution on Schooling by Gender

	(1)	(2)	(3)	(4)
Average monthly inversions...	Years of Schooling	Years of Schooling	HS Completion	HS Completion
FEMALE				
Trimester 1	0.0296 (0.0217)	0.0254 (0.0213)	0.00375 (0.00352)	0.00363 (0.00352)
Trimester 2	-0.0165 (0.0208)	-0.0232 (0.0202)	-0.00773** (0.00305)	-0.00792*** (0.00298)
Trimester 3	-0.0140 (0.0210)	-0.0109 (0.0196)	0.000748 (0.00307)	0.00179 (0.00311)
N	5634	5634	5634	5634
Dependent variable mean	9.521	9.521	0.288	0.288
MALE				
Trimester 1	-0.000200 (0.0194)	-0.00251 (0.0192)	-0.000848 (0.00298)	-0.000607 (0.00304)
Trimester 2	0.00665 (0.0183)	0.00771 (0.0196)	-0.000714 (0.00261)	-0.000211 (0.00270)
Trimester 3	-0.00524 (0.0178)	-0.00777 (0.0174)	0.00315 (0.00320)	0.00291 (0.00333)
N	5081	5081	5081	5081
Dependent variable mean	9.199	9.199	0.241	0.241
MALE - FEMALE DIFFERENCE				
Trimester 1	-0.0298 (0.0304)	-0.0279 (0.0293)	-0.00460 (0.00476)	-0.00423 (0.00478)
Trimester 2	0.0231 (0.0281)	0.0309 (0.0295)	0.00702* (0.00393)	0.00771* (0.00405)
Trimester 3	0.00874 (0.0256)	0.00314 (0.0243)	0.00240 (0.00429)	0.00112 (0.00431)
N	10715	10715	10715	10715
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year	None	state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. Separate regressions are conducted for men and women.

Table A6. Effects of Pollution on Income by Gender

	(1)	(2)
Average monthly inversions...	Annual income	Annual income
FEMALE		
Trimester 1	-140.1 (545.4)	10.54 (763.1)
Trimester 2	-1131.4** (570.1)	-1067.9* (585.7)
Trimester 3	-38.51 (735.5)	95.39 (847.6)
N	946	946
Dependent variable mean	24314.0	24314.0
MALE		
Trimester 1	-404.8 (430.1)	-766.2 (521.2)
Trimester 2	-465.2 (372.4)	-653.8 (415.4)
Trimester 3	-225.6 (305.6)	-508.7 (312.1)
N	1833	1833
Dependent variable mean	31101.5	31101.5
MALE - FEMALE DIFFERENCE		
Trimester 1	-264.7 (649.7)	-776.7 (783.3)
Trimester 2	666.2 (656.4)	414.1 (698.4)
Trimester 3	-187.1 (666.6)	-604.1 (747.9)
N	2779	2779
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year
Standard errors in parentheses (clustered at municipality level)		
* p<0.1 ** p<0.05*** p<0.01		

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01. All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. Separate regressions are conducted for men and women.

Table A7. Effects of Pollution on Early Educational Attainment, by Gender

	(1)	(2)	(3)	(4)
Average monthly inversions...	Elementary School Completion	Elementary School Completion	Junior High School Completion	Junior High School Completion
FEMALE				
Trimester 1	-0.000130 (0.00207)	-0.000326 (0.00203)	0.00264 (0.00326)	0.00210 (0.00325)
Trimester 2	0.00217 (0.00179)	0.00167 (0.00187)	0.000556 (0.00343)	-0.000529 (0.00342)
Trimester 3	-0.00173 (0.00180)	-0.00129 (0.00179)	-0.00285 (0.00361)	-0.00324 (0.00344)
N	5634	5634	5634	5634
Dependent variable mean	0.929	0.929	0.709	0.709
MALE				
Trimester 1	-0.000277 (0.00187)	-0.000402 (0.00197)	-0.00475 (0.00344)	-0.00493 (0.00351)
Trimester 2	0.00185 (0.00208)	0.00122 (0.00221)	0.00229 (0.00364)	0.00188 (0.00373)
Trimester 3	-0.00252 (0.00209)	-0.00303 (0.00208)	-0.00175 (0.00283)	-0.00162 (0.00293)
N	5081	5081	5081	5081
Dependent variable mean	0.908	0.908	0.664	0.664
MALE - FEMALE DIFFERENCE				
Trimester 1	-0.000147 (0.00300)	-0.0000758 (0.00307)	-0.00739 (0.00470)	-0.00702 (0.00486)
Trimester 2	-0.000324 (0.00262)	-0.000447 (0.00283)	0.00174 (0.00550)	0.00241 (0.00545)
Trimester 3	-0.000790 (0.00256)	-0.00174 (0.00243)	0.00111 (0.00450)	0.00162 (0.00441)
N	10715	10715	10715	10715
Additional Fixed Effects	None	state-by-season, state-by-quadratic-year	None	state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period. Separate regressions are conducted for men and women.

