THE EFFECT OF POLLUTION ON LABOR SUPPLY:
EVIDENCE FROM A NATURAL EXPERIMENT IN MEXICO CITY

Rema Hanna
Paulina Oliva

The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City
Rema Hanna and Paulina Oliva
NBER Working Paper No. 17302
August 2011
JEL No. O0,Q0,Q5,Q53

ABSTRACT

Moderate effects of pollution on health may exert an important influence on labor market decisions. We exploit exogenous variation in pollution due to the closure of a large refinery in Mexico City to understand how pollution impacts labor supply. The closure led to an 8 percent decline in pollution in the surrounding neighborhoods. We find that a one percent increase in sulfur dioxide results in a 0.61 percent decrease in the hours worked. The effects do not appear to be driven by labor demand shocks nor differential migration as a result of the closure in the areas located near the refinery.

Rema Hanna
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
rema_hanna@ksg.harvard.edu

Paulina Oliva
UC, Santa Barbara
oliva@econ.ucsb.edu

The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City*

Rema Hanna

Harvard Kennedy School, NBER, BREAD, and IZA

Paulina Oliva

University of California Santa Barbara

August 2, 2011

Abstract

Moderate effects of pollution on health may exert an important influence on labor market decisions. We exploit exogenous variation in pollution due to the closure of a large refinery in Mexico City to understand how pollution impacts labor supply. The closure led to an 8 percent decline in pollution in the surrounding neighborhoods. We find that a one percent increase in sulfur dioxide results in a 0.61 percent decrease in the hours worked. The effects do not appear to be driven by labor demand shocks nor differential migration as a result of the closure in the areas located near the refinery.

1   Introduction

In this paper, we estimate the relationship between pollution and labor supply in Mexico City. High

levels of pollution may cause temporary illness, which in turn may cause lost working hours. Under-

standing the relationship between pollution and lost working hours is vital for assessing the benefits of more stringent environmental regulation. This question is of particular importance for developing countries, which are often hesitant to enter international climate change treaties (or are lax about enforcing them) due to fears that tougher regulations may lead to reductions in productivity and business growth. Furthermore, an understanding of the effects of pollution reduction on hours worked is necessary to design optimal pollution taxation policies (Schwartz and Repetto 2000, Williams 2003).

The effect of pollution on work hours is theoretically ambiguous. On one hand, if pollution damages the health of an individual and his or her dependents, reductions in pollution will decrease the disutility from work causing work hours to rise. In other words, with better air quality, individuals may be less likely to miss work because either they or their children are home sick.[1] However, there are also many reasons why reductions in pollution would not increase work hours. First, it is possible that the effect of pollution on adult health is not large enough to interfere with attendance at work, or that individuals already practice mitigating behaviors on high pollution days (such as staying indoors) to minimize illness.[2] Second, pollution may additionally affect the utility derived from the consumption of leisure and health-related goods. If individuals enjoy leisure more due to better health or individuals substitute away from consumption of health-related goods when pollution is reduced, they might adjust their hours of work downwards. If these effects are large enough, the overall effect of pollution on work hours may even be negative. Finally, improvements in worker productivity as a result of reduced pollution may increase wage earning potential, which would have ambiguous effects on the number of hours worked.[3] In light of these opposing theoretical impacts, the relationship between pollution and hours worked is ultimately an empirical question.

Our study estimates the short-run elasticity of work hours with respect to pollution.[4] Estimating this effect is challenging due to the presence of confounding factors: for example, a decline in business

---

[1] Currie, Hanushek, Kahn, Neidell, and Rivkin (2009), Ransom and Pope (1992), Gilliland et al. (2001), and Park et al. (2002) have found significant effects of pollution on student absenteeism. This may translate to higher absenteeism among workers that are responsible for child care.

[2] We will discuss the evidence on the health effects of refineries in Section 3.

[3] Crocker and Horst (1981) and Graff and Neidell (2011) both show empirically that changes in pollution are linked to increases in productivity for farm workers. Given our empirical strategy, we cannot isolate what fraction, if any, of the observed changes in hours worked can be explained by changes in productivity. Nonetheless, we will discuss the theoretical channel through which productivity changes may affect hours worked in Section 2. We also test for changes in observed wages in Section 5.

[4] Pollution can also have long-run health impacts (e.g. by increasing the prevalence of lung or digestive diseases), which in turn reduce employment and hours worked. In this paper, we will be unable to look at long-run impacts, but we hope to pursue this in future work.

activity may affect both pollution levels and employment patterns.[5] To overcome this problem, we exploit exogenous variation in pollution that resulted from the closure of a large oil refinery in the Mexico City Metropolitan Area (MCMA) in March of 1991. The closure was mandated to reduce the high levels of pollution experienced in MCMA during the 1980s. It successfully reduced pollution: neighborhoods located within a five-kilometer radius to the refinery experienced, on average, an eight percent reduction in sulfur dioxide (SO_2) relative to other neighborhoods. Using a fixed effects strategy, we exploit these changes to understand whether pollution levels affected labor supply. Specifically, we compare the changes in work outcomes for individuals who lived in neighborhoods located near the refinery (that experienced a fall in pollution) with those who lived in neighborhoods far from the refinery (that did not experience a fall).

We find robust evidence that the refinery closure lead to increased hours of work. After the closure, the areas within 5 kilometers of the refinery experienced about a 5 percent increase in hours worked, controlling for week fixed effects and neighborhood fixed effects. Using the refinery closure as an instrument for pollution, we find that a one percent increase in SO_2 results in about a 0.43 to 0.67 percent decrease in hours worked the following week. The results are robust to the inclusion of demographic controls and neighborhood specific year trends.

Our analysis rules out two sources of potential bias in these estimates. First, many pollution regulations have a direct effect on business conditions, and thus the labor market.[6] As such, it is difficult to disentangle the effect of reduced pollution on labor market activities from the direct effect of the regulations on the labor market. The refinery closure presents a case where we can rule out the direct effect of the policy on the labor market. Theoretically, it is possible that the closure could have affected the labor market independently of health if there were a reduction in labor demand both from the refinery itself and businesses located in the surrounding neighborhoods. If labor supply is upward sloping, this would cause wages to fall and unemployment to increase near the refinery, leading us to *underestimate* the elasticity.[7] This is unlikely to have occurred, as we do not observe differential changes in wage in neighborhoods located near the refinery after the closure, and we find

---

[5]Several studies (such as Pönka, 1990) have looked at the association between pollution and worker absenteeism, and have found that SO_2 levels are associated with illness-related absenteeism.

[6]Recent papers, such as Walker (2011), study the effect of environmental regulations that impose additional costs on polluting firms, which would have direct implications for the labor market.

[7]While less likely, one can also imagine cases where labor supply is downward sloping, which would cause our estimate to be upwardly biased. We discuss this further in Section 4.

a small increase to no change in unemployment near the refinery after the closure. Nonetheless, we estimate an additional empirical model that exploits the variation in pollution levels produced by differences in altitude and wind patterns within areas that are located at the same distance from the refinery. Since areas at a similar distance from the refinery would experience similar labor market shocks related to the refinery closure, the pollution effects that we find within these areas are not susceptible to bias from a labor market mechanism. Using this triple-difference approach, we find a strong relationship between pollution levels and hours worked (similar in magnitude and significance to the primary results), despite the fact that this model is very restrictive. Therefore, we can rule out that the observed results are due to an independent effect of the refinery closure on the labor market.

Second, changes in air quality may result in selective sorting. More specifically, wealthier or more educated individuals may move into the newly clean neighborhoods and these individuals may be more likely to be employed or work long hours. This would cause us to overestimate the effect. To reduce the probability of bias, we restrict our analysis to the five years after the closure. In addition, we test for sorting using information on migration and demographic characteristics. We do not observe increased migration rates near the refinery after the closure, nor do we observe differences in demographic characteristics in these areas after the closure. The lack of change in migration and demographic characteristics increases our confidence that the labor supply outcomes we observe are not driven by sorting.

Finally, we test for the heterogeneity of the treatment results. We find suggestive evidence that parents with children under five have a larger labor supply response to pollution, but a limited sample size precludes conclusive evidence. While we find no evidence that blue and white collar workers respond differently to changing pollution, most of our estimated effects appear driven by workers on a fixed salary, who are likely to have the smallest opportunity cost of work absenteeism.

Our findings support the conclusion that the economic gains from increased worker attendance helped to offset the costs of the refinery closure. Our preferred estimates imply that the refinery closure led to about a 5-percent increase in hours worked for those who lived near the refinery. Given an average annual wage of 13,700 1993 Mexican Pesos (USD 3,600), this translates to a 756 Peso (USD

198) gain in work over the course of a year.[8] This substantial increase demonstrates the importance of accounting for work effort gains when calculating the economic benefits of tougher environmental regulation.

The paper proceeds as follows: Section 2 provides a conceptual framework for understanding how changes in pollution may affect work. Section 3 provides a background on the refinery closure, discusses the data that we compiled for the study, and provides a brief overview of refinery emissions on health. In Section 4, we describe our empirical strategy. We present our findings in Section 5, and then discuss their policy implications in Section 6. Section 7 concludes.

2 Conceptual Framework

This paper seeks to document the importance of non-mortality costs of air pollution by providing a credible estimate of how pollution levels affect labor supply decisions. In this section, we outline a partial equilibrium model to explore the underlying mechanisms that drive the relationship between pollution and hours worked.

2.1 Labor Supply and Air Quality

We consider a partial equilibrium framework where individuals maximize utility with respect to consumption, c, and hours worked, e. We assume utility is given by $u = u(c, e; \alpha)$ and is separable in c and e. Note that e is a "bad," so $u_e < 0$, while $u_c > 0$. We assume that the utility function is concave, i.e. that $u_{cc} < 0$ and $u_{ee} < 0$.

We model air quality, α, as an argument of the utility function that affects the marginal utility of both consumption and hours worked. We focus on the partial effects of air quality on hours worked, and therefore we assume that exposure to air quality is exogenous.[9] Specifically, we hypothesize that

---

[8]This number corresponds to the national average yearly earnings from the National Accounting System (SCN).

[9]Relaxing this assumption would affect the observed results in the following manner: Suppose that there are two neighborhoods, and one of the neighborhoods experiences a permanent reduction in pollution. The area that experienced the air quality shock could attract individuals whose marginal utility for clean air is higher than the marginal utility of the previous residents of the affected area. The resulting observed change in hours worked depends on whether the newcomers have a high marginal disutility of work with respect to air quality or whether they have a high marginal value for consumption given a certain level of air quality. The newcomers could experience stronger changes in hours worked than the leavers if the strong preferences for air quality are predominantly linked to a reduction in the disutility of work. If newcomers have higher consumption value for air quality, but do not experience additional

better air quality increases a worker's health as well as the health of his or her dependents, making it less costly for the worker to devote time to work. In other words, better air quality lowers the disutility of work: $u_{e\alpha} > 0$.[10] The relationship between air quality and consumption is ambiguous. For example, if improved air quality increased the utility derived from additional consumption (i.e. better air quality may improve the utility derived from amusement park visits or shopping outdoors), then consumption and air quality would be complements: $u_{c\alpha} > 0$. However, better air quality could also reduce the marginal utility derived from consumption ($u_{c\alpha} < 0$). This would be the case if, for example, asthma medication or gym memberships were substitutes for clean air.

We can rewrite the optimization problem faced by the individual using the indirect utility function:

$$\max v(e) = \lambda(\alpha) \cdot we - g(e; \alpha) \tag{1}$$

where w is the wage rate at time t. Note that for tractability, we assume the individuals are wage-takers, and the wage is unaffected by pollution; we discuss the implications of these assumptions in Section 2.2. The marginal utility of lifetime income along the optimal path is represented by $\lambda(\alpha)$; this term captures changes in the value of income stemming from individual reoptimization decisions on both consumption and hours worked. Finally, $g(e; \alpha)$ can be interpreted as the disutility of hours worked given a certain level of air quality:

$$g(e; \alpha) = -\int_0^e u_e(x; \alpha) dx \tag{2}$$

Note that the assumption $u_{e\alpha} > 0$ translates into $g_{e\alpha} < 0$. Similarly, the concavity assumption for the utility translates into a convexity assumption for $g(e; \alpha)$, since the marginal disutility of work increases with hours ($g_{ee} > 0$).

Optimal work hours for a given level of air quality, α, are the solution to the first order condition

---

reductions in the disutility of work, the change in hours of work could be smaller for the newcomers. We revisit this issue in the empirical section, and test for the presence of differential migration into the areas surrounding the refinery after the closure.

[10]You could also argue the opposite assumption, that better air quality produces a net increase in the disutility of work ($u_{e\alpha} < 0$). This may result, for example, if individuals who work in offices would rather be outside in a low pollution environment. This is unlikely to happen in our empirical framework since we measure the effect of pollution on the hours of work in the following week, and only contemporaneous air quality makes leisure more attractive. The model predictions under $u_{e\alpha} < 0$ are further discussed in Appendix 1. As we discuss in Section 5, our empirical results are largely consistent with the predictions yielded by $u_{e\alpha} > 0$.

for (1):

$$g_e(e; \alpha) = \lambda(\alpha)w$$

Given additive separability of hours of work and consumption, the change in hours worked as air quality increases is given by (see the Appendix for a formal derivation):

$$\frac{de}{d\alpha} = \frac{-g_{e\alpha} + \frac{\partial\lambda}{\partial\alpha}w}{g_{ee}} \tag{3}$$

The change in hours of work with respect to pollution depends on the sign of the two terms in the numerator of (3). The first term captures the "substitution effect" stemming from a reduction in the disutility of work. This term will be unambiguously positive since the disutility of work is decreasing in air quality.

The second term in the numerator of (3) can be interpreted as an "income effect." On net, we expect this effect to be negative. Intuitively, this term incorporates several effects (see the Appendix for a formal derivation of $\frac{\partial\lambda}{\partial\alpha}$). First, as income rises with more hours worked, we would expect the disutility of work to increase relative to the utility of consumption. Second, as air quality improves, the marginal utility of income $(\lambda(\alpha))$ will change as a result of a new optimal level of consumption. Better air quality may reduce the marginal utility gain from consumption if consumption is predominantly a substitute for air quality (individuals would work less, as consumption goods are less valuable with better air quality). If the income effect is large enough to dominate the substitution effect, the overall effect of pollution on work hours may even be negative.

It is important to note, however, that there are also cases where the income effect may be positive, thereby amplifying the substitution effect. If air quality is predominantly a complement of consumption, the marginal utility of income $(\lambda(\alpha))$ could increase with improved air quality, thus increasing work hours (intuitively, individuals would work more in order to purchase consumption goods, as the value of the goods rises with improved air quality). A positive income effect would result in an unambiguous increase in hours worked.

In sum, the partial equilibrium model predicts that high air quality should result in an increase in hours worked, unless a negative income effect dominates.[11] Next, we relax the simplifying assumption

---

[11]Note that, in a general equilibrium framework, aggregate labor supply may be mechanically affected as individuals

above that pollution does not affect worker productivity, and explore how changes in productivity may affect hours worked.

