Paper for Presentation next week: "Does Management Matter? Evidence from India" by Bloom et al.

Intro Comments:

The goal of today's lecture is to get you started in the part of the class on identification. Because we don't run experiments in social science (usually), and because a lot of the time we want to know the effect of policies given the firms that choose to react to them, we are going to have to deal with the *Selection Problem*. This is the fact that we observe a selected sample since firms and individuals make choices of beeing in or out of a treatment. The class will start with a bit of notation, but the goal is to get ourselves to a point where we can talk about identification.

The Extrapolation Problem:

Suppose we want to know about $P[y|x]$, the conditional probability of y given x (data about the situation we are in). Say we want to know about the effect of head-start, an intense early schooling program, mainly targeted to children from poor families.

The data gives us observations about $P[y|x = x_0]$ and we want to extend this analysis to $P[y|x = x_1]$, such as another state, does the program work for hispanic children versus black children (think of language issues for instance), if we change it from 30 hours to 20 hours a week, does the program still work?

We need a type of "Invariance Assumption":

$$P[y|x = x_0] = P[y|x = x_1]$$

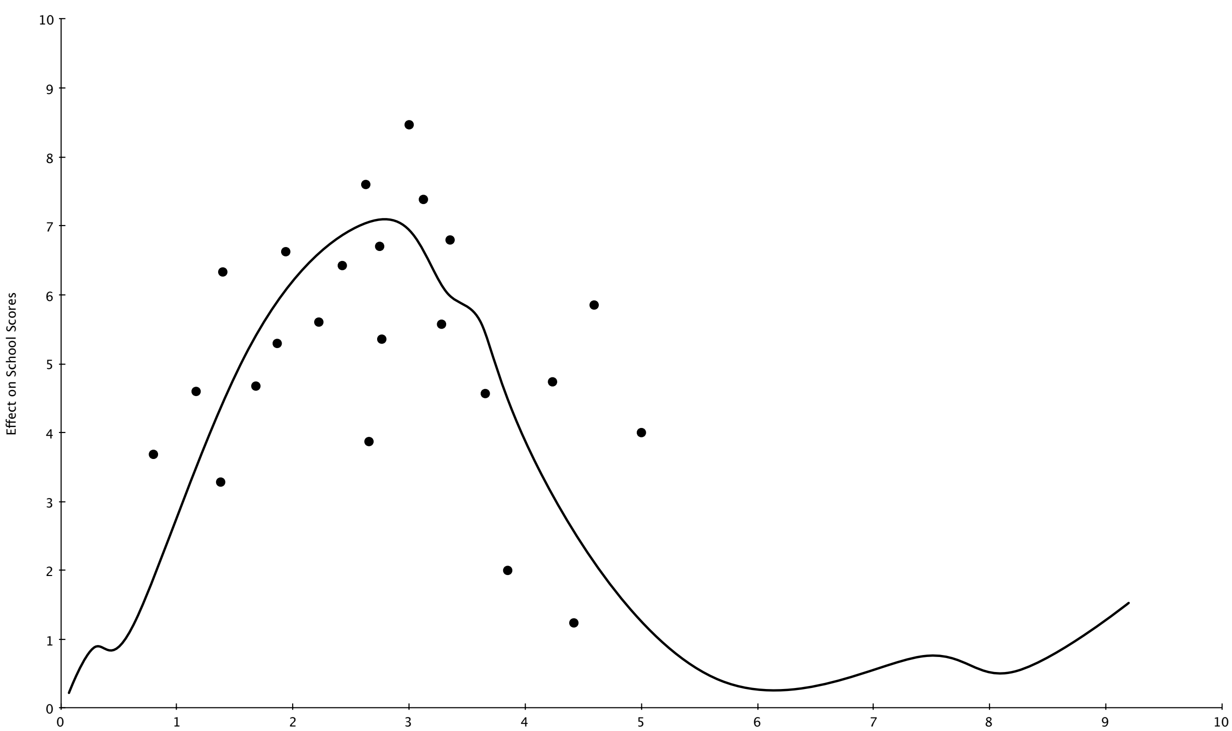
for values $x = x_0$ and $x = x_1$.

