

# The Transitional Costs of Sectoral Reallocation: Evidence From the Clean Air Act and the Workforce\*

W. Reed Walker<sup>†</sup>

*Columbia University*

September 2011

JOB MARKET PAPER

## Abstract

New environmental regulations lead to a rearrangement of production away from polluting industries, and workers in those industries are disproportionately affected. This paper uses linked worker-firm data in the United States to estimate the transitional costs associated with reallocating workers from newly regulated industries to other sectors of the economy. Using panel variation induced by the 1990 Clean Air Act Amendments (CAAA), I find that the reallocative costs of environmental policy are significant. Workers in newly regulated plants the year prior to the regulations experienced, in aggregate, more than \$9 billion in foregone earnings for the 6 years after the change in policy. Most of these costs are driven by non-employment and lower wages in future employment, highlighting the importance of longitudinal data when characterizing the cost and incidence of labor market innovations. Relative to the estimated benefits of the 1990 CAAA, these one time transitional costs are small. However, the estimated costs far exceed the workforce compensation policies currently in place designed to mitigate some of these earnings losses.

---

\*I would like to thank Janet Currie, Bernard Salanie, Wolfram Schlenker, and Till von Wachter for invaluable advice and discussions. I would also like to thank Alex Chinco, Lucas Davis, Olivier Deschenes, Walker Hanlon, Solomon Hsiang, Wojciech Kopczuk, Matt Kotchen, Todd Kumler, Erin Mansur, Matt Nodowidigdo, Johannes Schmieder, Eric Verhoogen, Jonathan Vogel, and seminar participants at the Census Bureau, the NBER Environmental and Energy Economics meetings, and Columbia University for useful comments. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following NSF Grants SES-9978093, SES-0339191 and ITR-0427889; NIA Grant AG018854; and grants from the Alfred P. Sloan Foundation. Support for this research from the EPA Grant 834259010 is gratefully acknowledged.

<sup>†</sup>Address: 420 West 118th Street, New York, NY 10027; email: [rw2157@columbia.edu](mailto:rw2157@columbia.edu).

# 1 Introduction

Government intervention into markets is controversial, and perhaps no other federal program embodies this controversy greater than the Clean Air Act (CAA) in the United States. Environmental policy pertaining to air pollution has been estimated to have large health benefits (Environmental Protection Agency 1999). However, these policies also come with costs. Production is typically reallocated from newly regulated industries and locations elsewhere within the economy (Henderson 1996, Greenstone 2002). Rearrangement of production from pollution-intensive to other industries creates a broad set of private and social costs. In terms of labor inputs, this reallocation is often framed in terms of “jobs lost”, and the distinction between “jobs versus the environment” is one of the more politically salient aspects of these regulations.<sup>1</sup> However, workers often find new jobs elsewhere, perhaps in different locations and/or industries. If workers simply transition from one employer to the next without any significant earnings loss, then job loss should not be a cost when evaluating policy. If workers lose job or industry specific skills and/or experience long periods of non-employment following layoffs, the cost of reallocating the workforce could be quite large. Moreover, there are also costs to workers who remain in these potentially less productive industries.

This paper uses new linked worker-firm data to estimate the magnitude and duration of adjustment costs pertaining to worker reallocation in the context of the Clean Air Act Amendments of 1990. The 1990 CAA Amendments created a new class of pollution standards and strengthened existing standards so that many areas of the United States fell under a new regulatory regime. Polluting firms in these areas were forced to reduce emissions by installing new pollution abatement technologies, increasing the cost of production. The empirical framework estimates the impact of this regulatory shock on the evolution of wages for the workers in newly regulated sectors. In doing so, I offer a new approach to understanding the costs and incidence of sector specific shocks in an economy where production and workers are not costlessly reallocated. Using confidential linked worker-firm data from the United States Census Bureau, I am able to follow workers across their job histories over time. This allows me to account for two substantive features of labor market innovations: the costs borne by workers who remain in a less productive sector and the long run earnings losses for those who leave the sector.

Prior research on the labor market impacts of environmental regulation has primarily focused on employment in manufacturing industries (Berman & Bui 2001, Greenstone 2002, Deschenes 2010, Kahn & Mansur 2010, Walker 2011). For example, Greenstone (2002) finds that the Clean Air Act Amendments of the 1970’s led to a loss of more than 500,000 jobs in regulated sectors relative to unregulated sectors.

---

<sup>1</sup>For representative examples from the popular press see the *Wall Street Journal* (July 26, 2011) op-ed entitled “The Latest Job Killer from the EPA”, “Smoke Signals” from *The New Republic* (April 7th, 2011), or “A Debate Arises on Job Creation and Environment” from *The New York Times* (Sep 4th, 2011).

Monetizing these effects from “jobs lost” is difficult if we do not know the actual costs borne by these displaced workers. As such, the appropriate measure of regulatory costs to the workforce should not be characterized by jobs lost but by any transitional costs associated with reallocating production or workers (Jaffe et al. 1995, Arrow et al. 1996, Greenstone 2002, Environmental Protection Agency 2005, Fullerton 2008, Congressional Budget Office 2009). Even though this fact has been pointed out in numerous papers, little to no work has attempted to understand the magnitudes of these frictions in the context of environmental regulation, or indeed in terms of other types of policies that also lead to labor reallocation.<sup>2</sup> Moreover, these costs have important distributional implications for regulatory policy (Fullerton 2008).

The impact of the regulations on workers is interpreted using a model of occupational choice. Workers self-select into industries based on their comparative advantage. Reductions in sectoral productivity reduce wages in equilibrium, causing least able workers to leave the sector. Both inframarginal workers who remain in the sector and workers who leave the sector bear a cost in terms of foregone earnings. The model also highlights the econometric challenges of estimating either of these wage components in isolation; namely, workers who remain in the sector are positively selected and workers who leave sector are negatively selected.

Longitudinal linked worker-firm data are critical to estimating the wage costs of these sector specific shocks. Since declines in labor demand are generally accommodated by laying off the least senior workers (Abraham & Medoff 1984, Farber 1993), industry wage estimates are biased by compositional changes in the newly regulated workforce. Moreover, without longitudinal data, it is difficult to observe any future earnings losses associated with job transitions. In contrast, by following the same group of workers over time, cohort wage estimates are not biased by compositional changes and any future earnings losses are explicitly observed for all workers in the newly regulated sector.

The primary estimation framework follows cohorts of workers in newly regulated counties and sectors over time, before and after specific regulatory changes. Since the Clean Air Act is enforced annually at the county level, regulations vary across both time and space. This panel variation facilitates the use of econometric models that control for both fixed unobservable attributes specific to these groups of workers as well as any national or sectoral trends that may covary with these county level regulatory changes. The cohort style analysis is meant to address concerns pertaining to compositional biases while also implicitly incorporating potentially costly job transitions into the average wage estimate. This

---

<sup>2</sup>There is a related empirical literature in international trade seeking to estimate the labor market implications associated with reductions in trade barriers. See Menezes-Filho & Muendler (2007), Artuc, Chaudhuri & McLaren (2010), Ebenstein et al. (2011), Autor, Dorn & Hanson (2011), and Dix-Carneiro (2011) for recent examples. The approach taken here is complimentary to this literature and offers a new approach to estimating the costs of labor market adjustment costs in the context of trade policy.

means that the baseline estimates consist of a weighted average of the earnings changes for workers who stayed at the firm and the earnings changes for those who separated.

The results in this paper suggest that the reallocative costs to the workforce from the 1990 Clean Air Act Amendments are significant. For those workers in the regulated sector prior to the change in regulation, the average earnings declined by more than 6 percent in the 3 years after the regulation. The earnings penalty persists for at least five years in most estimated models, with the total foregone wage bill of the economy estimated to be more than 9 billion dollars (1990\$).

To better understand the mechanisms generating these wage changes, I stratify the sample into those employees who stay with their initial firm versus those who leave to a new job in another location or sector. While a job separation is endogenous in this setting, comparisons across the two groups provide striking contrasts in the wage adjustment process for these sets of workers. Workers that remain at their firm in the newly affected industry experience no discernible wage change, whereas those who separate see both sizable and persistent earnings losses. These results suggest that previous research focusing on wages in affected industries misses important aspects of labor-market adjustment, namely long-term earnings losses associated with job separations. Lastly, in contrast to existing work on worker responses to innovations in labor demand, estimates here suggest that (in the context of this particular shock) most of the reallocation occurs through movements within rather than across local labor markets.

The rest of the paper is organized as follows: The subsequent section discusses the details of the Clean Air Act more fully with a brief discussion of the previous literature. To motivate the empirical analysis, Section 3 presents a conceptual framework that characterizes the incidence of the regulations on the affected workforce. Sections 4 and 5 discuss the research design and data while Sections 6 and 7 present an econometric framework and estimation results. Section 8 offers a short discussion of the results. Section 9 concludes.

## 2 The Clean Air Act and Environmental Regulation

Air pollution regulation in the United States is coordinated under the Clean Air Act, the largest environmental program in the United States. With the passage of the 1970 Clean Air Act Amendments, the EPA established national ambient air quality standards (NAAQS) which specify the minimum level of air quality acceptable for six criteria air pollutants.<sup>3</sup> These standards are enforced at the “air region” level where air regions broadly follow county boundaries. Counties or air regions exceeding air quality standards in a given year are designated as “nonattainment” for each particular pollutant exceeding the

---

<sup>3</sup>These pollutants consist of sulfur dioxide (SO<sub>2</sub>), particulates (TSP, PM<sub>2.5</sub>, and PM<sub>10</sub>), nitrogen dioxide (NO<sub>2</sub>), carbon monoxide (CO), ozone, and lead

threshold. Counties may be in nonattainment for more than one pollutant at a given time, and regulations may differ for each particular pollutant class. In any given year, some counties find themselves over these thresholds, while others do not.

Ambient air pollution is measured by EPA pollution monitors that take hourly/daily pollution readings for various pollutants. Monitor location is not subject to manipulation by local authorities. When a county is out of attainment for one of the regulated pollutants, the EPA requires states to adopt regulatory plans, known as State Implementation Plans (SIPs), which are then reviewed by the EPA to ensure that they meet statutory requirements. The EPA can impose sanctions in areas that fail to comply such as banning federal highway funding. The plans generally require polluting plants that emit the regulated pollutant to adopt “lowest achievable emission rates” (LAER) technologies without regard to costs. Furthermore, the EPA forces any new polluting plants that wish to locate in that particular county to offset their emissions from another polluting source within the county. In contrast, in areas designated as “attainment”, large polluting plants must use “best available control technology” (BACT), which is significantly less costly than LAER technology (Becker 2005). In summary, in nonattainment areas, firms are subject to regulations designed to reduce emissions without regard to costs; counties in attainment are more lightly regulated.

The Clean Air Act Amendments of 1990 introduced a new NAAQS standard (PM-10)<sup>4</sup> and formally required all polluters to obtain an operating permit for operation.<sup>5</sup> The EPA also formally evaluated their existing nonattainment designations for each air region so that 177 new counties found themselves in nonattainment for at least one pollutant (relative to 392 the year before the 1990 CAAA), a 45% increase. The requirement that polluting sources obtain operating permits is crucial for this study, as it allows one to observe regulatory status at the plant level, something that has not been possible in previous work in this literature.

In no small part due to these regulations, pollution levels have declined considerably from 1970 levels. More recently, between 1990-2007 pollution levels have declined even further despite GDP rising 63 percent, vehicle miles traveled increasing 45 percent, the United States population growing 21 percent, and energy consumption rising 20 percent (Environmental Protection Agency 2008). Achieving these reductions in air pollution are not without costs, and firms necessarily respond to the various regulatory changes. In some circumstances, this may involve installing more expensive pollution abatement technologies (Becker 2005). In other circumstances, plants on the margin of production may opt to close or

---

<sup>4</sup>The PM-10 standard replaced the existing TSP standard. Both of these are particulates, with the former being a smaller particulate thought to be more harmful to human health. In 2005, the PM-10 standard has been replaced by PM-2.5 for similar reasons.

<sup>5</sup>Title IV of the 1990 Clean Air Act Amendment established a tradeable market for sulfur dioxide emissions. These markets apply nationally, are primarily for electric utilities, and do not correspond to the variation in nonattainment designation used in this study.

relocate in response to these regulations (Henderson 1996, Becker & Henderson 2000). Regulations may also increase barriers to entry, reducing plant entry rates, increasing market concentration, and reducing output (List et al. 2003, Ryan 2011). In general, production and plant inputs are reallocated away from newly regulated industries and locations after these regulations go into effect (Greenstone 2002). More recent work, drawing upon better data and/or more recent time periods, supports this previous literature (Kahn & Mansur 2010, Walker 2011).

### 3 A Mapping Between Regulatory Change and Reallocation Incidence

This section presents a simple conceptual framework designed to motivate and interpret the subsequent empirical analysis. New regulations force plants to install pollution abatement technologies, so that for a given level of output plants must now use more inputs. Therefore, a straightforward way to model environmental regulations is to view these regulations as affecting plant level productivity.<sup>6</sup> This is consistent with recent empirical evidence suggesting that CAA nonattainment designation lowers plant level productivity (Greenstone, List & Syverson 2010). This modeling decision foregoes potential complementarities between pollution abatement and capital or labor while allowing for a more general treatment of labor reallocation in the context of productivity shocks.<sup>7</sup> In a competitive manufacturing industry, price is set on an international market, and a reduction in sectoral productivity will lead to a reduction in sectoral output as well as a reduction in the demand for inputs.<sup>8</sup>

In order to map these changes in factor demands into worker outcomes, consider the following stylized model occupational choice. A population of  $N$  workers supply an indivisible unit of labor to a sector  $s \in \{p, o\}$  (i.e. the polluting sector  $p$  or the outside sector  $o$ ). The outside sector is large relative to the polluting sector so that population movements from the polluting sector to the outside sector do not generate general equilibrium changes in outside sector wages. Workers are heterogeneous with respect to their sector specific ability, where ability  $\epsilon_{is}$  is drawn from a mean zero distribution  $F_s$  with full support over  $(-\infty, \infty)$ . In each sector, workers are compensated according to their ability so that workers self select into the sector that gives them the highest expected earnings.<sup>9</sup> Denote log earnings

---

<sup>6</sup>There are many ways to model pollution or pollution abatement within a plant level production decision (see e.g. Copeland & Taylor (2004)).

<sup>7</sup>If environmental regulations are factor biased (e.g. by increasing the rental rate of capital), the total effect of regulation on labor demand is ambiguous depending on both the marginal rate of technical substitution between capital and labor and any “scale effect” associated with changes in industry output.

<sup>8</sup>In a dynamic setting, there may also be interactions between plants’ endogenous production decisions in the context of these regulations (e.g. Jovanovic (1982) and Hopenhayn (1992)). Moreover, shifts in the costs of entry and investment can lead to markets with fewer firms and lower production (Ryan 2011).

<sup>9</sup>Similar to the canonical Roy Model (Roy 1951), workers have skills in each sector, but they can only use one skill so that workers self select into the sector that gives them the highest expected earnings.

for a worker  $i$  in sector  $s$  as

$$\omega_{is} = w_s + \epsilon_{is}$$

where  $w_s$  is the average ability for a worker in sector  $s$  and  $\epsilon_{is}$  is worker specific ability. Assuming the worker consumes their earnings in each period and utility is logarithmic, the present discounted value of future utility for a worker  $i$  choosing to work in sector  $s$  may be written as

$$U_{is} = \sum_{t=0}^{\infty} \beta^t (w_s + \epsilon_{is}).$$

where  $\beta$  is the common discount factor. The worker chooses a sector  $s$  that maximizes their present discounted value of future utility with full knowledge of their ability component in each sector. This maximization problem implies a decision rule whereby a worker chooses to work in sector  $p$  if and only if

$$U_{ip} > U_{io}.$$

Workers self select into the sector for which they have comparative advantage. Define the indicator variable  $D_{is} = 1$  when sector  $s$  satisfies individual  $i$ 's maximization criterion. The total utility of the workers in sector  $s$  may be written as

$$W_s = \sum_{i=1}^N \sum_{t=0}^{\infty} \beta^t (D_{is} \times (w_s + \epsilon_{is}))$$

which is simply equal to the total discounted wage bill in sector  $s$  from  $t = 0$  to the infinite future.

### 3.1 The Incidence of Regulation for Polluting Sector Workers

Consider the effects of environmental regulations on those workers in the polluting sector. I model regulations as a factor neutral productivity shock to the polluting sector. As a result of this shock, output falls and input demand falls. In a competitive labor market where wages are set on a demand curve, a reduction in labor demand will reduce the wage in equilibrium from  $\omega_{ip}$  to  $\omega'_{ip}$ . As skill prices vary, individuals shift sectors in pursuit of their comparative advantage.