2.2 Productivity effects

Worker productivity may improve as air quality improves. If wages reflect the marginal productivity of hours worked, higher work productivity would result in higher wages. Assume wages depend on air quality, α, through the following function:

$$w(\alpha) = f_1(e \cdot \phi(\alpha), K)\phi(\alpha) \tag{4}$$

where $f_1(\cdot, K)$ is the marginal productivity of labor efficiency units for given capital K, and $\phi(\alpha)$ is the number of labor efficiency units per hour of work. Assume further that air quality increases the productivity of hours worked, which yields $\frac{\partial}{\partial \alpha}\phi(\alpha) > 0$.[12] For simplicity, assume homogeneity of degree one in the production function. Then, in equilibrium, $f_1(e \cdot \phi(\alpha), K) = s_L$ and

$$\frac{\partial w(\alpha)}{\partial \alpha} = s_L \frac{\partial}{\partial \alpha}\phi(\alpha) > 0 \tag{5}$$

Higher wages will have two additional effects on labor supply besides the ones discussed above: a wage substitution effect and a wage income effect. The first effect is unambiguously positive, while the second effect is unambiguously negative (see the Appendix for a mathematical expression and further discussion of these two effects). Unless labor supply is downward sloping (which implies a dominating income effect), the productivity effect will increase labor supply beyond what the Equation 3 describes.

---

supply more hours of work as air quality improves. This shift would lead to a new labor market equilibrium with lower wages and more aggregate hours of work. However, because the equilibrium wage would fall, the increase in hours of work would be smaller than in the case where wages did not change.

[12]Crocker and Horst (1981) and Graff and Neidell (2011) find a negative empirical relationship between air pollution and productivity. Frankenberg et al. (2005) estimate that haze from fires in Indonesia caused older adults to be more likely to report having difficulty carrying a heavy load than older adults in non-haze areas. In addition, there is a large medical literature linking iron supplements, which improve respiratory functioning, to productivity (for example, see Davies et al., 1984; Haas and Brownlie, 2001; Zhu and Haas, 1998; Woodson et al., 1978). In turn, there is evidence from experimental studies that iron supplements can improve work output and productivity (Basta et al., 1979; Thomas et al., 2003; Li et al., 1994).

3 Background and Data

3.1 The Refinery Closure

In this study, we exploit the exogenous variation in pollution that resulted from the closure of a large oil refinery in the Mexico City Metropolitan Area (MCMA) in March of 1991. The Azcapotzalco Refinery opened in 1933 in the Northwestern quadrant of MCMA (see Figure 1). It was taken over by Pemex (originally Petromex) during the oil expropriation in Mexico that started in 1932. In 1946, the refinery's capacity was expanded from 7.5 thousand to 50 thousand barrels per day in response to growing demand, mainly within the MCMA.[13] The expansion in production continued and by 1976 the refinery had doubled its capacity to 105 thousand barrels per day. At its peak, the refinery complex occupied 174 hectares and included fourteen refining facilities, three petrochemical facilities, 218 storage tanks, auxiliary service facilities, a water treatment plant, laboratories, administrative buildings, recreational facilities and a residential area for employees. The refinery represented around 35 percent of the total refining capacity in the country. It produced and distributed final products mainly to the MCMA, such as gas, turbosine, gasoline, kerosene, diesel and other fuels as well as petrochemicals.

When the refinery first opened, the land use in the area was mainly agricultural. Between the 1940's and the 1980's, the MCMA had grown considerably, and by the late 1980s, the refinery was surrounded by residential areas. The closure of the refinery on March 18th of 1991 was mandated to reduce the high levels of pollution experienced in the MCMA during the 1980s.[14] The winter of 1990-1991, just before the refinery was closed, was considered one of "the most toxic in Mexico City history, triggering a 16% to 20% jump in the incidence of respiratory infections, nosebleeds and emphysema."[15] In fact, the Ministry of Environment estimated that, at its peak, the Azcapotzalco

---

[13]Around this date, the refinery was renamed "18 de Marzo" in commemoration of the oil expropriation date.

[14]Complaints regarding refineries emissions still persist in Mexico today. For example, families living near the Francisco I. Madero Refinery in the state of Tamaulipas, which is currently operating, filed a complaint in 2009 due to pungent smells, throat irritation and skin infections caused by refinery fumes (Milenio, 2008). Moreover, when cleanup activities of the Azcapotzalco Refinery began in 2009 (to convert the land to a public space), health concerns were once again raised as the excavation reintroduced buried pollutants back into the air. Complaints were made regarding increased headaches, dizziness, nausea, diarrhea, irritation of the respiratory tract, allergies, asthmatic crisis, etc. (http://www.kaosenlared.net/noticia/mexico-vecinos-azcapotzalco-luchan-contra-danos-salud-contaminacion-ex; accessed on June 6, 2011).

[15]"Mexico City's Menacing Air," Time Magazine, April 1, 1991.

Refinery emitted about 85 thousand tons of air pollutants per year, or about two percent of total air pollution across the MCMA (García Villanueva et al., 2009).

3.2 Sulfur Dioxide and Health

This study will focus on SO_2, a U.S. criteria pollutant, and one of the main gases emitted by oil refineries.[16] Due to the strong link between SO_2 and refinery emissions, the two air monitoring stations closest to the Azcapotzalco Refinery began monitoring SO_2 in 1986. We devote this section to summarizing some of the literature results on the health effects of exposure to SO_2.

Subjects exposed to SO_2 showed decreased lung functioning and increased respiratory symptoms in most controlled human experiments. The effects were stronger for high levels of exposure. It is worth noting, however, that there are also a few studies that find limited to no results.[17] While the trials provide evidence for potential impacts of short-term (5-10 minutes) exposure, individuals may not necessarily experience these adverse effects in the real world if they take precautions on high pollution days, mitigating the effect of pollution on both health and work outcomes.[18] Evidence from the epidemiology literature on the effects of ambient SO_2 on respiratory symptoms of adults are mixed, and therefore, it is not clear from this literature that the SO_2 reductions that we study would be large enough to affect work outcomes (EPA, 2008). However, there is at least some suggestive evidence that the health outcomes attributed to SO_2 may be large enough to induce absenteeism: Peel et al. (2005) and Ito (2007) document an association between short-run exposure to SO_2 and emergency room visits, while Schwartz et al. (1996) documents a marginal association between short-run exposure and increases in hospital admissions due to asthma for those aged 65 and above. Some of the respiratory symptoms associated with SO_2 may not appear immediately after exposure: Peel

---

[16]Bevilacqua and Braglia (2002) measure emissions for seven refineries over a four year period, and find that SO_2 comprises the largest share of emissions. García Villanueva et al. (2009) documents that the emissions from the Azcapotzalco Refinery, which we study in this paper, were composed primarily of sulfur dioxide, carbon monoxide and nitrogen oxides; the refinery also emitted particulate matter (PM), volatile organic compounds (VOC), fluorhidrics acid and sulfuric acid.

[17]Trials have documented decreased lung function (increased respiration rates, decreases in peak flow, bronchoconstriction, and increased airway resistance), particularly for asthmatic subjects with SO_2 exposure (for a summary of findings, see EPA, 2008). Moreover, a series of studies have observed increase respiratory symptoms (coughing, chest tightness, throat irritation) among exercising asthmatics who are exposed to SO_2 (for example, see among others, Linn et al., 1983; Linn et al., 1987).

[18]For example, Neidell (2009) provides evidence that individuals practice avoidance behaviors (e.g. stay indoors) to prevent illness on high pollution days.

et al. (2005) study the lag structure between pollution levels and emergency visits and find that the risk ratios for asthma were the strongest for lag of 5 to 8 days. [19]

In addition, as we discuss in the conceptual framework, adults may work fewer hours if their children are sick due to high levels of pollution. Children and infants are among the most susceptible to air pollution (including SO_2), as lung development continues throughout adolescence and a developing lung is particularly at risk from exposure to toxins (Dietert et al., 2000). Countless papers in epidemiology literature document the association between SO_2 and respiratory symptoms in children (for example, Schwartz et al. (1994) document associations between the incidence of respiratory symptoms and SO_2; Schildcrout et al. (2006) show that the sum of SO_2 over the past three days is related to asthma symptoms; etc). In the economics literature, Lleras-Muney (2005) found that high SO_2 concentrations led to increased hospitalization rates for children, although the effect is not significant. In addition, Jayachandran (2009) documents that smoke from the Indonesian wild-fires, a large source of natural SO_2, had large effects on infant mortality rates.[20]

3.3   Data

Pollution Data

The pollution data are drawn from the Automatic Network of Atmospheric Monitoring (RAMA), which consists of 25 stations placed throughout the MCMA. Hourly sulfur dioxide (SO_2) concentrations are available for 15 stations in the MCMA from 1986 to today. During the sample period, data for other pollutants were not systematically collected.[21]SO_2 is one of the main pollutants generated by the refineries, and, as set out above, is believed to have substantial impacts on respiratory health (see Section 2). However, it is important to note that SO_2 measurements are correlated with other pollutants and will thus proxy for refinery-originated air pollution in our elasticity estimates. In addition to looking at the effect of SO_2 on work outcomes, we will also present reduced-form estimates of the effects of the refinery, which should provide a measure of the "full" policy effect of the refinery

---

[19]SO_2 also contributes to acid rain, but acid rain in itself should not have adverse effects on health.

[20]Other key papers in the economics literature that document significant effects of pollution on infant mortality include Currie and Neidell (2005) and Chay and Greenstone (2003a,b), but these studies do not explicitly measure the effect of SO_2, which is the focus of our study.

[21]During this time, CO and NO_2 were only measured in several stations, but not in enough of them to conduct a similar analysis (e.g. NO_2 was measured only in 5 stations. CO was measured in 13 stations, but not in the two stations closest to the refinery).

closure.

Labor Market Data

We obtained labor market data from the National Employment Panel Survey of Mexico (ENEU). ENEU is a moving quarterly panel in which each individual is surveyed for five consecutive quarters and then exits the sample. We restrict the sample to individuals in the Mexico City Metropolitan Area, which includes two states (Distrito Federal and Estado de Mexico).

We focus on the 1989 to 1993 surveys, the years surrounding the refinery closure.[22] As we are interested in short-run labor market responses, we restrict the sample to individuals in the adult module (aged 12 and over) who are currently in the labor force (54 percent of sample, Appendix Table 1).[23]

We focus on two key work variables. Our primary focus is on the number of hours worked in the week prior to the survey.[24] Individuals who were unemployed are assigned zero hours worked in order to account for selection. We additionally consider the unemployment rate as an outcome for two reasons. First, as we discuss below, we would like to test whether the closure led to lower employment levels near the refinery. Second, lower pollution may lead to health improvements, and thus a higher probability of finding employment in a given week.

It is important to note several key facts regarding the data. First, in each quarter, individuals are randomly assigned a week to be interviewed and most of the labor market questions refer to the week prior to that week. Therefore, we can precisely link the daily measures of pollution to these weekly employment measures. Second, the data contain information on the census block that the individual lives in (in Mexico, this is known as the "basic geo-statistic area"). This allows us

---

[22]Focusing on the years around the closure reduces the probability of selective migration. Moreover, the 1994 data are not fully comparable to the earlier years of data. First, randomly selected census blocks comprise the sampling units of the survey. The rotating nature of the panel demanded a new set of households was selected from within the census blocks every quarter; this was done until all households within each census block were exhausted. This happened for nearly all census blocks between 1993:Q4 and 1994:Q1. After that, new census blocks were selected. Second, the survey questionnaire changed in 1994:Q2, compromising the comparability of key variables over time. Finally, 1994 was marked by political unrest that culminated in one of the largest economic crisis in Mexico.

[23]The labor force variable was constructed as follows: Individuals were assigned to be in the labor force if they worked in the last week, owned a business, owned a business but did not work in the last week, were starting a business, were looking for paid work, were trying to carry out a job on their own, were in the process of starting a business, or were currently trying to sell something to support themselves or their family.

[24]The survey collects the number of hours worked *each day* in the last week. Unfortunately, only the weekly total is available in the public database.

to link each individual's residence to the closest pollution measurement station. The survey lacks data on the census block where the individual works. However, individuals presumably experience substantial morning and evening exposure when they are at home. Moreover, the closure may have affected morning pollution disproportionately. Figure 2 graphs average SO_2 by hour of day, both prior to and after the refinery closure. Prior to the refinery closure, SO_2 peaked between 6 and 11 in the morning. While average pollution fell at all hours of the day after the closure, the morning peak was particularly reduced.[25]

Third, the survey includes rich data on both demographics and work characteristics. This is important for two reasons. First, this allows us to test whether demographic characteristics changed differentially for households in the neighborhood of the refinery after the closure, allowing us to observe whether the closure induced selective migration. Second, it allows us to test for the heterogeneity of the treatment effect by a number of key characteristics, such as the physical nature of the job (i.e. blue or white collar work), the type of payment scheme, and the presence of young dependents.

Finally, the data also include information on the weekly wage. As we discuss in Section 4, we would like to understand if the estimated effect is being driven by changes in wages as a result of the closure. Note that there are several concerns with the wage data. First, missing data appeared to be entered as "0," with many individuals having "0" weekly wage despite having a fixed salary and positive hours worked; these individuals were not included in the wage regressions. Second, for those who were currently unemployed, the questions ask for the worker's typical wage. Since we cannot know what time periods the individuals used to estimate their typical wage, we again did not include them in the wage regressions. Finally, the survey asked for the weekly wage, but did not ask for the specific wage schedule. This means that for workers with variable earnings, the wage per unit of work is unobservable. Thus, we consider the wages only for individuals who have a fixed salary in the wage regression. This restriction means that our wage measure will capture changes in wages due to individuals moving into and out of fixed salary jobs (about 70 percent of the employed), but may miss changes in wages for laborers with variable weekly earnings.

---

[25]We estimated the models in Tables 3 and 4 using SO_2 from 6 AM to 11AM rather than over the full day. The empirical results are similar in both sign, magnitude, and significance.

Wind and Altitude Data

We will estimate models that exploit variation in the change in pollution from the closure that is generated from altitude and wind patterns. Data on the altitude of the census block centroid come from topology maps. Data on wind speed and wind direction, by station, were not collected for the period of our study. Therefore, we use the 10-year averages of wind speed and wind direction information from 1997 to 2006 for 10 of the air monitoring stations used in our study. Wind patterns are very persistent over time: the correlation of wind-speed across years for a particular week-of-the-year and location is around 0.8. Thus, the measures of wind speed and wind direction should be informative of the wind patterns during the period of study.