- Can we extrapolate without theory: School Treatment Problems (like class size). We can argue that nothing changes in different schools, i.e. there really isn't any x that varies across schools.

- What theory says to allow us to extrapolate. So for a moral hazard problem, we know that technology that changes monitoring has an effect, but stuff that varies the worker's outside option won't effect the optimal incentives (just the wage I get paid in the low and high state).

- This is the main problem with Experimental Work in General: Details of the experiment matter so much that it is often hard to extrapolate an experiment to other situations.

- Finally, do we get enough variation of x to identify the effect we are looking for.



The Selection Problem:

The selection problem is the main issue in social science. Let's start with an example where there is some probability than people respond to a survey $z = 1$ or not $z = 0$. Let's think about why this might be a problem:

- When the U.S. Census counts the American population, they typically send out questionnaires using the mail. There's been a problem in the response of people who don't have mailing addresses, i.e. the homeless. If I only count people who have mailing addresses, I am liable to overestimate incomes in the population, since I'm not counting the homeless in this average. What Census does to correct for this problem, is have one day where they send out teams to survey the number of homeless people in the streets.

- Many surveys are done using retrospective analysis. For example suppose I want to look at the effect of firm size on the firm growth rates. One way to do this is to have a questionnaire which asks firms what their growth rate has been in the last two years and what their size was at the start of the period, and at the end of the period. However this means that I'm leaving out the firms that don't respond, in this case firms that failed in the last few years. You can't survey firms that went bankrupt. Now this is a problem because small firms may have the same growth rate on average as large firms. The problem is that small firms are also more likely to go bankrupt than large firms, so the set of firms that responds to the survey and can make retrospective evaluations is more selected for small firms than it is for large firms. This is one of the reasons why you might worry that you find evidence of a lot of data sets at small firms grow more quickly, when what you really finding as evidence that small firms grow more quickly conditional on firm survival.

- In social science, there are many large longitudinal surveys that are conducted i.e. surveys that try to follow particular firms or individuals over several years. One of the difficulties with these surveys is that you need to find and interview the same and individual repetitively. It is often true that it is hard to find these individuals or have them answer the survey year after year and this also creates a selection problem. In particular individuals with more stable jobs and places of residence are more likely to answer the survey, which means that there's going to be a strong selection for older and wealthier individuals as the survey gets sent out year after year.

The reason that you have to think hard about these selection problems, is that there is a lot of fairly well-known work which suffers from this type of issue, and in many cases it mechanically generates a result that might not be true when we don't condition on selection.

Censoring:

$$P(y|x) = P(y|x, z = 1)P(z = 1|x) + \underbrace{P(y|x, z = 0)}_{\text{unobserved}} P(z = 0|x)$$

we can see everything except for $P(y|x, z = 0)$, which is the issue here. One way took compute $P(y|x)$ given the available data, is to assume ignorable selection or nonresponse, i.e.:

$$P(y|x, z = 0) = P(y|x, z = 1)$$

This implies that

$$\begin{aligned} P(y|x) &= P(y|x, z = 1)P(z = 1|x) + P(y|x, z = 1)P(z = 0|x) \\ &= P(y|x, z = 1) \end{aligned}$$

Bounds Approach:

Can we still learn something given we never see $P(y|x, z = 0)$? Suppose we can bound $y \in [\underline{y}, \bar{y}]$ then we can say something about $P(y|x)$. So the worse case scenario is that $P(y|x, z = 0) = \underline{y}$ and the best case scenario is that $P(y|x, z = 0) = \bar{y}$. Thus we can create bounds on $P(y|x)$:

$$P(y|x) \geq P(y|x, z = 1)P(z = 1|x) + \underline{y}P(z = 0|x)$$

and

$$P(y|x) \leq P(y|x, z = 1)P(z = 1|x) + \bar{y}P(z = 0|x)$$

So for instance in binary response models, $y \in [0, 1]$ which make simple bounds to compute. This idea of bounds to get around selection problems is starting to percolate more and more into thinking in economics (and should infect other disciplines soon).