A fraction of workers who originally decided upon the polluting sector are now better off in the outside sector. In the new regulatory regime, some workers who originally chose sector  $p$  will now choose sector  $o$ . The structure of the model also tells us which workers switch, namely those who are least able. This implies the ability of workers in the polluting sector is *higher* after the regulations go into place:

$$E(\epsilon_{ip} | w'_p + \epsilon_{ip} > w_o + \epsilon_{io}) > E(\epsilon_{ip} | w_p + \epsilon_{ip} > w_o + \epsilon_{io}).$$

The average ability of workers in the polluting sector rises since the least productive workers switch to the outside sector. Therefore, while wages fall in the newly regulated sector due to lower productivity, the *average* wage may actually rise because the remaining workers are, on average, more productive than before. This reflects the well known empirical issue pertaining to composition bias in industry wage regressions (see e.g. Solon, Barsky & Parker (1994)).

Now consider the effect of the change in regulations on the workers in the polluting sector  $p$ . Denote the set of workers in the polluting sector prior to the policy change as  $N_p$ :

$$N_p = \{\epsilon_{ip} : w_p + \epsilon_{ip} > w_o + \epsilon_{io}\}.$$

Write the wage bill for the set of workers better off in sector  $o$  after the regulations as  $W_o|_{i \in N_p}$  and the wage bill of those who remain in sector  $p$  after the regulations as  $W'_p|_{i \in N_p}$ . Then we can write the change in welfare for workers in the polluting sector prior to the regulations as:

$$\begin{aligned} \Delta W &= (W'_p|_{i \in N_p} + W_o|_{i \in N_p}) - W_p \\ &= \sum_{t=0}^{\infty} \beta^t \left[ \underbrace{\sum_{i \in N_p} D'_{ip} \times (\omega'_{ip} - \omega_{ip})}_{\text{Wage Cost for Stayers}} + \underbrace{D'_{io} \times (\omega_{io} - \omega_{ip})}_{\text{Wage Cost for Leavers}} \right] \end{aligned} \quad (1)$$

where  $D'_{is}$  is the re-optimized decision rule pertaining to sector  $s$  under the new regulatory regime. Equation (1) suggests that we can write the change in welfare for workers in the polluting sector as a sum of two components: the first component represents the wage costs for the inframarginal workers (i.e. those workers not induced to switch sectors by the productivity shock), and the second component represents the wage costs for those workers who are now better off in sector  $o$ . The latter term consists of the difference between the individual's wage in the outside sector  $\omega_{io}$  and the wage the worker received in the previous sector  $\omega_{ip}$ . Note that both terms inside the brackets are negative; the sectoral wage has fallen in the polluting sector  $\omega'_{ip} < \omega_{ip}$ , and we know that  $\omega_{io} < \omega_{ip}$  since any worker now in the outside sector  $o$  had optimally chosen sector  $p$  to begin with.

This framework highlights two important components pertaining to reallocative incidence in the context of any external labor market shock. First, the total change in welfare consists of both the wage changes for those who remain in the sector as well as any change in wages experienced by those who leave the sector. Second, even if the wage costs of leaving are small (i.e.  $\omega_{io} - \omega_{ip} \approx 0$ ), average industry wages are insufficient for characterizing incidence due to compositional changes in the workforce.

With this in mind, I develop an econometric framework to match the main features of equation (1). Namely, in this highly simplified model, equation (1) suggests that the present discounted foregone wage

bill is a sufficient statistic for the reallocative cost to workers in the polluting sector. The empirical analysis then focuses on all workers who had optimally chosen to work in the polluting sector *prior* to the change in regulations and then estimates the change in earnings for this subset of workers after the policy change. Since the theoretical framework explicitly suggests that workers who remain in the sector are positively selected and those who leave are negatively selected, the baseline regression estimates refrain from separating these two components, instead focusing on the average wage change for stayers and leavers. Lastly, since I am following the same group of workers before and after the policy change, concerns pertaining to composition bias no longer apply (i.e. the composition is explicitly held constant).

### 3.2 Regulatory Effects in General Equilibrium

While regulations affect workers in the polluting sector of the manufacturing industry, employment flows away from the regulated industry may affect factor prices in other, non-regulated industries or locations. This section highlights these relevant factor price adjustments in the context of a stylized general equilibrium model.

Consider an economy with two locations (regulated and unregulated) and two sectors in each location (polluting and non-polluting) that produce freely tradable goods using a single input (labor). Labor is heterogeneous in the sense implied by the previous section; workers differ in their sector specific productivity. Workers consume production goods. Assuming perfectly competitive markets, all prices (wages and prices for polluting and non-polluting goods) are set such that each market clears.

Regulations lower productivity in the polluting sector which will lower polluting sector wages in equilibrium. Marginal workers leave the polluting sector, finding work either in the non-polluting sector of the same location or in either sector of the unregulated location. This will increase the sectoral labor supply in all sectors, which will reduce the equilibrium wage in those sectors (depending on the total increase in labor supply and the labor demand elasticity of each sector). Since labor is relatively less expensive in the newly regulated sector, this should encourage new polluting firms to enter the market. However, EPA regulations explicitly prevent new polluting firms from entering, so that all wage adjustment occurs through worker out-migration (rather than new firms entering). These effects suggest that relative geographic comparisons (i.e. within county comparisons) will overstate changes in labor demand (i.e. double counting) and understate wage declines in the regulated location. These will be important considerations for interpreting the empirical estimates to follow, and I will present a series of tests speaking to the relative importance of these spillovers in the context of the research design.

## 4 The Clean Air Act as a Research Design

In testing the predictions outlined above, the main econometric challenge is that firms are not randomly subject to CAA regulations. In addition, workers in those firms are not randomly assigned to the firm. Instead, firm and worker decisions are made to maximize the present discounted value of future profits or utility. Wages vary tremendously across both firms and locations due to reasons observed and unobserved to the econometrician. Moreover, this variation in the wage structure is likely correlated with a firm's status as a polluter. For example, polluting firms tend to be both older and larger (Walker 2011), and larger plants tend to pay significant wage premia (Brown & Medoff 1989). Workers may also demand some form of compensating wage differential associated with potentially harmful working conditions. Therefore, a naive comparison of wages in polluting firms of nonattainment counties is likely to yield biased estimates. Credible estimates of the effects of new environmental regulations requires the identification of a group of firms and workers that are similar to those affected by the regulations in ways observable and unobservable to the econometrician.

The regulatory variation inherent in the design of the CAA may provide a solution to this identification problem. Due to the way in which the CAA is implemented, regulations vary over both time and space. Figure 1a shows counties in the United States that were in nonattainment for any pollutant in 1991. Since only some counties are regulated in a given year, it is possible to include flexible controls for any nationwide or industry specific shocks to employment and/or earnings such as the recession in the early 1990's or the phase in of the North American Free Trade Act throughout the 1990's. Temporal variation in regulations exists from counties that go in and out of nonattainment based on annual pollution levels, allowing for pre-post comparisons within counties, industries, plants, or workers. This means that any time-invariant unobservables unique to these groups may be controlled for by including a set of group-specific fixed effects while still controlling for nationwide trends as detailed above. Accordingly, the inclusion of group-specific fixed effects ensures that estimates are derived only from those sources that experience a *change* in the regulation - comparing outcomes before and after the change.

A potential issue with time series variation in light of nonattainment designation is that pollution is correlated with economic activity (Chay & Greenstone 2003), so that counties that switch into nonattainment in a given year may be quite different than those counties that do not.<sup>10</sup> To address this issue, this study relies on the discrete policy change induced by the 1990 CAA Amendments, which should be more or less orthogonal to confounding economic factors after controlling for national trends and time-invariant unobservables. Figure 1b plots only those counties that switched into nonattainment

---

<sup>10</sup>Note that the direction of bias here is clear, and that this type of confounding variation would lead me to understate the effects of the regulations.

status in 1991, the first year that the 1990 CAAA were adopted. The figure shows that a considerable number of counties were newly regulated with the passage of the 1990 CAAA.

Lastly, within any nonattainment county, only polluting plants are regulated and only if they emit the specific pollutant for which the county is in violation. Even within 2-digit industry SIC codes, there is a considerable amount of variation in the fraction of plants that are classified as polluters. Figure 2 plots the fraction of establishments in nonattainment counties that are classified as a polluter by the EPA, split by 2-digit manufacturing SIC codes. Since only the polluting firms within a given county-industry are regulated, it is possible to control for unobservable county level trends.

All of these sources of variation amount to a research design which examines the earnings outcomes of workers in polluting plants of newly regulated counties, before and after the introduction of the plant level regulations - all while controlling for fixed, unobservable attributes of these workers, national variation in the polluting/non-polluting sectors, and county level trends. The following section details how observable characteristics of affected and non-affected sectors may differ across these sources of regulatory variation while also informing the choice of econometric model.

## 5 Data Sources, Research Design, and Summary Statistics

This project relies upon the most detailed and comprehensive data currently available on both local labor markets as well as the program components of the CAA. This subsection briefly describes the data sources. Appendix A provides additional details.

**The Longitudinal Employer Household Dynamics Files (LEHD)** The primary source of data used in this project comes from the Census Bureau’s LEHD file which provides administrative, quarterly earnings records for a large percentage of the United States workforce.<sup>11</sup> I observe the complete employment history and earnings profile for each worker in the LEHD, conditional on the worker remaining within the LEHD states over the course of the sample. Since the administrative earnings records are based on firms’ reports that are used to calculate tax liabilities they are presumably free of measurement error, compared to existing survey data (Duncan & Hill 1985, Bound & Krueger 1991). The LEHD also provides important demographic information on workers such as age, race, and education as well as time-varying information from the firms at which they work. This provides an incredibly detailed snapshot of local labor markets at any given point in time.

The entire LEHD file consists of over 2.8 billion quarterly earnings records. To limit the computational burden of working with the complete file, I make some important sample restrictions. I begin by

---

<sup>11</sup>The full LEHD file currently contains earnings records from over 30 states, with some states containing records as far back as 1985. See Abowd et al. (2008) and McKinney & Vilhuber (2008) for a comprehensive discussion of both the construction and contents of the LEHD files.

restricting the analysis to states in the LEHD that have data prior to the 1991 implementation of the CAAA. This limits the analysis considerably, as the only states that contain 1990 quarterly earnings are Illinois, Maryland, Washington, and Wisconsin.<sup>12</sup> I then limit the sample to workers who worked in the manufacturing and electric/gas industries (i.e. 2 digit SIC 20-39, 49) in 1990.<sup>13</sup> This leaves me with a balanced panel of 2.7 million workers in 1990 that I track over the course of their next 10 years irrespective of whether or not they remain with their employer, transition outside of the manufacturing sector, or move across state lines.<sup>14</sup>

While the LEHD data are incredibly detailed in some regards, there are several important limitations to the data that bear mention. As is true with most linked worker-firm administrative datasets, it is not possible to distinguish between unemployment and non-participation.<sup>15</sup> Furthermore, the data do not allow the researcher to assess whether a missing earnings record is due to unemployment/non-participation or whether the worker moved outside of the states covered in the data. This complicates earnings estimates in that replacing a missing earnings record with a zero will tend to overstate the estimates pertaining to worker reallocation. In practice, I use annual earnings rather than quarterly earnings, and since the average duration of non-employment between jobs is around 3 quarters, there are much fewer missing earnings observations in the data relative to using quarterly observations. Nevertheless, to reduce biases due to sample attrition, I follow Jacobson, LaLonde & Sullivan (1993) in requiring workers to have at least some positive earnings in a particular year to contribute to the cohort earnings averages. I explore the sensitivity of the results to reclassifying missing data as zeros, while treating those estimates as a bound on the true value. See Appendix A.1 for further details.

**Longitudinal Business Database (LBD)** Since the LEHD contains wage records only as far back as 1990, and the CAA Amendments went into place in 1991, I draw upon another administrative dataset from the Census Bureau to assess whether there exists significant pre-trends across the main sources of identifying variation. The Longitudinal Business Database (LBD) is a plant level database that covers the universe of establishments in the United States from 1975-2005.<sup>16</sup> Included in the database is annual information on employment, payroll, and firm age. The database also includes information on detailed industry, location, and entry/exit years for the respective establishments. Thus, I can observe trends in

---

<sup>12</sup>The LEHD was assembled piecemeal from the individual states' UI and ES202 records. Thus, there is significant variation in when states start reporting to the LEHD. New states enter throughout the 1990's so that by 2000, the LEHD contains data from 30 states. The full list of states and temporal coverage is detailed in Appendix A and Table A.1.

<sup>13</sup>These industries constitute over 95% of the total number of plants regulated by the EPA.

<sup>14</sup>Over time, more states are covered in the LEHD up until 2002, when the last participating state (Arkansas) finally enters the sample. As states come into the LEHD, I include wage records from any individuals in the "1990 states" so that I can implicitly follow these workers across state lines (conditional on that state reporting).

<sup>15</sup>Using administrative data linked to survey data, Frijters & van der Klaauw (2006) find this distinction not very important for men, but more so for women. Accordingly, I estimate robustness specifications using a sample of only prime-age males 25-55.

<sup>16</sup>I only have access to the LBD from 1985-2005.

employment and earnings per worker before and after these changes for newly affected sectors relative to unaffected sectors. The LBD also allows me to revisit the previous literature pertaining to CAA regulation and sectoral employment (Greenstone 2002, Walker 2011) to see how sector size (as measured by employment) is related to variation in the regulations. I limit the LBD sample to plants in the manufacturing industry (SIC 20-39), the electric/gas industries (SIC 49), and plants residing in states covered by the LEHD sample.<sup>17</sup>

**EPA Air Facility Subsystem (AFS)** I match both the LEHD and LBD to plant-level, regulatory micro data from the EPA. An important aspect of this project is the ability to observe *plant* level regulatory status over time. Using the Census Bureau’s Standard Statistical Establishment List (SSEL), I am able to link the LEHD and LBD to plant level regulatory and permit data from the EPA’s Air Facility Subsystem (AFS) using a name and address matching algorithm. The exact details of the matching algorithm as well as the match rates are described in Appendix A.4.

The EPA’s AFS database is a plant level database that provides information detailing the regulatory programs for which the plant is permitted (and regulated) as well as the specific pollutants for which the permit is issued.<sup>18</sup> One limitation of the AFS data is that it does not provide any information as to when these permits were issued. Fortunately, the regulatory structure of the CAA allows one to infer the timing of the regulations based on the county nonattainment status. Specifically, I define a plant as regulated if the plant has an operating permit in the AFS database and that plant resides in a county that is in nonattainment for the specific pollutant on the permit. To the best of my knowledge, this is the first time a researcher has been able to explicitly observe plant level regulatory status in a longitudinal, nationwide setting. This plant level definition of regulation is in contrast to previous research for which regulation was proxied by 2- and 4-digit SIC level national pollution estimates (Greenstone 2002, List et al. 2003, Becker 2005, Kahn & Mansur 2010, Greenstone, List & Syverson 2010). See Appendix A.5 for a more detailed discussion of how this regulatory measure relates to those in the previous literature. Lastly, I match the EPA’s annual county level nonattainment designations to each dataset.

## 5.1 Aggregation

The linked worker-firm sample from the LEHD consists of a balanced panel of 2.3 million workers for 12 years in the data so that the baseline sample consists of 27.6 million observations. In order to limit the computational burden, I collapse all of the data to the county×sector×year level, or in the case of

---

<sup>17</sup>In related work, I use the LBD to examine plant level employment, job creation, and job destruction rates as a function of environmental regulations across the entire United States manufacturing industry from 1985-2000. See Walker (2011) for details.

<sup>18</sup>Note that these permits are simply operating permits and are unrelated to “cap and trade” permits from other pollution abatement programs.

cohort analysis, the cohort×year level (where cohorts are defined by the county and sector of work in 1990). Each county has two observations in a given year, one observation each for the polluting sector and non-polluting sector of the manufacturing industry in that particular county/year. Collapsing the data to the county×sector×year level eases the computational burden while also taking seriously issues pertaining to inference when the identifying variation occurs at a more aggregate level (Bertrand, Duflo & Mullainathan 2004).

