The Transitional Costs of Policy: Evidence From the Clean Air Act and the Workforce

W. Reed Walker *

April 2011

Very Preliminary and Incomplete: Please do not circulate

Abstract

This paper characterizes and estimates the transitional costs of new environmental regulations in terms of costly workforce reallocation. New environmental regulations lead to a rearrangement of production away from polluting industries, and workers in those industries are disproportionately affected. To characterize and estimate the costs incurred by the affected workforce, I draw upon linked worker-firm data in the United States to follow workers over time and across employers, before and after new environmental regulations take place. I exploit panel variation induced by the 1990 Clean Air Act Amendments to estimate the change in earnings and sectoral transition rates for workers in newly regulated firms. I find that the reallocative costs of policy are significant, accounting for more than $9 billion in foregone wages for the 6 years after the change in policy. Most of these costs are driven by non-employment rather than reductions in worker productivity, and the effects differ quite dramatically by destination sectors. This work highlights the importance of accounting for frictions in the labor market when characterizing the cost and incidence of labor market adjustments.

*Walker: Columbia University, Department of Economics, 420 West 118th Street New York, NY 10027, rw2157@columbia.edu. I would like to thank Janet Currie, Wojciech Kopczuk, Bernard Salanie, Wolfram Schlenker, and Till von Wachter for invaluable advice and discussions. I would also like to thank Emek Baskar, Lucas Davis, Olivier Deschenes, Walker Hanlon, Matt Kotchen, Eric Verhoogen, and seminar participants at both the Census Bureau 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.

1 Introduction

Environmental policy pertaining to air pollution has been estimated to have large health benefits. However, these policies also come with costs. Rearrangement of production from pollution-intensive to other industries creates a broad set of private and social costs. Labor and capital are typically reallocated from newly regulated industries elsewhere within the economy. In terms of labor, this reallocation is often framed in terms of "jobs lost", and the distinction between "jobs versus the environment" is one of the most politically salient aspects of these regulations. 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, it is not clear that job loss should be a net cost when evaluating policy. Conversely, 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.

With this in mind, this paper estimates the degree to which changes in environmental regulation lead to *costly* labor reallocation in affected industries. In doing so, I provide the first estimates of the magnitude and duration of adjustment costs pertaining to worker reallocation in the context of the Clean Air Act Amendments of 1990. Prior research on the labor market impacts of environmental regulation has primarily focused on employment in manufacturing industries (Berman and Bui 2001, Greenstone 2002, Deschenes 2010, Kahn and 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. 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 (Arrow, Cropper, Eads, Hahn, Lave, Noll, Portney, Russell, Schmalensee, Smith, and Stavins 1996, Greenstone 2002). Even though this fact has been pointed out in numerous papers (Jaffe, Peterson, Portney, and Stavins 1995, Arrow, Cropper, Eads, Hahn, Lave, Noll, Portney, Russell, Schmalensee, Smith, and Stavins 1996, Greenstone 2002, Environmental Protection Agency 2005, Fullerton 2008, Congressional Budget Office 2009), 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. Moreover, these costs have important distributional implications for regulatory policy (Fullerton 2008).

To help frame the empirical analysis, I develop a micro-founded econometric model that characterizes the incidence of these reallocative costs. The theoretical framework specifies how workers are allocated across sectors, while allowing for simple characterizations of the incidence of the regulations. The heart of the model is a dynamic discrete choice decision for a worker choosing a sector, where worker

heterogeneity and costly mobility generate frictions in the reallocation process. The model suggests that both the speed of adjustment as well as the long run costs of labor reallocation are explicitly linked to costly mobility. That is, workers with higher mobility costs are reallocated back into the labor market slower and bear more of the incidence of the regulatory shock. The costs and incidence of workforce reallocation can be fully characterized by a set of reduced form parameters from the structural model in the spirit of Marschak (1953) and Chetty (2009).

I then draw upon detailed plant level regulatory data and longitudinal microdata on workers to estimate the reallocative costs associated with changes in environmental policy and workforce transitions. Using confidential linked worker-firm data from the United States Census Bureau, I am able to track workers across their job histories over time. I observe workers in their pre-regulation jobs, and I am able to compare their earnings with the earnings profiles of similar workers in unregulated firms before and after the change in environmental regulations. By matching plant level regulatory data from the EPA to this database, I can explicitly observe *plant level* regulatory status over time, something which has not been possible in previous work in this literature.[1]

The empirical framework estimates the impact of the regulatory shock on the evolution of wages for the workers in the newly regulated sector over time. Since the Clean Air Act is enforced at the county level, there is a tremendous amount of both temporal and spatial variation in the regulations. The primary regression framework follows groups of workers in newly regulated counties and sectors over time, before and after specific regulatory changes. This allows me to control both for fixed unobservable attributes specific to these groups of workers as well as any important national or sectoral trends that may covary with these county level regulatory changes.

Estimates 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

[1] Berman and Bui (2001) is an exception. However, the only focus on the Southern California Air Basin and Oil Refineries.

focusing on wages in affected industries misses important aspects of labor-market adjustment, namely costly non-employment between jobs.

Since previous literature suggests there exists important heterogeneity in the cost of job loss based on destination sectors (Jacobson, LaLonde, and Sullivan 1993, Neal 1995), I first estimate the relative likelihood of job transitions to other sectors of the economy as a function of the regulatory changes. I then estimate the wage effects associated with these various sectoral transitions. The results suggest that workers in the newly regulated sector disproportionately transition from their existing firm into a completely different industry, either within the same county or a new location entirely. Consistent with both the predictions of the model as well as previous research, I find that those workers who switch industries face significantly larger earnings penalties than workers who switch jobs but remain within the same industry.

This paper also contributes to previous work on the incidence, duration, and adjustment process of local labor demand shocks in the context of labor market frictions. Previous literature on regional adjustment to labor-market shocks to U.S. cities and states suggests that adjustment is sluggish, taking as long as 10 years to return to equilibrium wage levels, with most of this adjustment occurring through worker migration to other labor markets (Blanchard and Katz 1992, Bound and Holzer 2000, Greenstone and Looney 2010). There are several reasons why this work may not generalize into policy induced demand shocks. For example, environmental regulation limits the ability of new polluting firms to locate in a regulated area, effectively cutting off firm entry as a means for labor market adjustment and potentially lengthening the adjustment process relative to a freely functioning market. By focusing on this relatively well-defined, sector-specific shock and using highly-disaggregated data, I show that most of the reallocation occurs through movements *within* the local labor market rather than across labor markets as the previous literature suggests. The approach in this paper combines the relatively rich temporal dynamics found in Blanchard and Katz (1992) and Topel (1986) with the more local, disaggregated analysis of Bound and Holzer (2000) and Notowidigdo (2010) to provide a clear picture of local labor market adjustment to sector specific shocks.

A further contribution of this paper is to provide a new set of estimates for the cost of job loss. The best and most recent work pertaining to the costs of job loss comes from research analyzing the incidence of "mass layoffs" on workers long run earnings (Jacobson, LaLonde, and Sullivan 1993, Von Wachter, Song, and Manchester 2009). Mass layoffs, defined roughly as the occurrence of a 30 percent or greater reduction in employment at the firm level in a given year, are one particular form of displacement, and it is not clear these estimates generalize into a setting for which we see a more gradual reallocation of the workforce. Evidence here suggests that the costs of job loss are typically much lower than that found in

the "mass layoff" literature.

The rest of the paper is organized as follows. The following 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 theoretical 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 explores possible mechanisms for these results, and Section 9 offers a short discussion of the earnings results in the context of these mechanisms. Section 10 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. Regulation under the Clean Air Act is primarily administered at the county level. While regulation varies across counties, it also varies across 7 "criteria air pollutants". With the passage of the 1970 Clean Air Act Amendments, the EPA established national ambient air quality standards (NAAQS) that set the minimum level of air quality that all counties are required to meet for these criteria pollutants. Areas with high levels of a specific air pollutant are regulated only for that pollutant. The threshold for excessive pollution varies for each pollutant but is applied uniformly across the United States. In any given year, some counties find themselves over these thresholds, while others do not. Regulations primarily affect densely population areas where people live and work and where the potential benefits from abatement are larger.

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 forces local plants that emit that 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.

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 costly. Consequently, there is a body of research showing that regulations within the Clean Air Act have reduced the stock of polluting establishments (Henderson 1996), reduced the size of existing establishments (as measured through employment, capital, and output) (Greenstone 2002), and reduced the entry rate of new polluting establishments (Becker and Henderson 2000, List, Millimet, Fredriksson, and McHone 2003) within regulated counties. More recent work, drawing upon both better data and more recent time periods, supports this previous literature (Greenstone, List, and Syverson 2010, Kahn and Mansur 2010, Walker 2011).

Thus, while we see evidence of firm and industry level responses to environmental regulations, it is not clear how to interpret these responses in the context of economic incidence or welfare analysis, especially if resources are reallocated elsewhere within the economy. The next section presents a framework for analyzing a portion of these costs in the context of an equilibrium model of labor market adjustment.

3 A Mapping Between Regulatory Change and Reallocative Incidence

New regulations force plants to install expensive abatement technologies, so that for a given level of output plants must now use more inputs. Thus, a straightforward way to model environmental regulations is to view these regulations as affecting plant level productivity.[2] This is consistent with recent empirical evidence suggesting that CAA nonattainment designation lowers plant level productivity Greenstone, List, and 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. In the highly 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 labor demand.

With this in mind, I develop a spatial equilibrium model of labor market adjustment in the context of local productivity shocks. The model is similar in spirit to recent work by Artuc, Chaudhuri, and McLaren (2010), characterizing how workers are allocated across sectors, while providing a useful framework for evaluating the incidence of the regulation induced reallocation on the workforce. The heart of the model is a dynamic discrete choice labor supply decision, in which workers choose a sector to work based on both costly mobility constraints and expectations of future wage outcomes in all sectors.[3]

[2]There are many ways to model pollution or pollution abatement within a plant level production decision (e.g. see Copeland and Taylor (2004)).

[3]The canonical spatial equilibrium model of Rosen (1979) and Roback (1982) model includes an equilibrium condition in the housing market. I exclude changes in the housing market from this analysis while noting that changes in air quality are capitalized into housing values, increasing the value of housing in these areas (Chay and Greenstone 2005). This creates an interesting interaction between shocks to labor demand which reduce the relative cost of housing through worker outmigration (Roback 1982, Rosen 1979) and the capitalization effects pertaining to cleaner air. The implications of these interactions are

In order to characterize the distinctions implicit in the CAA regulations, the economy is modeled as consisting of different location-sector pairs (s). The sectors consist of the polluting sector (p), the non-polluting sector (g), and the residual sector (u) which can be thought of as as home production or an "out of the labor force" decision. The locations are divided into attainment counties (a) and nonattainment counties (n). This gives 5 independent location-sector pairs $S \in \{ag, ap, ng, np, u\}$ where, for example, the first term represents the non-polluting sector of attainment counties. Each sector potentially requires different skills and the locations are geographically distinct, making it costly for workers to move between them.

3.1 Production and Labor Demand

The quantity of aggregate output in production sector s in period t is given by:

$$q_t^s = F(L_t^s, \theta^s) = \theta^s H(L_t^s)$$

where L_t^s denotes labor used in location-sector s in time t and θ^s is the sector specific productivity parameter or demand shifter. The production function H is assumed to be increasing, strictly concave, and continuously differentiable in its first and only argument. Prices are exogenous in the model, where the price of a good in sector s is denoted p_s.[4] Let $L_t = \{L_t^{ag}, L_t^{ap}, L_t^{ng}, L_t^{np}, L_t^u\}$.

Assuming that wages are set on a demand curve, the wage is competitively determined as

$$w_t^s = p_s \theta^s H'(L_t^s) \tag{1}$$

so that workers are paid their marginal product.