As an example, suppose that there are two groups in the populations, the housed ($h = 1$) and the homeless ($h = 0$). We know from a previous Census Survey that 5% of the population is homeless. We send out a mailed survey on the fraction of households which have a cell phone ($P(c)$) for which we obtain that 60% of responding households have a cell phone. However, no homeless households respond to this survey (since it is mailed to a physical address). What are the bounds on the number of households with cell phones?

The Heckman Selection Problem:

The model (data generating process) is the following:

$$w^* = X\beta + \epsilon$$

and people see their wage and decide wether to work or not:

$$z = 1(w^* > R(y))$$

where R is a reservation wage.

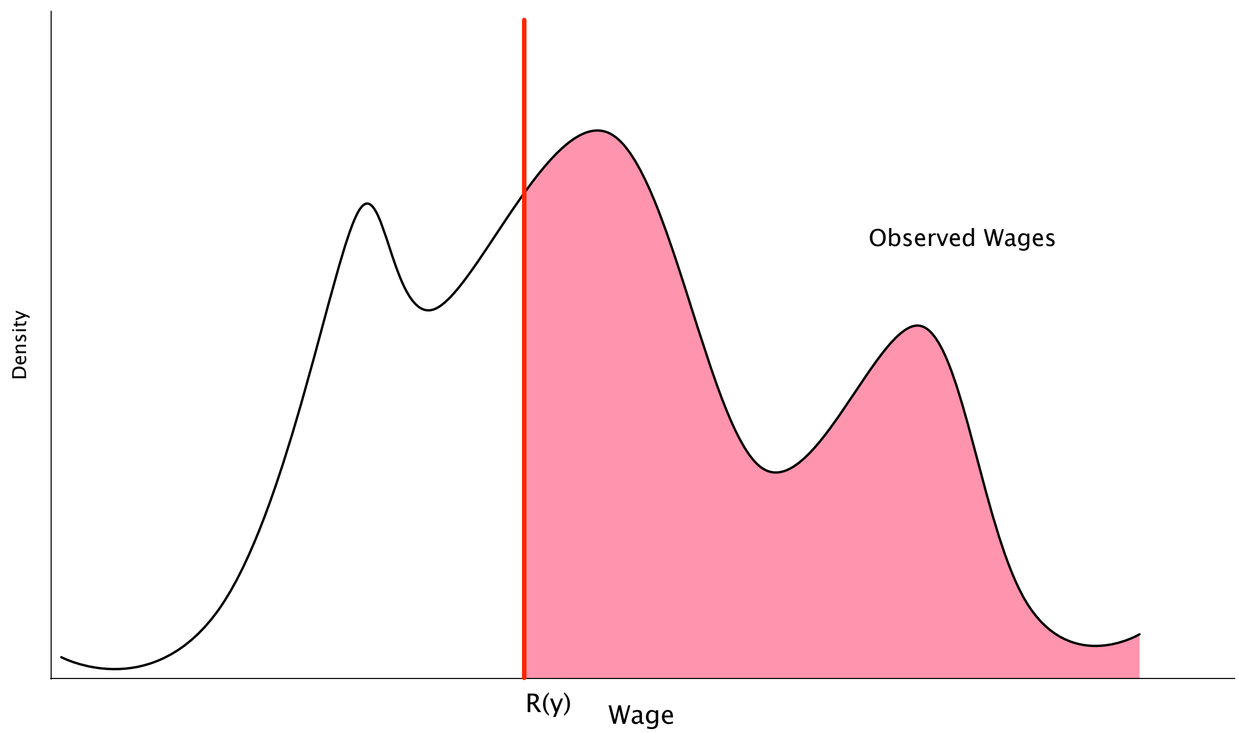
Okay, what we see as the econometrician is:

$$w^* = \begin{cases} X\beta + \epsilon & \text{if } y = 1 \\ 0 & \text{if } y = 0 \end{cases}$$