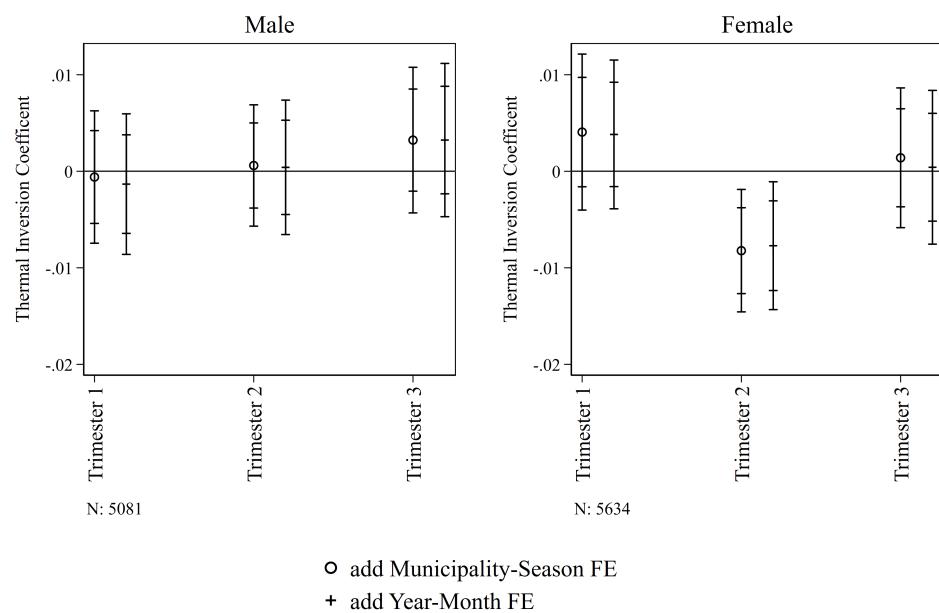
Table A8. Effects of Pollution on High School Graduation, by Alternative White-Collar Variables

	(1)	(2)	(3)
Average monthly inversions...	HS Completion	HS Completion	HS Completion
Trimester 1	0.00363 (0.00353)	0.00155 (0.00498)	0.000103 (0.00503)
Trimester 2	-0.00792*** (0.00299)	0.000500 (0.00441)	-0.00509 (0.00436)
Trimester 3	0.00179 (0.00311)	0.00216 (0.00546)	-0.000695 (0.00475)
Trimester 1 x 1(Male)	-0.00423 (0.00478)	-0.00337 (0.00584)	0.000263 (0.00606)
Trimester 2 x 1(Male)	0.00771* (0.00405)	0.00112 (0.00481)	0.00510 (0.00515)
Trimester 3 x 1(Male)	0.00112 (0.00431)	0.00000468 (0.00531)	0.00266 (0.00556)
Trimester 1 x White Collar Variable		0.00238 (0.00389)	0.0165 (0.0156)
Trimester 2 x White Collar Variable		-0.00971** (0.00378)	-0.0152 (0.0142)
Trimester 3 x White Collar Variable		-0.000640 (0.00435)	0.00828 (0.0137)
N	10715	10677	10572
Dependent variable mean	0.266	0.265	0.264
Additional Fixed Effects		state-by-season, state-by-quadratic-year	
White Collar Variable	None	Discrete, assigned by census	Continuous, predicted

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01
All regressions control for the following variables and their interactions with a male indicator (as well as the main effect of gender): birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period, as well as inversions in all other three-month periods. In columns 2 and 3, the main effect of the white collar variable and the interactions with inversions in all other three month periods are also included.

Figure A2. Effects of Pollution on High School Completion by Gender, with Additional Fixed Effects



Notes: Separate regressions are conducted for men and women. Intervals represent 95% and 83% confidence intervals, useful for a rough visual detection of significant differences across groups (at the 5% level). Controls include birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, state-by-quadratic year trends, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period.

Table A9. Effects of Pollution on Cognitive Ability, by White Collar Opportunities

	(1)	(2)	(3)	(4)
Average monthly inversions...	Raven's test z-score	Raven's test z-score	Raven's test z-score	Raven's test z-score
Trimester 1	0.00464 (0.00777)	0.0195 (0.0122)	0.00221 (0.00793)	0.0181 (0.0126)
Trimester 2	-0.00971 (0.00835)	-0.0134 (0.0115)	-0.0107 (0.00853)	-0.0198* (0.0115)
Trimester 3	0.00392 (0.00772)	0.0144 (0.0133)	0.00257 (0.00786)	0.0153 (0.0133)
Trimester 1 x 1(Male)	-0.00272 (0.0101)	-0.0143 (0.0132)	-0.000663 (0.0109)	-0.0134 (0.0140)
Trimester 2 x 1(Male)	-0.00422 (0.0117)	-0.000782 (0.0138)	-0.00199 (0.0122)	0.00564 (0.0143)
Trimester 3 x 1(Male)	0.000465 (0.0120)	-0.00880 (0.0154)	0.000376 (0.0120)	-0.0105 (0.0153)
Trimester 1 x 1(Predicted white collar proportion in top quartile)		-0.0154 (0.00961)		-0.0165* (0.00988)
Trimester 2 x 1(Predicted white collar proportion in top quartile)		0.00535 (0.00896)		0.0115 (0.00918)
Trimester 3 x 1(Predicted white collar proportion in top quartile)		-0.0116 (0.0108)		-0.0140 (0.0108)
N	10320	10171	10320	10171
Dependent variable mean	0.0164	0.0201	0.0164	0.0201
Additional Fixed Effects	None			state-by-season, state-by-quadratic-year

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01

All regressions control for the following variables and their interactions with a male indicator (as well as the main effect of gender): birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period, as well as inversions in all other three-month periods. In columns 2 and 4, the main effect of the white collar variable and the interactions with inversions in all other three month periods are also included. Predicted white collar proportions calculated using census data and annual industry growth rates from ENIGH. See Data Appendix for details on the construction of predicted white collar proportions.