A more subtle reason for aggregation in this particular context is due biases induced by selective attrition in log earnings models. When estimating models using log earnings, any cohort year in which all individuals are non-employed will drop out of the regression sample (i.e.  $\log(0)$  is undefined). Therefore, more granular definitions of cohorts leads to selective attrition from the sample.<sup>19</sup> The reason why log earnings models are important in this study is to reduce biases induced from cost of living differences across locations; the same proportional wage change in an area with higher median earnings will be larger than in areas with lower median earnings, giving undue weight to workers in high cost of living locations. Accordingly, I aggregate the data to the county×sector×year to minimize this type of selective attrition while also estimating log earnings models. As before, sectors are defined as either polluting or non-polluting so that each county implicitly contributes two observations to the data in a given year. Estimates using more granular group definitions yield similar results.<sup>20</sup>

## 5.2 Baseline Industry and Worker Characteristics by Regulatory Status

Table 1 presents characteristics of the data in terms of plant, worker, and sectoral characteristics. The columns are indexed first by county (i.e. attainment, nonattainment, and those that switch into nonattainment following the CAAA of 1990) and then by the polluting/non-polluting sector. Since this table presents statistics only from 1990, the sample only consists of the manufacturing sector. As workers move from their initial 1990 employer, gradually other industries work their way into the analysis. When comparing polluting to non-polluting sectors in Panel A, it becomes immediately clear that polluting plants tend to be older and larger than their non-polluting counterparts. This distinction becomes important when comparing worker outcomes across these groups, as younger and smaller firms

---

<sup>19</sup>Selective attrition is exacerbated by the dynamic nature of the estimation framework. In the distributed lag models estimated in this project, if a cohort has zero earnings in *any* lagged year, the entire cohort drops out in every year the zero earnings lagged variable remains in the model. These issues are reduced dramatically by using more aggregate cohort definitions (i.e. county×sector). Any worker with missing earnings for a given year simply does not contribute to the cohort average in the year they are missing (or in the case of models which replace non-employment with zero earnings, these workers will lower the respective cohort average wage in that year).

<sup>20</sup>An alternative sample selection criteria employed in Jacobson, LaLonde & Sullivan (1993) requires workers to have at least some positive earnings at every year in their data. While this sample of workers is substantially different than the full sample, estimates using this criteria also yield similar results at the expense of a substantial increase in computational burden (i.e.  $\approx 15$  million worker/year observations).

experience both faster growth (Dunne, Roberts & Samuelson 1989) and tend to pay higher earnings conditional on size (Schmieder 2010).

Turning to Panel B of Table 1, we see that workers in the polluting sector are paid considerably more than their non-polluting counterparts. There are several reasons why workers in the polluting sector may be paid more such as higher rates of unionization, compensating wage differentials pertaining to worker health, and/or any skill differentials in the production process. This last point can be seen when comparing education across polluting/non-polluting sectors. Workers in the polluting sector are, on average, older and more educated than their non-polluting counterparts.

Table 1 highlights two of the major forms of selection that need to be accounted for in any research design evaluating the differences in earnings profile across these groups. Younger firms tend to pay higher wages, and they experience higher separation rates relative to their older counterparts. In contrast, workers in the polluting sector tend to be older and more educated, leading them to have higher earnings ex-ante. The primary research design relies upon within sector comparisons before and after the changes in regulatory status. Accordingly, the main source of identifying variation comes from Column 6 of Table 1, which constitutes the polluting sector in counties that switched into nonattainment following the CAAA of 1990. The exact details pertaining to estimation and identification are described in the following section.

## 6 Measurement Framework

### 6.1 Estimation Setup: A Two Period Example

The estimation strategy is first illustrated in a two period example and then extended to incorporate dynamics. There are three margins of variation inherent in this empirical framework: county nonattainment status ( $c \in \text{Attain, Nonattain}$ ), polluter status ( $s \in \text{Polluting, Non-Polluting}$ ), and two time periods ( $\tau \in \text{Pre, Post}$ ). EPA regulations apply only to polluting firms in nonattainment counties after the nonattainment designation goes into place. Let  $N_c$  be an indicator equal to 1 for counties that were newly designated as nonattainment. Let  $P_s$  be an indicator for the polluting sector, and let  $1(\tau_t > 0)$  be an indicator for the years after the introduction of the new regulations. Then  $N_c \times P_s \times 1(\tau_t > 0)$  is an indicator equal to 1 for those sectors that change regulatory status with the introduction of the 1990 CAAA (i.e. polluting sectors in newly designated nonattainment counties after the change in regulation). This term will serve as the reduced form proxy corresponding to the regulations induced by nonattainment designation in the econometric model described below.

Consider the data generating process for outcome  $Y_{cst}$  (i.e. wages and employment) in the *unregulated*

county-sectors

$$Y_{cst} = \rho_{cs} + \gamma_t + n_{ct} + p_{st} + \epsilon_{cst}$$

where  $\rho_{cs}$  are county×sector indicators meant to model time invariant observed or unobserved characteristics that govern outcome  $Y_{cst}$ ;  $\gamma_t$  are year fixed effects to capture overall trends in outcomes;  $n_{ct}$  is a vector of nonattainment×year effects to model aggregate shocks common to nonattainment counties in a given year; and  $p_{st}$  is a vector of polluting sector×year fixed effects. The error,  $\epsilon_{cst}$ , represents unobserved county×sector shocks to outcomes that would have occurred in the absence of treatment so that

$$E[\epsilon_{cst} \mid \rho_{cs}, \gamma_t, n_{ct}, p_{st}] = 0.$$

Thus, the change in  $Y_{cst}$  due to a change in a county×sector’s nonattainment designation may be inferred from the coefficient on  $\eta_1$  in the equation below:

$$Y_{cst} = \eta_1 [N_c \times P_s \times 1(\tau_t > 0)] + \rho_{cs} + \gamma_t + n_{ct} + p_{st} + \epsilon_{cst} \quad (2)$$

where  $\eta_1$  represents the reduced form effect of nonattainment designation on outcome  $Y_{cst}$ . Equation (2) is simply a difference-in-difference-in-difference estimate of the change in outcome  $Y_{cst}$  attributable to changes in nonattainment designation for polluting sectors. For this estimate to be unbiased, outcome  $Y_{cst}$  must depend additively on these unobserved factors. The identifying assumption is that there are no other factors generating a difference in differential trends between production decisions in regulated and unregulated manufacturing firms.

While the identifying assumption of the model is inherently untestable, data from other time periods and alternative specifications permit several indirect tests. First, data from years prior to the change in regulations permit the analysis of trends in covariates and outcomes prior to the change in policy. Second, since regulations in the polluting sector may lead to indirect effects on workers outside the polluting sector, an alternative specification provides a test as to the importance of these general equilibrium spillovers in contributing to bias in relative wage comparisons. Nonattainment designation varies at the county level so that a model which estimates wage changes for *all* manufacturing workers in a county (i.e. workers in both the polluting and non-polluting sector) will encompass both the direct regulatory induced wage effects as well as any indirect local general equilibrium effects in the unregulated manufacturing

sector. This test is given by the difference-in-difference estimator:

$$Y_{ct} = \eta_G [N_c \times 1(\tau_t > 0)] + \rho_c + \gamma_t + n_{ct} + \epsilon_{ct}. \quad (3)$$

This test relies on the assumption that general equilibrium spillovers are important within a county but are not important across counties. Since the LEHD offers information on location and industry of any subsequent jobs, it is possible to observe the magnitudes of within county versus across county transitions, informing the strength of this assumption. If labor market spillovers into other markets do not appreciably affect the wages in those sectors, either due to inelastic labor demand or imperfectly competitive wage setting, the estimate of equation (3) should be exactly proportional to the size of the polluting sector in newly regulated counties. That is,  $\hat{\eta}_G$  from equation (3) should be equal to  $\hat{\eta}_1$  from equation (2) *times* the fraction of workers in the polluting sector newly regulated counties.

## 6.2 Estimation Setup: Incorporating Dynamics

To allow for important dynamics in how regulations may affect sectoral outcomes, I generalize equations (2) to allow the changes to evolve incrementally:

$$Y_{cst} = \sum_{k=-5}^{10} \eta_1^k [N_c \times P_s \times 1(\tau_t = k)] + \rho_{cs} + \gamma_t + n_{ct} + p_{st} + \epsilon_{cst}. \quad (4)$$

Equation (4) is simply a generalization of a triple-difference estimator that allows any regulatory effect to evolve over time as opposed to assuming that the effect occurs immediately and lasts forever. Equation (4) serves as the principal econometric framework for the rest of the empirical analysis. Using this framework, I examine the dynamic relationships between regulatory changes, employment, earnings, and gross job flows in affected sectors.<sup>21</sup>

## 6.3 A Cohort Based Approach

If labor markets are geographically integrated, fully competitive, and in continuous equilibrium, then all worker transitions are voluntary, and examining changes in average sectoral wages in this setting would be sufficient for estimating regulatory incidence. However, frictions in the labor market and involuntary

---

<sup>21</sup>In theory, since nonattainment designations are pollutant specific, it is possible to estimate models that jointly estimate the pollutant specific effect of nonattainment status (Greenstone 2002). In practice, this is difficult due to the dynamic nature of the estimation strategy and the large number of parameters for each pollutant class. Specifically, for each pollutant there are 45 parameters to be estimated; the baseline triple difference parameters (15), the polluter by year fixed effects for each pollutant (15), and the nonattainment by year fixed effects (15). Degrees of freedom issues prevent this from being a viable econometric test, although models stratifying by pollutant yield similar estimates to those produced in the baseline estimates below.

job transitions imply that industry wages may not fully characterize the effects of innovations in labor demand on the affected workforce. As the theoretical discussion in Section 3 suggests, ex-post industry wages may be subject to composition biases from non-random worker exits.

One possible solution to these issues is to follow the same group of workers over time, explicitly observing both the wage costs associated with reductions in average sectoral productivity as well as any future wage costs associated with transitions between jobs. This is the approach taken in this paper. Specifically, I group workers based upon their sector and location of work in the year prior to the 1990 CAA Amendments. I then follow these groups over time, allowing workers to change both industries and locations while still contributing to their “cohort’s” wage. The cohort wage then explicitly nests both the wage costs associated with reductions in sectoral productivity in the newly regulated sector with any wage costs associated with job transitions for workers who leave the sector.

Before proceeding to the results, a few final estimation details bear mention. First, estimates using the LBD (employment and industry earnings) rely on samples for every year from  $\tau = -5$  to  $\tau = 10$ , whereas the LEHD samples (cohort earnings) are limited to work histories from  $\tau = -1$  to  $\tau = 10$ . Because the LEHD begins in the year prior to the change in regulations, this is the earliest period for which worker earnings records exist. Second, cluster robust standard errors are used for inference to account for correlation between sectors and cohorts in the same county, both within periods and over time. Third, the specifications are weighted by the the sector or cohort employment size in the years before the change in regulations to account for heteroskedasticity associated with differences in group sizes. The weighting specification also ensures that estimates reflect the wage costs for the average worker as opposed to the average cohort/sector.<sup>22</sup>

## 7 Results

The results are presented in the various subsections below. I begin by examining employment at the industry level (Section 7.1). This is done to motivate the baseline empirical framework by first demonstrating that industry employment measures respond to changes in regulations. Section 7.2 presents the central findings of the paper, consisting of the wage responses of workers in newly regulated sectors over time. Subsequent sections (Sections 7.3-7.7) offer various robustness checks and explore both the mechanisms generating the observed wage change as well as heterogeneity in the central estimates.

---

<sup>22</sup>For samples using the LBD, the weights correspond to sectoral employment in year  $\tau = -5$ . In samples using the LEHD, the weights correspond to the size of cohort in year  $\tau = -1$ .

## 7.1 Regulation Leads to a Reduction in Sectoral Employment

While previous literature has shown that the CAA regulations lead to industry downsizing (Henderson 1996, Greenstone 2002, Kahn & Mansur 2010, Walker 2011), it is not clear how those results generalize into this particular setting. If reductions in labor demand (and/or labor supply) are costly to the affected workforce, results from this exercise should motivate the incidence analysis using workers' earnings histories. I draw upon the Longitudinal Business Database (LBD) to construct measures of average sectoral employment for each county $\times$ sector from  $\tau = -5$  to  $\tau = 10$ . The focus on sectoral employment rather than plant employment is done so that regression estimates reflect employment flows on both the intensive and extensive plant operating margin.<sup>23</sup>

Figure 3 plots the coefficients from a version of equation (4) using  $\log(\text{employment})$  as the dependent variable. Specifically, Figure 3 plots event time indicators (5 years before and 10 years after) for the polluting sector of counties that switch into nonattainment with the introduction of the 1990 CAAA. The plotted coefficients represent the time path of employment relative to the year before the change in regulations conditional on county $\times$ sector fixed effects  $\rho_{cs}$ , polluting sector by year fixed effects  $p_{st}$ , nonattainment county by year fixed effects  $n_{ct}$ , and year fixed effects  $\gamma_t$ . In the presence of county $\times$ sector fixed effects not all the event-time indicators are identified. For this reason, I normalize the coefficient on indicators in  $\tau = -1$  to be equal to zero.

There are two important features from Figure 3. First, employment in years prior to the change in regulations is nearly constant and not significantly different from zero. The lack of significant employment effects in the years prior to the change in policy provide an important test as to the validity of the identifying assumption; trends in outcomes in across comparison groups evolve smoothly except through the change in policy. The second important feature of Figure 3 is that beginning with the year of the regulatory change, employment of polluting sectors in newly regulated counties begins to fall, reaching levels nearly 20 percent below 1990 employment levels in the 5 years after the regulatory change. These estimates control for any national trends in the polluting sector, any common shocks to newly regulated counties, as well as any national economic trends in manufacturing employment. One remaining source of bias would come from any unobserved economic shock that affected polluting firms in newly designated nonattainment counties in years after changing nonattainment status.

The magnitudes and dynamics of the estimates are similar to those of Walker (2011), which uses the same dataset for the entire United States relative to this restricted subsample. See Walker (2011) for further discussion of these results as they pertain to plant and sectoral job destruction and job creation rates.

---

<sup>23</sup>If a plant were to exit in any period  $\tau \in [-5, 10]$ , that plant would drop out of a plant level regression of equation (4). Plant level models yield results similar to those presented here. See Walker (2011) for further plant level evidence.

## 7.2 The Wage Costs of Sectoral Reallocation: Evidence from Cohorts

The theoretical framework suggests that industry wages are not sufficient for characterizing the incidence of labor market reallocations since composition biases may affect wages and changes in wages may not reflect the costs born by those workers who switch sectors. In addition, gross labor market flows from a sector are insufficient, unless we know the costs born by those labor market transitions. Here, I present the results using cohort wages, where cohorts are defined by the place of work in year  $\tau = -1$ . By following the same group of workers over time before and after the changes in regulations, I can fix the composition of the workforce to be constant while explicitly observing any future earnings losses (or gains) associated with job transitions.

Table 2 presents the central findings of the paper using cohort wages as the dependent variable in various versions of equation (4). Each column in the table corresponds to a different regression. Column (1) of Table 2 suggests that new environmental regulations lead to a reduction in earnings for the cohort affected by the regulations. The earnings decrease steadily over time, beginning to level off after three years, at which point they begin to recover to their pre-regulation level. The last row of Table 2 presents estimates of the total present discounted value of the earnings change over the 9 reported years, discounting the future earnings changes using a 4 percent annual discount rate. After 9 years the average worker in the affected cohort experienced a present discounted wage loss of around 26.1 percent of their pre-regulatory earnings. Note that these estimates do not condition on job separation and thus comprise a weighted average of wages for those workers who leave the sector and those workers who stay. In contrast to the employment estimates from the previous section, these earnings estimates contain wage information for only a single year prior to the change in policy. This is because the cohort earnings data come from the LEHD which only contain data for periods  $\tau \in [-1, 10]$ .

Column (2) of Table 2 includes controls for the average experience of the cohort in a given year.<sup>24</sup> While including cohort fixed effects will control for mean differences in wages across these sectors, they do not account for differences in the growth rate of earnings across the various groups. Thus, a linear and quadratic term in potential experience is meant to capture some of this heterogeneity across cohort earnings profiles in the spirit of Mincer (1974). Controlling for average cohort experience reduces the size of the estimated effect slightly, but the 9 year present discounted value still exceeds 22 percent of pre-period earnings. Column (3) adds a common trend for wages in each cohort county.<sup>25</sup> Here, estimates are identified by comparing wage changes before and after the policy, after netting out any

---

<sup>24</sup>Sectoral experience is proxied by “potential experience” which is defined here as: average sectoral age - average sectoral education - 6.