3.2 Labor Supply

Workers are assumed to exist along a continuum of measure one, working in one of the five location-sectors. Worker i chooses to work in sector s in each period t. The utility of working in the sector depends on the wage in that sector w_t^s as well as an idiosyncratic benefit for working in a locations-sector s in a given year ϵ_{it}^s. The idiosyncratic term is meant to represent heterogeneity in people's preferences for working in a particular sector in a particular year. For example, a worker may become bored of work

left for future work.

[4]The assumption of price exogeneity can be relaxed, and the implications from the model remain the same. However, given that I have no information on prices, I maintain this assumption throughout. Furthermore, since I am examining local changes in regulations in the manufacturing industry in relatively small geographic areas of the economy, it is not clear that these particular manufacturers have any pricing power relative to the rest of the United States or the much larger international market.

or need to move to another location for personal reasons. On the other hand, the worker may really enjoy her work in a given period or have children who are attached to their school friends, making a move more costly. The terms ϵ_{it}^s are assumed to i.i.d. and drawn from distribution $G(\cdot)$ (with support $(-\infty, \infty)$). Denote the vector of ϵ_{it}^s terms in each period for each worker as ϵ_{it}. If a worker chooses to switch from location-sector j to location-sector s, she must pay a switching cost $C^{s(j)}$. This can be thought of as any costs pertaining to either migration or employment training associated with switching sectors and/or locations. The switching cost incurred by staying in the same sector $C^{s(s)}$ is normalized to zero. Thus, the flow utility for a worker i in sector s (who switched from sector j) at time t may be written as:

$$u_{it}^{s(j)} = w_t^s - C^{s(j)} + \epsilon_{it}^s. \tag{2}$$

where $C^{s(j)}$ is the cost incurred for a worker switching from industry j to industry s.

The worker's decision is dynamic, choosing which sector to work based upon their expectations of future wage trajectories in all other sectors. It is useful to think of the timing of the problem as follows: (i) in the beginning of the period a worker in sector j realizes their ϵ_{it}'s, (ii) the worker makes their sectoral choice decision and moves to their desired sector s, and (iii) upon arriving in sector s the worker pays their migration cost $C^{s(j)}$ and enjoys the flow utility in the current period $u_{it}^{s(j)}$. Let $\Omega_t = \{w_t, \epsilon_{it}\}$ be the state variable that contains all of the relevant information for the worker's decision, where w_t is the vector of wages in each sector and ϵ_{it} is the vector of person-specific, idiosyncratic shocks pertaining to each sector in time t. This dynamic optimization problem for a worker in sector j is described by the following equation:

$$V^j(\Omega_t) = \max_s \{u_{it}^{s(j)} + \beta E[V^s(\Omega_{t+1}) \mid \Omega_t]\}$$

where β is a common discount factor. Thus, in each period a worker chooses the sector s that maximizes the present discounted value of utility, conditional on the current wages in all sectors and the idiosyncratic preferences for each sector. For each state, the worker has an optimal decision rule over the sectors $s \in S$ which yields a choice-specific payoff

$$v_i^{s(j)}(\Omega_t) = u_{it}^{s(j)} + \beta E[V^s(\Omega_{t+1}) \mid \Omega_t] \tag{3}$$

In expectation, this choice specific value function is the same across all workers who moved from sector j to sector s:

$$E_\epsilon[v_i^{s(j)}(\Omega_t)] = v^{s(j)}(\Omega_t) = w_t^s - C^{s(j)} + \beta E[V^s(\Omega_{t+1}) \mid \Omega_t] \tag{4}$$

Consider a worker's decision to move from the nonattainment, polluting sector ($s = np$) to the attainment, non-polluting sector ($s = ag$). They incur a pecuniary cost $C^{ag(np)}$ as well as a non-pecuniary cost (or benefit), $\epsilon_{it}^{np} - \epsilon_{it}^{ag}$ pertaining to individual i's idiosyncratic preferences for sector ag relative to sector np in time t. Thus, the total moving costs for a worker moving from sector np to sector ag in year t can be written as

$$\epsilon_{it}^{np} - \epsilon_{it}^{ag} + C^{ag(np)}$$

where the first two terms constitute the net idiosyncratic cost/benefit to moving (which can be either positive or negative) and the last term is the common cost component. Worker i would move from sector np to sector ag if

$$v_i^{np(np)}(\Omega_t) < v_i^{ag(np)}(\Omega_t). \tag{5}$$

Using equations (2) and (3), we can rewrite this condition as:

$$\epsilon_{it}^{np} - \epsilon_{it}^{ag} < (w_t^{ag} - w_t^{np}) - C^{ag(np)} + \beta(E[V^{ag}(\Omega_{t+1}) \mid \Omega_t] - E[V^{np}(\Omega_{t+1}) \mid \Omega_t]). \tag{6}$$

The right hand side of equation (6) decomposes the decision to switch sectors into the three components: (i) the difference in wages between the two sectors, (ii) the pecuniary switching costs, and (iii) the difference in option value pertaining to being in each of the respective sectors. Let $\epsilon_{it}^{np} - \epsilon_{it}^{ag}$ be distributed according to the cdf $H(\cdot)$. Then the switching rule from equations (5) and (6) maps into aggregate choice probabilities via the following expression:

$$\Pr(s_t = ag \mid s_{t-1} = np) = H[(w_t^{ag} - w_t^{np}) - C^{ag(np)} + \beta(E[V^{ag}(\Omega_{t+1}) \mid \Omega_t] - E[V^{np}(\Omega_{t+1}) \mid \Omega_t])]$$

which I will denote as $\pi_t^{np,ag}$. Thus, the choice probability of switching to sector ag in period t conditional on being in sector np in period $t-1$ is defined by the probability that the difference in wages plus the difference in option values plus the mobility costs are sufficiently large enough to offset any differences in idiosyncratic sectoral preferences.

3.3 Welfare and Incidence

To characterize the equilibrium response to changes in environmental regulatory stringency at the sector level I model the regulations as a marginal change in the productivity of the polluting sector in the nonattainment counties (θ^{np}). Thus, the incidence of the regulatory shock on the workforce in the newly regulated sector (np) is simply the derivative of equation (4) with respect to the productivity of

the polluting sector (θ^{np}):

$$\frac{dv^{np}}{d\theta^{np}} = \frac{dw_t^{np}}{d\theta^{np}} + \frac{d}{d\theta^{np}}\beta E[V^{np}(\Omega_{t+1}) \mid \Omega_t]$$

$$= \frac{dw_t^{np}}{d\theta^{np}} + \beta \sum_{s \in S} \pi_{t+1}^{np,s} \frac{dV_{t+1}^{s}}{d\theta^{np}}$$

where the second line follows from an application of the envelope theorem.[5] This equation can be solved forward as

$$\frac{dv_t^{np}}{d\theta^{np}} = \sum_{n=0}^{\infty} \beta^n \sum_{s \in S} \pi_{t+n}^{np,s} \frac{dw_{t+n}^{s}}{d\theta^{np}} \tag{7}$$

where $\pi_{t+n}^{np,s}$ is the probability of ending up in sector s having started in sector np in period 0. Note that even though the microeconomic choices of choosing a sector are discrete, the value function is smooth because of aggregation (Chetty 2009). Equation 7 is the key equation that motivates the rest of the empirical analysis. It tells us that irrespective of mobility costs, the incidence of the regulation can be fully characterized by the present discounted value of future wage changes as a function of the regulatory shock. What is clear from equation (7) is that wage changes in a single sector or from a single point in time are insufficient for identifying the incidence of regulation induced labor reallocation. The remaining portion of the paper seeks to credibly estimate the parameters from equation (7) providing estimates of policy induced workforce reallocation costs.

4 The Clean Air Act as a Research Design

Due to the way in which the CAA is implemented, there is a tremendous amount of regulatory variation in the data. This section briefly describes the main sources of variation and how they relate to the paper's primary research design.

Figure 1a shows counties in the United States that were in nonattainment for any pollutant in 1991. While there is a considerable amount of variation in the cross-section, a notable feature of nonattainment designation is that it is primarily concentrated in large metropolitan areas. This suggests that relying purely on the cross sectional aspect of the regulations might confound regulations with unobserved features of urban labor markets. The advantage of cross sectional variation is that since only some

[5] Assuming agents are optimizing with respect to their location-sector choice, the envelope theorem allows us to ignore any changes in the switching probabilities (π_t^{js}) resulting from a marginal change in the productivity parameter. See Appendix A.1 for additional details of the derivation.

counties are regulated in a given year, it is much easier to control 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.

Since counties go in and out of nonattainment based on annual pollution levels, important time series variation exists, allowing for pre-post comparisons within counties or within plants. By including county or plant fixed effects I can control for unobserved determinants of employment or wages within a labor market. Accordingly, the inclusion of county, county×sector, or plant 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. Figure 1b plots only those counties that switched into nonattainment 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. Unlike previous studies, I observe both polluting plants and what those plants emit. Within a specific 2-digit industry SIC code, 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. This figure shows that relying on industrywide classifications of "polluting industries" to proxy for regulatory status is problematic, as only a small fraction of any given industry in a nonattainment county is regulated in a given year.

In practice, counties often keep their attainment/nonattainment status for long periods of time. Moreover, when counties do switch into (out of) nonattainment it is most often due to strong (weak) economic conditions which might lead to bias in my estimates.[6] With these two caveats in mind, I use an arguably cleaner source of variation in county level nonattainment status coming from revisions to the national level pollution standards (i.e. the 1990 Clean Air Act Amendments). In 1990, Congress passed a set of amendments to the Clean Air Act. Among other things, the 1990 Amendments created a new criteria pollutant class for particulates (PM-10) for which all counties where evaluated according to the new standard.[7] Furthermore, the EPA formally reviewed all county nonattainment designations for all criteria pollutants, adding 166 more counties to nonattainment designation across all pollutant classes.[8]

Taking all of these sources of variation into account, the baseline research design examines the wage

[6]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.

[7]Previously, the EPA regulated particulates in the form of Total Suspended Particulates (TSP's) which are larger and thought to be less pernicious than the smaller forms of particulates.

[8]Title IV of the 1990 Clean Air Act Amendment established a tradeable market for sulfur dioxide emissions. These markets apply nationally and do not correspond to the variation in nonattainment designation.

and employment trajectories of workers in these newly regulated plants, comparing them to similar workers in the same county and/or industry. To do this, I classify workers into "cohorts" based on where they worked the year prior to the introduction of the law, and I examine how the wages and employment histories of the newly regulated cohorts compare to unaffected cohorts over time. The identifying assumption is that the cohorts of workers in unaffected sectors or locations form a valid counterfactual for those workers in the newly regulated sector, after conditioning on various control variables such as cohort fixed effects and various sectoral trends. The following sections will discuss both the research design and the data in more detail, exploring further the sources of variation used in the analysis.

5 Data Sources and Summary Statistics

5.1 Data

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 B provides additional details.

The primary source of data used in this project comes from the Census Bureau's Longitudinal Employer Household Dynamics (LEHD) file which provides administrative, quarterly earnings records for a large percentage of the United States workforce.[9] I observe the complete employment history and earnings profile for each worker in the LEHD, conditional on the worker remaining in the sample. 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 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.[10] I then limit the sample to workers who worked in the manufacturing industry (i.e. 2 digit SIC 20-39) in 1990.

[9]The full LEHD file currently contains earnings records from over 30 states, with some states containing records as far back as 1985. See Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer, and Woodcock (2008) and McKinney and Vilhuber (2008) for a comprehensive discussion of both the construction and contents of the LEHD files.

[10]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 section B and Table B.1.

This leaves me with a sample 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.[11] In the event that a worker has more than one job in a year, I assign the worker to their highest paying job. 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.[12] I then collapse the data to the "cohort" by year level, so that I have a measure of the average earnings of a cohort over time. The research design compares the earnings profiles of workers in these various cohorts. Note that by creating cohorts, I allow workers to move to new jobs, sectors, and locations while still contributing to the average earnings statistics for their base-year cohort. Collapsing the data to the cohort by 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, and Mullainathan 2004).

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.[13] 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. Since I am relying on 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, and 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.[14]

[11] 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).

[12] From this point forward, sector is defined as either the "polluting sector" or the "non-polluting" sector. Definitions of what constitutes polluting and non-polluting plants/sectors are detailed later in this section.