Matching the Labor and Pollution Data

To estimate the effect of pollution on labor supply, we matched the households in the ENEU survey with pollution readings. To do so, we linked the census blocks in the ENEU to GIS maps and obtained the latitude and longitude of the center of each block. There were 481 census blocks in the ENEU (see Appendix Table 1). However, the earliest available georeferenced maps correspond to census blocks in 2000, and census block identification numbers are only concorded back to 1995. We obtained non-georeferenced maps for the 1990 census blocks from the Instituto Nacional de Estadística y Geografía (INEGI) and manually concorded the data. As shown in Appendix Table 1, we were able to ascertain the locations of 389 census blocks (or 81 percent of them).[26] However, as Columns 1 and 2 of Appendix Table 2 shows, the census blocks that we were unable to match to a location on the maps do not appear to be significantly different than those that we were able to match.

Following Currie and Neidell (2005), we created census block-specific measures of pollution using the inverse of the distance to the nearby stations as weights. There is a trade off between including areas that are far from a station where pollution estimates are imprecise and including more observations in the analysis. We only included census blocks that were within eight kilometers (or five miles) of a station.[27] Therefore, 265 census blocks (or 55 percent of them) were included in the regression

---

[26]There are two reasons why we could not fully concord the identification numbers of the census blocks across years. First, in some cases, the areas were very small and/or the maps were blurry and we could not fully read the identification numbers in the 1990 maps. Second, if a census block was split into two before 1990, we would be unable to locate the original on the 1990 map.

[27]We tested whether the primary results are sensitive to this restriction. We first estimated the main regressions

analysis (Appendix Table 1). Most of the census blocks that were near a station are located within Distrito Federal, which comprises the center part of the city and is also where the refinery is located (190 out of 199 census blocks in Distrito Federal were within eight kilometers of a station). Only 75 out of 190 matched census blocks in Estado de Mexico (the outer areas of the city) were near a station and thus included. Areas in the inner part of the city are quite different than the outer areas, and thus, we expected that the census blocks located near at least one station would be different than those located farther away from the stations. As Appendix Table 2 confirms, the included areas have higher unemployment and lower hours worked than those that are not included, suggesting that the stations may be placed predominantly in poorer neighborhoods.

4   Empirical Strategy

Our objective is to estimate the elasticity of the number of hours (Y_{int}) that individual i worked at time t in census block n with respect to pollution at time $t-1$, $P_{in(t-1)}$, or the parameter β_1 in equation (6):

$$\ln(Y_{int}) = \beta_0 + \beta_1 \ln\left(P_{in(t-1)}\right) + u_{int} \tag{6}$$

where u_{int}, the error term, condenses all unobserved determinants of the hours worked. We focus on pollution levels from the week prior to the interview, as respiratory diseases may take several days

with a shorter range (5km), in which we would have more precise pollution data, but a smaller sample size. Although the sample size falls, most of the main results on hours worked and unemployment stay the same sign, magnitude, and remain significant. However, while the magnitude of the models that include the location-specific trend is similar (if anything, a bit larger in magnitude), we lose power due to the reduced sample size. We then estimated the main regressions where there is larger range (10km), when the pollution data are noisier but we have a larger sample size. In this case, more census blocks from Estado de Mexico are included. We were concerned that these neighborhoods would have differential trends in employment characteristics than the main sample as they were experiencing high levels of growth during this time period. In this case, the basic fixed effects model is of similar magnitude and sign for hours worked, but the estimate is no longer significant at conventional levels. The model including census block-specific year trends (our preferred model) is of the same sign, slightly larger in magnitude and significant at the 5 percent level.

to develop.[28] To consistently estimate β_1, we require the following identification assumption to hold:

$$\mathbb{E}\left(u_{int} \cdot \ln\left(P_{in(t-1)}\right)\right) = 0 \qquad (7)$$

However, there are many reasons to believe that (7) does not hold. For example, both worker hours and pollution levels may vary with business cycles (Chay and Greenstone, 2003). Thus, omitted income variables in u_{int} could generate a spurious negative correlation between pollution and hours worked. On the other hand, the bias could be positive if, for example, wealthier households live in areas with low levels of pollution.

Instead, we take advantage of the plausibly exogenous variation in pollution from the refinery closure. Specifically, we estimate the following model using an instrumental variables approach:

$$\ln\left(P_{in(t-1)}\right) = \theta_0 + \alpha_{t-1} + \delta_n + \theta_1(Post_{(t-1)} \times D_{in}) + \epsilon_{in(t-1)} \qquad (8)$$

$$\ln(Y_{int}) = \beta_0 + \beta_1 \ln\left(P_{in(t-1)}\right) + \alpha_{t-1} + \delta_n + \nu_{int} \qquad (9)$$

where $Post_{(t-1)}$ is an indicator variable for weeks after the refinery closure, D_{in} is the distance in kilometers between the census block n and the refinery, α_{t-1} are week fixed effects, δ_n are census-block fixed effects, and ν_{int} represents all unobserved determinants of labor supply. We instrument for pollution in Equation (9) with $(Post_{(t-1)} \times D_{in})$. In this model, the estimate for β_1 is unbiased if, conditional on week and census block fixed effects, there are no unobserved determinants of labor supply that change differentially by distance after the closure.

$$\mathbb{E}\left((Post_{(t-1)} \times D_{in}) \cdot \nu_{int} | \alpha_{t-1}, \delta_n\right) = 0 \qquad (10)$$

Contrary to assumption (7) above, assumption (10) would hold if there are income shocks that affect

---

[28]The empirical evidence suggests a lag of up to a week: for example, Peel et al. (2005) study the lag structure between pollution levels and emergency visits and find that the risk ratios for asthma were the strongest for lag of 5 to 8 days. The employment data only includes the weekly measure of work (and not daily measures), and therefore, we cannot include a separate lag per work day. Instead, we explore how pollution levels in the previous week affects employment outcomes this week. As an additional test, we also estimated a model that measures the elasticity for pollution in the concurrent week. The observed effects are similar in magnitude, but have higher standard errors. Results are available upon request.

both labor supply and pollution as long as these shocks are common to all areas in the MCMA. In addition, assumption (10) is robust to permanent differences in labor supply outcomes across different census blocks of the MCMA.

Since our instrument is an interaction between distance and timing, the identification assumption, (10), relies on the fact that areas both close to and far from the refinery have similar trends in labor supply. This may not be the case if, for example, areas on the outskirts of the city (far from the refinery) are growing at a faster rate than areas closer to the center of the city (near the refinery). However, the panel structure of the data allows for additional controls that purge the effect of the refinery closure of any bias associated with differential neighborhood trends. The augmented model includes census block-specific linear year trends, $\delta_n \times T_{(t-1)}$, in both our first stage and second stage equations:

$$\ln\left(P_{in(t-1)}\right) = \theta_0 + \alpha_{t-1} + \delta_n + \theta_1(Post_{(t-1)} \times D_{in}) + \delta_n \times T_{(t-1)} + \epsilon_{in(t-1)} \tag{11}$$

$$\ln(Y_{int}) = \beta_0 + \beta_1 \ln\left(P_{in(t-1)}\right) + \alpha_{t-1} + \delta_n + \delta_n \times T_{(t-1)} + \nu_{int} \tag{12}$$

In this model, the exclusion restriction would be violated if other events that affected labor supply occurred at the exact time of the refinery closure and also disproportionally affected workers that lived near the refinery. This is unlikely, as we are confident that no other policies or economic events coincided with the exact timing and location of the refinery closure.

The exclusion restriction may still, however, be violated if the refinery closure itself impacted the labor market independently of the health channel. Moreover, the exclusion restriction could also be violated if differential migration into the areas surrounding the refinery occurred after it closed (i.e. sorting). We next discuss how these factors could theoretically affect the estimated elasticity of hours worked with respect to pollution levels and the additional specifications that we will test for these factors.

Direct Effects on the Labor Market

We first explore the sign and plausibility of a direct effect of the refinery closure on observed labor supply outcomes. By causing the lay-off of refinery staff, the closure could have mechanically lowered hours worked for the refinery workers. Refinery workers were often provided with housing in the neighborhoods surrounding the refinery as part of their employment benefits, and many of them remained living in the area after the refinery closed (Bazán 1999). Thus, the presence of refinery workers in the neighborhoods around the refinery would lead us to *underestimate* the relationship between pollution and labor supply that is due to health.

This downward bias would be large if the refinery workers comprised a large share of the labor market. However, employment at the refinery was not large relative to the overall labor market. The Azcapotzalco Refinery employed about 5,000 people at the time of its closure. This represented only about 0.56 percent of the population in the municipalities located in a five kilometer range of the Refinery, and 0.03 percent of the entire population of Mexico City (1990 Census). After the closure, 2,600 of these employees were forced into early retirement. About 1,500 were transitory employees and were laid off according to the minimum requirements mandated by the labor law. The rest were laid off according to the union's collective contract, which offered better terms than the labor law (García Villanueva et al., 2009).[29]

Even if the refinery jobs themselves did not comprise a large share of the labor market, it is still possible that the closure could have affected the labor market if businesses located near the refinery were affected by closure and labor markets were very localized (i.e. individuals work in the neighborhood in which they live). Specifically, both the increased unemployment due to the refinery and a slowdown of business activity near the refinery could have resulted in a localized labor demand shock that would have reduced equilibrium wages. If labor supply is upward sloping, this fall in wages should have led to reductions in work hours. Thus, the IV method would *underestimate* the effect of pollution on labor supply.

While less likely, there are also reasons to believe a negative demand shock could lead us to

---

[29]Total oil workers constitute less than 0.5 percent of our pre-closure sample (note that this includes anyone who worked in oil-connected industries and not just those who worked at the refinery). Appendix Table 3 confirms that employment in the oil industry fell from 0.48 percent in 1989:Q1 to almost zero percent in 1993:Q4 within our sample. The fall began prior to the closure, and the reduction of oil workers was concentrated near the refinery (see Figure 3).

overestimate the effect of pollution on labor supply. For example, if labor supply is downward sloping, workers may have responded to lower equilibrium wages by increasing hours of work. However, the empirical evidence does not support this: using data for the 1990-2000 period, Arceo and Campos (2010) estimate that the labor supply elasticity is 0.17 for men and 0.61 for women in the Mexican labor force. Non-wage effects in local labor markets could also bias our estimates upward. If real wages are rigid, a negative labor supply shock could increase the threat of unemployment for those employed and lead to increased work effort, some of which could be in the form of reduced absenteeism. However, rigidity of real wages is unlikely to happen in a context of high inflation (Fehr and Goette, 2005), and the annual inflation rate in Mexico in 1991 was 22 percent.

Existing data suggest that the labor market within Mexico City is quite integrated across neighborhoods; and therefore, the demand shock from the refinery was likely spread across a wide area. Data from the MCMA travel survey of 1994 reveal that, on average, individuals commute 57 kilometers to go to work, while the average census block in our sample is 0.36 square kilometers. Nevertheless, we conduct additional analysis to understand whether the closure had an effect on hours worked that is independent of the health channel.

First, we use our difference-in-difference design to directly test for localized wage effects as a consequence of the refinery closure. This method will only provide suggestive evidence for two reasons. First, wages may also be capturing productivity changes. Second, even with no effects on wages, an increased threat of unemployment could lead to less absenteeism in the presence of wage rigidity.

Instead, we exploit additional within-neighborhood variation in pollution to isolate the health channel from other localized effects of the refinery closure. More specifically, we conduct a triple-difference approach using data on wind patterns and altitude. The intuition is as follows: Assume the closure had an effect on hours worked (and employment in general) that was independent of the health gains from reduced pollution levels, and that this effect also faded away as a function of the distance to the refinery. Take the census blocks that lie between d and $d+1$ miles from the refinery. The health-independent effect of the refinery closure on hours of work should have been nearly identical for all individuals inside this mile-wide ring. Now assume that the census blocks north of the refinery have east-west fast winds, whereas the census blocks south of the refinery have north-south mild winds.

For a small d, the refinery closure would have had a larger effect on the pollution faced by those census blocks that lie south of the refinery than on those that lie north of the refinery (for a very large d, however, the refinery closure should have no effect on pollution within the mile-wide ring). Therefore, the information on wind patterns interacts with the distance from the refinery and the timing of the closure to produce variation in pollution levels that is independent from any other localized direct effects of the refinery closure.[30] In addition, the interaction of altitude with distance and the timing of the refinery closure may also provide exogenous variation in pollution: pollution concentrations vary at different levels from the ground due to thermal inversions and pollution exposure from wind patterns.

In a regression framework, we estimate:

$$
\ln\left(P_{in(t-1)}\right) = \theta_0 + \theta_1(Post_{(t-1)} \times D_{in} \times W_{in(t-1)}) + \theta_2(Post_{(t-1)} \times D_{in}) + \theta_3(Post_{(t-1)} \times W_{in(t-1)}) \\
+ \theta_4(D_{in} \times W_{in(t-1)}) + \theta_5 W_{in(t-1)} + \alpha_{t-1} + \delta_n + \epsilon_{in(t-1)}
$$

(13)

$$
\ln(Y_{int}) = \beta_0 + \beta_1 \ln\left(P_{in(t-1)}\right) + \beta_2(Post_{(t-1)} \times D_{in}) + \beta_3\left(Post_{(t-1)} \times W_{in(t-1)}\right) \\
+ \beta_4(D_{in} \times W_{in(t-1)}) + \beta_5 W_{in(t-1)} + \alpha_{t-1} + \delta_n + \eta_{int}
$$

(14)

Specifically, we instrument pollution in Equation 14 with $(Post_{(t-1)} \times D_{in} \times W_{in(t-1)})$. The estimate of β_1 is unbiased if:

$$
\mathbb{E}\left((Post_{(t-1)} \times D_{in} \times W_{in(t-1)})\eta_{int}|\mathbf{W}, \alpha_{t-1}, \delta_n\right) = 0
$$

where $\mathbf{W} = [(Post_{(t-1)} \times D_{in}), (Post_{(t-1)} \times W_{in(t-1)}), (D_{in} \times W_{in(t-1)}), W_{in(t-1)}]$, and $W_{in(t-1)}$ includes wind speed, wind direction and altitude. Note that this model only exploits the variation in pollution due to wind patterns and topography within areas that are equidistant from the refinery, while holding constant all the differential changes in pollution that occur across close and distant areas, $(Post_{(t-1)} \times D_{in})$. As such, this is a very restrictive model.

---

[30]Schlenker and Walker (2010) use a similar methodology to generate variation in pollution in a given distance from an airport.

Selective Migration/Sorting

The closure may have altered the attractiveness of surrounding neighborhoods for individuals with strong preferences for air quality (or any other neighborhood attributes linked to the refinery presence). If wealthier or healthier people moved closer to the refinery after it closed, the estimates may simply be capturing the differences in labor supply between the old and new residents of the refinery neighborhood.[31] We address this issue in two ways. First, our estimates are based on a sample that spans the five years around the closure to reduce the time frame for differential sorting. Second, we explore whether migration rates as well as time-fixed demographics (such as gender, education and marital status) changed differentially for workers who lived close to the refinery.

5 Results

5.1 Sample Statistics and Basic Correlation

Table 1 provides sample statistics for the full sample (Column 1), and the periods before (Column 2) and after (Column 3) the closure. We present the mean of each variable with the standard deviation in parenthesis below the mean, and the number of observations below that. Panel A provides information on SO_2, while Panel B provides information on labor market outcomes. Panels C and D provide information on the characteristics that we explore in the heterogeneity results.