A.1 Pollutants

CO is a colorless and odorless gas that binds more readily to hemoglobin than oxygen and hinders the body's ability to carry oxygen. CO is produced in combustion, and its main source (especially in urban areas) is vehicle emissions. In a pregnant woman, CO can hinder the delivery of oxygen to the fetus, leading to long-term neurological and skeletal damage (Aubard and Magne, 2000).

Particulate matter refers to a mixture of solid and liquid particles in the air, which includes fine particles known as PM-2.5 (with diameters less than 2.5 micrometers) and inhalable coarse particles known as PM-10 (with diameters less than 10 but greater than 2.5 micrometers). These particles can be emitted directly from a source, like fires or construction sites. They can also form as a result of chemical reactions in the atmosphere. When inhaled by a pregnant woman, particulate matter can cause inflammation or infection. This can thicken blood and plasma, hindering blood flow and glucose transport to the placenta (Lacasaña et al., 2005). The effects of one particular component of particulate matter, polycyclic aromatic hydrocarbons (PAHs), can be especially dangerous. PAHs are thought to increase the prevalence of DNA adducts, which are associated with negative birth outcomes like low birth weight and decreased head circumference (Perera et al., 1998; Le et al., 2012; Lacasaña et al., 2005). Moreover, PAHs can cross the placental barrier and damage the fetal brain by causing inflammation, oxidative stress, or damaging blood vessels. Recent evidence has shown this can result in lower cognition later in childhood (Peterson et al., 2015).

A.2 Thermal Inversions

There are three common types of inversions that are associated with worsened air quality; they form in slightly different circumstances but all result in a warm layer of air above a cooler layer. Radiation inversions take place at night, as the surface cools by emitting thermal infrared radiation. Unlike during the day, when radiation from the sun tends to have a stronger opposing effect, this results in cooler air near the surface than further above ground. Radiation inversions are more common during long, calm, and dry nights, when there is more time for the cooling to

Table A10. Effects of Pollution on High School Graduation, by White Collar Opportunities and Agricultural Shares

	(1)	(2)
Average monthly inversions...	HS Completion	HS Completion
Trimester 1	0.00454 (0.00551)	0.00152 (0.00495)
Trimester 2	-0.000317 (0.00439)	0.00143 (0.00442)
Trimester 3	0.000787 (0.00583)	0.00162 (0.00553)
Trimester 1 x 1(Male)	-0.00502 (0.00619)	-0.00399 (0.00582)
Trimester 2 x 1(Male)	0.00225 (0.00483)	0.00107 (0.00478)
Trimester 3 x 1(Male)	0.000563 (0.00582)	-0.000221 (0.00528)
Trimester 1 x 1(White collar variable in top quartile)	-0.000625 (0.00481)	0.00247 (0.00392)
Trimester 2 x 1(White collar variable in top quartile)	-0.00943** (0.00377)	-0.0105*** (0.00384)
Trimester 3 x 1(White collar variable in top quartile)	0.000744 (0.00459)	0.00000423 (0.00438)
Trimester 1 x 1(Agricultural share variable in top quartile)	0.00401 (0.00994)	0.00211 (0.00986)
Trimester 2 x 1(Agricultural share variable in top quartile)	-0.00623 (0.00561)	-0.00806 (0.00641)
Trimester 3 x 1(Agricultural share variable in top quartile)	-0.000157 (0.00659)	0.00638 (0.00957)
N	10572	10677
Dependent variable mean	0.264	0.265
Additional Fixed Effects		state-by-season, state-by-quadratic-year
White Collar Variable	Predicted	Assigned by census
Agricultural Share Variable	Assigned by census	Assigned by census

Notes:

Standard errors (clustered at municipality level) in parentheses. * p< 0.1 ** p< 0.05 *** p< 0.01
All regressions control for the following variables and their interactions with a male indicator (as well as the main effect of gender): birth month, birth year, municipality of birth, and survey wave by birth year fixed effects, mother's education, father's education, cubic functions of average monthly mean, minimum, and maximum 2m temperatures, monthly relative humidity, monthly precipitation, and monthly cloud coverage during each relevant 3-month period, as well as inversions in all other three-month periods. The main effect of the white collar variable, the agricultural share variable, and each one's interactions with inversions in all other three month periods are also included. See Data Appendix for details on the construction of predicted white collar proportions.

take place, less mixing in the air, and little water vapor to absorb the thermal infrared energy. Subsidence inversions take place when air descends and warms as it compresses, creating a warm layer above cooler air. This can happen in mountainous regions, when air flows down the side of a slope, or in high pressure systems,³⁸ which are characterized by this descending movement and compression of air. Over coastal areas, marine inversions take place when air above the sea, which is cooler than the air above land, flows inland and pushes the warm inland air upward.

A.2.1 NARR Validation Checks

I use the NARR data to identify inversions. Detailed validation exercises have concluded that the NARR data closely matches observational data and offers a considerable improvement over prior global reanalysis data sets (Mesinger et al., 2006). Because all of these checks have included the United States and Canada, which may dominate the validation exercises due to their size, I verify that these conclusions are still valid when I restrict to only Mexico. First, using temperature data that is available on the same INECC pollution database described above, I find a very high correlation (0.87) between the NARR 2-meter temperature and these ground-level measurements. Secondly, I compare my measure of inversions to a measure calculated using temperature readings from satellite data: NASA's Atmospheric Infrared Sounder (AIRS), used by Jans et al. (2014) to identify thermal inversions in Sweden. Because the AIRS was launched in 2002, the data is too recent to use as my measure of inversions or to instrument for my current measure of inversions, but for two overlapping years (2002 and 2003), I find correlations between the NARR and AIRS inversions measures of around 0.7.³⁹

³⁸High pressure systems are associated with high temperatures, clear skies, and light winds at the surface

³⁹It should be noted that there are several factors that complicate the comparison between the NARR and AIRS data. First of all, the times at which the AIRS and NARR data recorded temperatures do not match up exactly. Secondly, the AIRS data has a 1 by 1 degree resolution, substantially larger than the NARR's 0.3 by 0.3 degree resolution. Finally, the AIRS data records temperatures at fewer pressure levels than the NARR. If anything, these factors are likely to weaken the correlation between the two measures, suggesting that a correlation of 0.7 may be an underestimate.