[13] Using administrative data linked to survey data, Frijters and 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.

[14] In practice, sample selection issues pertaining to missing earnings are reduced dramatically by using annual earnings rather than quarterly earnings and by collapsing the data to the cohort county×sector year level. The records with missing data for a given year simply do not contribute to the cohort county×sector× year average in the year they are missing but any positive earnings reported by the individuals in future or past periods *do* count in the respective cell averages.

The LEHD only reports with certainty the firm at which the person worked (not the establishment) for each state. If there is only one establishment pertaining to a particular firm within a state, then the worker's plant is known with certainty. 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.[15] This 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.[16] 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).[17]

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.[18] 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 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. Similar to above, I collapse the data to the county×sector×year level to examine sector wide trends in employment before and after the regulations. Collapsing the data into sector level aggregates incorporates plant employment variation on both the intensive and extensive margins.

[15] The model imputes establishment of work using distance between place-of-work and place-of-residence and the distribution of employment across establishments of multi-unit employers. See McKinney and Vilhuber (2008) for the full details pertaining to the imputation procedure.

[16] 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.

[17] Multiple imputation is used in other popular datasets such as the National Health and Nutrition Examination Survey (NHANES).

[18] I only have access to the LBD from 1985-2005.

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 B.

The EPA's AFS database is a plant level database that provides information detailing the regulatory program(s) for which the plant is permitted (and regulated) as well as the specific pollutants for which the permit is issued.[19] 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, Millimet, Fredriksson, and McHone 2003, Becker 2005, Kahn and Mansur 2010, Greenstone, List, and Syverson 2010). See Appendix B 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.

5.2 Summary Statistics

Table 1 presents characteristics of the data - in terms of worker, firm, and sectoral characteristics. Panel A presents characteristics of plants in terms of total employment and age. 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 experience both faster growth (Dunne, Roberts, and Samuelson 1989) and tend to pay higher earnings conditional on size

[19] Note that these permits are simply operating permits and are unrelated to "cap and trade" permits from other pollution abatement programs.

(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 see 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 earnings profile across these groups. Namely, 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 Econometric Framework

6.1 The Incidence of Regulatory Reallocation

The model from Section 3 suggests that the incidence of the regulation is characterized by the conditional choice probabilities and wage changes associated with the regulations. The underlying structural parameters in the model are the mobility cost parameters which pin down the conditional choice probabilities and hence wages in equilibrium. In theory, one could estimate the structural model to recover the switching cost parameters and then simulate the model to calculate incidence using equation (7). This comes at the cost of having to make relatively strong functional form assumptions as well as having to specify the form of worker's expectations. I seek an alternative approach in the spirit of Marschak (1953) and Chetty (2009) that relies on the fact that I am primarily interested in the combination of parameters in equation (7) rather than the individual mobility cost parameters and conditional choice probabilities themselves.

Let N_c be an indicator equal to 1 for counties that were newly designated as nonattainment after the 1991 changes in nonattainment status. 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 for the latent productivity/labor demand shifter (θ_{st}) in the econometric model described below.

Consider the data generating process for wages in the *unregulated* sectors

$$\ln(w_{st}) = \rho_s + e_{st} + \gamma_t + n_{ct} + p_{st} + \epsilon_{st}$$

where ρ_s are sector level indicators meant to model time invariant observed or unobserved characteristics that govern wages within a sector; e_{st} is a quadratic in average sectoral experience[20]; γ_t are year fixed effects to capture overall trends in wages; n_{ct} is a trend variable for those counties that switch regulatory status; and p_{st} is a trend variable for the polluting sector. The error, ϵ_{st}, represents unobserved county×sector shocks to wages that would have occurred in the absence of treatment so that

$$E[\epsilon_{st} \mid \rho_s, e_{st}, \gamma_t, n_{ct}, p_{st}] = 0.$$

Thus, the change in wages due to a change in a county×sector's nonattainment designation may be written as

$$\ln(w_{st}) = \eta_1[N_c \times P_s \times 1(\tau_t > 0)] + \rho_s + e_{st} + \gamma_t + n_{ct} + p_{st} + \epsilon_{st} \tag{8}$$

where η_1 represents the heterogeneous reduced form effect of nonattainment designation mediated through changes in sector specific productivity. To allow for important dynamics in how regulations may affect wages, I generalize equation (8) to allow the changes to evolve incrementally

$$\ln(w_{st}) = \sum_{k=0}^{10} \eta_1^k[N_c \times P_s \times 1(\tau_t = k)] + \rho_s + e_{st} + \gamma_t + n_{ct} + p_{st} + \epsilon_{st} \tag{9}$$

Equation (9) is simply a generalization of a triple-difference estimator that allows for the effects to evolve over time as opposed to assuming that the regulatory effects occur immediately and last forever.

6.2 A Cohort Based Approach

To recover meaningful estimates of equation (7), I draw upon average wages in cohorts as opposed to sectors. Recall that a cohort is defined by where the worker worked in 1990, the year before the

[20] Sectoral experience is proxied by "potential experience" which is defined here as: average sectoral age - average sectoral education - 6.

implementation of the CAA Amendments. I then follow these cohorts over time, irrespective of whether or not they still work with their 1990 employer. As I show below, the cohort wage approximates the right hand side of equation (7). The intuition for this approximation comes from the fact that the cohort wage is simply the weighted average of the sectoral wages where the weights are the aggregate switching probabilities from the dynamic discrete choice problem.[21]

To formalize this argument, denote the average wage of a cohort from county-sector s, in year t as $\tilde{w}_{st}$, where the tilde henceforth represents a cohort specific variable. Letting $\tilde{L}_s$ denote the total number of workers in cohort s, the average cohort wage may be written as:

$$\tilde{w}_{st} \equiv \left[\sum_{i \in \tilde{L}_s} w_{ist}\right] \frac{1}{\tilde{L}_s}$$

where w_{ist} is the wage for worker i in period t. Notice that worker i can be in any sector s in period t and is not constrained to having the wage from their base-year sector, $\tilde{s}$. Then assuming, as in the model of Section 3, that all workers in a sector are paid the same wage:

$$\begin{aligned}\tilde{w}_{st} &= \left[\sum_{s \in S} \sum_{i \in \tilde{L}_s} 1(i \in s \mid \tau = t) \times w_{st}\right] \frac{1}{\tilde{L}_s} \\ &= \left[\sum_{s \in S} \pi_t^{\tilde{s},s} \times w_{st}\right]\end{aligned}$$

where $\pi_t^{\tilde{s},s}$ is the probability of being in sector s in time t, conditional on starting in sector $\tilde{s}$ in the base year. This result simply says that the wage of the cohort in period t approximates the weighted average of the sectoral wages, where the weights are equal to the conditional choice probabilities of switching from sector $\tilde{s}$ to sector s in the model outlined in Section 3.

Substituting $\ln(\tilde{w}_{st})$ for $\ln(w_{st})$ in equation (9), I arrive at the baseline regression specification:

$$\ln(\tilde{w}_{st}) = \sum_{k=0}^{10} \eta_1^k (\tilde{N}_c \times \tilde{P}_s \times 1(\tau_t = k)) + \tilde{\rho}_s + \tilde{e}_{st} + \gamma_t + \tilde{n}_{ct} + \tilde{p}_{st} + \epsilon_{st} \tag{10}$$

where the tildes represent cohort specific wages and controls as opposed to specific county-sector variables. Estimates of the η_1^k's correspond to the change in wages for those workers affected by the regulations, irrespective of whether or not they remain in their original sector. Appendix A.2 formalizes how the η_1^k's approximate the incidence derivation from the dynamic discrete choice model in Section 3.

The inclusion of the cohort specific fixed effects, $\tilde{\rho}_s$, forces the identification to come from within

[21] This follows from the modeling assumptions that all workers in the same sector are paid the same wage. The model is easily extended to allow for sector level heterogeneity in wages.

a cohort over time for those cohorts that were affected by the change in regulations. The lower order interaction terms essential to identification of the triple-difference estimator are included in the "switching cohort" by year fixed effects ($\tilde{n}_{ct}$) and the polluting cohort by year fixed effects ($\tilde{P}_s$). The other terms in equation (10) control for average cohort experience ($\tilde{e}_{st}$) and year fixed effects (γ_t) to account for national trends that might influence wages. Other reported specifications allow for cohort-county specific trends (i.e. a separate trend variable for each 1990 county in the data).

I focus on weighted versions of equation (10), where the weights are equal to the size of the cohort to account for heteroskedasticity associated with these size differences. All reported standard errors are clustered at the cohort-county level to account for the correlation in wages among cohorts in the same county, both within periods and over time.

7 Results

7.1 The Incidence of Regulation: Earnings Estimates

Table 2 presents estimates from various versions of equation (10) for the 6 years following the change in regulations. Each column in the table corresponds to a different regression. Column (1) of Table 2 constitutes the baseline estimate, suggesting 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. As the model in Section 3 makes clear, the total reallocative costs should be characterized by the present discounted value of future sectoral wage changes weighted by the probability of entering that sector. Since the cohort wage approximates this weighted average, the last row of Table 2 presents estimates of the total present discounted value of the earnings change over the 6 reported years, discounting the future earnings changes using a 4 percent annual discount rate. After 6 years the average worker in the affected cohort experienced a present discounted wage loss of around 25.6 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. By including cohort fixed effects, all of these estimates are identified off of within cohort comparisons of earnings before and after the change in policy. Moreover, the model also controls for nationwide trends in earnings, any shocks to the overall polluting sector, and any unobserved shocks to the "switching" counties. A potential confounding source of variation would be any unobserved shock to the polluting sector of counties that switched regulatory status in those years after the change in policy. Since there are significant number

of polluting and non-polluting firms within a county×4-digit industry code (i.e. see Figure 2), the unobserved shock would have to be specific to polluting firms and not to a specific industry in a county.

Column (2) of Table 2 includes a quadratic term for the average experience of the cohort in a given year. Since workers of the various sectors are paid different wages, the cohort fixed effect should control for the differences in levels across these sectors. The cohort fixed effect cannot account for differences in the growth rate of earnings across the various groups. Thus, the 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 6 year present discounted value still exceeds 22 percent of pre-period earnings. Column (3) adds a common trend for wages in each cohort county.[22] Here, estimates are identified by comparing wage changes before and after the policy, after netting out any earnings trends common to both the polluting and non-polluting sector of a given "county cohort". The results remain very similar to before.

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)[23], the total foregone wage bill is approximately 9 billion dollars.

The model in Table 2 assumes that the effect of the reallocative shock ends after 6 years. Figures 3a and 3b present graphical estimates of equation (10) for the 10 years following the policy shock. Specifically, the plotted coefficients in Figure 6a are event time indicators for the polluting and non-polluting sectors in counties that switched nonattainment status in 1991, controlling for county by sector 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 3b is simply the difference in the coefficients reported in Figure 3a. Thus, Figure 3b implicitly replicates the regression estimates presented in Table 2. The results in both figures are rather striking and correspond to the estimates seen in Table 2. The figures 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.

Recall that the baseline estimates ignore any missing *annual* earnings when calculating the average cohort earnings. Since missing earnings may occur through either unemployment or sample attrition, it is not a priori clear how to address such an issue. To understand the implications of missing earnings

[22] Recall that cohorts are defined by the 1990 county×sector. Thus, the "county trend" in Table 2 corresponds to a trend for each 1990 "county cohort".

[23] This last statistic comes from Table 1 of Walker (2011).

in my data, I replace any annual earnings data that is missing with a zero. In theory, this should serve as a lower bound on the earnings estimates as I am inaccurately classifying some sample attritors as unemployed. In practice, this distinction is not as clear because the zero-earnings classification will affect both treatment and control groups. This issue is not unique to this paper, and other papers have dealt with these limitations in similar fashions (e.g. Jacobson, LaLonde, and Sullivan (1993)). Columns (4) of Table 2 suggests that the losses are not significantly different from the baseline estimates.