So the $P[w|x]$:

$$\begin{aligned} P[w|x] &= P[w|x, z = 1]P[z = 1|x] + P[w|x, z = 0]P[z = 0|x] \\ &= (X\beta + P[\epsilon|\epsilon > R(y) - X\beta])\, P[z = 1|x] \\ &+ (X\beta + P[\epsilon|\epsilon < R(y) - X\beta])\, P[z = 0|x] \end{aligned}$$

and from here it is clear why $P[w|x] \neq P[w|x, z = 1]$. There are many ways to get around this problem, but this is the selection problem.



Note that to really get around this problem, you need to use some variable y that affects the reservation wage, but not the wages in the market (like having children, savings, distance to work or something like this). The fact that y changes $R(y)$ but not $w(x)$ is called an exclusion restriction.

Simultaneity is Selection:

The simultaneity problem is often encountered. Let me do the version with production functions. Consider the following production function:

$$Y_{it} = \Omega_{it} K_{it}^{\alpha_k} L_{it}^{\alpha_l}$$

where K is capital and L is labor input and Ω is productivity. Taking logs we get.

$$\begin{aligned} y(\text{log value added}) = \alpha_0 + \alpha_l l(\text{log salaries}) \\ + \alpha_k k(\text{log assets}) + \epsilon \end{aligned}$$

Suppose that a firm can't change it's capital stock in the short run, however it can change it's labor l. There it will choose higher labor when there is a high ϵ, i.e. $E[l\epsilon] > 0$. Indeed, the input demand function for labor here is given by maximizing the profit function $\pi = \Omega_{it} K_{it}^{\alpha_k} L_{it}^{\alpha_l} - wL_{it}$ with the first-order condition:

$$\frac{\partial \pi}{\partial L} = \alpha_l \Omega_{it} K_{it}^{\alpha_k} L_{it}^{\alpha_l - 1} - w = 0$$

Yielding:

$$L_{it} = \left(\frac{\Omega_{it} K^{\alpha_k}}{w} \right)^{\frac{1}{1-\alpha_k}}$$

or in logs:

$$l_{it} = \log(\frac{1}{1-\alpha_l})\,(\epsilon + \alpha_k k_{it} - \log(w))$$

This is a selection problem, since it is saying that:

$$P[y|k,l] = P[y|k, l = l_0]P[l = l_0|k] + P[y|k, l \neq l_0]P[l \neq l_0|k]$$

Again, the issue is that we don't see l unconditionally, we see l conditionally on a certain realization of ϵ.

Treatment Effects:

Some examples of treatment effects are:

- The effect of a job training program for unemployed workers on their future earnings.
- The effect of the merger on prices within a market.
- The effect of adopting ERP (enterprise resource planning) on firm profits.
- And finally the most classic one, the effect of a drug on patients health outcomes.

There is a censoring problem for treatment effects $d_i \in \{0, 1\}$ since we only observe one of:

$$y_i = y_{i0}(x)(1 - d_i) + y_{i1}(x)d_i$$

where y_{i0} is the individual's outcome if she doesn't get the treatments, and y_{i1} is individual i's outcome if she does get the treatment. Moreover, d_i if the assignments to the treatment for individual i.

There are several reasons why we might believe that the treatments are not assigned randomly:

- Administrators of Randomness:

  In the head start program, it might be the case that administrators are more likely to admit children with highly motivated parents. One of the natural reasons for this if parents know that there's a program that may give their children a better educational outcomes, the most motivated of those parents can put strong pressure to be selected into that group. So often the administrators don't really allocate the treatment randomly, even if often in papers researcher will say that they do.