<sup>25</sup>Recall that cohorts are defined by the 1990 county $\times$ sector. Thus, the “county trend” in Table 2 corresponds to a trend for each 1990 “county cohort”.

earnings trends common to both the polluting and non-polluting sector of a given “county cohort”. The results remain very similar to before.

Column (4) of Table 2 presents estimates where the dependent variable has been replaced with the average cohort earnings (as opposed to the log earnings), and the results do not change. Lastly, Column (5) presents estimates from models that replace any missing earnings observations with zeros. Recall that the baseline estimates ignore any missing *annual* earnings when calculating the average cohort earnings (i.e. up to three quarterly earnings can be missing per year). Since missing earnings may occur through either unemployment or sample attrition, it is not a priori clear how to address such an issue.<sup>26</sup> To understand the implications of missing earnings in my data, I replace any annual earnings data that is missing with a zero, conditional on that worker having positive earnings in the final year of the data. In theory, this should serve as a lower bound on earnings estimates as we are attributing all sample attrition to unemployment rather than other worker transitions (such as moving to non LEHD states or switching into self-employment). The estimates presented in Column (5) confirm this intuition.

A useful feature of these estimates is that it permits a direct monetization of these costs in terms of foregone earnings. The average present discounted earnings loss estimate from Table 2 is around 23 percent of 1990 earnings. Multiplying this number by the average annual earnings in that sector ( $\approx$  \$39,000) and then by the total number of employees in the polluting sector of all “switching” counties in the United States (approximately 1 million workers),<sup>27</sup> the total foregone wage bill is approximately 9 billion dollars.

To facilitate interpretation, Figures 4a and 4b present graphical estimates of equation (4). Specifically, the plotted coefficients in Figure 4a are event time indicators for the polluting and non-polluting sectors in counties that switched nonattainment status in 1991, controlling for cohort fixed effects and polluting sector by year fixed effects (identified by polluters outside the switching counties). The lines in both figures are from the same regression, where Figure 4b is simply the difference in the coefficients reported in Figure 4a. Thus, Figure 4b implicitly replicates the regression estimates presented in Table 2. The results in both figures correspond to the estimates seen in Table 2, and clearly show the wage profile in the newly regulated sector falling in the years after the policy change, only beginning to recover after 3-4 years.

---

<sup>26</sup>This issue is not unique to this paper, and other papers have dealt with these limitations in similar fashions See e.g. Jacobson, LaLonde & Sullivan (1993).

<sup>27</sup>This last statistic comes from Table 1 of Walker (2011).

### 7.3 “Local” General Equilibrium Effects

Any major shift in the sectoral labor supply or demand curves is likely to influence other sectors within the same local labor market. Assuming that polluting sector workers are substitutes across sectors, any workers leaving the polluting sector will increase the potential labor supply of the unregulated sector within that county. If wages are set on a downward sloping demand curve, this increase in the labor supply will drive down the equilibrium wage in that sector. Therefore, relative wage comparisons across these groups will understate the total wage costs.

To get a sense of the empirical magnitude of such effects, first consider the following back of the envelope calculation. Manufacturing accounts for about 15 percent of non-farm employment in the United States, and the polluting sector accounts for a 35 percent share of the manufacturing industry in newly regulated counties. The employment estimates from Section 7.1 and Figure 3 suggest that employment in newly regulated sectors falls by around 15 percent in the 5 years after the policy change. This amounts to 0.7 percent of the total employment in newly regulated counties. Using a labor demand elasticity of 0.5, the implied general equilibrium effect of a 0.7 percent increase in the labor supply would constitute a 0.35 percent decrease in equilibrium wages for workers in the non-polluting sector of newly regulated counties. This assumes that polluting sector workers are perfect substitutes in all other sectors of the county. Conversely, the other extreme case would occur where workers from the polluting sector transitioned only to the non-polluting manufacturing sector since they may share similar skills. In this case, the same calculation suggests that a reduction in labor demand in the polluting sector would increase labor supply for the non-polluting sector of the manufacturing industry by 8 percent. This would translate into a 4 percent reduction in the equilibrium wages of workers in the non-polluting sector. These two extreme assumptions serve as ex-ante bounds on the importance of local general equilibrium spillovers in the context of the research design.

Equation (3) from Section 6.1 offers a straightforward test as to empirical magnitude of any general equilibrium bias.<sup>28</sup> Table 3 presents results from a dynamic version of equation (3) which suggests that the average worker in the manufacturing industry of a county that switched regulatory status experienced a 9.1 percent (discounted) reduction in wages in the 9 years after the policy change. This consists of the direct wage effects of workers in the polluting sector and the indirect effects on workers in the non-polluting sector. If there are no general equilibrium wage effects in other sectors then these estimates

---

<sup>28</sup>Instead of the baseline triple-difference estimator from before, I use a difference-in-differences estimator to look at variation in county wage profiles of *all* manufacturing workers over time. The model then delivers estimates of the average change in wages for workers in the manufacturing industry of a county that switched into nonattainment with the introduction of the 1990 Clean Air Act Amendments. If a reduction in labor demand in the polluting sector drives down the equilibrium wage in the non-polluting sector of the same county, then estimates focusing on across county variation rather than within county variation should reflect these “local” general equilibrium wage effects.

should be directly proportional to the baseline wage effects multiplied by the fraction of workers in the polluting sector:

$$\hat{\eta}_G = \left( \frac{N_{Polluter}}{N_{Total}} \right) \times \hat{\eta}_1. \quad (5)$$

The polluting sector accounts for approximately 35 percent of the manufacturing industry in newly regulated counties, which implies that the right hand side of equation (5) is equal to -8.1 percent. This implies that wages fell by 0.95 percent in the non-polluting sector of the manufacturing industry for newly regulated counties.<sup>29</sup> This empirical test relies on the assumption that within county spillovers are of greater empirical importance than across county spillovers, but it suggests that the empirical magnitude of these spillovers are small.

#### 7.4 Possible Mechanisms: Regulation Increases the Rate of Separations

Presumably, the change in earnings is related to both reductions in sectoral productivity and/or costly job transitions. The subsequent empirical sections attempt to shed some light on this question by examining gross job flows and earnings changes associated with various labor market transitions.

Since sectoral employment from the LBD only measures net changes in employment, I turn to the LEHD data to estimate the degree to which changes in environmental regulations lead to excess labor reallocation in the years following the policy.<sup>30</sup> Here, I am primarily interested in the rate of separation from a worker’s “base-year” employer, and thus the following discussion parallels that of duration analysis, where a failure in this model is defined by a separation from the “base-year” firm.<sup>31</sup> Separation rates for a cohort are calculated at the county×sector level as the number of “base-year” firm separations in a cohort-year over the total number of workers in that cohort that remain at their pre-regulatory establishment. The data is constructed so that workers contribute to the cohort-year

---

<sup>29</sup>To see this formally, denote the difference-in-difference wage estimate for workers in the polluting sector of a newly regulated county as  $\hat{P}$ . Denote the difference-in-difference wage estimate for workers in the non-polluting sector of a newly regulated county as  $\hat{N}$ . The empirical estimate of  $\hat{\eta}_G$  is a combination of these two terms:

$$\hat{\eta}_G = (s_p \times \hat{P}) + (s_n \times \hat{N})$$

where  $s_p$  and  $s_n$  reflect the shares of workers in each of these sectors (i.e.  $s_n + s_p = 1$ ). Assuming that any across county wage spillovers are zero, then a triple difference estimate of the wage impact of regulations on polluting sector is given by  $\hat{\eta}_1 = \hat{P} - \hat{N}$ . This implies that

$$\hat{N} = \hat{\eta}_G - (s_p \times \hat{\eta}_1) = -0.09 - (0.35 \times -0.23) = -0.0095$$

<sup>30</sup>The employment results from Section 7.1 suggest that some of the wage changes are associated with reductions in labor demand. However, this tells us little about whether reductions in labor demand cause wages to fall from the standpoint of reducing the equilibrium wage in the newly regulated sector or by causing costly job transitions from that sector.

<sup>31</sup>A job separation is defined as equaling one in period  $t$  if earnings  $> 0$  at plant  $j$  in period  $t$  and either earnings=0 in period  $t + 1$  or  $\text{plant}_{jt} \neq \text{plant}_{jt+1}$

observation for every year they remain with their “1990” plant. The risk set evolves over time as workers leave their 1990 plant.

The empirical analysis estimates how this failure rate changes as a function of the new environmental regulations. The regression framework is the same as used in equation (4) where the dependent variable is the failure rate for workers remaining at their 1990 “pre-regulation” firm. A separation is defined as equal to 1 if the earnings of the worker in the *next* period are missing or if the worker’s employer in the *next* period is different than the current period employer. Given this definition, there are a positive number of separations in the “base-year”. The baseline estimates then measure the differences in the failure rate as a function of the change in regulatory status.

Figure 5 plots estimates of the  $\eta_1^k$ ’s from equation (4), normalized by  $\eta_1^{-1}$ . Specifically, Figure 5 plots the difference in the failure rate for the newly regulated sector relative to the year prior to the nonattainment designation. Accompanying Figure 5 is a bar graph representing the fraction of workers in polluting plants of newly designated nonattainment counties that remain in their 1990 firm  $t$  years after the CAA Amendments (i.e. the Kaplan-Meier survival probability). The figure shows that the rate of separation increases quite dramatically for the newly regulated cohort in the years following the regulations. After 4 years, the failure rate begins to peak, and after 8 years it is nearly back to the baseline level. Similar to the sectoral employment regressions from before, the estimates here suggest that the changes in regulations lead to a relative reallocation of labor away from the newly regulated sector.

## 7.5 Effects of Regulations for “Stayers” and “Leavers”

While, the temporal dynamics from job separations parallel those from the cohort wages (i.e. Figure 4b), it is still not clear what proportion of the observed wage changes is explained by costly job transitions relative to reductions sectoral productivity. Here, I examine any wage changes for those workers that remain with their initial employer and for those workers who separate from their initial employer. As the theoretical discussion in Section 3 suggests, these groups will be self-selected based on sector-specific ability. However, the form of selection has relatively clear predictions for the direction of bias in each of the regressions that I will highlight below. Moreover, by including group level fixed effects the regression model includes controls for any time-invariant unobservable characteristics of these “stayers” or “leavers”.

I stratify the treatment group based on whether the worker stayed with their firm or separated at some point within the first 4 years after the regulations. In each specification, I decompose the wages for the affected cohort *only*. That is, the wages of the various control group cohorts remain the same

as in the previous section, while the newly regulated sector’s wage consists only of “stayers” or only of “leavers”, depending upon the specification.<sup>32</sup> Column (1) of Table 4 presents results from a version of equation (4) where I compare earnings trends in wages of the affected cohort only if they remain at their “base-year” firm for more than 4 years after the change in regulations (i.e. they remain at their firm for at least the years  $\tau \in [0, 4]$ ). The results from column (1) suggests that the wages of “stayers” are essentially unaffected by the regulatory change. Column (2) of Table 4 presents the same model as above, except that I examine wage changes for those who separate from their “base-year” firm in the 4 years after the change in regulations (i.e.  $\tau \in [0, 4]$ ). Here we see much larger wage changes that are all statistically significant. The pattern of the estimates suggests that the average wage declines rapidly in the years following the change in regulations, and they begin to recover only in the later years. The present discounted earnings change for separators is more than 130 percent of their pre-regulatory earnings.<sup>33</sup>

In general, the wage cost of job separations in this context is smaller than estimates found in the literature on displacement induced by mass layoffs (Jacobson, LaLonde & Sullivan 1993, Von Wachter, Song & Manchester 2009). Notably, the earnings recovery of the average worker in this setting is much more rapid than that found in the displacement literature. There are a few possible explanations for these discrepancies. First, it seems possible that the rapid earnings recovery comes from the fact that most of these regulations occur in dense, urban labor markets. Workers are likely able to re-integrate themselves into the workforce quicker in these “thicker” labor markets than elsewhere where alternative job options are limited (Marshall 1920). Second, it is possible that some of the job transitions I observe are voluntary, job-to-job transitions for which workers often experience a *rise* in earnings (Bjelland et al. Forthcoming). In contrast, the job transitions from mass layoffs are often characterized by involuntary job loss and prolonged unemployment durations (Jacobson, LaLonde & Sullivan 1993, Davis, Faberman & Haltiwanger 2006, Von Wachter, Song & Manchester 2009). Lastly, an alternative explanation for this relatively quick recovery comes from the external validity of the respective research designs. Namely, the long run implications from job loss for those who were affected by mass layoffs may be fundamentally different from those affected by the more gradual sectoral changes that we see in this setting.

---

<sup>32</sup>An alternative comparison that gets around this form of selection bias is to compare the earnings of stayers to that of stayers in other sectors. Results (not reported) are consistent with those shown here.

<sup>33</sup>Another estimate of the average cost of job transitions comes from “scaling” the baseline cohort wage regressions (i.e. column (3) in Table 2) by the number of workers we see transitioning out of that sector (i.e. figure 3). The average sectoral earnings loss for a cohort from Table 2 is around 23 percent. Estimates from figure 3 suggest a more than 15 percent reduction in the sectoral workforce in the 5 years after the policy. This suggests that the average earnings loss for a worker who loses their job because of the policy shock is around 150 percent of their 1990 earnings. This is consistent with what we see for those who separate from their job in column (2) of Table 4.

## 7.6 Sectoral Transition Rates and the Wage Costs of Labor Reallocation

The ability to follow workers over their employment histories allows for decompositions of worker transitions based on destination sectors. This offers some degree of insight as to *how* workers respond to sector specific shocks of this nature. Similar to before, I estimate how the “failure rate” responds to changes in regulations, where a failure is defined as a job separation from the “base-year” firm *to* a specific location and/or industry. Figure 6 presents these estimates for workers who transition to a different industry in a different county (panel 6a), the same industry in the same county (panel 6b), a different industry in the same county (panel 6c), and the same industry in a different county (panel 6d). The figures suggest that workers in newly regulated sectors are disproportionately more likely to exit to a completely different industry after the regulations, relative to before. Furthermore, workers are more likely to transition to a different industry within the same county.

There is considerable heterogeneity in wage estimates based on this same set of destination industry-locations. I decompose the wage effects of separators (i.e. column (2) of Table 4) based on the location and sector where workers in the newly regulated sector find their subsequent job. The estimates suggest that the wage changes for workers who stay within the same industry and the same county (column (3) of Table 4) are significantly less than for those workers that change industries (even within the same county - i.e. column (4) of Table 4). This is consistent with previous literature suggesting a role for industry specific human capital as barriers to job mobility (Podgursky & Swaim 1987, Topel 1991, Neal 1995). Of course, these wage losses also reflect losses due to non-employment between jobs that may also be higher for workers who switch industries (Murphy & Topel 1987). Figure 7 displays the estimates from columns (3)-(6), summarizing the dynamics and incidence by destination sectors.

## 7.7 Additional Results

Additional results are reported in Appendix B. In particular Appendix Table B.1 presents results for different age categories in the data. Older workers (ages 50 to 59) experience dramatically larger wage declines in the affected sectors, consistent with the fact that older workers are less likely than younger workers to be working full-time after displacement (i.e. older displaced workers may “partially” retire) (Farber 1999).

## 8 Interpretation and Discussion

Increasing the cost of production through environmental regulations leads to a reallocation of production and workers away from newly regulated areas. Estimates suggest that the average worker in these newly

regulated sectors experienced a nearly 9 percent reduction in earnings in the years following the change in regulations. These earnings changes are driven by costly job transitions rather than a reduction in labor productivity in the regulated sector. This section further discusses the implications from these results in the context of the research design and external validity.

While the data used in this project are rich in many dimensions, there are certain limitations that affect how much we may be able to generalize these findings. The most important limitation comes from the fact that I am focusing on workers in only 4 states of the United States. While these counties comprise a significant proportion of the overall polluting sector amongst all counties affected by the 1990 CAA Amendments (20 percent as measured by employment), any significant difference between these counties or states compared to the rest of the country will make it difficult to scale these estimates to a national, representative setting. Table 5 presents various county characteristics from the 1990 Decennial Census for both in and out of sample counties. Columns (5) and (6) present sample statistics for the in and out of sample “treatment” counties, respectively. Column (7) presents t-statistics pertaining to the null hypothesis that columns (5) and (6) are statistically similar. The in-sample counties are on average smaller and less urban, but column (7) suggests that these differences are not statistically significant at conventional levels. Notably, the in-sample counties are statistically higher income and have a slightly higher fraction of durable manufacturing employment (although the proportion of non-durables seem quite similar).