Column (5) presents estimates where the dependent variable has been replaced with the average cohort earnings (as opposed to the log earnings), and the results remain qualitatively similar. Lastly, column (6) of Table 2 presents estimates using a subsample of the labor force that exhibits relatively high labor force attachment, male workers aged 25-55. Any biases pertaining to non-random exit of workers from the sample should be minimized in this sample (Jacobson, LaLonde, and Sullivan 1993, Von Wachter, Song, and Manchester 2009, Frijters and van der Klaauw 2006). While the estimates are somewhat smaller than the baseline estimates, the results remain qualitatively similar. Figure 4 plots the event study estimates for both the zero earnings sample (column (4) of Table 2) and the prime-age male workers sample (column (5) of Table 2), both of which display similar long run dynamics to the the baseline sample.

While I follow workers across sectors and over time, a potential source of bias may come from any equilibrium effects pertaining to the change in regulations that impact worker's wages. Any major shift in the sectoral labor supply or demand curves is likely to influence other sectors within the same county. If workers are suddenly forced to leave the newly regulated sector, this should increase the potential labor supply of the unregulated sector within that county (assuming the workers are substitutes), which should drive down the equilibrium wage in that sector. The earnings trends in Figure 3a shows some evidence of equilibrium effects. Consider the initial earnings loss of the affected cohort (i.e. the solid line in Figure 3a). If this is driven by costly job transitions *and* those workers move to the unregulated sector of the same county, then this would drive down the equilibrium wage in the unaffected sector. Notice that the earnings profile of the unregulated sector in Figure 3a (i.e. the dashed line in Figure 3a) dips slightly after the major decline in the regulated sector's earnings. While this may be over-interpreting the estimates, the figure still illustrates the point. However, Table 1 suggests that the size of the polluting sector in switching counties (i.e. column (6)) is small relative to the non-polluting sector. In such a case, any general equilibrium effects are likely to be small.

Up until this point, I have made no distinction as to the mechanisms generating the observed wage changes for the affected cohort. Presumably, these are 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 first decomposing these observed earnings changes into estimates for workers that remain at their firm and estimates for workers who leave. The latter earnings changes are then further stratified by destination location and sector.

7.2 Heterogeneity in Reallocation Costs by Destination Sectors

I begin by presenting earnings estimates for those workers in the affected sector who either 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 wages consist only of "stayers" or only of "leavers", depending upon the specification. Since switching decisions and destination sectors are endogenous decisions on the part of the agents, there is going to be selection into these samples. However, the form of selection has relatively clear predictions for the direction of bias in each of the regressions which I will highlight below. Moreover, the regression model controls for cohort fixed effects, so that any time-invariant unobservable characteristics of these "stayers" or "leavers" are controlled for.

Column (1) of Table 3 presents results from a version of equation (10) 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 1991-1994). The results from column (1) suggests that the wages of "stayers" are essentially unaffected by the regulatory change. One explanation for this zero estimate is one of wage rigidity. An alternative explanation comes from the fact that I am comparing stayers to other cohorts of stayers and leavers. Since stayers are on average higher quality workers than leavers, examining the wage path of stayers while using control groups that consist of both stayers and leavers will lead to upward bias in my earnings estimates.[24]

Column (2) of Table 3 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. 1991-1994). 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 bottom two rows of Table 3 show the present discounted earnings change relative to 1990 which suggests that the total foregone earnings for separators is nearly 130 percent of their pre-regulatory earnings. While these estimates are quite large in magnitude, they are significantly smaller than estimates found in the literature on displacement induced by mass

[24] 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. Namely, the regulatory effects for stayers seem unaffected.

layoffs (Jacobson, LaLonde, and Sullivan 1993, Von Wachter, Song, and 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. 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. In contrast, the job transitions from mass layoffs are very often thought to be characterized by involuntary job loss and prolonged unemployment durations (Jacobson, LaLonde, and Sullivan 1993, Davis, Faberman, and Haltiwanger 2006, Von Wachter, Song, and 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.

Since I observe where workers find their next job, I decompose column (2) of Table 3 based on the location and sector where they find their subsequent job. Specifically, I estimate the model for the following four destination groups: the same industry in the same county (column (3)), a different industry in the same county (column (4)), the same industry in a different county (column (5)), and a different industry in a different county (column (6)). The model in section 3 suggests that transitions with lower ex-ante mobility costs should experience a more rapid recovery relative to those with higher mobility costs, the latter needing a larger wage differential (and/or preference shock) to induce a move.

It seems plausible that the lowest ex-ante mobility costs should come from column (3), where a worker transitions into the same industry of the same county. The pecuniary costs of moving are likely negligible, and any losses pertaining to industry specific human capital (which is also a form of mobility cost in this model) would be minimized. The estimates suggest that the wage changes for these workers (i.e. those who stay within the same industry and the same county) are significantly less than those workers that change industries (even within the same county - i.e. column (4)). This is consistent with the previous literature suggesting that industry specific human capital matters in the cost of job transitions (Podgursky and Swaim 1987, Topel 1991, Neal 1995). Figure 5 displays the estimates from columns (3)-(6), summarizing the dynamics and incidence by destination sectors.

In results not reported here, I also estimate the incidence of workforce reallocation for different age categories in the data. I find that both young and old workers are disproportionately affected the industry reallocation compared to middle age workers. This is consistent with the fact that younger, less-skilled

workers bear the brunt of reductions in sectoral labor demand due to institutional arrangements in the firm (Farber 1999). Furthermore, older workers (ages 55 to 64) 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 Examining Possible Mechanisms

8.1 Regulation Leads to a Reduction in Sectoral Size

The discussions in the previous sections suggest that the earnings changes I observe have occurred through costly job transitions following a regulation induced reallocation of the workforce. While previous literature has shown that the CAA regulations lead to industry downsizing (Henderson 1996, Greenstone 2002, Kahn and Mansur 2010, Walker 2011), it is not clear how those results generalize into this particular setting. Moreover, no existing work has explicitly examined worker flows such as separation rates, which are key to interpreting the results from the previous section.[25] As the previous section has shown, job separations are the key component that determines incidence in the context of the earnings estimates from before. Therefore, understanding the relationship between changes in environmental regulations and the rate of job separation in the affected sectors is essential to interpreting the previous estimates.

To address the relationship between sectoral size and nonattainment designation, I draw upon a secondary data source, the Longitudinal Business Database (LBD), which allows me to construct industry aggregates for employment over time. An added benefit of the LBD is that the temporal coverage is much longer, allowing me to implicitly test the identifying "similar-trends" assumption from the previous sections. Differential pre-trends in the earnings data that correspond with the variation from the CAA may lead to biases in the estimates. Even though identification relies on within county×sector variation, other trends are pinned down by unregulated or "non-switching" cohorts. In the LBD, I observe plant level employment for the 5 years before and the 10 years after the change in regulations.

Consider the following generalization of equation (9):

$$\ln(L_{st}) = \sum_{k=-5}^{10} \eta_1^k (N_c \times P_s \times 1(\tau_t = k)) + \rho_{sc} + \gamma_t + n_{ct} + p_{st} + \epsilon_{cst} \tag{11}$$

where L_{st} is employment in sector s of county c in year t. Equation (11) incorporates pre-regulatory

[25] Previous work in this area has looked at plant entry (List, Millimet, Fredriksson, and McHone 2003), plant exit (Henderson 1996), employment (Greenstone 2002, Walker 2011), and gross job flows (Walker 2011).

lags where the excluded category is $k = -1$, so that trends in employment are measured prior to the change in regulations. The coefficients η_1^k for $k \in [-5, -2]$ provide an important validity test for the models in the previous section.

Figures 6a and 6b plot the coefficients from a version of equation (11) for the 5 years prior and 10 years after the changes to the 1990 CAA amendments using the LBD data. Specifically, Figure 6a plots event time indicators for the polluting sector and non-polluting sector of counties that switch into nonattainment with the introduction of the 1990 CAAA. Figure 6b plots the difference in the Figure 6a lines, corresponding to the η_1^k's from equation (11). The regressions are weighted by the 1985 county/sector level employment, and all estimates report robust standard errors clustered at the county level.

There are two important features from these figures. First, the trends in employment in the polluting and non-polluting sectors of "switching" counties for the years prior to the change are remarkably similar (as reflected by the zero pre-trend differences). The similar pre-trends of the non-polluting sector provide one piece of evidence for a research design using within county trends as a counterfactual. Secondly, beginning with the year of the regulatory change, the 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. Recall that these estimates control for any national trends in the polluting sector as well as any trends specific to those counties that switched regulatory status, and thus any cyclical differences pertaining to the recession in the early 1990's should be accounted for (conditional on the control group evolving similarly). The magnitudes and dynamics of the estimates are similar to those of Walker (2011) (specifically see Figure 1 of Walker (2011)), which uses the same dataset for the entire United States relative to this restricted subsample. There is a significant relationship between changes in regulatory stringency and employment in the affected sectors, where these changes occur through plant level adjustment on both the intensive (plant employment) and extensive (plant entry and exit) margins. See Walker (2011) for further discussion of these results as they pertain to plant and sectoral job destruction and job creation rates.

While Table 3 and section 7.2 presented estimation results for those who separate from their job, another implicit check on the estimates in the paper is to "scale" the cohort wage estimates from Table 2 by the number of workers we see transitioning out of the sector. This generates a back of the envelope approximation to the cost of job loss pertaining to policy induced reallocations. The average sectoral earnings loss for a cohort from Table 2 is around 23 percent (averaging across the columns). Estimates from Figure 6b suggest a more than 20 percent reduction in the sectoral work force in the 5 years after the policy. This suggests that the average earning loss for a worker who loses their job because of the

policy shock is around 115 percent of their 1990 earnings. Roughly speaking, this is consistent with what we see for those who separate from their job in Table 3.

8.2 Regulation Increases the Rate of Separations in Early Years

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. 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.[26] 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 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 empirical framework is essentially the same as used in equation (9) from before. I estimate various forms of the following regression model:

$$\tilde{S}_{1990,st} = \sum_{k=0}^{10} \eta_1^k (\tilde{N}_c \times \tilde{P}_s \times 1(\tau_t = k)) + \tilde{\rho}_{sc} + \gamma_t + \tilde{n}_{ct} + \tilde{p}_{st} + \epsilon_{cst} \tag{12}$$

where $\tilde{S}_{1990,st}$ is the failure rate for workers remaining at their 1990 "pre-regulation" firm in cohort county c, sector s, and year t. As before, this model includes cohort fixed effects, $\tilde{\rho}_{sc}$, so that identification comes from within a cohort over time. 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". I measure the change in the failure rate relative to the base year by excluding $k = -1$ from the analysis. The η_1^k's measure the differences in the failure rate as a function of the change in regulatory status. Similar to before, my ability to follow workers over their employment histories, allows me to decompose any changes in worker flows based on destination sectors to assess the degree to which workers remain in the same sector versus transitioning to other sectors/locations.

Figure 7 plots estimates of the η_1^k's from equation (12). All regressions are weighted by the pre-treatment, 1990 "cohort" size, and robust standard errors are clustered at the county level. Specifically,

[26] 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* $\text{plant}_{jt} \neq \text{plant}_{jt+1}$

Figure 7 plots the difference in the failure rate for the newly regulated sector relative to the year prior to the nonattainment designation. Accompanying Figure 7 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-sector 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 reallocation of labor away from the newly regulated sector. The results here parallel those from the cohort wage regression (i.e. Figure 3b) in that reallocation process evolves slowly over many years. Notice that the transitional dynamics here do not precisely match those in the cohort wage regressions. The reason for this is that the cohort wages are the average earnings across all individuals; whereas the failure rate is calculated as the rate of failures conditional on survival up to time t. Since most earnings losses occur through separations (as suggested above), we should weight the failure rate estimates by the total number of failures in a given year (i.e. the vertical difference between the blue bars across years) which will skew this picture to match the transitional dynamics of the cohort wage regressions in Figure 3b.