Pollution concentrations are high (Panel A). The mean SO_2 level was 4.46 parts per hundred million (Column 1); mean SO_2 in the United States during the same period was 0.82 parts per hundred million, and the 90th percentile was 1.05. Pollution fell by about a fifth in the post-closure period from 5.02 (Column 2) to 4.04 parts per hundred million (Column 3).[32]

On average, the labor market outcomes were fairly constant over the sample period (Panel B).

---

[31]For a sample in California, Coffey (2003) shows that households with weaker health are located in areas with cleaner air, and ignoring this affect may cause one to underestimate infant mortality rates.

[32]Of course, this overall fall in pollution may have also been due to other policies, as the closure of the refinery was part of a set of policies aimed at improving the air quality in the MCMA. However, all other policies had either no effects or a city-wide effect, and therefore, should not affect our fixed effects strategy for measuring the impact of the refinery closure. For instance, the Driving Restrictions program (Hoy No Circula program), which affected all vehicles in circulation, was introduced in 1989 (see Davis (2008) for an evaluation of this program). In 1990, Pemex introduced unleaded gasoline. Low SO_2 diesel and three-way catalytic converters were both introduced in 1993, but in both cases the introductions were gradual and their effects were not localized (Secretaría de Medio Ambiente, DF).

About four percent of those in the labor force were unemployed, with no change in the post-closure period. Individuals typically worked 38.07 hours per week, with slightly higher hours worked in the post period (38.27 as compared to 37.79 prior to the closure).

Panels C and D provide data on the demographic and job characteristics, respectively. Thirty-six percent are female, 15 percent are above the age of 50, and 37 percent have at least one child who is under the age of 5. Eighty-three percent of workers who are employed have blue collar positions (Column 1); this figure remained constant over the sample period (Columns 2 and 3). Of those employed, 69 percent reported to be exclusively on a fixed salary contract; the remaining 30 percent received at least part of their earnings in the form of incentive payments: the worker was paid by the hour, by the day, by the task, by commission, received tips, or a share of the profits (henceforth this group will be referred to as "variable earnings workers").

In Table 2, we provide the correlation between pollution and the labor market outcomes. In Column 1, we present the coefficient estimate of β_1 from Equation 6, where the outcome variable of interest is the natural log of hours worked and the dependent variable is the natural log of the mean of SO_2 during the last week.[33] In Column 2, we present this estimate, conditional on census block fixed effects and week fixed effects, while we additionally include demographic controls in Column 3. Columns 4 - 6 replicate the analysis in Columns 1-3, but instead focuses on unemployment. Note that unemployment is divided by 100 for ease of interpretation (as it is a level-log model).

We observe a negative relationship between the labor market variables and pollution. A one percent increase in pollution is associated with a 0.01 percent fall in hours worked, but the coefficient is insignificant (Column 1). Controlling for the fixed characteristics of neighborhoods and weeks, a one percent increase in pollution is associated with a 0.02 percent fall in hours worked; this is significant at the 10 percent level (Column 2). Including demographic variables does not significantly alter the coefficient estimate (Column 3). Pollution is also negatively associated with unemployment, but the coefficient estimate cannot be distinguished from zero when census block and week fixed effects are included (Columns 4 - 6).

The inclusion of week and census block fixed effects should help reduce the bias in the OLS estimate. Specifically, the inclusion of week fixed effects removes all time-varying shocks to work

---

[33]We take the natural log of hours worked plus one to account for zeros in hours worked.

outcomes that are common to the entire city. This is important if we believe time-varying shocks (such as exchange rate fluctuations, changes in domestic policies, etc.) have had large effects on both work outcomes and pollution. Census block fixed effects absorb the unobserved heterogeneity in the determinants of labor outcomes that are common to a particular neighborhood in the city. However, these estimates may still be biased for two reasons. First, these estimates would still be biased if there were shocks that affect a particular location at a particular time. For example, suppose there is a large factory in a neighborhood, and that due to increased work orders that week, the factory increases work hours. The longer work hours at the factory may also result in higher pollution levels. Alternatively, suppose that certain neighborhoods have increased traffic due to road construction patterns that week; increased traffic may increase pollution, but it might also reduce the hours that individuals have available for work. Second, the OLS estimates will almost surely be attentuated due to classical measurement error. With relatively few stations, and with estimated pollution measures per census block, we expect measurement error is an important concern in this case. This is particularly a problem in developing countries where stations are relatively sparse as compared to developed country settings. The fixed effects approach will exacerbate the measurement error problem, biasing the estimated parameters towards zero. Our instrumental variables approach, to be discussed next, addresses these potential confounding effects, as well as classical measurement error in the pollution measures.

5.2   First Stage Results

Figure 4 provides a graph of SO_2, by week, for census blocks within and beyond a five kilometer radius from the refinery, controlling for census block fixed effects. After the refinery closes (March 1991), pollution falls in the census blocks close to the refinery relative to census blocks farther away. The spikes in pollution around winter time, like the ones observed at the beginning of 1989 and the end of 1991, are a common phenomenon in Mexico City. At this time of the year, thermal inversions and wind direction conspire to prevent pollution from flowing out of the valley. Due to the high variance in pollution, the seasonal patterns, and the falling pollution trends for both groups, Figure 4 illustrates the need to use more precise distance information, control for seasonal effects, and also

23

control for census block-specific trends to reduce the variance in the estimates.[34]

In Table 3, we estimate the effect of the refinery closure on pollution. These results are interesting in their own right — given the cost imposed by closing the refinery, the success of the closure in reducing pollution levels is an important outcome to measure. The results are also important, for the purpose of this study, because they form the first stage estimates of the subsequent instrumental variables model. In Column 1, we present the estimate of θ_1 from Equation (8), where we replace the distance variable with a dummy variable for whether the census block is located within five kilometers of the refinery. While this does not take full advantage of the variation in distance, it provides a simple interpretation of the effect of the closure. In Column 2, we present the estimate of θ_1 from Equation (8). Column 3 replicates the analysis in Column 2, but additionally includes demographic controls — an indicator variable each for gender and age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. Finally, Column 4 reports on a specification that additionally controls for year trends in each census block (Equation 11). Note that all specifications include both week and census block fixed effects and are clustered by census block.

The refinery closure substantially reduced pollution. As shown in Column 1, after the refinery closed, areas located within five kilometers of the refinery experienced a 8.28 percent decrease in mean SO_2 (significant at the one percent level). We observe relatively higher SO_2 within a census block after the closure for those located farther away (Column 2). The estimated effect is not driven by standard demographic characteristics (Column 3) nor by neighborhood-specific trends (Column 4). For example, controlling for census block-specific year trends, the magnitude of the effect increases from 0.65 percent to 0.94 percent increase in relative pollution per km and the estimate remains significant at the one percent level (Column 4).

5.3   Policy Impact Evaluation (Reduced Form Estimation)

We begin by estimating the reduced form impact of the closure on work decisions. This is important because it allows us to easily interpret the benefits of the closure, which also imposed significant

---

[34]Note that pollution falls over time. This is due in part to other environmental regulations that occurred in this time period. However, this should not affect our estimates as the policies were city-wide and we are controlling for week fixed effects.

economic costs in the form of lost profits and wages. We estimate the direct effect of the policy on the log of hours worked, with (Equation 16) and without (Equation 15) census block-specific year trends:

$$\ln(Y_{int}) = \pi_0 + \alpha_{t-1} + \delta_n + \pi_1(Post_{(t-1)} \times D_{in}) + \epsilon_{in(t-1)} \tag{15}$$

$$\ln(Y_{int}) = \pi_0 + \alpha_{t-1} + \delta_n + \pi_1(Post_{(t-1)} \times D_{in}) + (\delta_n \times T_{(t-1)}) + \epsilon_{in(t-1)} \tag{16}$$

We present these results in Panel A of Table 4. In Columns 1 - 4, the outcome variable of interest is the natural log of hours worked. The number of hours worked is zero for those who are unemployed, and thus this measure captures both the intensive and extensive margins of work.[35] In Column 1, we present the estimate of π_1 from Equation (15), where we replace the distance variable with an indicator variable for whether the census block is located within five kilometers of the refinery. In Column 2, we present the coefficient estimate of π_1 from Equation 15. Column 3 adds controls for demographic characteristics, while Column 4 adds year-specific trends in each census block (Equation 16). All specifications are clustered at the census block level, and include both week and census block fixed effects. Columns $5 - 8$ replicate the analysis in Columns 1 - 4 with unemployment status as the dependent variable.

On average, after the closure, we observe a 5.5 percent increase in hours worked for individuals who lived within five kilometers of the refinery (Column 1). After the refinery closed, hours of work were 0.32 percent relatively lower per each km from the refinery (Column 2). Adding demographic controls (Column 3) slightly reduces the magnitude of the estimated coefficient (0.32 to 0.29 percent). As shown in Column 4, the effect is larger in magnitude when controlling for census block-specific year trends, but we cannot reject that it is significantly different from the earlier estimates.

After the refinery closed, the unemployment rate fell by about 1 percentage point in the area surrounding the refinery (Column 5). This has two important implications, as we discussed in Section 4. First, the refinery closure may have generated a negative demand shock. In this case, if labor supply is upward-sloping, we would expect a spurious negative correlation between unemployment

---

[35]We take the natural log of hours worked plus one to account for zeros in hours worked.

and distance, which would cause us to underestimate the effect of pollution on hours worked. Second, 30 percent of the employed are variable earnings workers. Some of the variable earning workers may have short term contracts (i.e. by the day or by the hour) and, if they become ill due to respiratory disease, it is possible that they do not find work the following week and are counted as unemployed. The estimated effects are positive and thus more consistent with the latter story. However, it is important to note that when we control for census block-specific trends in unemployment (our preferred specification), the reduced form estimate is not statistically significant (Column 8).

5.4 Estimated Effects of Pollution on Labor Outcomes

Panel B of Table 4 presents the IV estimates of the elasticity of hours worked with respect to SO_2. Specifically, we present coefficient estimates of β_1 from Equation (9). In the first four columns, the log of hours worked is the outcome of interest. The hours worked is zero for those who are unemployed, and thus this measure captures both the intensive and extensive margins of work.[36] The first stage estimates for each of these columns correspond to those presented in Table 3. In Columns $5-8$, we replicate the analysis in Columns $1-4$, but we now have an indicator variable for being unemployed as the outcome variable.[37] For these models, unemployment is divided by 100 for ease of interpretation (as it is a level-log model). All specifications include both week and census block fixed effects and are clustered at the census block level.

Higher SO_2 leads to reduced hours worked in the following week (Panel B, Columns 1-4).[38] Using $(Post_{(t-1)} \times D_{in})$ as the instrument for SO_2 (Column 2), a one percent increase in SO_2 leads to a 0.48 percent fall in hours worked (significant at the 5 percent level).[39] The coefficient estimates

---

[36]We take the natural log of hours worked plus one to account for zeros in hours worked. In these models, we have a censored dependent variable (one cannot work less than zero hours per week). Due to our fixed effects approach, a tobit model is subject to the incidental parameter problem. As a robustness check, we have estimated Equation 9 using a two-stage Tobit model (equivalent to Table 4, Panel B, Column 2). The coefficient remains negative and significant.

[37]Despite an indicator variable as the outcome variable, we present the estimates of a linear probability model. Given the fixed effects approach, a logit or probit model is subject to an incidental parameter problem , especially when trends are included. Nonetheless, we estimated the two-stage probit model for Equation 9 (equivalent to Table 4, Panel B, Column 6). The coefficient remains positive and significant.

[38]As discussed in the data section, SO_2 may be a proxy for other refinery-generated pollutants that are not collected in this setting, such as PM, NO_x and CO. It is possible that other pollutants are correlated with SO_2 and work hours, and are also moved by the refinery closure. In this case, the estimated elasticity would additionally capture the effect of other refinery-generated pollutants.

[39]Note that this effect size falls within the range of effect sizes of smoking on absenteeism: Leigh (1995) observed a 40 minute to 3 hour increase in work hours for non-smokers, while we observe a 2 two hour total effect of the closure.

are similar in magnitude and significance when controlling for demographic characteristics (Column 3). The estimated effect is not due to differential trends: controlling for census block-specific yearly trends, a one percent increase in SO_2 results in a 0.61 percent fall in hours worked (Column 4).[40]

We next explore the effect of pollution on unemployment. In the basic fixed effects model using distance as an instrument (Column 6, Panel B of Table 4), we observe that a one percent increase in mean SO_2 leads to a 0.10 percent increase in unemployment (significant at the 5 percent level). Controlling for census block-specific year trends, the effect size falls to 0.07 percent and is no longer significant at conventional levels (Column 8). Thus, while we observe a substantial effect of pollution on hours worked per week, we observe only a small increase or no effect on overall unemployment levels.

Ruling Out the Direct Effect of the Closure on the Labor Market   As outlined in Section 4, the closure itself could mechanically lower hours of work in the short-run (through lay-offs of refinery workers). While the refinery itself did not comprise a large share of the labor market surrounding the refinery, it may also have had effects on local businesses. If labor markets are localized by neighborhood (i.e. individuals work in the same neighborhoods that they live), the resulting change in labor demand would likely result in an underestimate of the estimated elasticity. While the existing qualitative evidence (discussed in Section 4) suggests that the labor market in Mexico City is fairly integrated, we formally test for this effect in two ways.

We first estimate the reduced form effect of the refinery closure on the natural log of wages (Table 5). The table format is similar to Table 3. All models are estimated using OLS, with standard errors clustered at the census block level. Across all four specifications, we cannot reject that the effect of the refinery on the log of wages is different from zero. In addition, the magnitudes of the estimated coefficients are very small across all four models. However, given that the theoretically predicted change in wages is ambiguous (the demand shock could have lowered equilibrium wages, but the

---

Van Ours (2004) also finds large productivity effects for smokers, with a 10 percent reduction in wage for male smokers.

[40]Due to the nature of the rotating panel, only about 10 percent of our sample is present both before and after the refinery closure. As such, we lack statistical power to include individual fixed effects. We also experimented with further restricting the sample to the period to a narrower period around the refinery closure, but again run into problems with statistical power. If we restrict only to the years 1990 and 1991 (only 56, 137 observations), the effects are negative (-0.36) and marginally significant (p=0.10) using post times the distance variables as the instrument (this compares to -0.48 in Table 4, Panel B, Column 2). If we restrict further to only those present before and after the closure (only 13, 145 observations), the estimated effect is negative, but cannot be distinguished from zero.

productivity shock could have raised them), this test in itself does not fully rule out the direct effect of the closure on the labor market.