A.2.2 Construction of Thermal Inversion Variables

As described in the main body of the paper, the NARR dataset provides temperature values on a 0.3 by 0.3 degree grid for 29 pressure levels (extending vertically into the atmosphere), every three hours. For each latitude-longitude grid point and for each recorded hour, I create an indicator equal to 1 if the 2-meter temperature (equivalent to what is usually reported in weather reports) is higher than the temperature at the first pressure level above the surface, which lies roughly 300 meters above the surface. Because surface pressure varies across space, I use the temperatures from different pressure levels depending on the altitude at a particular grid point. For a municipality at sea level (1000 hPa), I use the temperature at 975 hPa, but for a higher altitude location in Mexico City, for example, I use the temperature at 700 hPa (if surface pressure is 725 hPa). After creating inversion indicators for every recorded hour, I then collapse to two indicators per day – one for any daytime inversion and one for any nighttime inversion. I then match each Mexican municipality to its four closest grid points and assign each municipality with the inverse-distance weighted average of the nighttime inversion indicator for each day. I sum this indicator over the month and then average over three-month periods.

I assign inversions to individuals based on municipality of birth, a restricted use variable obtained from the migration module of the MxFLS, which is directed to individuals aged 15 and older. For the individuals who are missing this variable,⁴⁰ I assign them the inversions in their municipality of residence. I do this instead of dropping these individuals because over 80% of individuals who move municipalities between birth and the survey date report that they are currently living in their state of birth and about 70% report living in their municipality of birth. Therefore, for over half of the individuals with missing birth municipalities, using municipality of residence is the correct imputation, while a majority of the remainder only moved short distances (i.e. within state).

To create trimester-specific inversion variables for each individual, because I do not have the

⁴⁰Less than 5% of individuals in each wave of the migration module listed either no municipality at all or a municipality name that could not be mapped to a unique municipality code. A slightly larger percentage of individuals in my final sample were missing this variable simply because they had not completed a migration module in any wave, despite being older than 15 by the most recent wave.

actual date of conception or date of birth, I simply count backwards, in three month increments, from an individual’s month of birth and average over each three month period.

A.3 Construction of Individual-Level Variables

Both my reduced form and structural analyses only require one cross-section of data. I therefore merge all waves of data for each individual and extract the relevant information from the relevant waves. For variables that should be consistent across waves (gender, birth year), I use data from all available waves and resolve inconsistencies by prioritizing values that are consistent across at least two waves. For other variables, I pick one wave for each individual. In particular, I use the Raven’s test score from the first wave the individual took the test (to avoid capturing any learning effects). For all other reduced-form outcome variables (schooling and income), I use the variable from the most recent survey wave available, with one important exception.

In 2009, the share of individuals in the MxFLS who report being a technician in their main job rises by 11 percentage points, from 1% in 2005 and 2% in 2002. This dramatic increase does not show up when comparing the share of technicians in the 2000 and 2010 censuses and therefore seems to be driven by a change in coding, rather than an actual increase in the share of individuals in this occupation. In order to avoid using variables that are coded differently across survey waves, I ignore all work-related variables for individuals who report working as a technician in 2009. This does not mean that I drop these individuals from the sample – this simply means that their work-related variables (income and occupation category) are taken from the most recent available wave prior to 2009.

To represent occupation types, the MxFLS uses a different categorization system from the ISCO codes that are used in the census (summarized in Appendix Table A1). Appendix Table A11 lists how I map these Mexican Classification of Occupations (CMO) codes to the white-collar and blue-collar categories. This mapping was fairly straightforward, based simply on comparing CMO descriptions to ISCO descriptions (and then using the Vogl (2014) classification to categorize into white-collar and blue-collar).

Table A11. Mexican Classification of Occupation (CMO) Codes

CMO Code and Description
White-Collar ("Brains")
11 Professionals
12 Technicians
13 Education Workers
14 Arts, sports, performance, and sports workers
21 Employees and directors of the public, private, and social sectors
61 Department chiefs, coordinators and supervisors of the administrative activities and services
62 Workers in the support of the administrative activities
Blue-Collar ("Brawn")
41 Agricultural, cattle activities, forestering, hunting, and fishing workers
51 Chiefs, supervisors, and other control workers in craft and industrial manufacture and in maintenance and repairing activities
52 Craftsmen and manufacturers in the transformation industry and workers of maintenance and repairing activities
53 Operators of fixed machinery of continuous movement and equipment in the process of industrial production
54 Assistants, laborers, and similar in the process of artisan and industrial manufacture and in repairing and maintenance activities
55 Conductors and assistants of conductors of movable machinery and means of transport
71 Retailers, employees in commerce, and sales agents
72 Street sales and services workers
81 Workers in personal establishments
82 Workers in domestic services

Notes: CMO codes are first matched to ISCO codes. Brain and brawn categorizations from Vogl (2014).