Another concern pertaining external validity comes from the fact that these estimates are identified off of the CAA Amendments of 1990, which happened in the middle of the recession in the early 1990’s. Even though the richness of the data allows me to control for national and sectoral trends, it seems plausible that the cost of job loss during a recession is higher than in non-recessionary periods. There might also be important interactions between the 1990’s recession and plant level production decisions. If the recession increased the number of plants on the margin of exit, a further cost shock from the new regulations could lead to more plant exits than in a more stable economic period.<sup>34</sup>

## 8.1 Labor Reallocation and Welfare

If workers are paid their marginal product and firms earn zero profits, then all wage losses are social costs, and the estimates allow for simple approximations of the welfare costs pertaining to policy induced job reallocation. Namely, if utility is linear in wages, workers do not value leisure, and there are no government transfers in the form of unemployment insurance, then the total foregone wage bill associated with industry reallocation is a sufficient statistic for the welfare costs to workers.

---

<sup>34</sup>This logic stems from the literature on endogenous production. See e.g. Jovanovic (1982), Hopenhayn (1992), Melitz (2003).

These are obviously strong assumptions, but they also provide a useful benchmark for assessing the degree to which the estimates presented here serve as upper or lower bounds on the true welfare costs. For example, if individuals have concave utility and are liquidity constrained then the welfare losses from unemployment might be much larger than what is implied by mean earning losses. If workers value leisure, these estimates will overstate the welfare costs.

In terms of aggregate welfare and the social costs of workforce reallocation, it is important to realize that, in the United States, workers are often compensated for involuntary job loss in the form of unemployment insurance (UI). In this case, UI is a transfer that primarily reflects income redistribution among citizens via taxation and transfers and should not be treated as an economic loss. This implies that the welfare cost is no longer the foregone wage bill, but instead should be measured by the social cost of public funds (or the deadweight loss of taxation) associated with increased UI payments. In addition, if a workers market wage (prior to the shock) exceeds their value of leisure, then any increases in non-participation or non-employment after the regulations should be interpreted as a social cost (Autor, Dorn & Hanson 2011).

Consider the following thought experiment pertaining to UI payments in the context of this particular setting. Average UI payments consist of 20 weeks of compensation at 36% replacement. Assuming that 200,000 jobs were lost (based on a 20% reduction in sectoral employment multiplied by the number of workers in the polluting sector of newly regulated counties), this implies an increase in UI payments of approximately \$1 billion.<sup>35</sup> Then using an estimate of 40% for the deadweight loss of taxation from Gruber (2010), the change in worker's welfare amounts to \$8.4 billion rather than \$9 billion.<sup>36</sup> This suggests that the costs of sectoral reallocation in terms of foregone earnings far exceed any transitional assistance programs currently in place in the United States.

## 8.2 Costs Relative to Benefits and Policy Prescriptions

The EPA estimates that the present discounted value (in terms of health benefits) of the 1990 Clean Air Act Amendments from 1990-2010 to be between 160 billion and 1.6 trillion dollars (1990\$) (Environmental Protection Agency 1999).<sup>37</sup> The lower bound estimate is similar to the estimated benefits from various hedonic studies which estimate marginal willingness to pay for improvements in air quality (see e.g. Chay & Greenstone (2005) and Bayer, Keohane & Timmins (2009)). This suggests that the estimated

---

<sup>35</sup>This number is calculated by multiplying average earnings (39,000) by the 36% replacement rate for 20 weeks and then times the reduction in employment in the newly regulated sector.

<sup>36</sup>\$8.4 billion is calculated as the estimated foregone wage bill (\$9 billion) minus UI transfers (\$1 billion) plus the social cost of public funds (\$400 million).

<sup>37</sup>Note that these estimates do not include the estimated benefits from the SO<sub>2</sub> trading markets, which are not analyzed in this study.

benefits are at least one order of magnitude larger than the adjustment cost estimates presented here.

Obviously, there are many more economic costs associated with the regulations, notably plant productivity effects (Greenstone, List & Syverson 2010), losses to foreign direct investment (Hanna 2010), as well as any adjustment costs associated with reallocating the capital stock (Goolsbee & Gross 1997, Ramey & Shapiro 2001).<sup>38</sup>

While the results here suggest that the adjustment costs in terms of labor reallocation are relatively small in magnitude compared to the estimated benefits of the policy, these costs are still borne by those affected by the reallocation. Accordingly, the EPA included a component of transition assistance in the 1990 CAAA under the Clean Air Employment Transition Assistance Program administered by the Department of Labor. This program was designed to aid workers who were displaced by environmental regulations and was allocated \$50 million for the 1990 CAAA. Estimates here suggest this program was severely underfunded in terms of compensating workers for any earnings losses pertaining to regulatory induced reallocation.

## 9 Conclusion

Results from this paper suggest that the transitional costs of policy can be significant. In the context of the Clean Air Act Amendments of 1990 and labor reallocation, the adjustment costs are estimated to be nearly 9 billion dollars. Prior research on the labor market impacts of environmental regulation has primarily focused on employment in manufacturing industries (Berman & Bui 2001, Greenstone 2002, Deschenes 2010, Kahn & Mansur 2010, Walker 2011). These results suggest that the predominant focus of the previous literature on labor demand misses important aspects of labor market adjustment to environmental regulations.

The focus of this paper has been on the wage effects associated with changes in environmental regulation. From the point of view of total employment changes, this paper has said relatively little. This being said, even if net job loss is zero, in the sense that every displaced worker finds a job in another sector, excess job reallocation is still costly to those workers affected. A separate, but related, question concerns the extent to which new environmental regulations lead to net job loss on an economy wide level. This is a difficult but important question where further research is needed.

The findings in this paper have important policy implications both for worker compensation in the context of environmental regulations and also more generally in terms of labor market adjustment.

The results suggest that firms respond to environmental regulations; costly job transitions from the

---

<sup>38</sup>While both Goolsbee & Gross (1997) and Ramey & Shapiro (2001) look at the adjustment costs of capital, these studies do not consider capital adjustment costs in the context of environmental regulation. However, this research suggests that these costs may be significant in the context of environmental regulations and suggest an interesting area for future research.

perspective of the workforce are a consequence of this response. The magnitude of these estimates calls into question the amount of transition assistance allocated under the CAA Amendments of 1990, which is an order of magnitude lower than what would be needed to fully compensate the affected workforce.

This paper offers a new approach to understanding the costs and incidence of labor market shocks when production and workers are reallocated elsewhere within the economy. In particular, the adjustment costs and incidence of labor reallocation for the affected workforce can be easily approximated by the present discounted value of future wage changes of cohorts as a function of the policy shock. This particular approach could presumably be applied to any area concerned with the magnitude and duration of reallocative frictions in the labor market, be it cost shocks to firms, local labor demand shocks to economies, trade shocks, or the welfare costs of natural disasters. The one limitation of this approach is the data necessary to implement it. However, with the growing availability of longitudinal microdata, this should be a fruitful area of future research.

## References

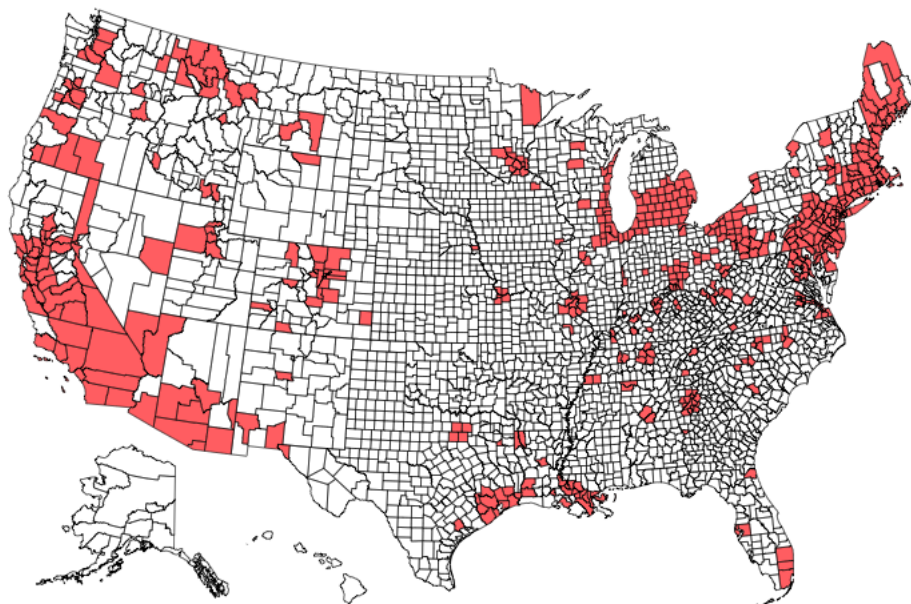
- Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2008. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." *LEHD Working Paper*.
- Abraham, K.G., and J.L. Medoff. 1984. "Length of Service and Layoffs in Union and Nonunion Work Groups." *Industrial and Labor Relations Review*, 87–97.
- Arrow, Kenneth, Maureen Cropper, George Eads, Robert Hahn, Lester Lave, Richard Noll, Paul Portney, Milton Russell, Richard Schmalensee, Kerry Smith, and Robert Stavins. 1996. "Benefit-Cost Analysis in Environmental, Health, and Safety Regulation." *Washington, DC: American Enterprise Institute*, 1–17.
- Artuc, Erhan, Shubham Chaudhuri, and John McLaren. 2010. "Trade Shocks and Labor Adjustment: A Structural Empirical Approach." *The American Economic Review*, 100(3): 1008–1045.
- Autor, David, David Dorn, and Gordon Hanson. 2011. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *Working Paper*.
- Bayer, P., N. Keohane, and C. Timmins. 2009. "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management*, 58(1): 1–14.
- Becker, Randy. 2005. "Air pollution abatement costs under the Clean Air Act: evidence from the PACE survey." *Journal of environmental economics and management*, 50(1): 144–169.
- Becker, Randy, and Vernon Henderson. 2000. "Effects of air quality regulations on polluting industries." *Journal of Political Economy*, 108(2): 379–421.
- Berman, Eli, and Linda Bui. 2001. "Environmental regulation and productivity: evidence from oil refineries." *Review of Economics and Statistics*, 83(3): 498–510.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1): 249–275.
- Bjelland, M., B. Fallick, J. Haltiwanger, and E. McEntarfer. Forthcoming. "Employer-to-Employer Flows in the United States: Estimates Using Linked Employer-Employee Data." *Journal of Business and Economic Statistics*.
- Bound, J., and A.B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, 1–24.
- Brown, C., and J. Medoff. 1989. "The Employer Size-Wage Effect." *The Journal of Political Economy*, 97(5): 1027.
- Chay, Kenneth, and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession\*." *Quarterly Journal of Economics*, 118(3): 1121–1167.
- Chay, Kenneth, and Michael Greenstone. 2005. "Does air quality matter? Evidence from the housing market." *Journal of Political Economy*, 113(2).
- Congressional Budget Office. 2009. "The Economic Effects of Legislation to Reduce Greenhouse-Gas Emissions."
- Copeland, Brian, and Scott Taylor. 2004. "Trade, Growth, and the Environment." *Journal of Economic Literature*, 42(1): 7–71.
- Davis, Steven, Jason Faberman, and John Haltiwanger. 2006. "The flow approach to labor markets: new data sources and micro-macro links." *The Journal of Economic Perspectives*, 20(3): 3–26.
- Deschenes, O. 2010. "Climate Policy and Labor Markets." *NBER Working Paper*.
- Dix-Carneiro, Raphael. 2011. "Trade Liberalization and Labor Market Dynamics." *Mimeo*.
- Duncan, G.J., and D.H. Hill. 1985. "An investigation of the extent and consequences of measurement error in labor-economic survey data." *Journal of Labor Economics*, 508–532.

- Dunne, Timothy, Mark Roberts, and Larry Samuelson.** 1989. "The growth and failure of US manufacturing plants." *The Quarterly Journal of Economics*, 104(4): 671–698.
- Ebenstein, Avraham, Ann Harrison, Margaret McMillan, and Shannon Phillips.** 2011. "Estimating the Impact of Trade and Offshoring on American Workers Using the Current Population Surveys." *Working Paper*.
- Environmental Protection Agency.** 1999. "The Benefits and Costs of the Clean Air Act 1990 to 2010." *EPA Report to Congress*.
- Environmental Protection Agency.** 2005. "Environmental Economics Research Strategy." 600(R-04): 195.
- Environmental Protection Agency.** 2008. "National Air Quality-Status and Trends through 2007."
- Farber, Henry.** 1999. "Mobility and stability: The dynamics of job change in labor markets." *Handbook of labor economics*, 3: 2439–2483.
- Farber, H.S.** 1993. "The incidence and costs of job loss: 1982-91." *Brookings Papers on Economic Activity. Microeconomics*, 1993(1): 73–132.
- Frijters, Paul, and Bas van der Klaauw.** 2006. "Job Search with Nonparticipation." *The Economic Journal*, 116(508): 45–83.
- Fullerton, Don.** 2008. "Distributional effects of environmental and energy policy: an introduction." *NBER working paper*, , (W14241).
- Goolsbee, Austan, and David Gross.** 1997. "Estimating adjustment costs with data on heterogeneous capital goods." *NBER working paper*, , (W6342).
- Greenstone, Michael.** 2002. "The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures." *Journal of Political Economy*, 110(6).
- Greenstone, Michael, and Adam Looney.** 2010. "An Economic Strategy to Renew American Communities." *Brookings: Hamilton Project*.
- Greenstone, Michael, John List, and Chad Syverson.** 2010. "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing." *mimeo*.
- Gruber, Jonathan.** 2010. *Public Finance and Public Policy*. . 3rd Edition ed., Worth Publishers.
- Hanna, Rema.** 2010. "US Environmental Regulation and FDI: Evidence from a Panel of US-Based Multinational Firms." *American Economic Journal: Applied Economics*, 2(3): 158–189.
- Henderson, Vernon.** 1996. "Effects of air quality regulation." *The American Economic Review*, 86(4): 789–813.
- Hopenhayn, Hugo.** 1992. "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica: Journal of the Econometric Society*, 60(5): 1127–1150.
- Jacobson, Louis, Robert LaLonde, and Daniel Sullivan.** 1993. "Earnings losses of displaced workers." *The American Economic Review*, 83(4): 685–709.
- Jaffe, Adam, Steven Peterson, Paul Portney, and Robert Stavins.** 1995. "Environmental Regulation and the Competitiveness of U.S. Manufacturing: Environmental Regulation and the Competitiveness of U.S. Manufacturing: What Does the Evidence Tell Us?" *Journal of Economic Literature*, XXXIII: 132–163.
- Jovanovic, Boyan.** 1982. "Selection and the Evolution of Industry." *Econometrica: Journal of the Econometric Society*, 50(3): 649–670.
- Kahn, Matthew, and Erin Mansur.** 2010. "How Do Energy Prices, and Labor and Environmental Regulations Affect Local Manufacturing Employment Dynamics? A Regression Discontinuity Approach." *NBER working paper*, , (16538).
- List, John, Daniel Millimet, Per Fredriksson, and Warren McHone.** 2003. "Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator." *Review of Economics and Statistics*, 85(4): 944–952.