Instead, we exploit additional within-neighborhood variation in pollution to isolate the health channel from other localized effects of the refinery closure. As explained in Section 4, we use information on wind patterns and altitude to conduct a triple-difference approach. The IV strategy used in these specifications requires substantial variation in the altitude and wind measurements within groups of census blocks that are located at a similar distance to the refinery. Altitude has moderate variation close to the refinery: the mean altitude is 2,256 m and the standard deviation is 15 m across census blocks within 5 km of the refinery, while the standard deviation is 63 m across census blocks that are further than 5 km. Wind patterns have considerably greater variation across census blocks and across weeks. The average wind direction in degrees is 60 with a standard deviation of 40 degrees in census blocks close to the refinery and 43 degrees in census blocks further than 5 km to the refinery, where 0 degrees corresponds to a position directly upwind from the refinery and a maximum of 180 degrees corresponds to a position directly downwind. Similarly, the standard deviation for wind speed is between 10 and 50 percent of the mean.

The results of this strategy are presented in Table 6. Panel A reports the estimated coefficient vector θ_1 from Equation (13), along with F-tests for joint significance when multiple instruments are used. Panels B - D present the IV estimates of β_1 for the natural log of hours worked, unemployment, and wages respectively (Equation 14). Column 1 reports on specifications that include both week and census block fixed effects, Column 2 reports on specifications that add demographic controls, and Column 3 reports on specifications that add census block specific year trends.

Panel A of Table 6 confirms that altitude and wind patterns provide additional variation in the change in pollution exposure that is generated from the refinery closure. In particular, after the closure, individuals that live downwind from the refinery and farther away from it experience relatively higher levels of pollution than individuals that live downwind and closer to the refinery (there is a a positive and significant coefficient on the triple interaction of Post, Distance and Wind Direction). Wind speed and elevation do not seem to have a differential impact on pollution for individuals that live far away from the refinery.

Despite having fewer observations because of missing wind information for five of our monitoring

stations (which forces us to drop 25 of the 265 census blocks), the IV results on hours of work from Columns 1 and 2 (Panel B) are remarkably similar to the results from the simple difference-in-difference model that is presented in Column 2 and 3 of Table 4. Given the restrictiveness of this model, we check whether weak instruments might be a source of concern. The models in Columns 1 and 2 perform well under a weak instruments test (the Kleibergen-Paap weak instrument F statistic is above the most stringent Stock-Yogo critical value,or have a maximum of 10% of the OLS bias). Note that when we try to introduce census block-specific year trends, the coefficient falls both in magnitude and significance (Column 3), which is not surprising given the weak instrument test results.

The IV results on unemployment rates (Panel C) follow a similar pattern to the difference-in-difference approach. Once again, there is no effect on wages for any of the specifications we consider (Panel D).

In summary, our results are robust to what we believe is a very strict test of endogeneity. This is all the more remarkable giving that this model gives up substantial variation in the instruments used compared to the model defined by Equations (11) and (12). As such, these findings indicate that the estimated elasticities do not appear to be driven by local labor demand shocks from the closure.

Ruling Out Selective Migration/Sorting   The refinery closure may have induced differential migration into the areas that saw the sharpest reductions in pollution. For example, wealthier or more educated individuals may have chosen to move into the newly "clean" neighborhoods. The fact that our estimates are robust to the inclusion of demographic controls is at least suggestive that selection is not an important source of bias in our estimates. Nonetheless, we test for selection bias in two ways.

First, we test whether migration rates changed differentially by distance after the closure. Figure 5 graphs out the propensity to migrate by quarter-year for households that live within 5 kilometers of the refinery versus those who live farther away. The migration rate is low: only 3.5 percent of households migrated during the 5 quarter period they remained in our sample. The rates are fairly constant before and after the closure for both groups (the rates are more volatile for those closer to the refinery, but this may be due to the smaller sample size of this group). In Table 7, we more formally test whether migration is correlated with the instrument. To do so, we organize the data

29

at the level of a household-quarter-year. We then create an indicator variable that equals 1 if the household was replaced in the survey in that quarter-year due to migration, and zero otherwise. We then estimate the reduced form equation of the closure on migration (controlling for week and census block fixed effects). The results are presented in Column 1 of Panel A of Table 7. The coefficient estimate is small in magnitude and not significantly different from zero. Including census block specific-year trends (Column 1 of Panel B) does not change the estimates.

Second, we test whether key demographic characteristics differentially change after the closure. We consider gender, age, marital status and education level. The results are presented in Columns 2- 6 of Panel A of Table 7. Of the five variables considered, only one is significant at the 10 percent level: after the refinery closed, individuals were relatively less likely to be married the farther they lived from the refinery. We conducted a joint test of the instrument across all five variables, and we fail to reject the null hypothesis that the instrument had no effect on any of the variables (p=0.27) cannot be rejected. Controlling for census block specific year trends (Panel B), none of the variables are significant at conventional levels and the joint test fails to reject the null hypothesis that there is no effect (p=0.77). Taken together, this evidence suggests that differential migration is not driving the estimated elasticities.[41]

5.5 Heterogeneous Treatment Effects

The extent to which workers miss work due to high pollution levels may vary by demographics and job types. In Table 8, we test for heterogeneous treatment effects. All specifications include census block fixed effects, week fixed effects, census block specific year trends, and demographic controls. All specifications are estimated using a two-stage least squares model, with standard errors clustered at the census block.

Whose Health? The elasticity may differ for men and women, especially if a household has young children. For example, women may be more likely to stay home if a young child is sick, and thus lower pollution levels could lead to increased hours worked if the children are healthier. Therefore,

---

[41]We also estimated the specifications in Table 4 restricting the sample to those families that do not move during the 5 quarter period that they remain in our sample. The results are virtually identical to those in Table 4 and are available upon request.

we test whether men and women exhibit different elasticities (Column 1 of Table 8), whether having a young child in the household increases the effect of pollution on hours worked (Column 2), and whether the effect of having a young child differs for men and women (Column 3). Note that we now have multiple endogenous variables. For example, in Column 1, we now have two endogenous variables—$\ln\left(P_{in(t-1)}\right)$ and $\ln\left(P_{in(t-1)}\right) \times Female_{int}$. As such, we create two instruments: the original instrument $(Post_{(t-1)} \times D_{in})$ and its interaction with an indicator for female $((Post_{(t-1)} \times D_{in}) \times Female_{int})$. This strategy results in a weak instrument problem, therefore the results should be considered only suggestive.

We find suggestive evidence that high pollution levels reduce hours worked at least in part through their effect on the health of workers' children, but the limited sample sizes (Column 2) preclude conclusive findings. Women have a smaller reaction to SO_2 than men (Column 1), but this effect is not significant at conventional levels. The elasticity is stronger for individuals with children under five (Column 2), but this is not significant at conventional levels. We find that having a child under 5 increases absolute value of the estimated elasticity (significant at the 10 percent value), but we cannot reject the null that the estimated effect is different for males with young children and females with young children (Column 3).

In addition, we may expect that pollution differentially affects the health of older workers. We test whether the elasticity is larger (in absolute value terms) for workers over the age of 50 (Column 4). While the elasticity is larger for older workers, we cannot conclude that this difference is greater than zero.

Employment and Contract Type

Next, we explore the effect of being in a blue collar position and having a fixed salary contract on hours worked. The theory on the effect of both of these characteristics on the elasticity of hours worked with respect to pollution is ambiguous. Blue collar positions may be more physical than white collar jobs, making it harder to work when sick. On the other hand, blue collar workers tend to be lower paid and perhaps have less power in their work relationships, making it hard to take days off.

Similarly, fixed salary workers may be more likely to receive the same pay whether or not they

attend work, while variable earnings workers will have reduced earnings if they stay home when sick. However, individuals with variable earnings contracts are often paid by the day or by the task. If their productivity is low due to illness, their potential "wage" per day may be quite low, and low enough that it would not be worth it for them to attend work.

In Column 5 of Table 8, we test whether the elasticity differs for blue collar and white collar workers. In Column 6, we test whether there is a difference in the elasticity for fixed salary workers versus those with variable earnings. Note that we only have data on industry and contract type for those who are employed that week. We observe no difference in the elasticities for white and blue collar workers (Column 5), although these results are also subject to a weak instrument problem. On the other hand, much of the effect of SO_2 on hours worked is driven by fixed salary workers (Column 6).[42]

6   Discussion

Interpreting our Results   In this paper, we demonstrate that an improvement in air quality substantially increases the amount of hours worked. The dollar value of hours worked is not by any means a complete measure of the overall welfare effect of better air quality. As discussed in Section 2, individuals may gain other benefits from improved air quality. First, workers may value consumption more when they and their families experience better health. Second, their expenditures on some goods (like gym memberships and medicines) may go down, releasing income for consumption of other goods. Finally, they may simply enjoy cleaner air and good health. However, the fact that we find such an important effect on labor supply behavior suggests that this component of the total welfare gains from cleaner air is large enough that it should not be ignored.

In addition to understanding welfare implications of environmental regulation, understanding whether air pollution affects labor supply—and, if so, the magnitude of the effect—has important

---

[42]Of course, this estimate may be biased upward as we only know the contract type for those who are employed. In the extreme case, some variable earnings workers may choose to take an entire week off when sick, and therefore would be classified as unemployed. In this case, we would underestimate the effect for those on incentive based contracts. To test for this possibility, we conducted a simple bounding exercise. Specifically, we estimated the effects of demographic characteristics and location within the city on contract status for those who were employed, and then predicted contract status for those who were unemployed. We then re-ran the regression including the predicted contract status for those who were unemployed. The results remained the same, suggesting that those on incentive-based contracts were not taking a full week off when sick. These result are available from the authors upon request.

implications for designing the optimal tax on pollution. [43] Pollution taxes are appealing relative to other forms of environmental regulation because the revenue they generate can be used to lower other distortionary taxes. However, there is a debate in the double dividend literature on how to set optimal pollution taxes. On one hand, it has been argued that pollution taxes are likely to result in lower labor supply (as a result of direct regulation or through substitution effects on leisure as the price of regulated goods goes up). This general equilibrium effect, known as the "interaction effect," implies that the optimal tax on pollution might be lower than the marginal damage of pollution and that other distortionary taxes may be necessary to maintain a balanced budget (Parry, 1995; Goulder and Bovenberg, 1997). Importantly, most of the early papers in this literature implicitly assume that air quality does not affect labor supply through health. Recent theoretical work shows that the interaction effect is mitigated when air quality independently increases labor supply (Schwartz and Repetto, 2000; Williams, 2003). Therefore, the fact that we observe a large and positive effect of pollution on labor supply implies that the optimal tax on pollution is therefore closer to (and potentially above) the marginal damage of air pollution.[44]

Finally, it is also important to note that the effect on labor supply that we measure is restricted to the short-term impacts of a change in air quality. If health effects of air quality are cumulative, we could expect even larger labor supply effects. Moreover, as discussed in Section 2, if better air quality contributes to work productivity per hour, the value of hours worked may increase and eventually translate into higher wages. Although we do not observe changes in wages in the period that we focus on, this may change over a longer time horizon.

Distribution of costs and benefits The closure was instigated to reduce pollution levels in Mexico City. It came, however, at a substantial cost. First, there was a loss of employment for 2,400 workers, with an additional 2,600 workers forced into early retirement. Second, the closure may have led to oil shortages in the short-run. The refinery was not replaced by another facility. The refining capacity of Pemex fell in 1990 from 3,466 to 3,331 thousand daily barrels. While the refining capacity

---

[43]This result also has important implications for the optimal level of any environmental regulation when maximizing welfare under a balanced budget constraint.

[44]For example, West and Williams (2007) estimate a demand system and find that the optimal tax on gasoline should be above the marginal damage of gasoline emissions. Their calculation does not take into account the fact that lower gasoline consumption may reduce emissions and increase labor supply. Such an effect would yield an even larger optimal gasoline tax.

started a slow recovery in 1992, fuel consumption increased at a faster pace. In 1989, the imports of refined oil products exceeded Pemex exports for the first time since 1980 and the gap continued growing through 2000. Most of the costs associated with short-term shortages were absorbed by the government since gasoline prices did not increase more than inflation between 1990 and 1993 (Pemex Statistical Yearbooks 1990 and 2001).

Our estimates suggest the closure led to substantial benefits in the form of labor income. Suppose we assume that all gains were concentrated in the 5 kilometers near the refinery. Then, our estimates imply that the refinery closure lead to about a 5-percent increase in hours worked for those who lived near the refinery. Given an average annual wage of 13,700 1993 Mexican Pesos (USD 3,600), there was a 756 Peso (or USD 198) per worker gain over the course of a year.

Note that we are unable to estimate whether these gains accrued to the workers or to their employers. If the refinery closure led to fewer sick days for fixed salary employees, and these employees enjoyed paid sick leaves, the benefits would have been accrued to the employer. However, if employees had a limited number of paid sick days or were able to cash out unused sick days, employees would have shared in the benefits of increased work ability. At the time of the closure, the work regulations in place did not specifically stipulate that employers were responsible for providing paid sick days to their employees, although some may have chosen to do so. The formally employed did enjoy an extended sick absence benefit paid through Social Security (IMSS), receiving 60 percent of forgone salary starting on the fourth day of an absence due to a documented illness. Thus, the financial benefits from reduced absences due to poor air quality were likely split between employers, employees, and the Social Security System.

7 Conclusion

Pollution is high in the developing world and increasing at a very fast rate. For example, with 2005 emissions estimated to be about 25 million tons, China is the world's largest polluter of sulfur dioxide. China became the world's largest emitter of greenhouse gases in 2006, while India is predicted to become the third largest emitter (after China and the United States) by 2030. Improvements in global environmental quality will therefore not occur without the cooperation of developing countries

to enforce stricter environmental regulations. However, many low income countries may be hesitant to potentially sacrifice economic growth for possible improvements in environmental quality. This study suggests that gains in worker effort may help offset some of the economic losses that may result from more stringent environmental regulation. Specifically, we find that a one percent increase in SO_2 results in a 0.43 to 0.67 percent decline in hours worked in the following week. Given an average annual wage of 13,700 1993 Mexican Pesos (USD 3,600), this resulted in an estimated 756 (USD 198) per-worker gain from reduced absenteeism over the course of a year.

These findings contribute to the literature that documents the costs of pollution associated with short-term effects on health, and are consistent with studies that have documented the effect of pollution on school absenteeism and performance (Currie, Hanushek, Kahn, Neidell, and Rivkin, 2009; Ransom and Pope, 1992; Gilliland et al., 2001). Moreover, the findings suggest pathways for future research. First, they raise questions about the distribution of gains. In the context of Mexico City, labor regulations were such that the gains from lower air pollution were likely split between workers, employers and the social securtiy sistem. More generally, it would be interesting to explore how labor regulations affect the propensity to miss work due to pollution-related illness, and how this then affects the distribution of benefits. Second, while the empirical model documents a substantial *short-run* effect of changes in pollution on labor market outcomes, there may be larger long-run impacts if there are cumulative effects of pollution on health or if wages change in response to pollution. Future research, thus, should address longer-run outcomes on employment capabilities and changes in work productivity and wages.

References

[1] Arceo Gómez, Eva and Raymundo Campos Vázquez, "Labor Supply of Married Women in Mexico: 1990-2000," Working Paper, Centro de Estudios Economicos, COLMEX, No. XVI, 2010.