A.4 Predicting White Collar Proportions

Conceptually, p_g represents the perceived likelihood of an individual working a white collar job. If parents or individuals use current conditions to inform this expectation, then this probability should vary with the local labor market opportunities at times when children are making important schooling transitions. I aim to match individuals to the relevant labor market variables in the year they turn 12 years old, in the municipality in which they are living at that age. For the vast majority of the sample, I know exactly where they are living at age 12. If individuals report that they are currently living in the same municipality in which they were living at age 12, I use

their current residence. If individuals report that they were living in their municipality of birth when they were 12 years old, I use their municipality of birth. For the remainder of individuals, who make up less than 10% of the sample, I also assign them to their municipality of birth, acknowledging that there will be some measurement error, because municipality of residence at age 12 is a restricted-use variable.

For the actual data on local labor market conditions, I use the 1990, 2000, and 2010 Mexican censuses, which span the decades during which individuals in my sample transitioned from elementary to junior high. I use the provided International Standard Classification of Occupations (ISCO) codes to categorize individuals as working in white-collar or blue-collar jobs using the same classification as in Vogl (2014). I then calculate the proportion of men and women in white-collar jobs, separately for each commuting zone. Following Atkin (2016), I use commuting zones instead of individual municipalities because these better represent local labor markets. For instance, large metropolitan areas are often composed of many municipalities, with individuals often working and residing in different ones. I combine all municipalities in the same Zona Metropolitan (according to the 2000 INEGI classification) into a single commuting zone and also combine municipalities where over 10% of the working population in one reports commuting to another for work (according to the more detailed version of the 2000 census, obtained from INEGI).

The census data provides me with gender-specific white-collar proportions for each commuting zone for 1990, 2000 and 2010. However, my empirical test requires knowledge about labor market conditions for each birth cohort at age 12. I use two different methods to assign values to the individuals who turn twelve during intercensal years. The simplest method involves assigning individuals with the relevant value from the census just prior to the decade in which they turned 12. This would be the 2000 census for those who were born in the years 1988 to 1997, for example. These results are reported in Table A8.

The results discussed in the main body of the paper (Table 3) combine census data with national-level growth rates from ENIGH to predict intercensal years. For each year y , I calculate

national-level growth rates of six major industries⁴¹ (subscripted by j) relative to the most recent census decade d . I denote these growth rates g_{yjd} . From the census, in addition to the gender-specific proportions of white collar jobs in each decade (p_{gd}), I also calculate the gender-specific share of brain-intensive jobs in each industry: s_{gjd} . My predicted proportion, \hat{p}_{gyjd} , is simply:

$$\hat{p}_{gyjd} = p_{gd} + \sum_{j=1}^5 s_{gjd} g_{yjd}. \quad (16)$$

⁴¹The six broad industry categories I use are: (1) agriculture, (2) oil, natural gas, and construction, (3) education, health, and government, (4) manufacturing, (5) service and hospitality, and (6) trade.

B Structural Estimation Details

In this section, I outline the decision rules, transition probabilities, and likelihood function used to estimate the structural model in section 6. The last sub-section discusses model fit.

B.1 Agent's Decision Rules

To simplify future notation, I collect all the non-stochastic terms of each state's net rewards into one parameter, so that they can now be written

$$\begin{aligned}
Y(hw) - c(hw) &= \gamma_{hw} + \sum_{t=0}^{20} \delta^t \epsilon(hw, t) - (\eta(hw) - \eta(hb)) \\
Y(lw) - c(lw) &= \gamma_{lw} + \sum_{t=0}^{20} \delta^t \epsilon(lw, t) - (\eta(lw) - \eta(lb)) \\
Y(hb) - c(hb) &= \gamma_{hb} + \sum_{t=0}^{20} \delta^t \epsilon(hb, t) \\
Y(lb) - c(lb) &= \gamma_{lb} + \sum_{t=0}^{20} \delta^t \epsilon(lb, t) \\
Y(hn) - c(hn) &= \gamma_{hn} - (\eta(hn) - \eta(hb)) \\
Y(ln) - c(ln) &= \gamma_{ln} - (\eta(ln) - \eta(lb)) \\
Y(h) - c(h) &= \gamma_h - (\eta(h) - \eta(l)),
\end{aligned}$$

where

$$\begin{aligned}
\gamma_{hw} &= \frac{1 - \delta^{21}}{1 - \delta} \left[\beta_{w0} + \beta_{w1} + (\beta_{w2} + \beta_{w3})\theta + \sum_{j=10}^{k_w} \beta_{wj} X_{wj} \right] + \frac{1 - \delta^5}{1 - \delta} \sum_{j=4}^6 (\beta_{wj} + \beta_{w(j+3)}) \delta^{5(j-3)+1} \\
&\quad - \left(c_{hw} + \alpha_{hw}\theta + \sum_{j=1}^{q_{hw}} \delta_{hwj} Q_{hwj} \right) \\
\gamma_{lw} &= \frac{1 - \delta^{21}}{1 - \delta} \left[\beta_{w0} + \beta_{w2}\theta + \sum_{j=10}^{k_w} \beta_{wj} X_{wj} \right] + \frac{1 - \delta^5}{1 - \delta} \sum_{j=4}^6 \beta_{wj} \delta^{5(j-7)+1} \\
&\quad - \left(c_{lw} + \alpha_{lw}\theta + \sum_{j=1}^{q_{lw}} \delta_{lwj} Q_{lwj} \right)
\end{aligned}$$