Decomposing these transition rates into destination sectors allows one to gauge the relative likelihood that an individual transitions out of the newly regulated sector into a new industry or location. These estimates are important in light of the significant earnings loss heterogeneity across destination sectors from Table 3. For each destination sector of interest, I construct separation rates so that a separation is only defined when a worker transitions to the sector of interest. This allows me to estimate the relative change in transition probabilities as a function of the policy change. Figure 8 presents these estimates for workers who transition to a different industry in a different county (8a), the same industry in the same county (8b), a different industry in the same county (8c), and the same industry in a different county (8d). What is immediately clear is 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 seem disproportionately more likely to transition to a different industry within the same county. As the earnings estimates suggested before, the fact that we see workers transitioning to other industries suggests that the incidence of the labor reallocation to the workforce will be larger than if workers transitioned within the same industry.

9 Interpretation and Discussion

The results from the previous sections suggest that changes in regulatory stringency lead to a reduction in the size of the sectoral labor force. Moreover, substantial frictions in the labor market lead to a costly reallocation of the workforce following the policy change. The earnings estimates suggest that the average earnings for a worker in a newly regulated sector declined by more than 6 percent in the years following the change in regulations. These earnings changes are driven by costly job transitions rather than a reduction in sectoral labor productivity. 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 4 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). A priori, it is hard to say how this may affect generalizability. A useful statistic for a future version of this paper is the employment and plant concentrations in the polluting sector for both in and out of sample counties.

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 that 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.[27]

[27] This logic stems from the literature on endogenous production. See e.g. Jovanovic (1982), Hopenhayn (1992).

9.1 Labor Reallocation and Welfare

If we take the model from section 3 seriously, namely that workers are paid their marginal product and firms earn zero profits, then all wage losses are social costs, and the estimates allow for an easy approximation of welfare costs pertaining to policy induced job reallocation. Of course, if there are general equilibrium effects pertaining to lower wages in the non-affected sectors, this measure will tend to understate the total welfare losses. Moreover, if wages are non-competitive and incorporate rents specific to the firm and/or worker, then this framework will tend to overstate the social costs of sectoral reallocations. With any job loss, both human capital and job specific rents are reflected in the wage lost, the former being a social cost, while the latter arguably is not. Disentangling the component of reallocation costs pertaining to non-competitive wage setting is beyond the scope of this work, but it is an interesting area for future research.

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). Consider the extreme assumption that all of the displaced workers in the affected sector were reimbursed 100 percent for any wage change associated with job loss. In this case, UI is a transfer that primarily reflects income redistribution among citizens via taxation and transfers, and should not be completely calculated as an economic loss. However, the social cost of public funds or the deadweight loss of taxation is non-trivial, with recent estimates amounting to almost 40 cents on the dollar (Gruber 2010). Thus, in the case of 100 percent reimbursement, we would multiply the aggregated wage estimates by 40 percent to get the deadweight loss pertaining to the increased UI transfers. Realistically, workers are not nearly compensated in full, and the true loss will be the weighted average of the social cost of public funds and the uncompensated wage sectoral wage effects. Moreover, we should also count the regulation induced rise in labor force nonparticipation as a deadweight loss if the workers market wage (prior to the shock) exceeds their value of leisure (Autor, Dorn, and Hanson 2011).

Lastly, the conceptual framework suggests that any option value associated with being in a new sector will tend to reduce the welfare losses of labor reallocation. Previous literature suggests that these welfare gains from option value can exceed that of the short term earnings losses that workers experience (See e.g. Artuc, Chaudhuri, and McLaren (2010)).

9.2 Costs Relative to Benefits and Policy Prescriptions

Given the caveats pertaining to welfare mentioned above, it is still useful to compare the costs estimated here to any monetized benefits of the policy. 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).[28] These benefits are many orders 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, and Syverson 2010) as well as any adjustment costs associated with reallocating the capital stock (Goolsbee and Gross 1997, Ramey and Shapiro 2001).[29]

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.

10 Conclusion and Future Work

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. While these estimates are large, they are small relative to the health and amenity benefits pertaining to cleaner air.

The findings 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, and costly job transitions from the 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. Moreover, one possible interpretation of the estimates presented here suggests that workers would benefit (in the form of higher earnings) by being able to move locations to find a job in their previous industry. For various reasons pertaining to costly mobility, this does not happen very often in the data. Since the costs of relocating for employment are front-loaded and difficult to finance, helping workers move in search of

[28] Note that these estimates are for Titles I-V of the CAA regulations and do not include the estimated benefits from the SO_2 trading markets, which are not analyzed in this study.

[29] While both Goolsbee and Gross (1997) and Ramey and 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.

a better job could both aid job matching and reduce the length of unemployment for affected workers (Greenstone and Looney 2010, Ludwig and Raphael 2010).

This paper presents a useful framework for thinking about the costs of job loss and labor reallocation in the context of policy changes. 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 pertaining to the policy shock. The particular approach could presumably be applied to any area concerned with the magnitude and duration of reallocative frictions, 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.

Future work pertaining to this project intends to link the LEHD to survey data from the Census to be able to more accurately assess transitions to informality and/or to distinguish voluntary from involuntary job transitions. Moreover, better understanding any heterogeneity at the industry level remains a high priority. Lastly, obtaining structural estimates of the underlying mobility cost parameters and seeing how counterfactual income transfers affect both the speed and incidence of labor reallocation remains a high priority in the future of this project.

References

Abowd, J. M., B. E. Stephens, L. Vilhuber, F. Andersson, K. L. McKinney, M. Roemer, and S. Woodcock (2008): "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," *LEHD Working Paper*.

Arrow, K., M. Cropper, G. Eads, R. Hahn, L. Lave, R. Noll, P. Portney, M. Russell, R. Schmalensee, K. Smith, and R. Stavins (1996): "Benefit-Cost Analysis in Environmental, Health, and Safety Regulation," *Washington, DC: American Enterprise Institute*, pp. 1–17.

Artuc, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *The American Economic Review*, 100(3), 1008–1045.

Autor, D., D. Dorn, and G. Hanson (2011): "The China Syndrome: Local Labor Market E!ects of Import Competition in the United States," *Working Paper*.

Becker, R. (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, R., and V. Henderson (2000): "Effects of air quality regulations on polluting industries," *Journal of Political Economy*, 108(2), 379–421.

Berman, E., and L. Bui (2001): "Environmental regulation and productivity: evidence from oil refineries," *Review of Economics and Statistics*, 83(3), 498–510.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119(1), 249–275.

Blanchard, O., and L. Katz (1992): "Regional evolutions," *Brookings papers on economic activity*, pp. 1–75.

Bound, J., and H. Holzer (2000): "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s," *Journal of Labor Economics*, 18(1), 20–54.

Busso, M., J. Gregory, and P. Kline (2010): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *mimeo*.

Chay, K., and M. Greenstone (2005): "Does air quality matter? Evidence from the housing market," *Journal of Political Economy*, 113(2).

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.

Congressional Budget Office (2009): "The Economic Effects of Legislation to Reduce Greenhouse-Gas Emissions," .

Copeland, B., and S. Taylor (2004): "Trade, Growth, and the Environment," *Journal of Economic Literature*, 42(1), 7–71.

Davis, S., J. Faberman, and J. Haltiwanger (2006): "The flow approach to labor markets: new data sources and micromacro links," *The Journal of Economic Perspectives*, 20(3), 3–26.

Deschenes, O. (2010): "Climate Policy and Labor Markets," *NBER Working Paper*.

Dunne, T., M. Roberts, and L. Samuelson (1989): "The growth and failure of US manufacturing plants," *The Quarterly Journal of Economics*, 104(4), 671–698.

Environmental Protection Agency (1999): "The Benefits and Costs of the Clean Air Act 1990 to 2010," *EPA Report to Congress*.

——— (2005): "Environmental Economics Research Strategy," 600(R-04), 195.

——— (2008): "National Air Quality-Status and Trends through 2007," .

Farber, H. (1999): "Mobility and stability: The dynamics of job change in labor markets," *Handbook of labor economics*, 3, 2439–2483.

Frijters, P., and B. van der Klaauw (2006): "Job Search with Nonparticipation," *The Economic Journal*, 116(508), 45–83.

Fullerton, D. (2008): "Distributional effects of environmental and energy policy: an introduction," *NBER working paper*, (W14241).

Goolsbee, A., and D. Gross (1997): "Estimating adjustment costs with data on heterogeneous capital goods," *NBER working paper*, (W6342).

Greenstone, M. (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, M., J. List, and C. Syverson (2010): "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing," *mimeo*.

Greenstone, M., and A. Looney (2010): "An Economic Strategy to Renew American Communities," *Brookings: Hamilton Project*.

Gruber, J. (2010): *Public Finance and Public Policy*. Worth Publishers, 3rd edition edn.

Henderson, V. (1996): "Effects of air quality regulation," *The American Economic Review*, 86(4), 789–813.

Hopenhayn, H. (1992): "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 60(5), 1127–1150.

Jacobson, L., R. LaLonde, and D. Sullivan (1993): "Earnings losses of displaced workers," *The American Economic Review*, 83(4), 685–709.

Jaffe, A., S. Peterson, P. Portney, and R. 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, B. (1982): "Selection and the Evolution of Industry," *Econometrica: Journal of the Econometric Society*, 50(3), 649–670.

Kahn, M., and E. 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, J., D. Millimet, P. Fredriksson, and W. 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, J., and S. Raphael (2010): "The Mobility Bank: Increasing Residential Mobility to Boost Economic Mobility," *Brookings: Hamilton Project*.

Marschak, J. (1953): "Economic measurements for policy and prediction," *Studies in econometric method*, pp. 1–26.

McKinney, K. L., and L. Vilhuber (2008): "LEHD Infrastructure files in the Census RDC - Overview," *LEHD Working Paper*, Revision : 219.

Mincer, J. (1974): *Schooling, experience, and earnings*. National Bureau of Economic Research: distributed by Columbia University Press.

Neal, D. (1995): "Industry-specific human capital: Evidence from displaced workers," *Journal of labor Economics*, pp. 653–677.

Notowidigdo, M. (2010): "The Incidence of Local Labor Demand Shocks," *Unpublished, University of Chicago, Graduate School of Business*.

Podgursky, M., and P. Swaim (1987): "Job displacement and earnings loss: Evidence from the Displaced Worker Survey," *Industrial and Labor Relations Review*, 41(1), 17–29.

Ramey, V., and M. Shapiro (2001): "Displaced capital: A study of aerospace plant closings," *Journal of Political Economy*, 109(5), 958–992.

Roback, J. (1982): "Wages, rents, and the quality of life," *The Journal of Political Economy*, 90(6).

Rosen, S. (1979): "Wage-based indexes of urban quality of life," *Current Issues in Urban Economics*, 3.

Rubin, D. B. (1987): *Multiple imputation for nonresponse in surveys*. John Wiley & Sons Inc.

Schmieder, J. (2010): "Labor Costs and the Evolution of New Establishments," *mimeo*.

Topel, R. (1991): "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 99(1).

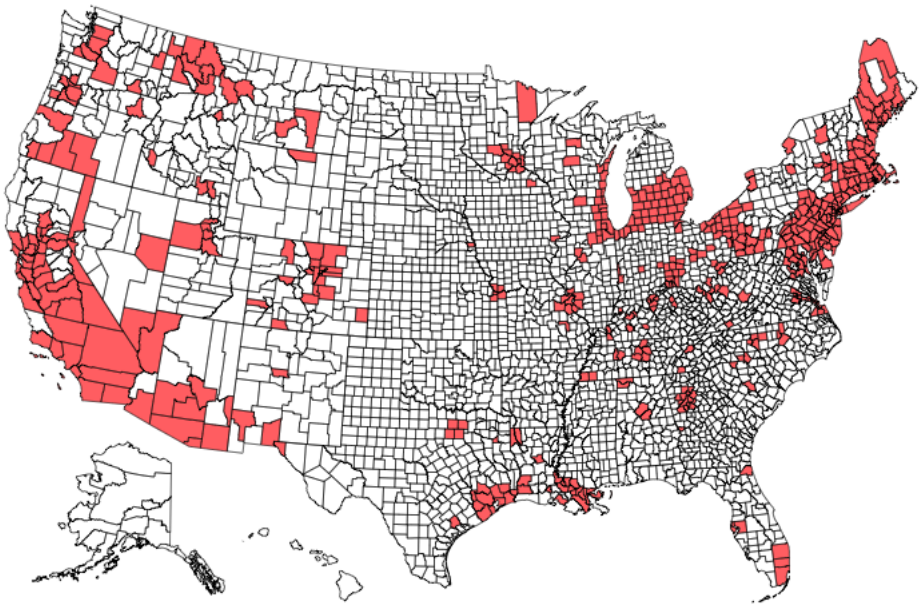
Topel, R. H. (1986): "Local Labor Markets," *The Journal of Political Economy*, 94(S3).