[2] Basta, Samir S., Soekirman, Darwin Karyadi, and Nevin Scrimshaw, "Iron Deficiency Anemia and Productivity of Adult Males in Indonesia." *American Journal of Clinical Nutrition*, 32 (1979), 916–925.

[3] Bazán, Lucía, "Casa y familia. Los Recursos de los Desempleados de Pemex en la Ciudad de México," Estudios Sociológicos, Colegio de México, XVII, 50, 1999.

[4] Bevilacqua, M., Braglia, M., "Environmental Efficiency Analysis for ENI Oil Refineries," *Journal of Cleaner Production*, 10 (2002), 85–92.

[5] Coffey, Bentley, "A Reexamination of Air Pollution's Effects on Infant Health: Does Mobility Matter?" Duke University mimeo. Durham, North Carolina: Duke University, 2003.

[6] Chay, Kenneth, and Michael Greenstone, "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 (2003), 1121–1167.

[7] Crocker, Thomas D. and Robert L. Horst Jr., "Hours of Work, Labor Productivity, and Environmental Conditions: A Case Study," *The Review of Economics and Statistics*, 63 (1981), 361–368.

[8] Currie, Janet, Eric Hanushek, E. Megan Khan, Matthew Neidell, and Steven G. Rivken, "Does Pollution Increase School Absences?" *The Review of Economics and Statistics*, 91 (2009), 682–694.

[9] Currie, Janet and Matthew Neidell, "Air Pollution and Infant Health: What Can We Learn From California's Recent Experience?" *The Quarterly Journal of Economics*, 120 (2005), 1003–1030.

[10] Davies, K.J., C.M. Donovan, C.J. Refino, G.A. Brooks, L. Packer, and P.R. Dallman, "Distinguishing the Effects of Anemia and Muscle Iron Deficiency on Exercise Bioenergetics in the Rat," *American Journal of Physiology – Endocrinology and Metabolism*, 246 (1984), E535–E543.

[11] Davis, Lucas, "The Effect of Driving Restrictions on Air Quality in Mexico City," *Journal of Political Economy*, 116 (2008), 38-81.

[12] Dietert RR; Etzel RA; Chen D; Halonen M; Holladay SD; Jarabek AM; Landreth K; Peden DB; Pinkerton K; Smialowicz RJ; Zoetis T.. "Workshop to Identify Critical Window of Exposure for Children's Health: Immune and Respiratory Systems Work Group Summary," *Environ Health Perspect*, 108 (2000), 483-490.

[13] EPA, "Integrated Science Assessment for Sulfur Oxides - Health Criteria," ISA: EPA/600/R-08/047F, 2008.

[14] Fehr, Ernst and Lorenz Goette, "Robustness and Real Consequences of Nominal Wage Rigidity." *Journal of Monetary Economics*, 52 (2005), 779-804.

[15] Frankenberg, Elizabeth, Douglas McKee, and Duncan Thomas, "Health Consequences of Forest Fires in Indonesia." *Demography*, 42 (2005), 109–129.

[16] García Villanueva, Luis Antonio, Georgina Fernández Villagómez, Rosario Iturbe Argüelles and Ulises Ruiz Saucedo, "Caso de Estudio: Antecedentes Históricos de la Ex-refinería "18 de marzo"," Ciudad de México, 2009.

[17] Gilliland et al., "The Effects of Ambient Air Pollution on School Absenteeism Due to Respiratory Illness," *Epidemiology*, 12 (2001): 43–54.

[18] Gorman, Christine, and Laura Lopez, "Mexico City's Menacing Air," *Time Magazine*, April 1, 1001.

[19] Goulder, L. H., and L. Bovenberg, "Costs of Environmentally Motivated Taxes in the Presence of other Taxes: General Equilibrium Analyses," *National Tax Journal*, 50 (1997), 59-87.

[20] Graff, Joshua S., and Matthew J. Neidell, "The Impact of Pollution on Worker Productivity," NBER Working Paper #17004, 2011.

[21] Haas, Jere D., and Thomas Brownlie IV, "Iron Deficiency and Reduced Work Capacity: A Critical Review of the Research to Determine a Causal Relationship," *The Journal of Nutrition*, 131 (2001), 676S-690S.

[22] Instituto Nacional de Estadistica y Geographia, National Mexico General Census of Population and Housing, 1990.

[23] Ito, Kazuhiko, George D. Thurston and Robert A. Silverman, "Characterization of PM2.5, Gaseous Pollutants, and Meteorological Interactions in the Context of Time-Series Health Effects Models." *Journal of Exposure Science and Environmental Epidemiology*, 17 (2007), 45–60.

[24] Jayachandran, Seema, "Air Quality and Early-Life Mortality: Evidence from Indonesias Wildres." *Journal of Human Resources*, 44 (2009), 916-954.

[25] Leigh, J Paul, "Smoking, Self-Selection and Absenteeism," *The Quarterly Journal of Economics*, 35 (1995), 365-386.

[26] Li, Ruowei, Xuecun Chen, Huaicheng Yan, P. Deurenberg, L. Garby, and J.G. Hautvast, "Functional Consequences of Iron Supplementation in Iron-Deficient Female Cotton Mill Workers in Beijing, China," *American Journal of Clinical Nutrition*, 59 (1994), 908–913.

[27] Lleras-Muney, Adriana, "The Relationship Between Education and Adult Mortality in the United States," *The Review of Economic Studies*, 72 (2005), 189-221.

[28] Linn WS; Venet TG; Shamoo DA; Valencia LM; Anzar UT; Spier CE; Hackney JD, "Respiratory Effects of Sulfur Dioxide in Heavily Exercising Asthmatics. A Dose-Response Study." *Am Rev Respir Dis*, 127 (1983), 278-283.

[29] Linn WS; Avol EL; Peng RC; Shamoo DA; Hackney JD, "Replicated Dose-Response Study of Sulfur Dioxide Effects in Normal, Atopic, and Asthmatic Volunteers." *Am Rev Respir Dis*, 136 (1987), 1127- 1134.

[30] Milenio, "Denuncian Emanación de Gases de Refinería Madero," December 12, 2008.

[31] Neidell, Matt. "Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations," *Journal of Human Resources*, 44 (2009), 450-478.

[32] Park H; Lee B; Ha EH; Lee JT; Kim H; Hong YC. "Association of Air Pollution with School Absenteeism due to Illness." *Arch Pediatr Adolesc Med*, 156(2002), 1235-1239.

[33] Parry, Ian. "Pollution Taxes and Revenue Recycling," *Journal of Environmental Economics and Management*, 29 (1995), S64-S77.

[34] Pönka A., "Absenteeism and Respiratory Disease Among Children and Adults in Helsinki in Relation to Low-level Air Pollution and Temperature." *Environ Res*, 52 (1990), 34-46.

[35] Peel, Jennifer L., Paige E. Tolbert, Mitchel Klein, Kristi Busico Metzger, W. Dana Flanders, Knox Todd, James A. Mulholland, P. Barry Ryan, and Howard Frumkin, "Ambient Air Pollution and Respiratory Emergency Department Visits." *Epidemiology*, 16 (2005), 164–74.

[36] Ransom, Michael and C. Arden Pope, "Elementary School Absences and PM_{10} Pollution in Utah Valley," *Environmental Research*, 58 (1992), 204-219.

[37] Schildcrout JS; Sheppard L; Lumley T; Slaughter JC; Koenig JQ; Shapiro GG., "Ambient Air Pollution and Asthma Exacerbations in Children: An Eight-City Analysis," *Am J Epidemiol*, 164 (2006), 505- 517.

[38] Schlenker, Wolfram and Reed Walker, "Air Pollution and Contemporaneous Health: Evidence from Random Variation in Pollution Shocks from Airports." Mimeo, 2010.

[39] Schwartz et al., "Acute Effects of Summer Air Pollution on Respiratory Symptom Reporting in Children." *American Journal of Respiratory and Critical Care Medicine*, 150 (1994), 1234–1242.

[40] Schwartz J; Spix C; Touloumi G; Bacharova L; Barumamdzadeh T; le Tertre A; Piekarksi T; Ponce de Leon A; Pönka A; Rossi G; Saez M; Schouten JP, "Methodological Issues in Studies of Air Pollution and Daily Counts of Deaths or Hospital Admissions," *J Epidemiol Community Health*, 50 Suppl 1 (1996), S3-11.

[41] Schwartz, J., Repetto, R., "Nonseparable Utility and the double dividend debate: reconsidering the tax-interaction effect," *Environmental and Resource Economics*, 15 (2000), 149–157.

[42] Thomas, Duncan et al."Iron Deficiency and the Well-Being of Older Adults: Early Results from a Randomized Nutrition Intervention," Paper presented at the Population Association of America Annual Meetings, Minneapolis, April 2003, and the International Studies in Health and Economic Development Network meeting, San Francisco, May 2003.

[43] van Ours, Jan C., "A Pint a Day Raises a Man's Pay; but Smoking Blows that Gain Away," *The Journal of Health Economics*, 23 (2004), 863–886.

[44] Walker, Reed, "The Transitional Costs of Policy: Evidence From the Clean Air Act and the Workforce." Mimeo, 2011.

[45] Williams, Roberton C., "Health Effects and Optimal Environmental Taxes," *Journal of Public Economics*, 87 (2003), 323-335.

[46] Williams, Roberton and Sarah West. "Optimal Taxation and Cross-Price Effects on Labor Supply: Estimates of the Optimal Gas Tax." *Journal of Public Economics*, 91 (2007), 593-317.

[47] Woodson, R.D., R.E. Wills, and C. Lenfant, "Effect of Acute and Established Anemia on O_2 Transport at Rest, Submaximal and Maximal Work." *Journal of Applied Physiology*, 44 (1978), 36–43.

[48] Zhu Y. Isabel, and Jere D. Haas. "Altered Metabolic Response of Iron-Depleted Non-Anemic Women During a 15-km Time Trial." *Journal of Applied Physiology*, 84, (1998), 1768–1775.

Appendix

An expression for $\frac{de}{d\alpha}$ and $\frac{dc}{d\alpha}$ can be obtained from differencing the first order conditions of:

$$\max u(c, e; \alpha) \text{ s.t. } we + y - pc = 0 \tag{17}$$

The first order conditions are:

$$u_c(c, e; \alpha) = \lambda p_t \tag{18}$$

$$-u_e(c, e; \alpha) = \lambda w \tag{19}$$

where w and p are the price of consumption and the wage rate. Differencing (18) and (19) and solving for $\frac{dc}{d\alpha}$ and $\frac{de}{d\alpha}$ yields

$$\frac{dc}{d\alpha} = \frac{\frac{\partial \lambda}{\partial \alpha} p - u_{c\alpha}}{u_{cc}} \tag{20}$$

$$\frac{de}{d\alpha} = \frac{\frac{\partial \lambda}{\partial \alpha} w + u_{e\alpha}}{-u_{ee}} \tag{21}$$

We can obtain the third equation we need to obtain expressions for $\frac{d\lambda}{d\alpha}, \frac{dc}{d\alpha}$ and $\frac{de}{d\alpha}$ by totally differentiating the budget constraint:

$$w \frac{de}{d\alpha} - p \frac{dc}{d\alpha} = 0 \tag{22}$$

The sign of $\partial \lambda / \partial \alpha$ can be inferred from substituting (20) and (3) in (25) and solving for $\partial \lambda / \partial \alpha$:

$$\frac{\partial \lambda}{\partial \alpha} = \frac{w \frac{u_{e\alpha}}{u_{ee}} - p \frac{u_{c\alpha}}{u_{cc}}}{-\left[\frac{w^2}{u_{ee}} + \frac{p^2}{u_{cc}} \right]} \tag{23}$$

It is easy to check that $\frac{\partial \lambda}{\partial \alpha} < 0$ holds if $u_{c\alpha} \leq 0$. First notice that the denominator in (23) is positive, since $u_{cc} < 0$, $u_{ee} < 0$. Second, notice that $u_{e\alpha} > 0$, hence the numerator is negative. Consumption and air quality are complements if $u_{c\alpha} > 0$ and strong complements if $u_{c\alpha} > \frac{w}{p} \frac{u_{e\alpha}}{u_{ee}} u_{cc} > 0$. In the later case, $\frac{\partial \lambda}{\partial \alpha} > 0$.

Finally, we consider the case in which leisure is complementary to air quality. In this case, $u_{e\alpha} < 0$. So, given expression (23), it is easy to show that $\frac{\partial \lambda}{\partial \alpha} > 0$ if $u_{c\alpha} \geq 0$ and $\frac{\partial \lambda}{\partial \alpha} \gtrless 0$ if $u_{c\alpha} < 0$.

The effect on hours of work, $\frac{de}{d\alpha}$, will then be negative unless the income effect, which would now be positive, dominates ($\frac{\partial\lambda}{\partial\alpha}w > -u_{e\alpha}$); or unless consumption and leisure are strong substitutes ($u_{c\alpha} < \frac{w}{p}\frac{u_{e\alpha}}{u_{ee}}u_{cc} < 0$).

Effect on hours of work with productivity effects When productivity increases with air quality, and this affects equilibrium wages, equations (21) and (25) become

$$\frac{de}{d\alpha} = \frac{\frac{\partial\lambda}{\partial\alpha}w + \frac{\partial w}{\partial\alpha}\lambda + u_{e\alpha}}{-u_{ee}} \tag{24}$$

$$w\frac{de}{d\alpha} + \frac{\partial w}{\partial\alpha}e - p\frac{dc}{d\alpha} = 0 \tag{25}$$

and $\partial\lambda/\partial\alpha$ becomes

$$\frac{\partial\lambda}{\partial\alpha} = \frac{w\frac{u_{e\alpha}}{u_{ee}} - p\frac{u_{c\alpha}}{u_{cc}}}{-\left[\frac{w^2}{u_{ee}} + \frac{p^2}{u_{cc}}\right]} + \frac{dw}{d\alpha}\frac{\left[\frac{w\lambda}{u_{ee}} - e\right]}{-\left[\frac{w^2}{u_{ee}} + \frac{p^2}{u_{cc}}\right]} \tag{26}$$

Expression (26) has two components. The first component is identical to (23), so its sign will be determined by the sign and size of $u_{c\alpha}$. The second component corresponds to the traditional income effect from increased wages: higher earnings per hour worked ($\frac{dw}{d\alpha}e$) and higher earnings from additional hours worked ($\frac{dw}{d\alpha}\frac{w\lambda}{u_{ee}}$) will reduce the marginal utility from income.

The total effect of air quality on hours worked is given by Equation (24). Higher wages will increase the substitution effect on the labor supply , which can be seen from the additional positive term in (26) compared to (23), $\frac{\partial w}{\partial\alpha}\lambda > 0$. However, as discussed in the previous paragraph, the income effect from higher wages will pull labor supply in the opposite direction. For high labor supply (high e), the income effect could even outweigh the substitution effect. However, this would require a negative labor supply elasticity, which has not been documented empirically (see Arceo and Campos 2010).