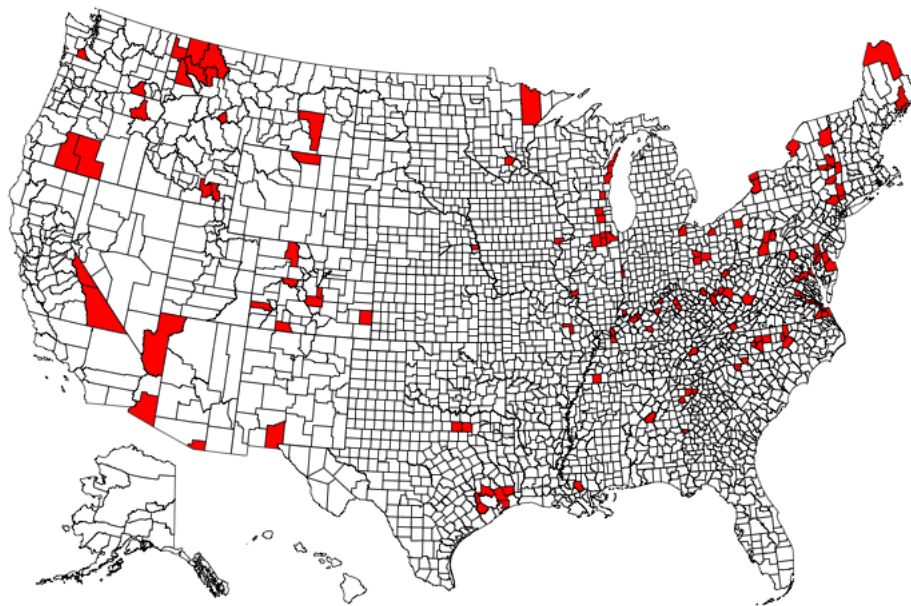
- Ludwig, Jens, and Steven Raphael.** 2010. "The Mobility Bank: Increasing Residential Mobility to Boost Economic Mobility." *Brookings: Hamilton Project*.
- Marshall, Alfred.** 1920. *Principles of economics: an introductory volume*. Macmillan London.
- McKinney, Kevin L., and Lars Vilhuber.** 2008. "LEHD Infrastructure files in the Census RDC - Overview." *LEHD Working Paper*, Revision : 219.
- Melitz, Marc.** 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica*, 71(6): 1695–1725.
- Menezes-Filho, Naercio Aquino, and Marc-Andreas Muendler.** 2007. *Labor reallocation in response to trade reform*. Vol. 1936, CESifo Working Paper.
- Mincer, Jacob.** 1974. *Schooling, experience, and earnings*. National Bureau of Economic Research: distributed by Columbia University Press.
- Murphy, K.M., and R.H. Topel.** 1987. "The Evolution of Unemployment in the United States: 1968-1985." *NBER Macroeconomics Annual*, 2(1): 11–58.
- Neal, Derek.** 1995. "Industry-specific human capital: Evidence from displaced workers." *Journal of Labor Economics*, 653–677.
- Podgursky, Michael, and Paul Swaim.** 1987. "Job displacement and earnings loss: Evidence from the Displaced Worker Survey." *Industrial and Labor Relations Review*, 41(1): 17–29.
- Ramey, Valerie, and Matthew Shapiro.** 2001. "Displaced capital: A study of aerospace plant closings." *Journal of Political Economy*, 109(5): 958–992.
- Roy, A.D.** 1951. "Some thoughts on the distribution of earnings." *Oxford Economic Papers*, 135–146.
- Rubin, Donald B.** 1987. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons Inc.
- Ryan, Stephen.** 2011. "The costs of environmental regulation in a concentrated industry." *Mimeo*.
- Schmieder, Johannes.** 2010. "Labor Costs and the Evolution of New Establishments." *mimeo*.
- Solon, G., R. Barsky, and J.A. Parker.** 1994. "Measuring the Cyclicalities of Real Wages: How Important is Composition Bias." *The Quarterly Journal of Economics*, 1–25.
- The New Republic.** April 7th, 2011. "Smoke Signals." *The New Republic*.
- The New York Times.** Sep 4th, 2011. "A Debate Arises on Job Creation and Environment." *The New York Times*.
- Topel, Robert.** 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1).
- Von Wachter, Till, Jae Song, and Joyce Manchester.** 2009. "Long-Term Earnings Losses due to Job Separation During the 1982 Recession: An Analysis Using Longitudinal Administrative Data from 1974 to 2004." *Columbia University Department of Economics Discussion Paper Series*, 0708–16.
- Walker, Reed.** 2011. "Environmental Regulation and Labor Reallocation." *American Economic Review: Papers and Proceedings*, 101(2).
- Wall Street Journal.** July 26, 2011. "The Latest Job Killer from the EPA."

Figure 1: Nonattainment Counties Across All Pollutants in 1991

(a) Nonattainment for Any Pollutant 1991

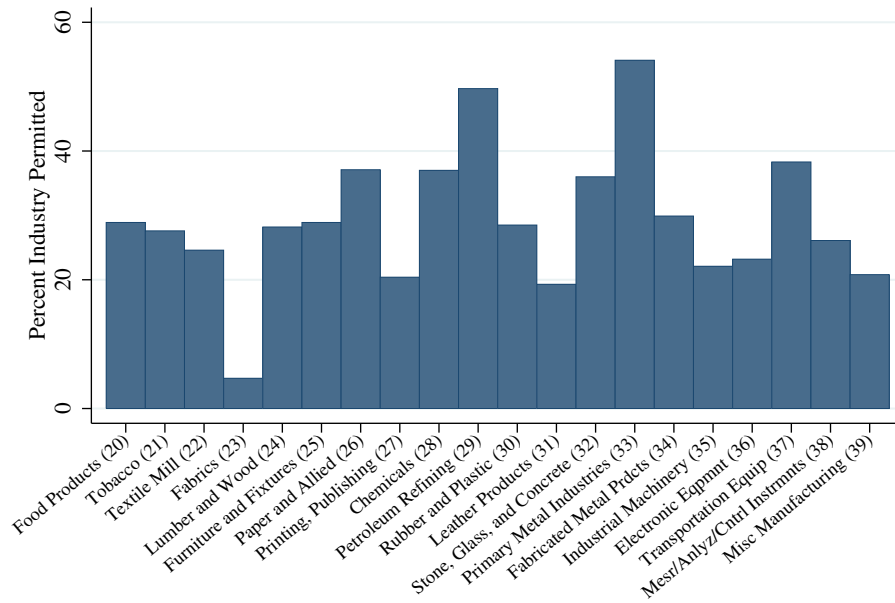


(b) New Nonattainment Counties 1991



NOTE: The top panel shows all counties in nonattainment for any pollutant in 1991. The bottom panel shows those counties that switched into nonattainment in 1991 (i.e. from zero nonattainment designations to one or more designations). The top panel shows that while there is considerable amount of variation in the cross-section, nonattainment designation is primarily concentrated in large metropolitan areas and likely to be correlated with other unobservable features of urban labor markets. The bottom panel shows only those counties that switched into nonattainment status in 1991, the first year that the 1990 CAAA were adopted, showing a considerable number of counties that were newly regulated with the passage of the 1990 CAAA. These counties serve as the primary source of identification in the paper.

Figure 2: Fraction of Permitted Plants By Industry in Nonattainment Counties



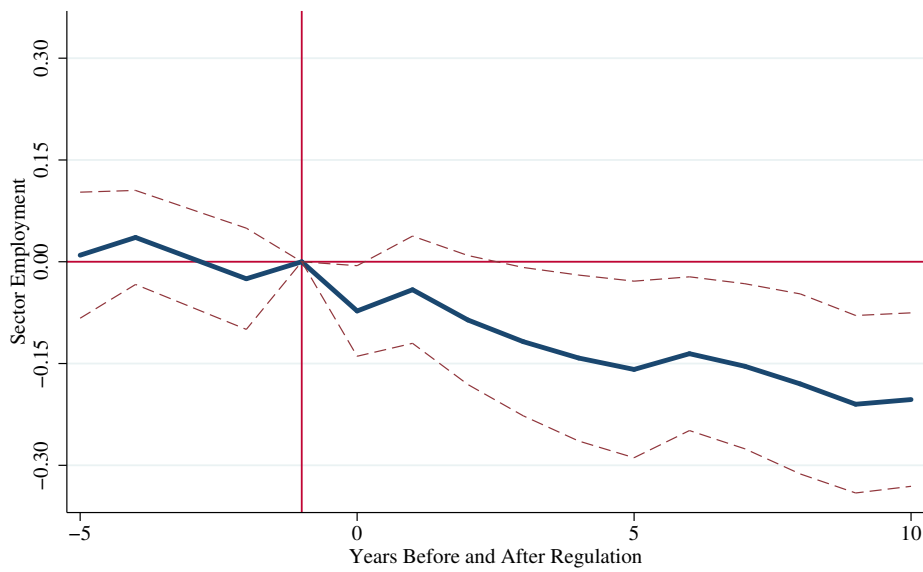
NOTE: This figure shows the fraction of plants by major industry in nonattainment counties that have air pollution operating permits as mandated by the CAA. The percentages are weighted by employment. This figure shows that within a specific 2-digit industry SIC code, less than half of most industries are classified as polluters (and hence regulated under nonattainment). This suggests that relying on industry wide classifications of “polluting industries” to proxy for regulatory status is likely to attenuate estimates of nonattainment regulatory effects, as only a small fraction of any given industry in a nonattainment county is regulated in a given year. Source LBD 1985-2005.

Table 1: Sample Statistics: Pre Regulation (1990)

	Attainment		Nonattainment		Switch Into Nonattainment	
	Non Polluting (1)	Polluting (2)	Non Polluting (3)	Polluting (4)	Non Polluting (5)	Polluting (6)
Panel A: Plant Characteristics (LBD)						
No. of Plants	6806	1707	5910	1549	7622	1452
Plant Employment	146.82	244.6	133.15	266.42	144.24	234.58
Plant Age	8.64	10.07	8.4	10.89	8.84	10.71
Panel B: Workforce Characteristics (LEHD)						
Age	36.88	38.61	36.67	38.3	37.04	38.1
Education	13.34	13.56	13.72	14.08	13.67	13.89
Earnings	26996.56	36950.09	31159.89	38030.07	31692.56	39307.95
Separation Rate	0.22	0.08	0.24	0.11	0.23	0.08
Same Ind, Same County	0.04	0.01	0.05	0.02	0.04	0.01
Same Ind, Diff County	0.02	0.01	0.05	0.02	0.01	0.00
Diff Ind, Same County	0.06	0.02	0.05	0.02	0.1	0.03
Diff Ind, Diff County	0.09	0.03	0.08	0.03	0.06	0.02
Leave Sample	0.01	0.01	0.02	0.02	0.02	0.02
Sep: Job to Job	0.19	0.06	0.2	0.08	0.18	0.05
Panel C: Sectoral Characteristics (LBD and LEHD)						
Total Employment	233994	210386	179590	146123	379553	209364
Total Earnings (in millions)	6317.03	7773.78	5596	5557.07	12029.01	8229.67

NOTES: Sample statistics for the “base” states from which the LEHD workforce sample is drawn (IL, MD, WA, WI). Since these are statistics for the “base” year, they only include the manufacturing industry. All dollar amounts are reported in 1990 dollars. Separation is defined as working in 1990 and working either at a different plant in 1991 or having missing earnings in 1991. The separation rate is further decomposed into five different categories listed in the table. LBD denotes characteristics derived from the LBD sample. LEHD denotes characteristics from the LEHD worker cohort sample. Source LEHD, LBD.

Figure 3: Sector Level Employment Before and After Regulation



NOTE: Plotted are the coefficient estimates from a version of equation (4) showing that employment in the polluting sector of counties newly affected by EPA’s nonattainment designation declines after the designation while earnings per worker are not affected. Moreover, both the polluting and non-polluting sectors share similar trends in employment prior to nonattainment designation. Specifically, this figure plots the event time indicators for employment in the polluting sector of counties that switched into EPA’s nonattainment designation in 1991 (normalized to zero in year  $\tau = -1$ . The first year of the nonattainment designation corresponds to year 0 in the graph. The dashed lines represent 95% confidence intervals. See text for details. Source: LBD.

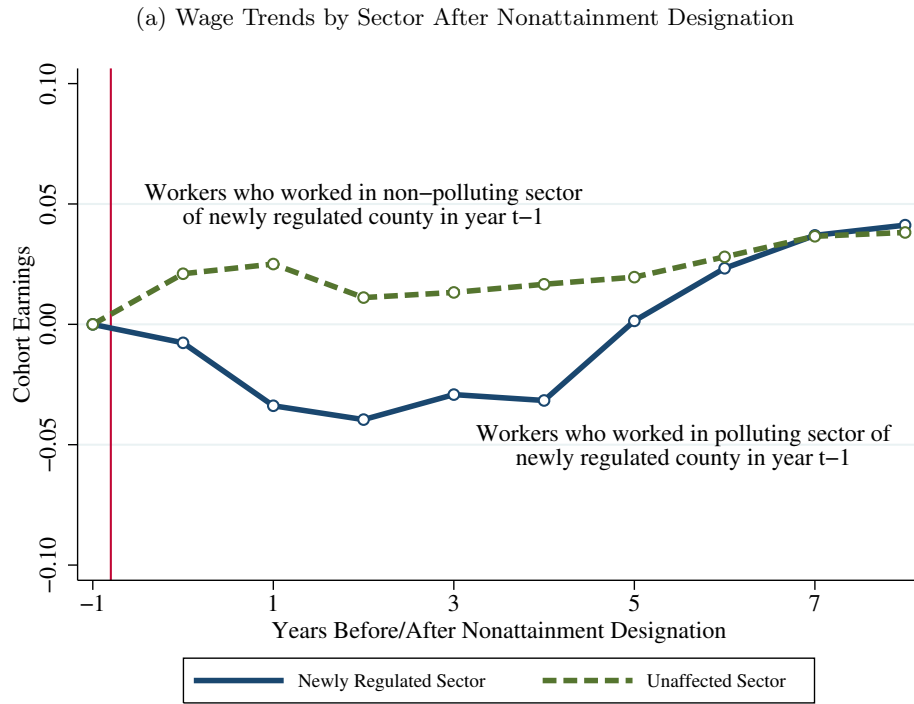
Table 2: Effect of Sector Level Regulation on Earnings

	(1)	Baseline (2)	(3)	Levels (4)	Zero Earnings (5)
Regulation (t+0)	-0.037 (0.029)	-0.029 (0.022)	-0.030 (0.018)	-1045.38 (812.10)	-1202.74 (831.06)
Regulation (t+1)	-0.064* (0.033)	-0.058** (0.026)	-0.059** (0.025)	-1889.78* (1071.37)	-1989.91* (1185.02)
Regulation (t+2)	-0.056** (0.024)	-0.050*** (0.019)	-0.051*** (0.018)	-1898.65** (819.29)	-2148.54** (848.44)
Regulation (t+3)	-0.047*** (0.015)	-0.042*** (0.012)	-0.043*** (0.013)	-1740.67*** (589.96)	-2015.65*** (650.99)
Regulation (t+4)	-0.052*** (0.016)	-0.048*** (0.018)	-0.048** (0.020)	-1992.42** (829.28)	-2039.14*** (750.97)
Regulation (t+5)	-0.022 (0.022)	-0.019 (0.019)	-0.019 (0.018)	-785.76 (703.65)	-785.56 (734.25)
Regulation (t+6)	-0.008 (0.019)	-0.006 (0.017)	-0.006 (0.017)	-275.51 (670.41)	-1197.87 (1212.38)
Regulation (t+7)	-0.002 (0.014)	-0.001 (0.014)	-0.001 (0.014)	39.49 (545.17)	6.89 (521.49)
Regulation (t+8)	0.002 (0.011)	0.002 (0.011)	0.002 (0.011)	307.51 (390.56)	441.14 (343.14)
9 Year PDV	-0.261** (0.120)	-0.229** (0.089)	-0.233*** (0.085)	-8477.32** (3811.51)	-9902.34** (3841.66)
N	6847	6847	6847	6847	6847
Cohort FE	X	X	X	X	X
Experience		X	X	X	X
County Trends			X	X	X

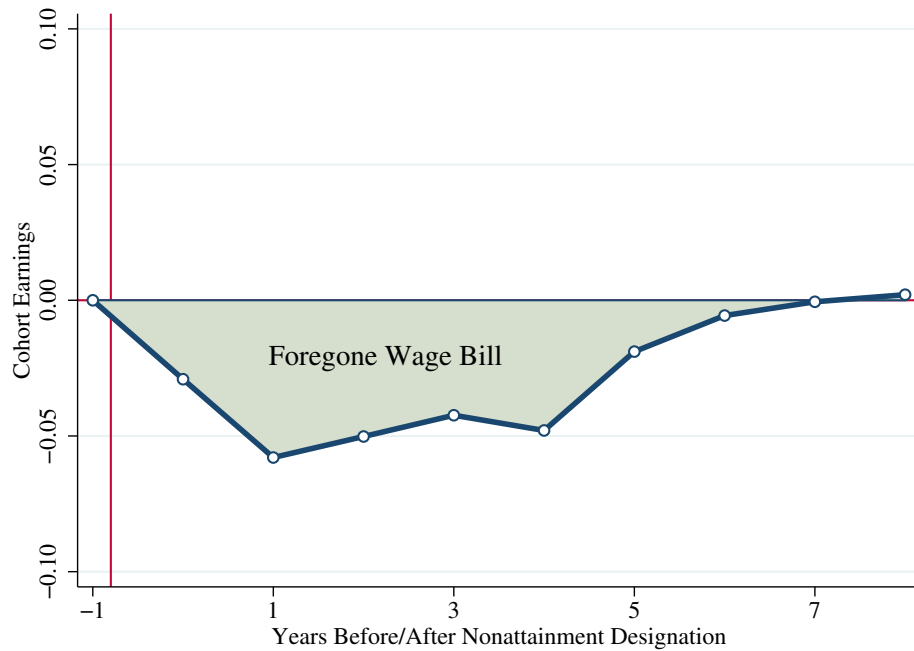
NOTE: This table reports regression coefficients from equation (4). Each column corresponds to a different regression, where the results are presented for three different samples as indicated in the table headings. An observation is a cohort $\times$ year so that each cohort county has two observations per year. Standard errors are in parentheses and are calculated allowing for an arbitrary variance-covariance matrix within each county. The baseline sample is estimated using three different log specifications (columns (1)-(3)), and exponentiated coefficients are reported using the translation  $(\exp(\eta_1^k) - 1)$ . The regressions are weighted by the 1990 cohort size. The final row of the table reports the discounted sum of the coefficients using a 4 percent annual discount rate (with discount factor  $\beta$ ). Columns (1)-(3) use the translation  $\sum_{t=0}^8 \beta^k (\exp(\eta_1^k) - 1)$ . Standard errors for the final row are calculated using the delta method. Earnings are reported in 1990(\$). See text for details. Source: LEHD.