$$\begin{aligned}
\gamma_{hb} &= \frac{1 - \delta^{21}}{1 - \delta} \left[\beta_{b0} + \beta_{b1} + (\beta_{b2} + \beta_{b3})\theta + \sum_{j=10}^{k_b} \beta_{bj} X_{bj} \right] + \frac{1 - \delta^5}{1 - \delta} \sum_{j=4}^6 (\beta_{bj} + \beta_{b(j+3)}) \delta^{5(j-3)+1} \\
\gamma_{lb} &= \frac{1 - \delta^{21}}{1 - \delta} \left[\beta_{b0} + \beta_{b2}\theta + \sum_{j=10}^{k_b} \beta_{bj} X_{bj} \right] + \frac{1 - \delta^5}{1 - \delta} \sum_{j=4}^6 \beta_{bj} \delta^{5(j-7)+1} \\
\gamma_{hn} &= -(c_{hn} + \beta_{hn}\theta + \sum_{j=1}^{q_{hn}} \delta_{hnj} Q_{hnj}) \\
\gamma_{ln} &= -(c_{ln} + \beta_{ln}\theta + \sum_{j=1}^{q_{ln}} \delta_{lnj} Q_{lnj}) \\
\gamma_h &= -(c_h + \alpha_h\theta + \sum_{j=1}^{q_h} \delta_{Hj} Q_{hj}).
\end{aligned}$$

The agent's value function at state s is

$$V(s) = Y(s) + \underbrace{\delta \max_{s' \in S^f(s)} \{-c(s') + \delta \mathbb{E}[V(s')|I(s)]\}}_{CV(s) = \text{Continuation Value}}, \quad (17)$$

where $I(s)$ denotes the agent's information set at state s , and $S^f(s)$ denotes the set of feasible states at s . The second term on the right-hand side is the continuation value of state s , $CV(s)$. This value function determines the agent's decision in each state. I solve for the optimal decision rules at each node, starting with the terminal nodes.

An agent in $s = h$ chooses hw if, given her high school degree, the expected lifetime net rewards from a white collar job exceed expected lifetime net rewards from a blue collar job and the expected net rewards from not working, i.e. if

$$\gamma_{hw} - \eta(hw) > \gamma_{hb} - \eta(hb) \text{ and} \quad (18)$$

$$\gamma_{hw} - \eta(hw) > \gamma_{hn} - \eta(hn). \quad (19)$$

An agent chooses hb if

$$\gamma_{hb} - \eta(hb) \geq \gamma_{hw} - \eta(hw) \text{ and} \quad (20)$$

$$\gamma_{hb} - \eta(hb) > \gamma_{hn} - \eta(hn), \quad (21)$$

and chooses hn otherwise.

Similarly, an agent in $s = l$ chooses lw if

$$\gamma_{lw} - \eta(lw) > \gamma_{lb} - \eta(lb) \text{ and} \quad (22)$$

$$\gamma_{lw} - \eta(lw) > \gamma_{ln} - \eta(ln), \quad (23)$$

chooses lb if

$$\gamma_{lb} - \eta(lb) \geq \gamma_{lw} - \eta(lw) \text{ and} \quad (24)$$

$$\gamma_{lb} - \eta(lb) > \gamma_{ln} - \eta(ln), \quad (25)$$

and chooses ln otherwise.

These decision rules also help to solve for the agent's optimal decision in $s = 0$. Here, an agent will choose h if the expected net rewards and continuation value of a high school degree exceeds the expected net rewards plus continuation value of dropping out.

$$\mathbb{E}[Y(h) - c(h) + CV(h)|I(0)] > \mathbb{E}[Y(l) - c(l) + CV(l)|I(0)], \quad (26)$$

where

$$\begin{aligned} \mathbb{E}[Y(h) - c(h) + CV(h)|I(0)] &= \mathbb{E}[Y(h) + CV(h)|I(0)] - c(h) \\ &= \mathbb{E}[CV(h)|I(0)] + \gamma_h - (\eta(h) - \eta(l)) \\ &= \delta \mathbb{E} \left[\max_{s' \in \{hw, hb, hn\}} \{-c(s') + \mathbb{E}[V(s')|I(h)]\} \right] + \gamma_h - (\eta(h) - \eta(l)) \\ &= \delta \rho_{\eta h} \ln(\exp(\frac{\gamma_{hw}}{\rho_{\eta h}}) + \exp(\frac{\gamma_{hb}}{\rho_{\eta h}}) + \exp(\frac{\gamma_{hn}}{\rho_{\eta h}})) + \gamma_h - (\eta(h) - \eta(l)) \end{aligned} \quad (27)$$

where the simplification in the final line is due to the Type 1 extreme value distribution assumption.

Similarly,

$$\begin{aligned}
\mathbb{E}[Y(l) - c(l) + CV(l)|I(0)] &= \mathbb{E}[CV(l)|I(0)] \\
&= \delta\rho_{\eta l} \ln(\exp(\frac{\gamma_{lw}}{\rho_{\eta l}}) + \exp(\frac{\gamma_{lb}}{\rho_{\eta l}}) + \exp(\frac{\gamma_{ln}}{\rho_{\eta l}})). \tag{28}
\end{aligned}$$

Combining equations 26, 27, and 28, we can derive a cutoff rule for the agent's first decision.