Von Wachter, T., J. Song, and J. 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*, pp. 0708–16.

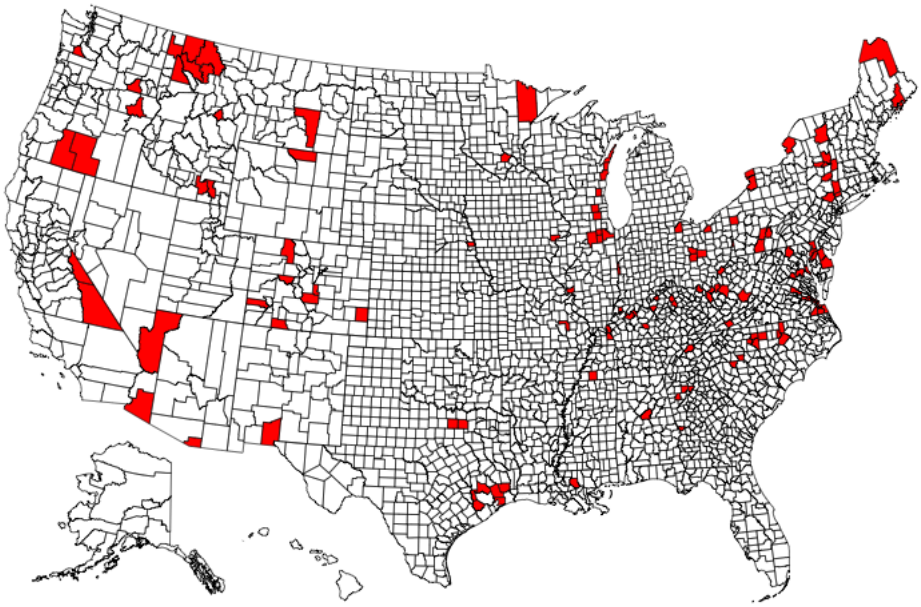
Walker, R. (2011): "Environmental Regulation and Labor Reallocation," *American Economic Review: Papers and Proceedings*, 101(2).

Figure 1: Nonattainment Counties Across All Pollutants in 1991

(a) Nonattainment for Any Pollutant 1991

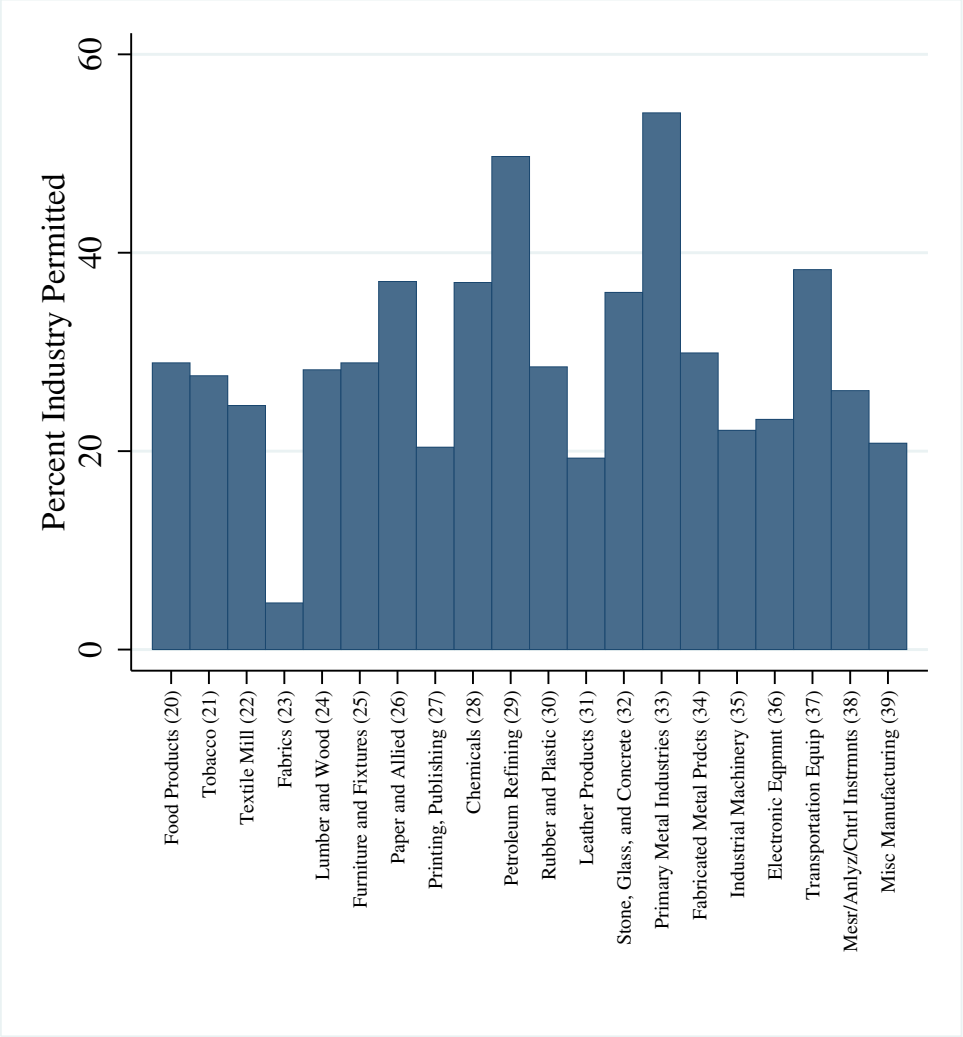


(b) New Nonattainment Counties 1991



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 industrywide 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	Polluting	Non Polluting	Polluting	Non Polluting	Polluting
	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.

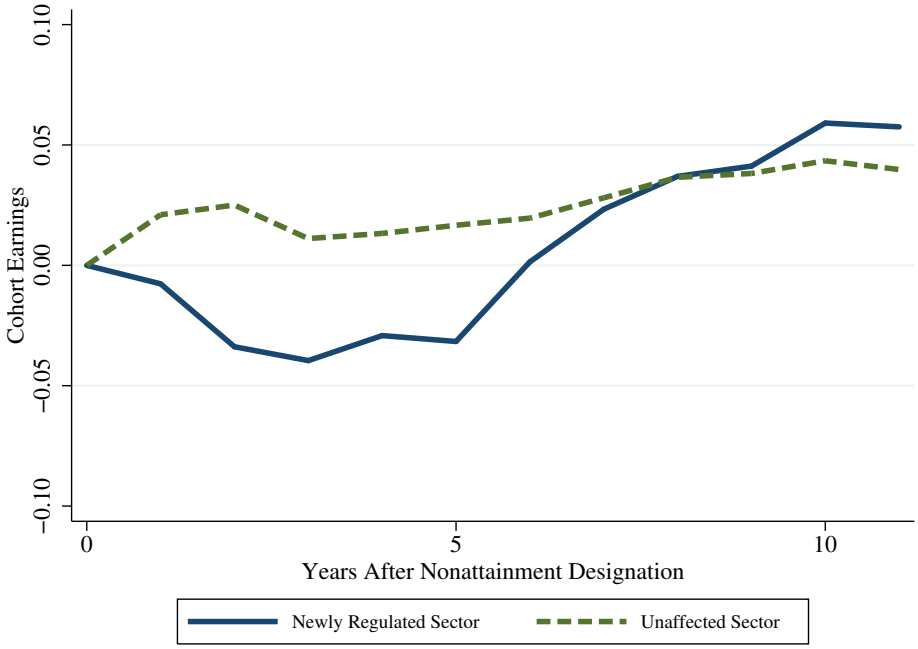
Table 2: Effect of Sector Level Regulation on Earnings

	Baseline			Zero Earnings	Earnings Levels	Prime-Age Males
	(1)	(2)	(3)	(4)	(5)	(6)
Regulation (t)	-0.036	-0.029	-0.03	-0.024	-1052.622	-0.014
	(0.029)	(0.022)	(0.018)	(0.022)	(768.412)	(0.024)
Regulation (t+1)	-0.066*	-0.059**	-0.060**	-0.057*	-1897.148*	-0.048
	(0.034)	(0.027)	(0.025)	(0.032)	(1004.522)	(0.033)
Regulation (t+2)	-0.057**	-0.051***	-0.052***	-0.052**	-1905.788***	-0.046*
	(0.024)	(0.018)	(0.017)	(0.021)	(731.365)	(0.024)
Regulation (t+3)	-0.048***	-0.043***	-0.043***	-0.045***	-1748.079***	-0.039**
	(0.014)	(0.012)	(0.012)	(0.013)	(529.402)	(0.016)
Regulation (t+4)	-0.052***	-0.049***	-0.049**	-0.045**	-2000.752**	-0.047**
	(0.017)	(0.019)	(0.020)	(0.018)	(831.568)	(0.019)
Regulation (t+5)	-0.021	-0.019	-0.019	-0.014	-794.225	-0.016
	(0.020)	(0.017)	(0.016)	(0.019)	(593.868)	(0.023)
6 Year PDV	-0.256**	-0.228***	-0.231***	-0.217***	-8555.893***	-0.190**
	(0.099)	(0.070)	(0.061)	(0.081)	(2665.785)	(0.096)
N	6847	6847	6847	6847	6847	6847
Cohort FE	X	X	X	X	X	X
Experience		X	X	X	X	X
County Trends			X	X	X	X

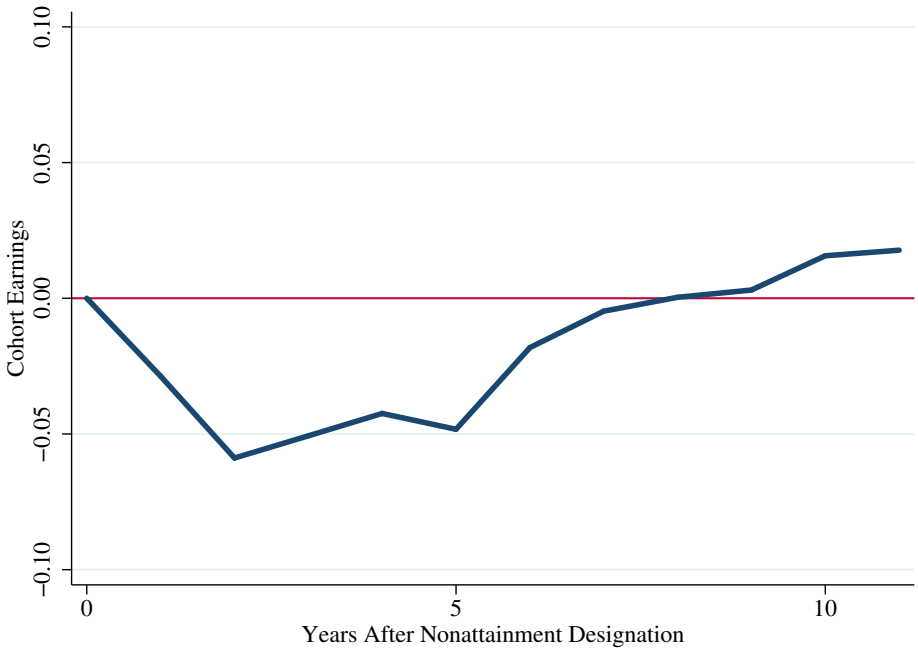
NOTE: This table reports regression coefficients from equation (10). Each column corresponds to a different regression, where the results are presented for three different samples as indicated in the table headings. The baseline sample is estimated using three different specifications (columns (1)-(3)). Cluster robust standard errors are reported where errors are clustered at the county level. An observation is a cohort×year so that each cohort county has two observations per year. The regressions are weighted by the 1990 cohort size. The final row of the table reports the discounted sum of the coefficients. Estimates are presented using a annual discount rate of 4 percent, and standard errors are calculated using the delta method. Earnings are reported in 1990($). See text for details. Source: LEHD.