- Expectations

  Another issue is about the expectations of the participants about the effectiveness of the program. For instance I might get different selection into the program if people think that it's effective versus if people don't think that it's effective. Thus the very results of the study might affect the selection into the program and hence the validity of the study itself.

- Self-Selection: Retraining programs for workers that have been laid off.

  We might worry that workers who have higher ability, or more motivation, might select into the retraining program, while the workers with lower ability, or less motivated workers might opt out of this treatment. If we compare the effect of the treatment program on workers wages we clearly have to account for the fact that be expectation of these wages are different for the two groups even net of the treatment program.

- Site Selection

  it is often the case fats you can only put a treatment program together in certain areas. For instance if I wanted to look at the effect of investments into schooling in India, I might want to assign the sites randomly, in order to understand the effect of investment across the entire country. However you might be worried that Maoist guerrillas are going to destroy the site in Bihar, and kidnap your workers. For this reason you end up having only sites in safe states, such as Kerala. Again this will create a selection effect of the treatment.

We will need to make some assumptions to be able to estimate the treatment effect $y_1 - y_0$. As well, there are different treatment effects that we are usually interested in. Let's go through some really stupid terminology (I think that its a bit of jargon for jargon's sake).

Average Treatment Effect (ATE):

$$ATE = E(y_1 - y_0)$$

The average treatment effect or ATE, is the expected effect of the treatment on the entire population. Notice that there are many cases were not just interested in the mean outcome of the treatment, but you might also worry about distributional effects of the treatment as well. For example if I was looking at the effect of head start on schooling outcomes, I might value outcomes very low in the schooling distribution much higher than outcomes up in the top tail of the schooling distribution, since we might believe that is children with very bad schooling outcomes that need to be helped the most by this program

Average Treatment Effect on the Treated (ATET):

$$ATET = E(y_1 - y_0|d = 1)$$

At first the average treatment effect on the treated or ATET might seem a bit bizarre, since why do we care about the effect only on one subpopulation? Isn't this what the selection problem was all about, fixing the fact that we don't observe the entire population, but only the section of the population that responds? But if I'm looking at the effects of a labor training program, and particular if the benefits of this program outweigh the costs of the program, what I want to know is not what's the effect of this labor training program if I assigned it arbitrarily to somebody in the population, for instance somebody who already has a job, or somebody who is permanently unemployed or on disability but I want to know what's the effect of this program on the specific population that would actually use it. This is the case where I care about the average treatment effect on the treated.

Let me give you an even more specific example that comes from some my work on the effects of mergers on prices. For the merger to be consummated, we need that the firms choose to merge, and then we need the regulator (in the United States either the FTC or the Department of Justice) to approve the merger. From the perspective of the regulator looking to evaluate the effect of a proposed merger, I don't want to look at the effect of a merger for two randomly chosen firms, but I want to look at the effects of the merger given that firms are proposing it. The effect given that firms are proposing it might lead to substantially higher price increases, if we believe that firms merge in order to raise prices, or the effect might go the other way if I expected firms only proposed mergers that they know won't have strong anticompetitive effects, and thus might be approved by the regulator. Either way, looking at mergers that were proposed or mergers that actually happened might be the exact treatment effect we are looking for.

Average Treatment Effects, Heterogeneous Response:

Sometimes we think that treatment effects might very be there by subpopulation or by some other characteristic x. This is known as the problem of heterogeneity in the response to a treatment. For instance, a famous case comes up in the treatment of heart disease, were some drugs may have positive effects for white patients, but high-risks for black patients. Alternatively, there is now evidence that the head start program is much more effective at a very young age, i.e. 1 to 2 years old rather than at a slightly older age, i.e. 3 to 4 years old. Let's define the average treatment conditional on characteristics x as:

$$ATE(x) = E(y_1 - y_0|x)$$

Now for some assumptions about what we need to identify average treatment effects. Clearly, the assumption of ignorable selection that we've seen above will do the job. But let's try to see if there's anything else that is slightly weaker that also works.