Table 1: Sample Statistics

	1989-1993 (1)	Pre-Closure (2)	Post-Closure (3)
Panel A: Pollution Concentration			
Sulfur Dioxide (SO$_2$)	4.46	5.02	4.04
	(2.13)	(1.49)	(2.42)
	143311	61438	81873
Panel B: Labor Outcomes			
Hours Worked	38.07	37.79	38.27
	(18.94)	(18.62)	(19.18)
	143311	61438	81873
Unemployed	0.04	0.04	0.04
	(0.20)	(0.20)	(0.20)
	143311	61438	81873
Weekly Wage	798.49	590.12	963.77
	(695.45)	(516.85)	(770.50)
	81246	35939	45307
Panel C: Demographics			
Female	0.36	0.36	0.37
	(0.48)	(0.48)	(0.48)
	143311	61438	81873
Age > 50	0.15	0.15	0.15
	(0.35)	(0.35)	(0.35)
	143311	61438	81873
Indicator for Children Under 5	0.37	0.37	0.36
	(0.48)	(0.48)	(0.48)
	143311	61438	81873
Panel D: Job Characteristics			
Blue Collar Work	0.83	0.83	0.83
	(0.38)	(0.38)	(0.38)
	137221	58745	78476
Paid by Salary	0.69	0.69	0.69
	(0.46)	(0.46)	(0.46)
	130793	56058	74735

Notes: This table provides means for the key variables in the regression analysis, as well as standard deviations (in parenthesis) and number of non-missing observations listed below each mean. Column 1 provides the sample statistics for the full sample, while Columns 2 and 3 provide sample statistics for the pre- and post-closure periods, respectively. Pollution data comes from the Sistema de Monitoreo Atmosferico de la Ciudad de Mexico (www.sma.df.gob.mx). Labor outcomes, job characteristics, and demographic data were obtained from the National Survey of Urban Employment (ENEU) conducted by the Instituto Nacional de Estadistica y Geografia (INEGI). For those who are unemployed, hours worked in the last week is defined as zero. Wage is defined as the weekly wage in pesos for salaried employees.

Table 2: Correlations Between Pollution and Labor Outcomes

	ln(Hours Worked)			Unemployment		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Sulfur Dioxide)	-0.01278	-0.02494*	-0.02357*	-0.00003***	-0.00003	-0.00003
	(0.00776)	(0.01341)	(0.01305)	(0.00001)	(0.00002)	(0.00002)
Week FE		X	X		X	X
Census Block FE		X	X		X	X
Demographic Controls			X			X

Notes: This table provides estimates of the association between sulfur dioxide and the labor market outcomes. All regressions are estimated using OLS, with standard errors clustered at the census block level. Demographic controls include an indicator variable for gender, age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. For those who are unemployed, hours worked in the last week is defined as zero. The unemployment dummy is divided by 100 to account for the level-log specification. All regressions have 143,311 observations. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 3: First Stage Regressions

	Natural Log of Sulfur Dioxide			
	(1)	(2)	(3)	(4)
Post * Close	-0.0828***			
	(0.0227)			
Post * Distance		0.0067***	0.0067***	0.0096***
		(0.0016)	(0.0016)	(0.0021)
Week FE	X	X	X	X
Census Block FE	X	X	X	X
Demographic Controls			X	X
Census Block Specific Year Trends				X

Notes: This table provides the coefficient estimates of the effect of the refinery closure on sulfur dioxide. All regressions are estimated using OLS, with standard errors clustered at the census block level. Demographic controls include an indicator variable for gender, age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. All regressions have 143,311 observations. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 4: Reduced Form and IV Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln(Hours Worked)				Unemployment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Reduced Form Equation								
Post * Close	0.0552***				-0.0110***			
	(0.0178)				(0.0035)			
Post * Distance		-0.0032**	-0.0029*	-0.0060**		0.0007***	0.0006**	0.0006
		(0.0015)	(0.0015)	(0.0024)		(0.0002)	(0.0002)	(0.0005)
Panel B: Instrumental Variables Estimation								
ln(Sulfur Dioxide)	-0.6672***	-0.4806**	-0.4275**	-0.6095**	0.0013**	0.0010**	0.0009**	0.0007
	(0.2466)	(0.2231)	(0.2179)	(0.2632)	(0.0006)	(0.0004)	(0.0004)	(0.0005)
Instrument: Post * Close	X				X			
Instrument: Post * Distance		X	X	X		X	X	X
Week FE	X	X	X	X	X	X	X	X
Census Block FE	X	X	X	X	X	X	X	X
Demographic Controls			X	X			X	X
Census Block Specific Year Trends				X				X

Notes: Panel A presents the coefficient estimates of the reduced form estimate of the effect of the refinery closure on the labor market outcomes; Panel B presents the instrumental variables estimates of the effect of sulfur dioxide on the labor market outcomes. The outcome of interest in Columns 1-4 is the natural log of hours worked in the last week (with '0' included for the unemployed); the outcome of interest in Columns 5-8 is an indicator variable for unemployment. For the estimates in Panel B, the unemployment dummy is divided by 100 to account for the level-log specification. Demographic controls include an indicator variable for gender, age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. All regressions have 143,311 observations. Standard errors (listed below each estimate in parenthesis) are clustered at the census block level. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 5: Effect on Wages

	ln(Weekly Wage)			
	(1)	(2)	(3)	(4)
Post * Close	-0.0137			
	(0.0259)			
Post * Distance		-0.0014	-0.0005	-0.0019
		(0.0015)	(0.0012)	(0.0015)
Week FE	X	X	X	X
Census Block FE	X	X	X	X
Demographic Controls			X	X
Census Block Specific Year Trends				X

Notes: This table presents the coefficient estimates of the reduced form estimate of the effect of the refinery closure on the natural log of weekly wages for salaried workers. Demographic controls include an indicator variable for gender, age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. All regressions have 81,246 observations. Standard errors (listed below each estimate in parenthesis) are clustered at the census block level. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 6: IV Estimates Exploiting Wind Patterns and Altitude

	(1)	(2)	(3)
Panel A: First Stage, ln(Sulfur Dioxide)			
Post * Distance * Wind Speed	0.0064	0.0064	-0.0160
	(0.0071)	(0.0071)	(0.0108)
Post * Distance * Wind Direction	0.0001***	0.0001***	0.0002***
	(0.0000)	(0.0000)	(0.0001)
Post * Distance * Elevation	0.0343	0.0342	-0.0396
	(0.0448)	(0.0448)	(0.0493)
Panel B: ln(Hours Worked)			
ln(Sulfur Dioxide)	-0.5472**	-0.5677**	-0.0951
	(0.2204)	(0.2212)	(0.2502)
Kleibergen-Paap rk Wald F statistic	8.263	8.27	5.36
Stock-Yogo weak ID test critical value 10%	6.46	6.46	6.46
Panel C: Unemployment			
ln(Sulfur Dioxide)	0.0007*	0.0008**	0.0004
	(0.0004)	(0.0004)	(0.0004)
Panel D: ln(Weekly Wage)			
ln(Sulfur Dioxide)	0.1645	0.1701	-0.0901
	(0.1927)	(0.1677)	(0.1328)
Week FE	X	X	X
Census Block FE	X	X	X
Demographic Controls		X	X
Year Trends			X

Notes: Panel A presents the coefficient estimates of the differential effect of the refinery closure on Sulfur Dioxide based on wind patterns and elevation. F-test and p-value for the joint test of the instruments is provided below the coefficient estimates. Panels B-D presents the instrumental variables estimates of the effect of sulfur dioxide on the labor market outcomes. The outcome of interest in Panel B is the natural log of hours worked in the last week (with '0' included for the unemployed). The outcome of interest in Panel C is an indicator variable for unemployment; the unemployment dummy is divided by 100 to account for the level-log specification. The outcome of interest in Panel D is the log weekly wage of salaried employees. Observations for all regressions in Panels A-C are 137,184, and for Panel D is 77,788. Demographic controls include an indicator variable for gender, age, a set of indicator variables for educational attainment, and a set of indicator variables for marital status. Standard errors (listed below each estimate in parenthesis) are clustered at the census block level. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 7: Migration

| | Migrated | Demographics | | | | |
| | | Female | Age | Married | Low Education | High Education |
	(1)	(2)	(3)	(4)	(5)	(6)
		Panel A: Week FE and Census Block FE				
Post * Distance	0.0001	0.0011	-0.0200	-0.0017*	0.0001	-0.0006
	(0.0003)	(0.0007)	(0.0199)	(0.0009)	(0.0008)	(0.0007)
Joint Test P-Value						0.2747
Observations	51,518	143,311	143,311	143,302	143,292	143,292
		Panel B: Week FE, Census Block FE, and Census Block Specific Year Trends				
Post * Distance	0.0002	0.0008	0.0101	-0.0017	-0.0003	0.0002
	(0.0005)	(0.0011)	(0.0290)	(0.0012)	(0.0010)	(0.0009)
Joint Test P-Value						0.7868
Observations	51,518	143,311	143,311	143,302	143,292	143,292

Notes: This table provides the coefficient estimates of the changes of the refinery closure on a dummy for migration and several key demographic characteristics. In Column 1, an observation is defined at the level of the household-quarter-year, and the outcome variable is an indicator for whether the household moved in that quarter, zero otherwise. In Columns 2-6, an observation is defined at the level of individual-quarter-year. Married is an indicator variable for being married or cohabitating with someone. Low education is defined as having completed at most junior high school level formal schooling. High education is defined as having had some university or higher formal schooling. The p-value for the joint test of the treatment on all demographic characteristics in Columns 2-6 is provided in Column 6. The estimates in Panel A are conditional on week and census block fixed effects; the estimates in Panel B are additionally conditional on census block-specific year-trends. All regressions are estimated using OLS. Standard errors are clustered at the census block level, and are reported in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

|	(1)	(2)	(3)	(4)	(5)	(6)
ln(Sulfur Dioxide)	-0.7746**	-0.3380	-0.3985	-0.5902**	-0.5716	0.0241
	(0.3394)	(0.3330)	(0.3787)	(0.2960)	(0.5744)	(0.3101)
ln(Sulfur Dioxide) * Female	-0.5141		-0.2470			
	(0.4839)		(0.4671)			
ln(Sulfur Dioxide) * Children Under 5		-0.4714	-0.9423*			
		(0.3815)	(0.4918)			
ln(Sulfur Dioxide) * Female * Children Under 5			-0.2470			
			(0.4671)			
ln(Sulfur Dioxide) x Age>50				-0.6067		
				(0.5441)		
ln(Sulfur Dioxide) x Blue Collar					0.1039	
					(0.4916)	
ln(Sulfur Dioxide) x Salaried						-0.6686*
						(0.3747)
Kleibergen-Paap rk Wald F statistic	8.687	16.39	2.258	8.826	4.642	19.20
Stock-Yogo weak ID test critical value 10%	7.03	7.03	NA	7.03	7.03	7.03
Observations	143,311	143,311	143,311	143,311	137,221	130,793

Notes: This table explores the heterogeneity of the treatment effects. All regressions include week fixed effects, census block fixed effects, demographic controls, and census block-specific year-trends. Standard errors are clustered at the census block level, and are reported in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.
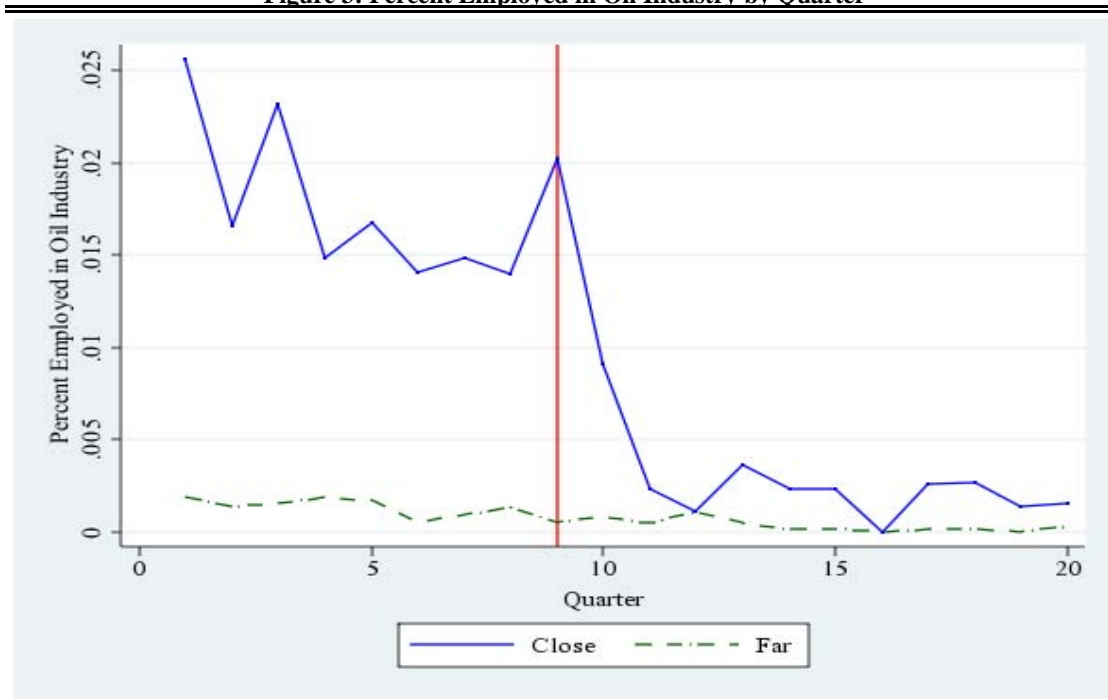
Figure 1: Maps of AGEBs and Refinery
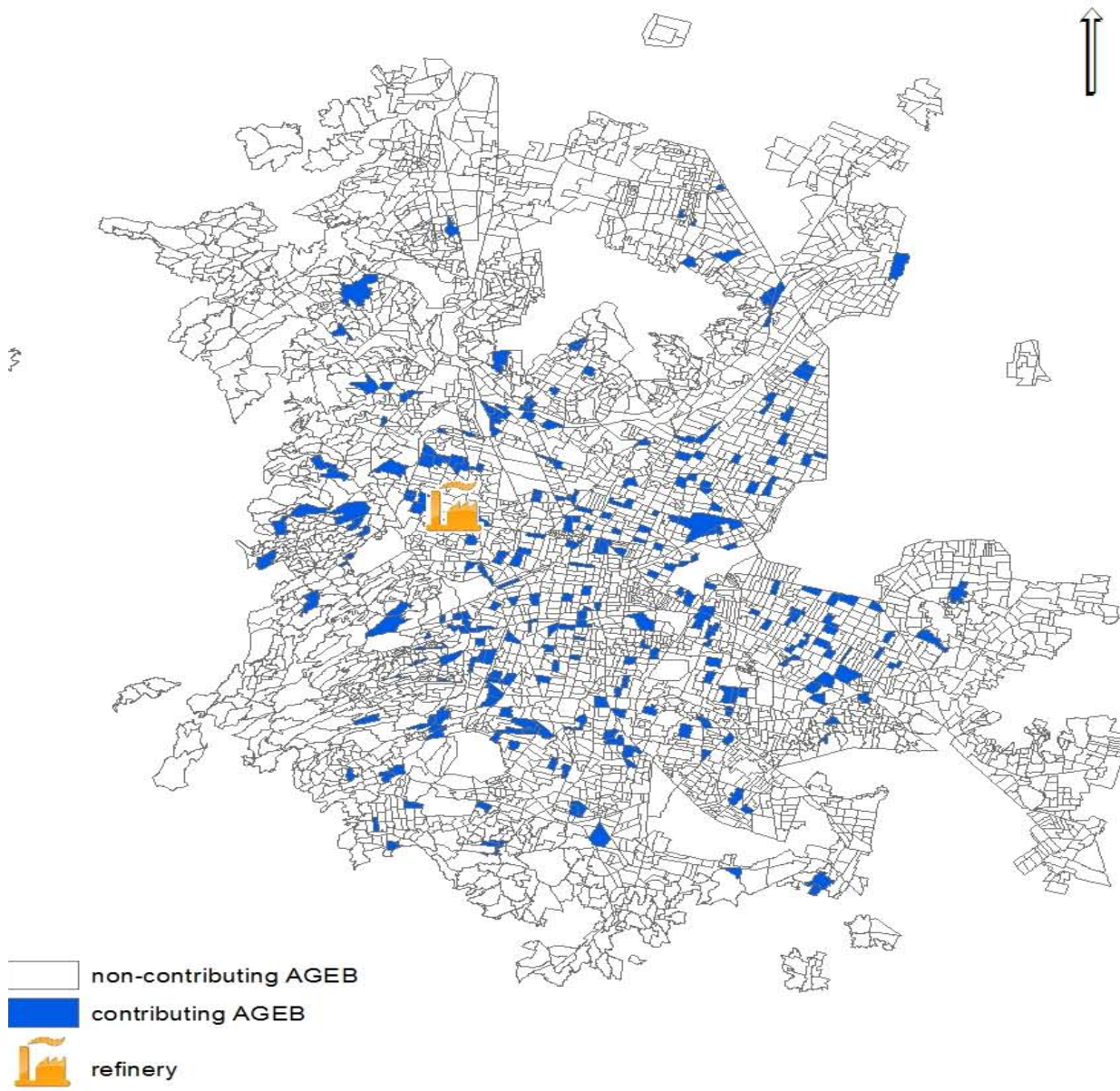
Notes: An AGEB is a census block. The blue colored census blocks were included in the analysis.