Figure 3: Sector Level Wage Trends After Nonattainment Designation

(a) Wage Trends by Sector After Nonattainment Designation



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



NOTE: Plotted are the coefficient estimates from a version of equation (10). 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 (10). See text for details. Source: LEHD.

Figure 4: Heterogeneity in Wage Trends After Nonattainment Designation



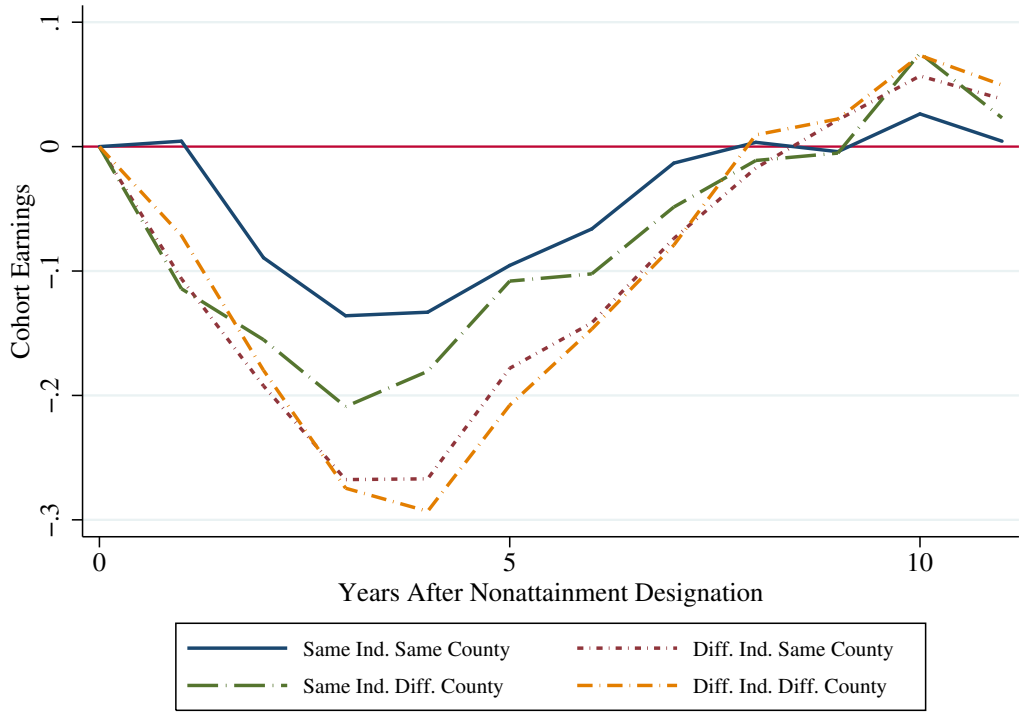
NOTE: Plotted are the coefficient estimates from a version of equation (10) for the "zero-earnings" and prime-age male subsamples of the data. Specifically, Panel (a) and (b) plot the event time indicators for two sets of cohorts in counties that switched into EPA's nonattainment designation in 1991. 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. Panels (c) and (d) plot the difference in the event time indicators from Panels (a) and (b) respectively, corresponding to the "triple-difference" regression coefficients in equation (10). See text for details. Source: LEHD.

Table 3: 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)	0.017	-0.122***	0.004	-0.106***	-0.114***	-0.071***
	(0.016)	(0.022)	(0.024)	(0.025)	(0.035)	(0.024)
Regulation (t+1)	-0.006	-0.233***	-0.089	-0.192***	-0.155***	-0.180***
	(0.019)	(0.018)	(0.062)	(0.014)	(0.049)	(0.020)
Regulation (t+2)	0.010	-0.341***	-0.136***	-0.268***	-0.209***	-0.275***
	(0.015)	(0.034)	(0.035)	(0.035)	(0.031)	(0.026)
Regulation (t+3)	-0.001	-0.334***	-0.133***	-0.267***	-0.181***	-0.293***
	(0.015)	(0.064)	(0.046)	(0.070)	(0.027)	(0.045)
Regulation (t+4)	-0.017	-0.243***	-0.096***	-0.178***	-0.108***	-0.208***
	(0.027)	(0.054)	(0.033)	(0.063)	(0.023)	(0.035)
Regulation (t+5)	0.016	-0.188***	-0.066*	-0.142***	-0.102***	-0.147***
	(0.014)	(0.034)	(0.034)	(0.040)	(0.024)	(0.021)
Regulation (t+6)	0.020*	-0.112***	-0.013	-0.074***	-0.049	-0.079***
	(0.011)	(0.011)	(0.020)	(0.012)	(0.030)	(0.014)
Regulation (t+7)	0.011	-0.039**	0.003	-0.017	-0.011	0.009
	(0.017)	(0.017)	(0.022)	(0.018)	(0.027)	(0.020)
Regulation (t+8)	0.008	-0.006	-0.004	0.022	-0.005	0.022
	(0.026)	(0.015)	(0.021)	(0.031)	(0.026)	(0.014)
Regulation (t+9)	0.009	0.050***	0.026	0.057**	0.075	0.073***
	(0.035)	(0.015)	(0.023)	(0.024)	(0.049)	(0.024)
Regulation (t+10)	0.018	0.032***	0.004	0.038***	0.023	0.049***
	(0.023)	(0.012)	(0.014)	(0.013)	(0.019)	(0.018)
6-Year PDV	0.018	-1.320***	-0.462***	-1.045***	-0.794***	-1.057***
	(0.052)	(0.137)	(0.143)	(0.156)	(0.142)	(0.090)
11-Year PDV	0.066	-1.386***	-0.451***	-1.035***	-0.776***	-1.011***
	(0.085)	(0.151)	(0.161)	(0.153)	(0.176)	(0.115)
N	6847	6847	6847	6847	6847	6847
Cohort FE	X	X	X	X	X	X
Potential Experience	X	X	X	X	X	X

NOTE: This table reports regression coefficients from equation (10). Each column corresponds to a different regression and a different sample, where the different samples are indicated in the table headings. Cluster robust standard errors are reported where errors are clustered at the county level. An observation is a cohort×year so that each cohort county has two observations per year. The regressions are weighted by the 1990 cohort size. The final two rows of the table report the discounted sum of the coefficients. Estimates are presented using a annual discount rate of 4 percent, and standard errors are calculated using the delta method. Earnings are reported in 1990($). See Section 7.2 for details. Source: LEHD.

Figure 5: Effect of Sector Level Regulation on Earnings: Decomposition Based on Sectoral Transitions



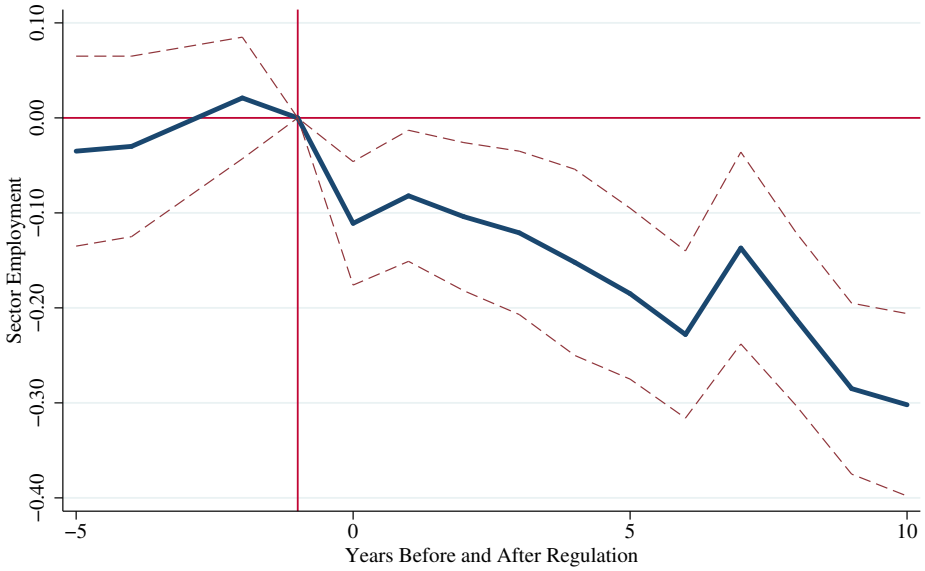
NOTE: Plotted are the coefficient estimates from a various versions of equation (10) fit for each sub-category of job transitions. The figure shows that worker transitions that result in a new job in the same industry experience a far lower earnings loss than those workers who switch industries. The lines correspond to the estimates reported in columns (3)-(6) of Table 3. See Section 7.2 for details. Source: LEHD.

Figure 6: Sector Level Employment Trends and Differences Before and After Regulation