She will choose h if

$$\begin{aligned}
&\delta\rho_{\eta h} \ln(\exp(\frac{\gamma_{hw}}{\rho_{\eta h}}) + \exp(\frac{\gamma_{hb}}{\rho_{\eta h}}) + \exp(\frac{\gamma_{hn}}{\rho_{\eta h}})) + \gamma_h - \eta(h) \\
&> \delta\rho_{\eta l} \ln(\exp(\frac{\gamma_{lw}}{\rho_{\eta l}}) + \exp(\frac{\gamma_{lb}}{\rho_{\eta l}}) + \exp(\frac{\gamma_{ln}}{\rho_{\eta l}})) - \eta(l) \tag{29}
\end{aligned}$$

B.2 Transition Probabilities and Likelihood Function

The individual likelihood contribution of a particular agent is the joint probability of observing that agent's schooling choice, occupational choice, and (for working individuals) income that is realized in the data.⁴² Beginning with the choice probabilities, I define for each agent an indicator function $d(s)$ which equals 1 if the agent visits state s , and calculate the conditional probability of visiting each state in $S^v(s)$, the set of visited states. Collecting all of the observed characteristics in D and structural parameters in a vector ψ , we can use the above cutoff rules to write out these transition probabilities as follows:

$$\begin{aligned}
\Pr(d(h) = 1|D, \psi, \theta) &= \left[1 + \exp \left(-\frac{1}{\rho_{\eta 0}} \left(\delta\rho_{\eta h} \ln(e^{\frac{\gamma_{hw}}{\rho_{\eta h}}} + e^{\frac{\gamma_{hb}}{\rho_{\eta h}}} + e^{\frac{\gamma_{hn}}{\rho_{\eta h}}}) + \gamma_h \right. \right. \right. \\
&\quad \left. \left. \left. - \delta\rho_{\eta l} \ln(e^{\frac{\gamma_{lw}}{\rho_{\eta l}}} + e^{\frac{\gamma_{lb}}{\rho_{\eta l}}} + e^{\frac{\gamma_{ln}}{\rho_{\eta l}}}) \right) \right) \right]^{-1} \\
\Pr(d(l) = 1|D, \psi, \theta) &= \left[1 + \exp \left(\frac{1}{\rho_{\eta 0}} \left(\delta\rho_{\eta h} \ln(e^{\frac{\gamma_{hw}}{\rho_{\eta h}}} + e^{\frac{\gamma_{hb}}{\rho_{\eta h}}} + e^{\frac{\gamma_{hn}}{\rho_{\eta h}}}) + \gamma_h \right. \right. \right. \\
&\quad \left. \left. \left. - \delta\rho_{\eta l} \ln(e^{\frac{\gamma_{lw}}{\rho_{\eta l}}} + e^{\frac{\gamma_{lb}}{\rho_{\eta l}}} + e^{\frac{\gamma_{ln}}{\rho_{\eta l}}}) \right) \right) \right]^{-1}
\end{aligned}$$

⁴²Although the agent observes the idiosyncratic shocks $\eta(s')$ (before deciding on their next state) and $\epsilon(s)$ (after making their decision), the researcher does not.

$$\begin{aligned}
\Pr(d(hw) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{hw}}{\rho_{\eta h}}\right) \left[\exp\left(\frac{\gamma_{hw}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hb}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hn}}{\rho_{\eta h}}\right) \right]^{-1} \\
\Pr(d(hb) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{hb}}{\rho_{\eta h}}\right) \left[\exp\left(\frac{\gamma_{hw}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hb}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hn}}{\rho_{\eta h}}\right) \right]^{-1} \\
\Pr(d(hn) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{hn}}{\rho_{\eta h}}\right) \left[\exp\left(\frac{\gamma_{hw}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hb}}{\rho_{\eta h}}\right) + \exp\left(\frac{\gamma_{hn}}{\rho_{\eta h}}\right) \right]^{-1} \\
\Pr(d(lw) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{lw}}{\rho_{\eta l}}\right) \left[\exp\left(\frac{\gamma_{lw}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{lb}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{ln}}{\rho_{\eta l}}\right) \right]^{-1} \\
\Pr(d(lb) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{lb}}{\rho_{\eta l}}\right) \left[\exp\left(\frac{\gamma_{lw}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{lb}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{ln}}{\rho_{\eta l}}\right) \right]^{-1} \\
\Pr(d(ln) = 1|D, \psi, \theta) &= \exp\left(\frac{\gamma_{ln}}{\rho_{\eta l}}\right) \left[\exp\left(\frac{\gamma_{lw}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{lb}}{\rho_{\eta l}}\right) + \exp\left(\frac{\gamma_{ln}}{\rho_{\eta l}}\right) \right]^{-1}
\end{aligned}$$

These transition probabilities are combined with the per-period wage functions (of which we only observe one per working individual) to construct the individual likelihood function for observation i . Recall that t_i is the age of the individual (in the year their income is observed) minus 30. Then, i 's contribution to the likelihood function is:

$$\int_{-\infty}^{\infty} \left[\prod_{s \in S} \left\{ \Pr(d_i(s) = 1|D_i, \theta_i(\nu); \psi) \prod_{t=0}^{20} f(Y_i(s, t)|D, \theta_i(\nu); \psi)^{1(t_i=t)} \right\}^{1(s \in S_i^v)} \right] dF_{\nu}(\nu), \quad (30)$$

where

$$\begin{aligned}
f(Y(hw, t)|D, \theta; \psi) &= \phi \left(\beta_{w0} + \beta_{w1} + (\beta_{w2} + \beta_{w3})\theta + \sum_{j=4}^6 (\beta_{wj} + \beta_{w(j+3)})A_{j-3}(t) - Y(hw, t) \right) \\
f(Y(lw, t)|D, \theta; \psi) &= \phi \left(\beta_{w0} + \beta_{w2}\theta + \sum_{j=4}^6 \beta_{wj}A_{j-7}(t) - Y(lw, t) \right) \\
f(Y(hb, t)|D, \theta; \psi) &= \phi \left(\beta_{b0} + \beta_{b1} + (\beta_{b2} + \beta_{b3})\theta + \sum_{j=4}^6 (\beta_{bj} + \beta_{b(j+3)})A_{j-3}(t) - Y(hb, t) \right) \\
f(Y(lb, t)|D, \theta; \psi) &= \phi \left(\beta_{b0} + \beta_{b2}\theta + \sum_{j=4}^6 \beta_{bj}A_{j-7}(t) - Y(lb, t) \right) \\
f(Y(s, t)|D, \theta; \psi) &= 1 \quad \forall s \in \{h, l, hn, ln\}.
\end{aligned}$$