1. Random assignment of treatments.

$$d_i \perp\!\!\!\perp y_{i0}, y_{i1}, x$$

This is essentially assuming that treatment is completely random, and it is a very strong assumption. Let's compute ATE:

$$E(y|d_i = 1, x) = E(y_1|d_i = 1, x) = E(y_1, x)$$

and also

$$E(y|d_i = 0, x) = E(y_0|d_i = 0, x) = E(y_0, x)$$

Combining these two results:

$$\begin{aligned} E(y|d_i = 1, x) - E(y|d_i = 0, x) &= E(y_1|x) - E(y_0|x) \\ &= E(y_1 - y_0|x) \end{aligned}$$

So we can get an estimate of the treatment effect by taking the difference in means between the treated group in the untreated group.

2. Mean or Median Independence

A weaker assumption that's frequently used in the literature, is mean independence. I also really like the assumption of median independence, since an assumption of mean independents really imply something about the shape of the entire function, while median independent is just saying something about a particular quantile of the distribution.

Mean Independence:

$$E[y_1|d_i,x] = E[y_1|x]$$

$$E[y_0|d_i,x] = E[y_0|x]$$

Clearly given the assumption mean independence, we can estimate the average treatment effect as:

$$E[y_1|d_i,x]-E[y_0|d_i,x] = E[y_1|x]-E[y_0|x] = E[y_1-y_0|x]$$

Median Independence:

$$Med[y_1|d_i,x] = Med[y_1|x]$$

$$Med[y_0|d_i,x] = Med[y_0|x]$$

Clearly given the assumption median independence, we can estimate the average treatment effect as:

$$\begin{aligned} Med[y_1|d_i,x] - Med[y_0|d_i,x] &= Med[y_1|x] - Med[y_0|x] \\ &= Med[y_1 - y_0|x] \end{aligned}$$

The reason median independent is so much weaker, is that suppose that for instance incomes in the top 30 percentile are top coded, and the 20% of people who don't work, for these people we never see their incomes either.

It is Going to be very hard to estimate an average treatment effect given these data limitations. But to estimate the median treatment effect is actually quite straightforward, since all I care about is being able to observe what happens near the median in the data.

One last cool fact about medians, and then we can move on to the rest of the lecture. The media of a nonlinear function is the function evaluated at the median.

$$Med[\phi(x)] = \phi(Med[x])$$

But this isn't true for the expectation:

$$E[\phi(x)] \neq \phi(E[x])$$

Since that's what Jensen's inequality is all about.

Does Indiscriminate Violence Cause Attacks?

This is a political science paper on uses of violence. It's interesting since it spends a lot of time talking about the selection problem.

Essentially the theory is does violence beget violence (spiral of recriminations theory). It is nicely illustrated by a quote:

> The elder of the village of Liakhovo, together with some villagers and German soldiers, robbed a partisan base. The next day the partisan detachment demanded that Liakhovo's peasants return all that had been taken. The elder promised, but the next day tried to hide and was caught on the road and killed. The German HQ sent soldiers to the village. . . The partisan detachment destroyed the German convoy with seven men. After this, German soldiers razed the settlement to the ground with tanks (Hill, 2005, 52)