\*\*\*, \*\*, \* denotes statistical significance at the 1, 5, and 10 percent levels, respectively.

Figure 4: Sector Level Wage Trends After Nonattainment Designation



(b) Differences in Wage Trends by Sector After Nonattainment Designation



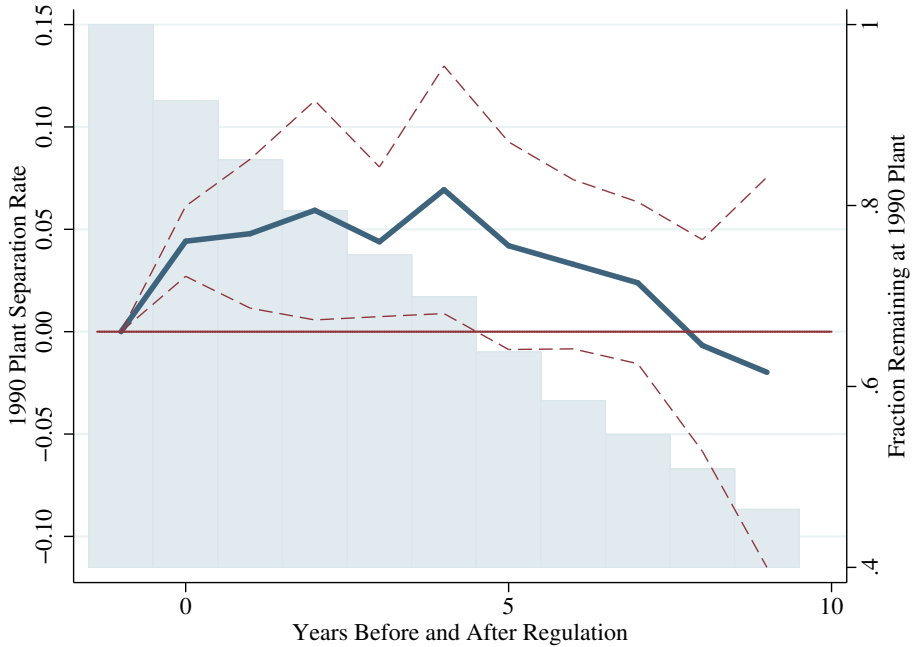
NOTE: Plotted are the coefficient estimates from a version of equation (4). Specifically, Panel (a) plots the event time indicators for two sets of cohorts in counties that switched into EPA’s nonattainment designation in 1991. The solid line represents the cohort that worked in the polluting sector in 1990, and the dashed line represents the cohort that did not work in the polluting sector in 1990. Panel (b) plots the difference in the event time indicators from Panel (a), corresponding to the “triple-difference” regression coefficients in equation (4). See text for details. Source: LEHD.

Table 3: Testing for “Local” General Equilibrium Effects: County Level Earnings Estimates

	(1)	Baseline (2)	(3)	Levels (4)	Zero Earnings (5)
9 Year PDV	-0.110*** (0.034)	-0.091*** (0.038)	-0.090*** (0.042)	-3597.67*** (1342.59)	-3837.45*** (1670.12)
N	3823	3823	3823	3823	3823
Cohort FE	X	X	X	X	X
Experience		X	X	X	X
County Trends			X	X	X

NOTE: This table reports regression coefficients from equation (4). Each column corresponds to a different regression, where the results are presented for three different samples as indicated in the table headings. An observation is a cohort×year so that each cohort county has two observations per year. Standard errors are in parentheses and are calculated allowing for an arbitrary variance-covariance matrix within each county. The baseline sample is estimated using three different log specifications (columns (1)-(3)), and exponentiated coefficients are reported using the translation  $(\exp(\eta_1^k) - 1)$ . The regressions are weighted by the 1990 cohort size. The final row of the table reports the discounted sum of the coefficients using a 4 percent annual discount rate (with discount factor  $\beta$ ). Columns (1)-(3) use the translation  $\sum_{t=0}^8 \beta^k (\exp(\eta_1^k) - 1)$ . Standard errors for the final row are calculated using the delta method. Earnings are reported in 1990(\$). See text for details. Source: LEHD. \*\*\*, \*\*, \* denotes statistical significance at the 1, 5, and 10 percent levels, respectively.

Figure 5: Job Transition Rates from the Newly Regulated Sector



NOTE: Plotted are the coefficient estimates from a version of equation (4), which pertain to the difference in separation probabilities for the newly regulated sector after the change in regulations. See text for details. The dashed lines represent 95% confidence intervals. The bars represent the survivor probability for workers in polluting plants of newly designated nonattainment counties. Source: LEHD.

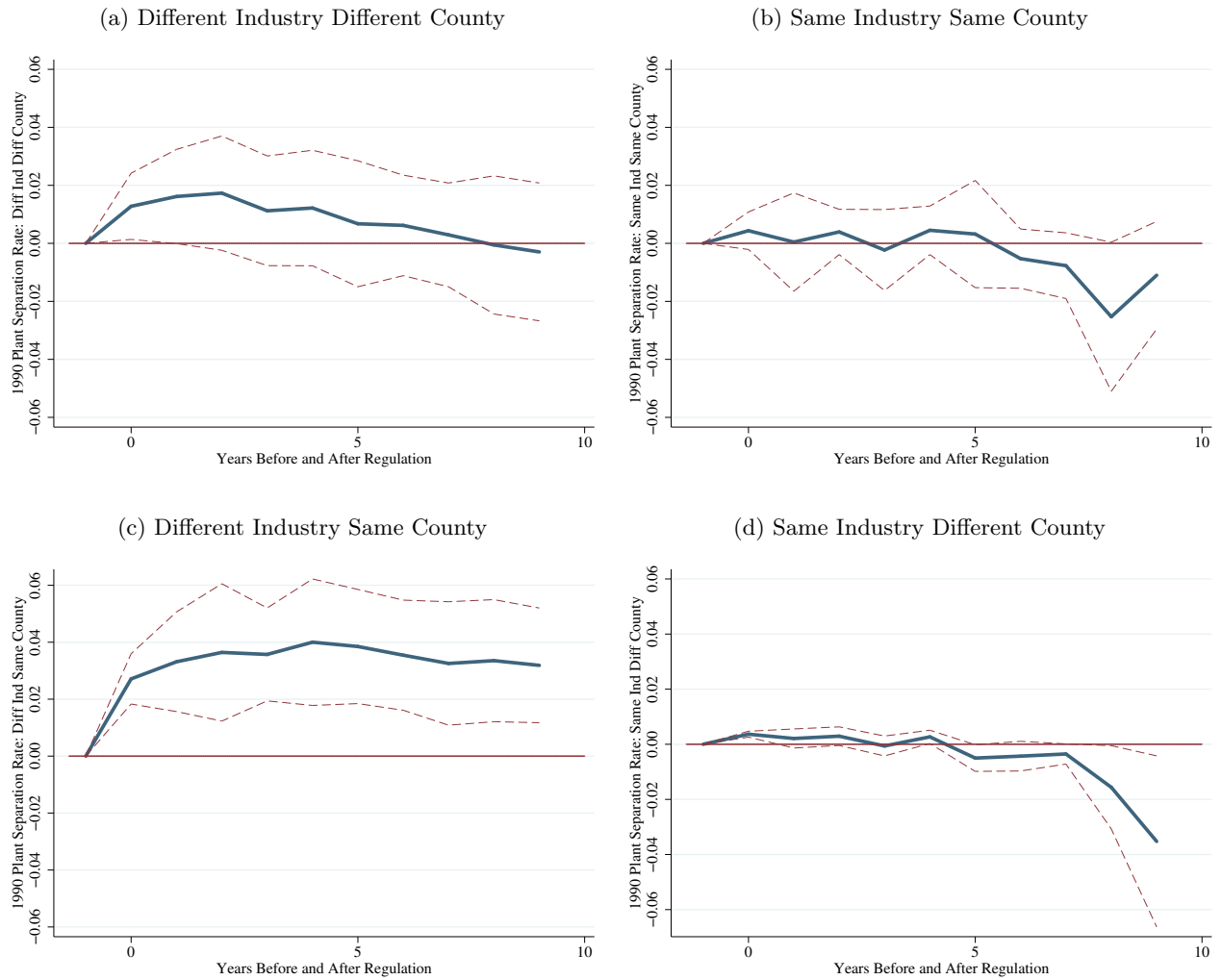
Table 4: Effect of Sector Level Regulation on Earnings: Decomposition Based on Sectoral Transitions

	Stayer (1)	Separator (2)	Separator: Same Industry Same County (3)	Separator: Diff. Industry Same County (4)	Separator Same Industry Diff. County (5)	Separator Diff. Industry Diff. County (6)
Regulation (t+0)	0.007 (0.022)	-0.125*** (0.038)	-0.056** (0.026)	-0.162*** (0.040)	-0.177*** (0.037)	-0.168*** (0.033)
Regulation (t+1)	-0.016 (0.028)	-0.218*** (0.028)	-0.151*** (0.032)	-0.235*** (0.024)	-0.196*** (0.050)	-0.245*** (0.022)
Regulation (t+2)	0.002 (0.022)	-0.308*** (0.016)	-0.189*** (0.030)	-0.298*** (0.018)	-0.241*** (0.026)	-0.313*** (0.013)
Regulation (t+3)	-0.006 (0.015)	-0.313*** (0.040)	-0.206*** (0.055)	-0.309*** (0.044)	-0.224*** (0.020)	-0.321*** (0.029)
Regulation (t+4)	-0.019 (0.021)	-0.244*** (0.038)	-0.143*** (0.041)	-0.230*** (0.046)	-0.153*** (0.021)	-0.249*** (0.024)
Regulation (t+5)	0.010 (0.020)	-0.194*** (0.024)	-0.117*** (0.043)	-0.185*** (0.028)	-0.140*** (0.021)	-0.188*** (0.014)
Regulation (t+6)	0.015 (0.014)	-0.118*** (0.011)	-0.011 (0.026)	-0.097*** (0.017)	-0.056 (0.035)	-0.116*** (0.013)
Regulation (t+7)	0.007 (0.014)	-0.050*** (0.014)	0.010 (0.027)	-0.043** (0.018)	-0.008 (0.031)	-0.017 (0.017)
Regulation (t+8)	0.006 (0.021)	-0.012 (0.016)	0.010 (0.026)	0.013 (0.031)	-0.003 (0.027)	0.008 (0.014)
9-Year PDV	0.008 (0.083)	-1.375*** (0.063)	-0.740*** (0.125)	-1.349*** (0.073)	-1.025*** (0.158)	-1.389*** (0.060)
N	6847	6847	6847	6847	6847	6847
Cohort FE	X	X	X	X	X	X
Experience	X	X	X	X	X	X

NOTE: This table reports regression coefficients from equation (4). Each column corresponds to a different regression, where the results are presented for six different samples as indicated in the table headings. An observation is a cohort×year so that each cohort county has two observations per year. Standard errors are in parentheses and are calculated allowing for an arbitrary variance-covariance matrix within each county. Exponentiated coefficients are reported using the translation  $(\exp(\eta_i^k) - 1)$ . The regressions are weighted by the 1990 cohort size. The final row of the table reports the discounted sum of the coefficients using a 4 percent annual discount rate (with discount factor  $\beta$ ) and the translation  $\sum_{t=0}^8 \beta^k (\exp(\eta_i^k) - 1)$ . Standard errors for the final row are calculated using the delta method. Earnings are reported in 1990(\$). See text for details. Source: LEHD.

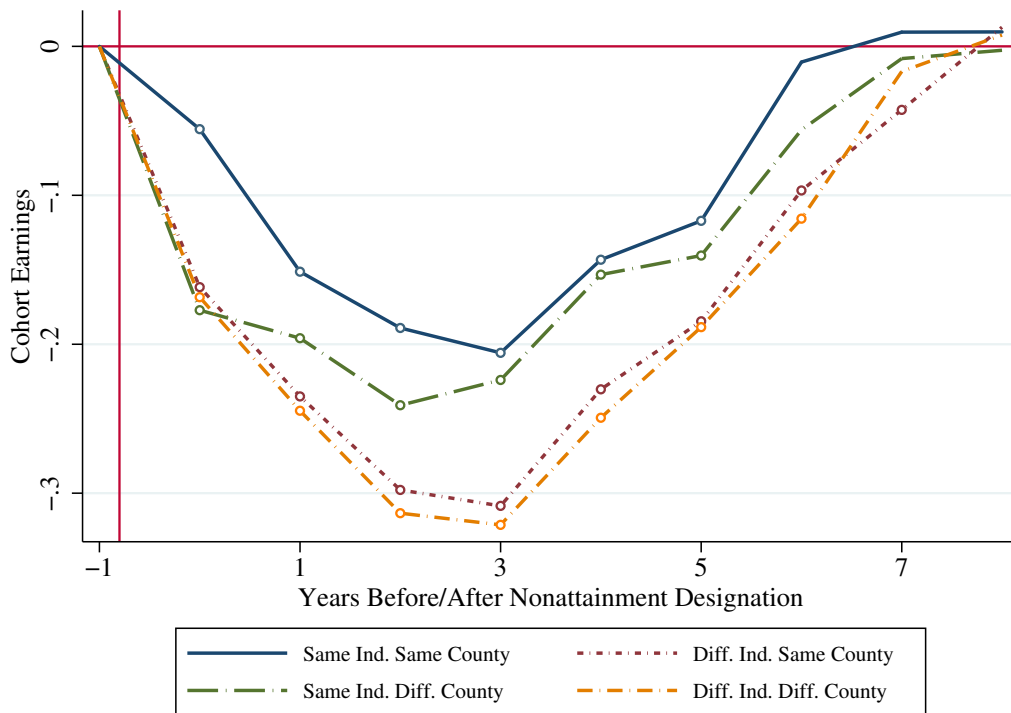
\*\*\*, \*\*, \* denotes statistical significance at the 1, 5, and 10 percent levels, respectively.

Figure 6: Decomposing Differences in Job Transition Rates



NOTE: Plotted are the coefficient estimates that decompose the separation probabilities into destination sectors. The plotted coefficients come from a version of equation (4), which pertain to the difference in transition probabilities for the newly regulated sector after the change in regulations. See text for details. The dashed lines represent 95% confidence intervals. Source: LEHD.

Figure 7: Effect of Sector Level Regulation on Earnings: Decomposition Based on Sectoral Transitions



NOTE: Plotted are the coefficient estimates from a various versions of equation (4) fit for each sub-category of job transitions. The figure shows that worker transitions that result in a new job in the same industry experience a far lower earnings loss than those workers who switch industries. The lines correspond to the estimates reported in columns (3)-(6) of Table 4. See Section 7.5 for details. Source: LEHD.