θ is equal to the individual's standardized Raven's test score plus a standard normal error

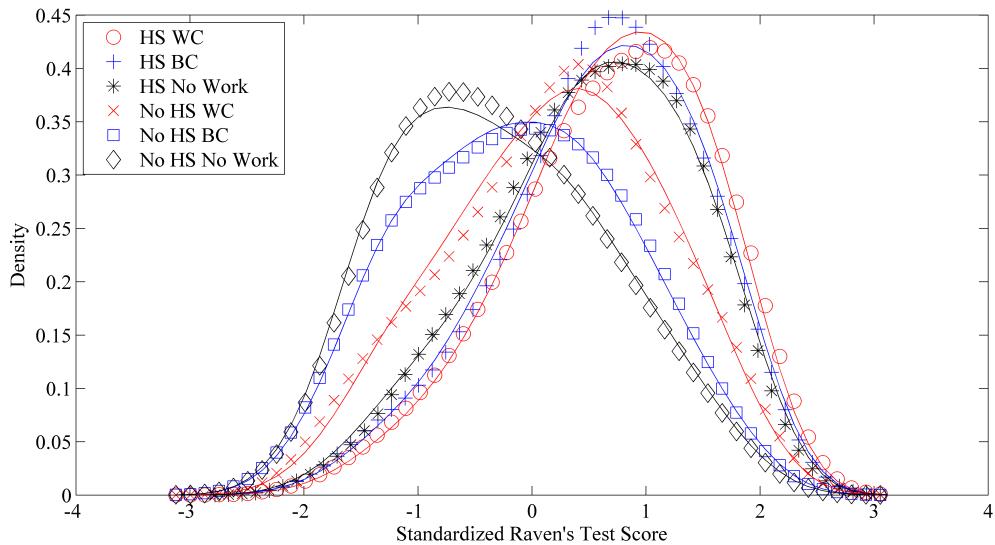
term ν . The discount factor is set to 0.04. I use maximum likelihood to estimate the structural parameters ψ , using 100 simulations to calculate the integral over ν for each individual. The wage parameters are reported in Table 7, and the cost parameters are reported in Table A12.

B.3 Model Fit

To evaluate the fit of the model, I simulate the schooling and occupational decisions of each individual using the estimated model parameters and 50 draws for the error terms. Because the main focus of this paper is on how ability affects the joint determination of schooling and occupation choice, I investigate how well the model simulations recreate the relationship between ability, schooling, and occupational choice.

In Figure A3, I plot the distribution of Raven's test scores separately for each schooling-occupation group in the actual data, using solid lines. On the same graph (using the discrete markers), I also plot the Raven's test score distributions for each schooling-occupation category, as predicted by model simulations. The distributions match up very closely, and Chi-squared tests cannot reject the equality of any pair of actual and simulated distributions.

Figure A3. Raven's Score Distributions, by Actual and Simulated Schooling-Occupation Choice



Notes: The solid lines represent the distribution of Raven's test scores across the individuals that fall into each schooling-test score combination in the actual data. The discrete markers represent the distribution of Raven's test scores across the individuals that fall into each schooling-occupation combination according to model simulations, which use model parameters to predict decisions.

Table A12. Cost Parameter Estimates

Panel A: Occupational Choice

	Estimate	Standard Error
White Collar - Blue Collar		
HS: Constant	-1.389	(5.172)
HS: Ability	0.163	(1.679)
HS: White Collar Proportion	2.960	(3.751)
HS: Male	6.082	(4.528)
HS: Urban	-6.030	(2.346)**
No HS: Constant	19.074	(2.734)***
No HS: Ability	-3.509	(1.220)***
No HS: White Collar Proportion	-8.158	(2.982)***
No HS: Male	-4.493	(1.270)***
No HS: Urban	-4.143	(1.486)***
No Work - Blue Collar		
HS: Constant	-156.885	(4.230)***
HS: Ability	0.774	(1.536)
HS: Unemployment	-1.497	(4.954)
HS: Male	27.094	(14.046)*
HS: Urban	-4.673	(1.803)***
No HS: Constant	-144.941	(1.006)***
No HS: Ability	2.035	(0.342)***
No HS: Unemployment	-2.645	(1.234)**
No HS: Male	3.453	(3.556)
No HS: Urban	-4.175	(0.820)***
HS: Scale Parameter	9.437	(3.538)***
No HS: Scale Parameter	2.915	(0.761)***

Panel B: Schooling Choice

	Estimate	Standard Error
Constant		
Ability	27.236	(10.574)***
Male	-3.363	(2.487)
Parental Education	-9.001	(4.176)**
Youth Employment	-0.850	(0.690)
Scale Parameter	3.050	(2.370)
Scale Parameter	3.386	(2.390)

Notes:

Standard errors are calculated analytically using the information matrix.

* p < 0.1 ** p < 0.05 *** p < 0.01