**Figure 2: SO₂ by Hour of Day, Pre and Post Closure**
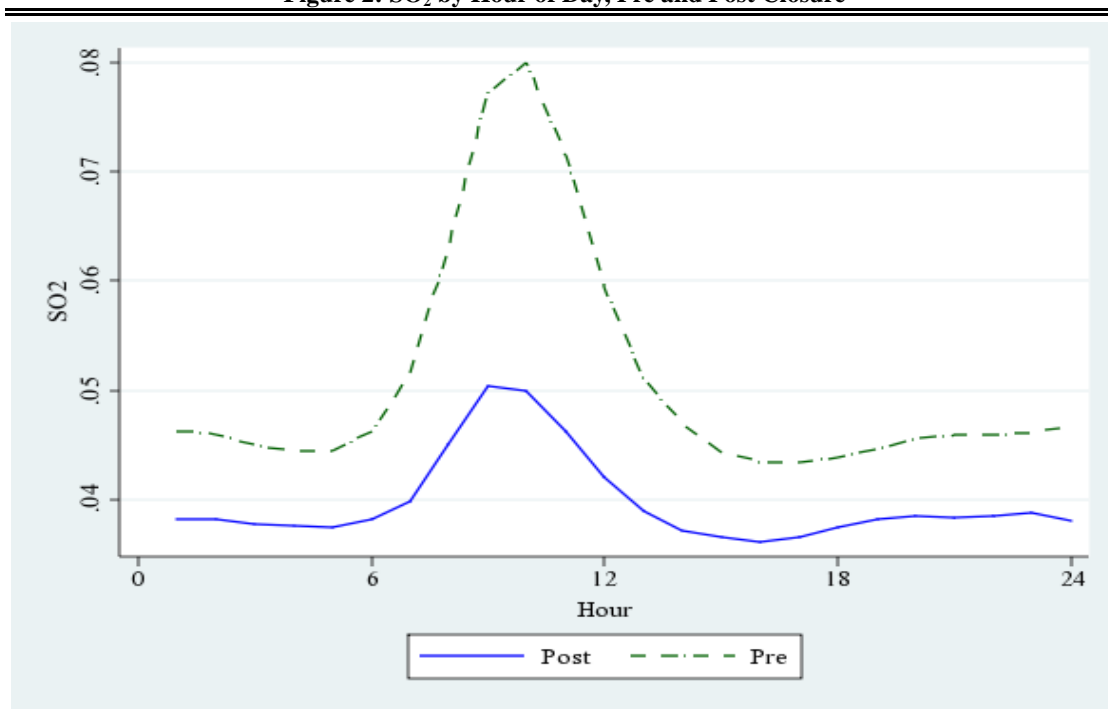
Note: Data is averaged at the hourly level. Equal weight is given to each station.

**Figure 3: Percent Employed in Oil Industry by Quarter**

Notes: Close was defined as the centroid of the respective census block being within 5km of the refinery, and far was correspondingly more than 5km from the refinery. Observations are at the quarter level, starting at Q1 1989. Employment in the oil industry was defined as all workers whose job coding corresponded to the oil industry, except for extraction.

Figure 4: Pollution over Time by Distance from the Refinery
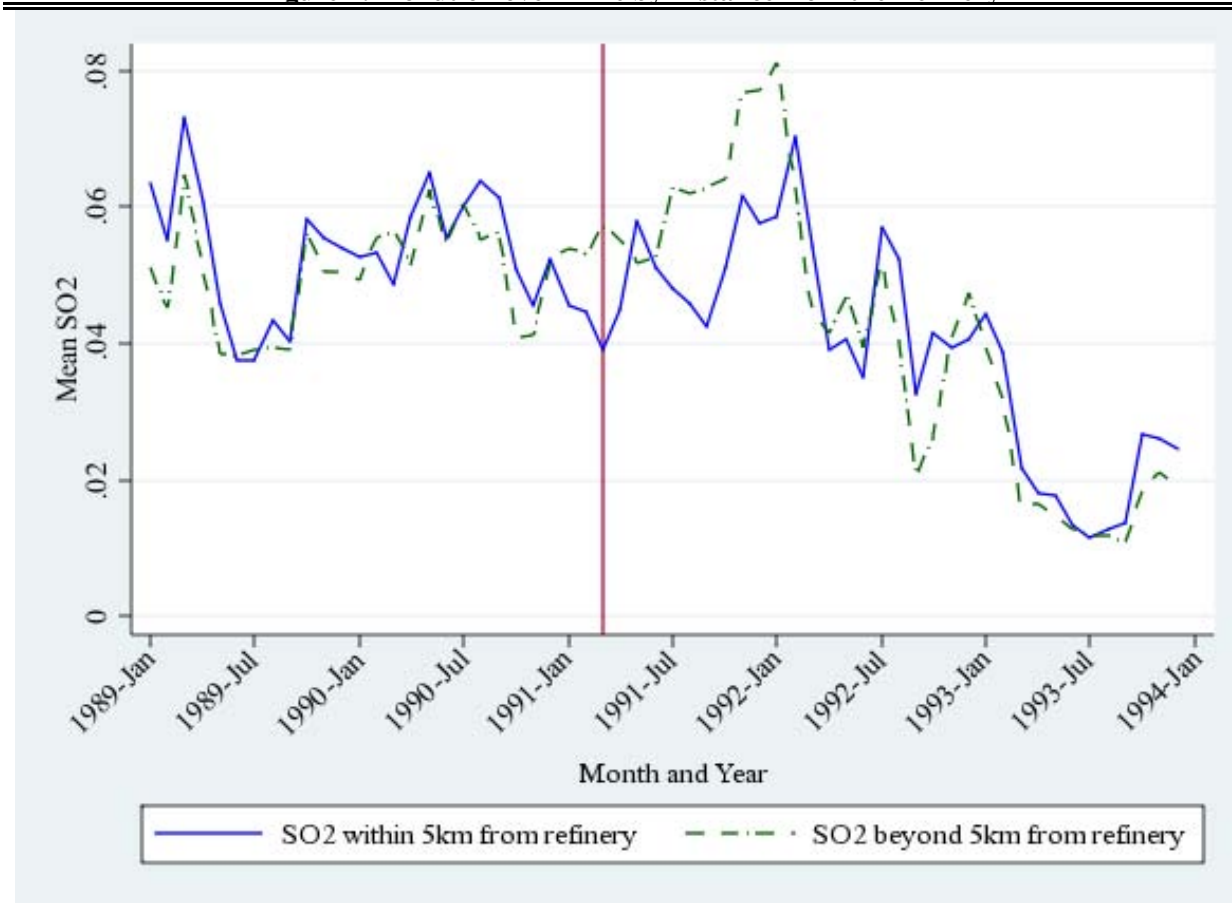
Note: Data in the solid line represents monthly averages across AGEB's that fall within 5km of distance near the refinery and data in the dashed line represents weekly averages of SO_2 across AGEB's that are beyond 5km from the refinery.

Figure 5: Migration Rates over Time, By Distance from the Refinery
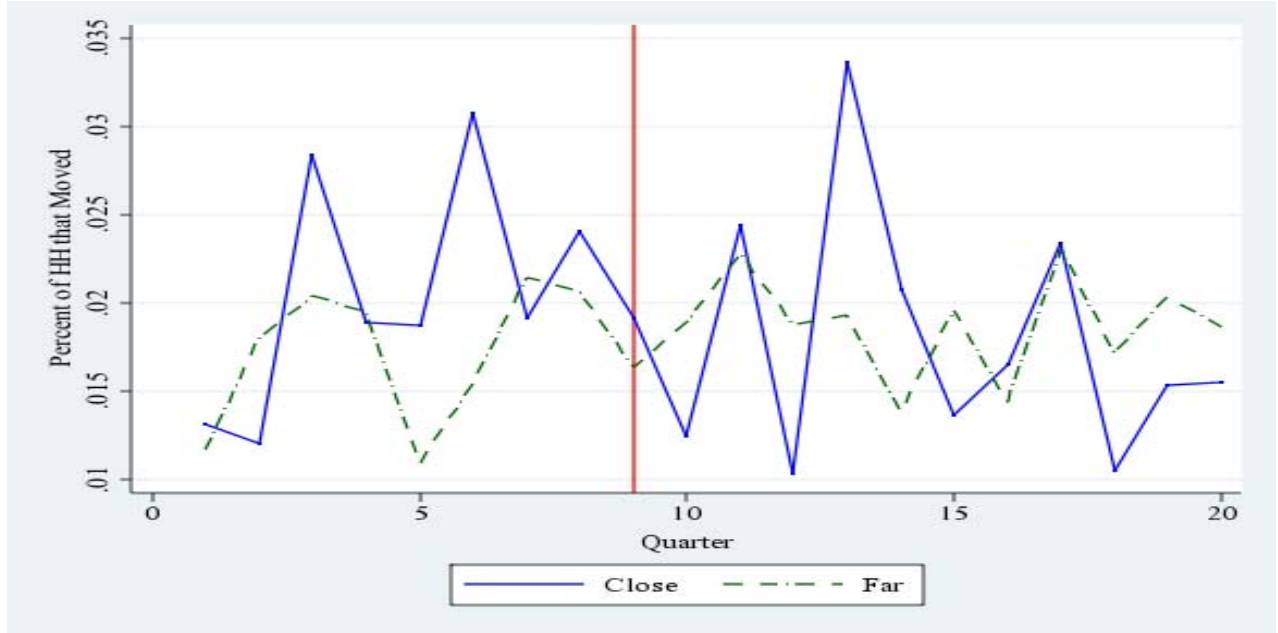
Appendix Table 1: Match Rate

	Total		Distrito Federal		Estado de Mexico	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Census Block Match Rate						
Total Census Blocks	481		242		239	
Matched to Location	389	80.9%	199	82.2%	190	79.5%
Within 8km of a Station	265	55.1%	190	78.5%	75	31.4%
Panel B: Observations in Labor Supply Dataset Matched to Location						
Number of Observations	385586		234813		150773	
Number in Labor Force	209303	54.3%	129839	55.3%	79464	52.7%
Within 8km of a Station	143311	37.2%	105545	44.9%	37766	25.0%

Notes: Census Blocks matched to a location were defined as those that we were able to generate accurate centroid coordinates.

Appendix Table 2: Comparing Included and Excluded Census Blocks

	Unemployment (1)	ln(Hours Worked) (2)	Unemployment (3)	ln(Hours Worked) (4)
Matched to Location	-0.0000	-0.0109		
	(0.0019)	(0.0312)		
Matched and Within 8KM of a Station			0.0039***	-0.0829***
			(0.0012)	(0.0162)
Year FE	X	X	X	X
Observations	1,946	1,946	1,946	1,946

Notes: This table compares the unemployment rate and the mean natural log of hours worked for included and excluded census blocks. An observation is a census block-year. In Columns 1 and 3, the variable of interest is an indicator for whether the census block was matched to GIS data; in Columns 2 and 4, the variable of interest is an indicator for whether the census block was both matched to the GIS data and within 8 kilometers of a station. All regressions are estimated using OLS, and include year fixed effects. Standard errors are clustered at the census block level, and are reported in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table 3: Employment in Oil Industry, by Quarter

Year	Quarter	Number in Labor Force (1)	Percent Employed (2)	Percent Employed in Oil Industry (3)
1989	1	6774	94.86%	0.48%
1989	2	6783	95.18%	0.31%
1989	3	7001	94.96%	0.42%
1989	4	6982	95.75%	0.36%
1990	1	6897	96.10%	0.36%
1990	2	6600	95.42%	0.25%
1990	3	6625	95.62%	0.30%
1990	4	7053	96.37%	0.29%
1991	1	6723	96.70%	0.29%
1991	2	7276	96.95%	0.10%
1991	3	7516	96.00%	0.06%
1991	4	7447	96.17%	0.10%
1992	1	7491	95.96%	0.06%
1992	2	7470	95.96%	0.04%
1992	3	7454	96.18%	0.04%
1992	4	7140	96.13%	0.00%
1993	1	7436	94.76%	0.03%
1993	2	7563	95.74%	0.04%
1993	3	7663	95.03%	0.01%
1993	4	7417	95.81%	0.04%

Notes: This table provides the employment rates in the labor force as a whole, as well as in the oil industry specifically, at the quarter level from 1989-1993. Employment in the oil industry was defined as all workers whose job coding corresponded to the oil industry, except for extraction. The dotted line signifies the refinery closure.