Table 5: 1990 Decennial Census Characteristics of In-Sample and Non-Sample Counties

	Attainment		Nonattainment		Switch Into Nonattainment		t-statistic
	In Sample (1)	Out of Sample (2)	In Sample (3)	Out of Sample (4)	In Sample (5)	Out of Sample (6)	Col. 5 - Col. 6 (7)
<b>Panel A: County Demographics</b>							
Population	70387.2	33998.6	515753.4	348729.8	87927.7	108540.5	-0.90
Population urban (percent)	0.529	0.307	0.791	0.671	0.470	0.517	-0.81
Median household income (dollars)	26549.8	22384.7	35505.1	31891.2	34893.9	28325.2	3.61
Median household rent (dollars)	331.4	298.6	465.1	453.7	440.2	389.1	1.84
Median home value (dollars)	52431.7	46115.8	90309.4	97000.8	84652.6	73951.7	1.36
<b>Panel B: County Industrial Composition</b>							
Agriculture, forestry, and fisheries	0.0237	0.0419	0.0113	0.0120	0.0247	0.0163	2.09
Mining	0.00350	0.00766	0.000826	0.00335	0.00117	0.00550	-4.30
Construction	0.0258	0.0290	0.0303	0.0308	0.0425	0.0335	2.52
Manufacturing, nondurable goods	0.0335	0.0355	0.0333	0.0336	0.0325	0.0323	0.05
Manufacturing, durable goods	0.0552	0.0432	0.0604	0.0538	0.0637	0.0459	2.10
Transportation	0.0194	0.0167	0.0214	0.0207	0.0192	0.0190	0.17
Communications and other public utilities	0.0112	0.0104	0.0124	0.0127	0.0148	0.0131	0.86
Wholesale trade	0.0165	0.0141	0.0228	0.0199	0.0188	0.0172	1.16
Retail trade	0.0802	0.0671	0.0818	0.0793	0.0809	0.0802	0.29
Finance, insurance, and real estate	0.0223	0.0169	0.0348	0.0303	0.0260	0.0251	0.44
Business and repair services	0.0153	0.0132	0.0233	0.0214	0.0183	0.0185	0.14
N							

NOTE: This table compares the various county characteristics for those counties in the LEHD base year files (MD, IL, WA, WI) to the rest of the counties in the United States, further split along the dimensions of variation in the dataset (i.e. Attainment, Nonattainment, and Switch into Nonattainment). Panel A displays basic demographic characteristics of the respective counties, whereas Panel B displays the industrial composition (as measured by employment). Industrial composition is measured as total number of workers in a given industry as a fraction of the total population. Source: 1990 Decennial Census.

## A Appendix: Data

### A.1 LEHD Earnings Data Construction

This section describes the basic data cleaning necessary to arrive at the analysis files used in this project. For a much more detailed description of the LEHD files see McKinney & Vilhuber (2008) and Abowd et al. (2008). I begin creating the data by creating a “base-year” file of workers who worked in the manufacturing industry in 1990 in the states for which I have data in 1990 (MD, IL, WA, WI). I then use the Employment History Files (EHF) from all states in all subsequent years to “follow” these workers over time, irrespective of whether or not they remain in their initial state. Using these earnings histories, I assign the worker to their highest paying job in a given year. I exclude any worker whose maximum annual earnings over their entire work history was less than \$10,000. Using this work history, I assign workers to “cohorts” where a cohort is defined by the county and sector that a worker worked in 1990.

Two important limitations of the LEHD data in the context of this project are the temporal coverage over time (detailed further in Table A.1) and the inability to observe the *unit* of work with certainty. To be more specific about the latter issue, state UI data records the *firm* that an employee works at in a given quarter. If a firm has more than one operating unit within the same state, then it becomes difficult to say with certainty which of the units the employee works. Approximately 30 to 40 percent of state-level employment is concentrated in employers that operate more than one establishment in that state. If there is more than a single plant belonging to a firm within a given state, the Census imputes place of work using a non-ignorable missing data model with multiple imputation. The model imputes establishment-of-employment using two key characteristics available in the LEHD Infrastructure Files: 1) distance between place-of-work and place-of-residence and 2) the distribution of employment across establishments of multiunit employers. This same imputation file is used in the creation of Census’s Quarterly Workforce Indicators. See McKinney & Vilhuber (2008) for more information.

This place of work distinction is important for my research design as CAA regulations only affect polluting plants in nonattainment counties. If a firm has a plant in a nonattainment county and another plant within an attainment county *of the same state*, then the only workers who should be affected are those in the nonattainment plant. Thus, I need the plant level imputation to distinguish between those workers who are affected and those who are not. In practice, around 60 percent of the workforce in my sample resides in a single-unit firm, and thus the remaining 40 percent of workers have to be matched using imputation.<sup>39</sup> The imputation procedure yields 10 (not mutually exclusive) implicates pertaining to a person’s place of work. This requires the creation of 10 separate datasets pertaining to the 10 implicates, estimating the model on each of these datasets, and combining the parameter estimates as in Rubin (1987).<sup>40</sup>

### A.2 Longitudinal Business Database (LBD)

The Longitudinal Business Database (LBD) is a longitudinal plant level database that covers the universe of establishments in the United States from 1975-2005. In each year, there consists of detailed industry information as well as information on employment and earnings for the particular plant. Also included in the data are years of entry and exit for the particular plant (and hence firm age). In the analysis, I limit the dataset to include only establishments within the manufacturing and utility sectors. I also exclude any establishments for which the maximum employment over the sample frame is less than 50 employees as well as any establishment with a lifespan of less than 3 years. Since EPA regulations primarily apply to major sources with potential to emit of more than 100 tons per year, excluding these smaller establishments has little effect on the estimates.

---

<sup>39</sup>I have also estimated the models in this paper using only workers from “single-unit” plants, and the results (not reported) are similar to those presented here.

<sup>40</sup>Multiple imputation is used in other popular datasets such as the National Health and Nutrition Examination Survey (NHANES).

Table A.1: LEHD State Year Coverage: Based on Years with Unit-to-Work Imputation Data

State	State Abbreviation	First Year	Last Year
Alabama	AL	2001	2004
Arkansas	AR	2002	2004
California	CA	1991	2003
Colorado	CO	1993	2004
Delaware	DE	1998	2004
Florida	FL	1992	2004
Iowa	IA	1998	2004
Idaho	ID	1991	2004
Illinois	IL	1990	2004
Indiana	IN	1998	2004
Kansas	KS	1993	2004
Kentucky	KY	2001	2004
Maryland	MD	1990	2004
Maine	ME	1996	2004
Minnesota	MN	1994	2004
Missouri	MO	1995	2004
Montana	MT	1993	2004
North Carolina	NC	1991	2003
North Dakota	ND	1998	2004
New Jersey	NJ	1996	2004
New Mexico	NM	1995	2004
Oklahoma	OK	1999	2004
Oregon	OR	1991	2004
Pennsylvania	PA	1997	2004
South Carolina	SC	1998	2004
Texas	TX	1995	2004
Virginia	VA	1995	2004
Vermont	VT	2000	2004
Washington	WA	1990	2004
Wisconsin	WI	1990	2004
West Virginia	WV	1997	2004

### A.3 EPA Air Facility Subsystem (AFS)

The EPA’s Air Facility System (AFS) contains compliance and enforcement data and permit data for stationary sources of air pollution regulated by EPA, state and local air pollution agencies. The permit data is the primary data element used in this analysis, detailing whether or not a particular plant has ever had a Title V Operating Permit, a New Source Review Permit, a “Prevention of Significant Deterioration” permit, or a “Pre-Construction” permit issued under the Clean Air Act. These are legally enforceable documents that authorities issue to air pollution sources whether or not they are in a nonattainment region. The permits are issued to all large sources (major sources) and a limited number of smaller sources (called area sources, minor sources, or non-major sources). Permits include pollution-control requirements from federal or state regulations that apply to a particular plant. Construction permits specify what construction is allowed, what emission limits must be met, and often how the source must be operated. Thus, for each AFS plant I match to the Census data (see section ??), I include an indicator for whether or not they have ever been issued one of these permits. This is my plant-level “polluter” indicator detailed in the text.

## A.4 Matching Algorithm and Match Results

To match plant level data from the EPA’s Air Facility Subsystem to the Longitudinal Business Database, I rely upon a name and address matching algorithm. Using information from the EPA and the Census on plant name, detailed address, and industry, I iterate over various combinations of these observables to construct unique matches between the two datasets. The name and address data from the Census comes from the Standard Statistical Establishment List (SSEL) which can be linked to the Longitudinal Business Database (LBD) and Longitudinal Employer Household Dynamics file (LEHD) after the matching process. Since I have annual plant level information from the Census and the EPA AFS database is only available as a cross-section, I construct a match candidate for each year in the Census data. That is, for every year in the SSEL, I perform a match of the SSEL to the AFS database using information on address, location, industry, and name. I then combine all possible matches across all years and select the modal match from the full distribution of potential matches. Thus, for each SSEL year in the loop, the basic matching algorithm is as follows:

1. Select a matching group of characteristics to match on (e.g. plant name, street number, street name, city, state)
2. Keep only the unique combination of these variables in each of the respective datasets
3. Perform a one-to-one merge across these characteristics
4. Save matched results
5. Loop back to beginning, removing matched results

This basic loop is performed for over 100 unique combinations of plant level characteristics for each SSEL year in the data, but a plant can only match once within the SSEL year loop. This creates a maximum of 20 possible matches for each plant. For various reasons some plants cannot be matched at all. Table A.2 details the various match statistics for the initial match to the SSEL (Column 2) and then after linking the SSEL to the LBD (Column 3). These statistics only correspond to the “1990 states” used to construct the cohorts (i.e. IL, MD, WA, WI). I then select the modal LBD match across all possible matches to construct the link. In the event that the mode is not unique, I randomize to select the match. Column (5) details the percentage of matches that contained a non-unique mode, and thus were randomly assigned to a modal match candidate.

Table A.2: AFS to Business Register Match Statistics: Multi Year Matching Algorithm

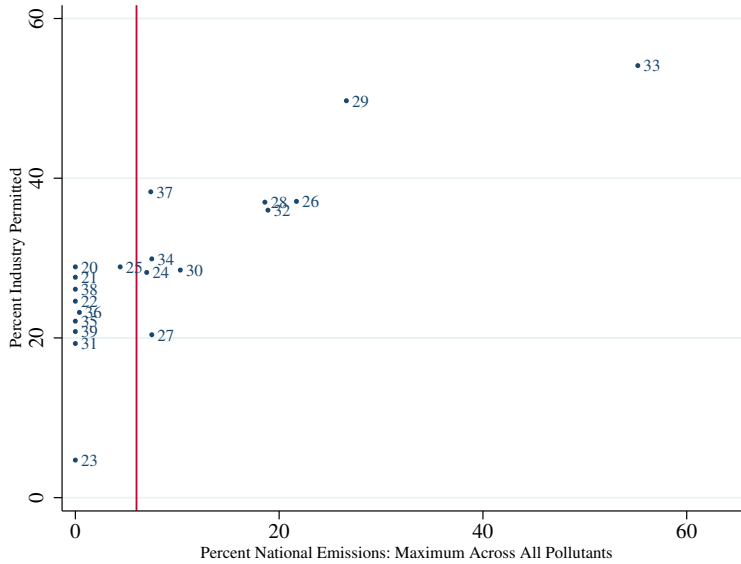
	(1)	(2)	(3)	(4)	(5)
	Total AFS Est.	SSL Match Pcnt.	LBD Match Pcnt.	Percent Non-Unique LBD	Percent Non-Unique Mode LBD
Match Statistics	8777	0.80	0.78	0.41	0.03

## A.5 Assessing the Validity of the “6 Percent Rule”

Greenstone (2002) examines the implications of CAA county level nonattainment designation on manufacturing activity as measured by employment, capital stock, and output. Since plant level regulation status was unobserved, the paper proxies for regulation status using a discrete cutoff for whether or not an industry was highly polluting relative to national industry emissions estimates. Specifically, Greenstone (2002) uses data from the EPA Sector Notebook Project (1985) which estimates the fraction of national criteria air emissions attributable to 2 digit industry classifications. These estimates are subsequently used to classify industries as “polluting” or “non-polluting” based on whether the industry was responsible for more than 6 percent of national emissions for a given pollutant. Since this serves as an important source of identifying variation in many subsequent studies (e.g. List et al. (2003), Becker (2005), Kahn & Mansur (2010), Greenstone, List & Syverson (2010)), it is worthwhile to assess how the regulation data used in this paper compare to the existing literature.

Figure A.1 plots the EPA Sector Notebook Project data against the fraction of the industry that is permitted (and hence regulated) in counties that are in nonattainment. The dots represent specific 2-Digit SIC codes which are listed to the right of the dot. For the “6 Percent Rule” to be a reasonable proxy for polluting/regulatory status we should see the fraction of plants permitted in industries to the right of the line nearing 100 percent. Moreover, the fraction of plants permitted an industry to the left of the line should be fairly close to 0.

Figure A.1: Fraction of Permitted Plants By Industry in Nonattainment Counties



NOTE: This figure shows the relationship between the fraction of an industry that is permitted in a nonattainment county (weighted by employment) plotted against the percent national emissions from the EPA Sector Notebook Project (1985). The red line represents the “6 Percent Rule” used in the previous literature (Greenstone 2002, List et al. 2003, Greenstone, List & Syverson 2010) whereby industries to the left are designated as unregulated, whereas industries to the right are designated regulated. Source LBD 1985-2005 and Greenstone (2002).

The figure suggests that not a single industry is classified near to 100 percent and there does not seem to be a strong discontinuity at 6 percent (x-axis). In order to more fully understand the implications of this discrepancy, I draw upon a cross-section of plant level data to relate observed plant level regulations with aggregate industry codes. The goal of this exercise is to understand the amount of variation in regulatory classifications explained by either 2-digit or 4-digit industry SIC codes. I estimate the following linear regression

$$\text{Permit}_{ij} = \alpha + \mathcal{J}'\Pi + \mu_{ij} \tag{6}$$

where  $\mathcal{J}$  is a vector of indicator variables, corresponding to either SIC 2-digit or 4-digit classifications. Table A.3 provides the adjusted R-squared for weighted and unweighted versions of equation (6), where the weights correspond to the average employment of a plant over the sample period (1985-2005).

Table A.3: How Well Do Aggregate Industry Codes Explain Plant-Level Criteria Air Operating Permits?

	2-Digit SIC	2-Digit SIC	4-Digit SIC	4-Digit SIC
Adjusted R-Squared	0.034	0.035	0.093	0.135
N	114426	114426	114426	114426
Weighted Regression		X		X

NOTE: This table reports adjusted R-square statistics from estimating equation (6). The sample consists of a single observation for every plant observed in a nonattainment county in the manufacturing industry from 1985-2005. Source: AFS, EPA Greenbook, and LBD.

## B Appendix: Tables

Table B.1: Baseline Regression Estimates: The Effect of Regulation on Cohort Earnings

	Prime-Age Male Age 20-55 (1)	Age 25-34 (2)	Age 35-44 (3)	Age 45-55 (4)
Regulation (t+0)	-0.022 (0.025)	0.007 (0.025)	-0.022 (0.024)	-0.115*** (0.038)
Regulation (t+1)	-0.055* (0.031)	-0.045 (0.032)	-0.061** (0.025)	-0.116*** (0.043)
Regulation (t+2)	-0.050** (0.023)	-0.035 (0.021)	-0.060*** (0.022)	-0.107*** (0.035)
Regulation (t+3)	-0.040** (0.016)	-0.022 (0.014)	-0.048*** (0.014)	-0.102*** (0.033)
Regulation (t+4)	-0.045** (0.018)	-0.041* (0.023)	-0.036** (0.014)	-0.105*** (0.019)
Regulation (t+5)	-0.011 (0.025)	0.004 (0.021)	-0.023 (0.023)	-0.067* (0.035)
Regulation (t+6)	-0.011 (0.028)	0.009 (0.020)	-0.021 (0.026)	-0.030 (0.024)
Regulation (t+7)	-0.000 (0.016)	0.017 (0.011)	-0.014 (0.017)	-0.022 (0.021)
Regulation (t+8)	0.006 (0.013)	-0.001 (0.011)	-0.011 (0.012)	0.025 (0.017)
9 Year PDV	-0.209* (0.121)	-0.101 (0.106)	-0.264** (0.126)	-0.583*** (0.193)
N	6847	6847	6847	6847
Cohort FE	X	X	X	X
Experience		X	X	X
County Trends			X	X

NOTE: This table reports regression coefficients from equation (4). Each column corresponds to a different regression, where the results are presented for four different samples as indicated in the table headings. An observation is a cohort×year so that each cohort county has two observations per year. Standard errors are in parentheses and are calculated allowing for an arbitrary variance-covariance matrix within each county. Exponentiated coefficients are reported using the translation  $(\exp(\eta_i^k) - 1)$ . The regressions are weighted by the 1990 cohort size. The final row of the table reports the discounted sum of the coefficients using a 4 percent annual discount rate (with discount factor  $\beta$ ) using the translation  $\sum_{t=0}^8 \beta^k (\exp(\eta_i^k) - 1)$ . Standard errors for the final row are calculated using the delta method. Earnings are reported in 1990(\$). See text for details. Source: LEHD.

\*\*\*, \*\*, \* denotes statistical significance at the 1, 5, and 10 percent levels, respectively.