(a) Sector Level Employment Trends Before and After Regulation

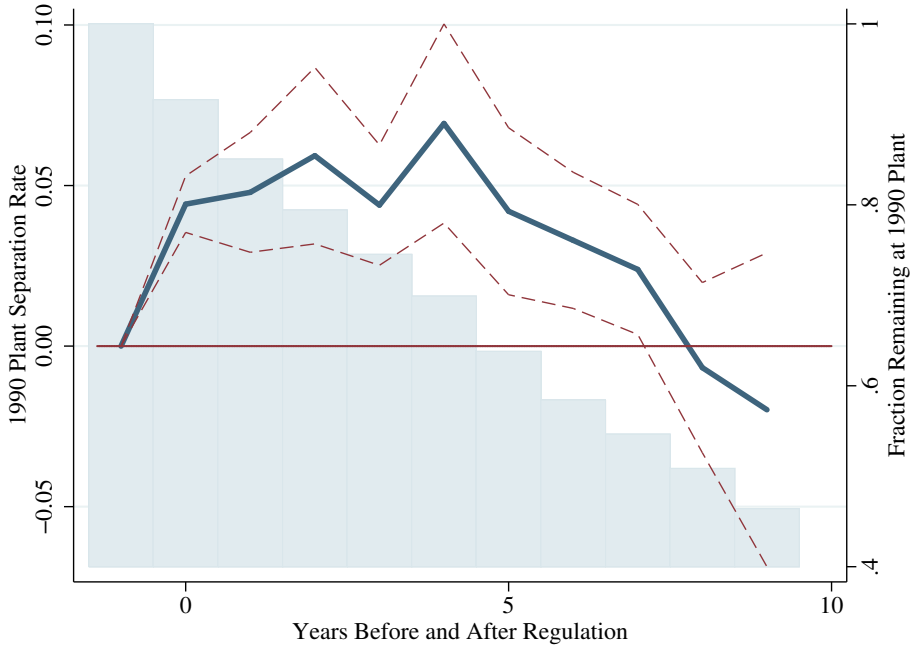


(b) Sector Level Employment Differences Before and After Regulation



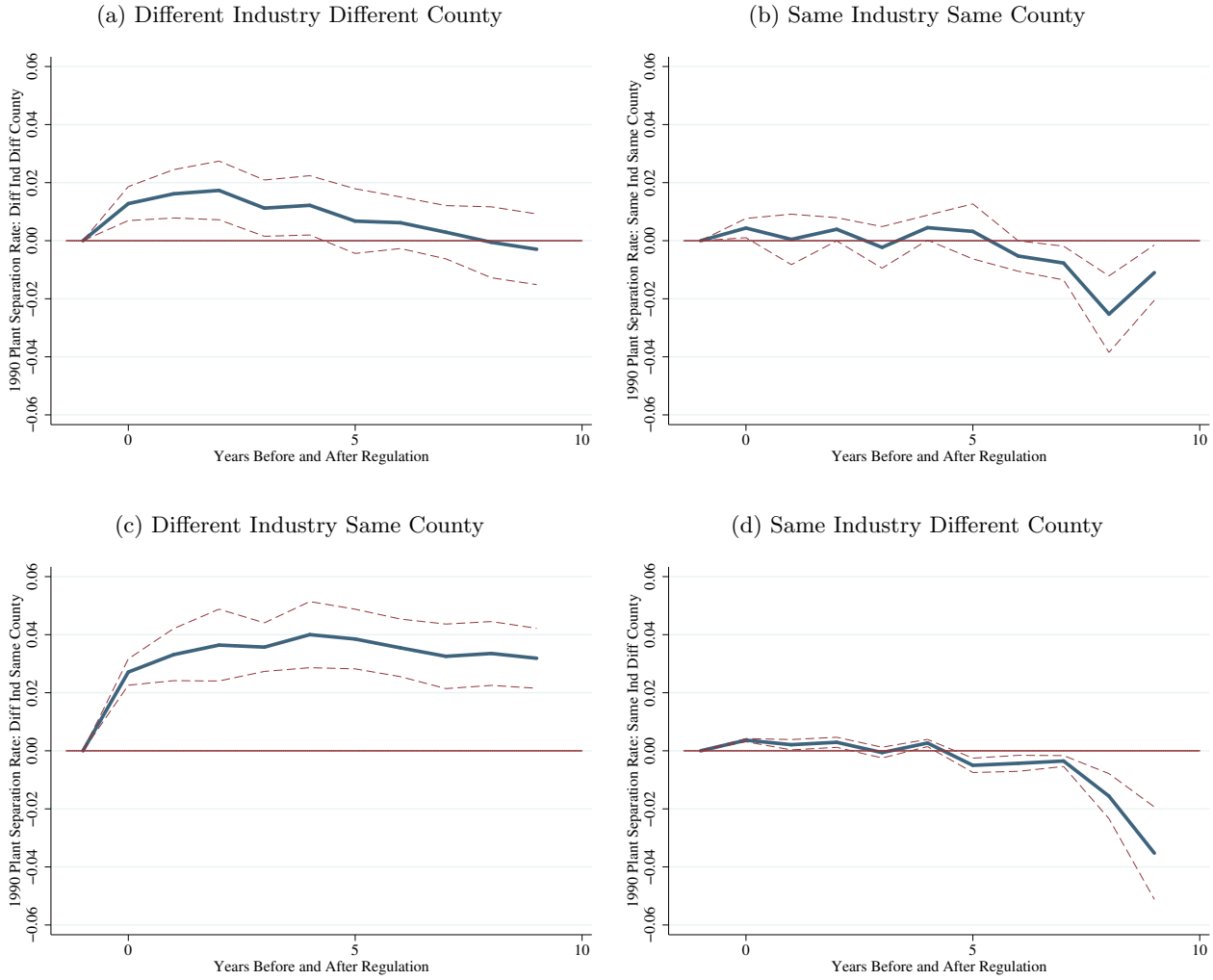
NOTE: Plotted are the coefficient estimates from a version of equation (11) showing that employment in the polluting sector of counties newly affected by EPA's nonattainment designation declines after the designation. Moreover, both the polluting and non-polluting sectors share similar trends in employment prior to nonattainment designation. Specifically, Panel (a) plots the event time indicators for the polluting and non-polluting sectors of counties that switched into EPA's nonattainment designation in 1991. The solid line represents the polluting sector, and the dashed line represents the non-polluting sector. The first year of the nonattainment designation corresponds to year 0 in the graph. Panel (b) plots the difference in the event time indicators from Panel (a), corresponding to the "triple-difference" regression coefficients in equation (11). The dashed lines represent 1 standard deviation above and below the estimated solid line. See text for details. Source: LBD.

Figure 7: Job Transition Rates from the Newly Regulated Sector



NOTE: Plotted are the coefficient estimates from a version of equation (12), 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 1 standard deviation above and below the estimated solid line. The bars represent the survivor probability for workers in polluting plants of newly designated nonattainment counties. Source: LEHD.

Figure 8: 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 (12), 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 1 standard deviation above and below the estimated solid line. Source: LEHD.

Table 4: 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)
	Panel A: County Demographics						
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
	Panel B: County Industrial Composition						
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: Derivations

A.1 Welfare and Incidence Derivation from Section 3.3

Section 3.3 presents the following result pertaining to the incidence of the regulatory shock:

$$\frac{dv^{np}}{d\theta^{np}} = \frac{dw_t^{np}}{d\theta^{np}} + \frac{d}{d\theta^{np}}\beta E[V^{np}(\Omega_{t+1}) \mid \Omega_t] \tag{13}$$

$$= \frac{dw_t^{np}}{d\theta^{np}} + \beta \sum_{s \in S} \pi_{t+1}^{np,s} \frac{dv_{t+1}^{s}}{d\theta^{np}}$$

To get from the first line to the second line, I follow Busso, Gregory, and Kline (2010). Consider a worker i who has a decision rule for each state. Define a set of indicator variables $\{D_{ist}\}$ equal to one if and only if

$$\max_{s'}\{v_i^{s'}(\Omega_t)\} = v_i^s(\Omega_t).$$

To simplify notation, let

$$v_i^s(\Omega_t) = v^s(\Omega_t) + \epsilon_{ist}$$

Then the choice probability of switching from sector j to sector s in period t can be written as

$$\pi_t^{j,s} = \Pr(D_{ist} = 1 \mid s_{t-1} = j,\ \{v^s(\Omega_t)\})$$

Denote the average utility for workers in sector j as $\bar{V}^j$:

$$\bar{V}^j = E_\epsilon[\max_s\{v_i^s(\Omega_t)\}]$$

Then

$$\begin{aligned}
\frac{d}{dv^s}\bar{V}^j &= E_\epsilon\left[\frac{d}{dv^s}\max_s\{v_i^s(\Omega_t)\}\right] \\
&= E_\epsilon\left[I\left[\max_s\{v_i^s(\Omega_t)\} = v_i^s(\Omega_t)\right]\right] \\
&= \Pr(D_{ist} = 1 \mid s_{t-1} = j,\ \{v^s(\Omega_t)\}) \\
&= \pi_t^{j,s}.
\end{aligned}$$

Using this result

$$\frac{d}{d\theta^{np}}\beta E[V^{np}(\Omega_{t+1}) \mid \Omega_t] = \beta\left(\frac{d}{dv_{t+1}^s}\bar{V}_{t+1}^{np}\right)\left(\frac{dv_{t+1}^s}{d\theta^{np}}\right) = \pi_{t+1}^{j,s}\left(\frac{dv_{t+1}^s}{d\theta^{np}}\right)$$

which can be used to get from the first to the second line in equation (13).

A.2 Cohort Wage Regressions and Incidence in a Dynamic Discrete Choice Model

Here I derive the relationship between the baseline cohort wage regression in equation (10) from Section 6.2 and the incidence derivation in equation (7). Recall that productivity in sector s is denoted θ^s. Moreover, recall that the reduced form productivity shifter for the polluting sector is the indicator variable $[N_c \times P_s \times 1(\tau_t = k)]$. Then, the coefficients, η_1^k, approximate the terms inside the brackets from equation (7) via the following relationship:

$$\begin{aligned}
\eta_1^k =& \frac{\Delta\tilde{w}_{t+k}^{np}}{\Delta[N_c \times P_s \times 1(\tau_t = k)]} \\
\approx& \left[\sum_{s \in S} \pi_{t+k}^{\tilde{np},s}\frac{dw_{t+k}^{np}}{d\theta^{np}} + \frac{d\pi_{t+k}^{\tilde{np},s}}{d\theta^{np}}w_{st+k}\right]
\end{aligned} \tag{14}$$

where the second term inside the brackets corresponds to the changes in switching behavior induced by the policy.[30] The envelope theorem tells us that even though the regulations might induce changes in behavior (i.e. $\frac{d\pi_{t+k}^{\tilde{np},s}}{d\theta^{np}}$), these behavioral responses cannot have a first-order effect on private welfare, and hence the second term drops out (i.e. $\frac{d\pi_{t+k}^{\tilde{np},s}}{d\theta^{np}} \approx 0$). Essentially, those workers on the margin who were induced to switch by the policy were indifferent between switching in the first place. Notice that the envelope condition only holds if workers are behaving optimally.

What this derivation makes explicit is that by examining the cohort wage, the estimates of the η_1^k's implicitly incorporate both the first order wage effect as well as any changes in the behavioral responses and do not rely upon the first order approximations implicit in the envelope theorem. By appealing to this cohort based approach, I do not have to invoke the envelope theorem to assume that the second term in equation (14) is approximately zero.

[30] The choice/switching probabilities are determined by differences in the wages and switching costs. Thus, the second term can be decomposed further as

$$\frac{d\pi_{t+k}^{\tilde{np},s}}{d\theta^{np}} w_{t+k}^{np} = \left[\left(\frac{d\pi_{t+k}^{\tilde{np},s}}{dv_{t+k}^{s}} \times \frac{dv_{t+k}^{s}}{d\theta^{np}} \right) - \left(\frac{d\pi_{t+k}^{\tilde{np},s}}{dv_{t+k}^{\tilde{np}}} \times \frac{dv_{t+k}^{\tilde{np}}}{d\theta^{np}} \right) \right] w_{t+k}^{np}$$

B Appendix: Data

B.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 and Vilhuber (2008) and Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer, and Woodcock (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.

Two important limitations of the LEHD data in the context of this project are the temporal coverage over time (detailed further in Table B.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.

I use the Census's Unit-to-Worker imputation file that imputes the worker establishment using a non-ignorable missing data model with multiple imputation.[31] 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 and Vilhuber (2008) for more information.

B.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.

B.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 B.4), I

[31] Note that I also have estimated models that restrict my sample to consist *only* of single-unit employers (not reported) and the results remain qualitatively similar.

Table B.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

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.

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

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

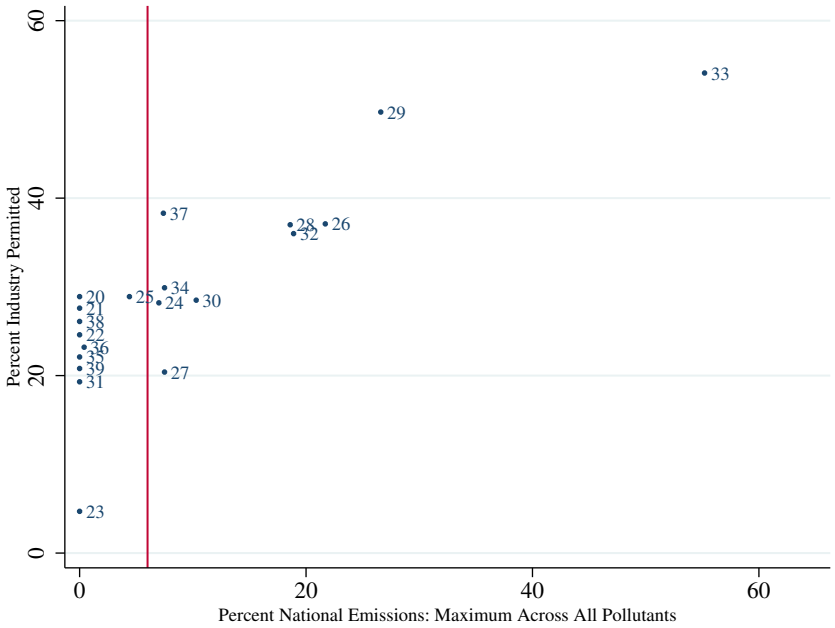
B.5 Assessing the Validity of the "6 Percent Rule"

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, Millimet, Fredriksson, and McHone (2003), Becker (2005), Kahn and Mansur (2010), Greenstone, List, and Syverson (2010)), it is worthwhile to assess whether or not this classification still makes sense more recent time periods.

Figure B.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.

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). Thus, relying on these classifications for more recent time periods may no longer make sense.

Figure B.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, Millimet, Fredriksson, and McHone 2003, Greenstone, List, and 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).