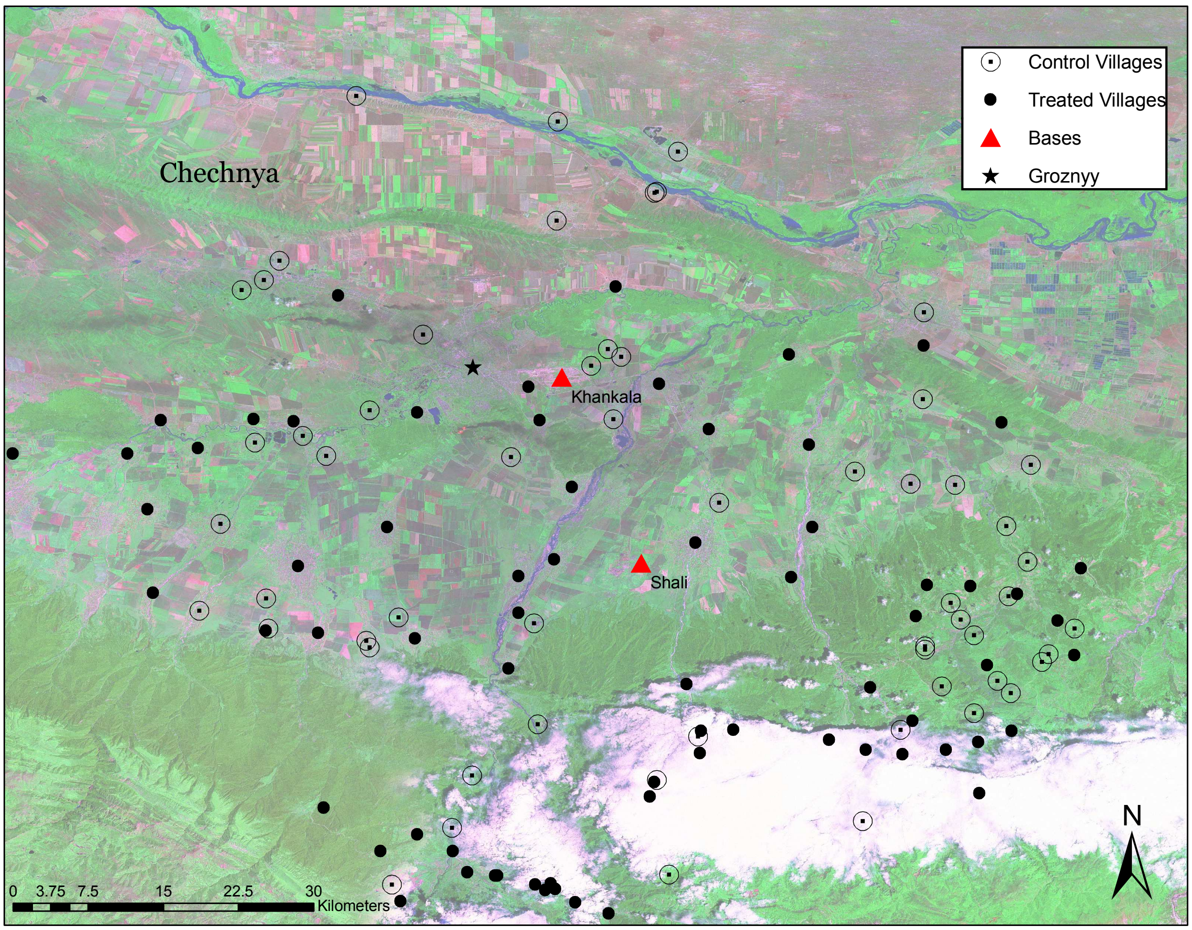
The issue is the role of selection effects in this story. Not all villages have insurgent attacks nor is there use of state violence. Let's look at the case in the paper which is about shelling villages in Chechnya (state violence) and subsequent insurgent attacks. The regression in the paper would look like:

$$\underbrace{Y_{it}}_{\text{Insurgent Attacks}} = \alpha \underbrace{T_{it}}_{\text{Shelling}} + X_i\beta + \epsilon_i \quad (1)$$

But the issue is are villages with more shelling also more likely to have a higher ϵ (say local organization of militias is better, or more violent ethnic groups, or more rugged terrain).

If you don't believe this, then you think that **Russian troops shell villages randomly**.

Figure 1: The Natural Experiment



NOTE. 147 populated settlements (73 treated, 74 control).

Data:

- Collects data on insurgent attacks and shelling from newspaper sources, NGOs and official press releases. (you should be thinking about the potential for missing data here).

- Demographic and Ethnic data on villages.

Table 3: Treatment and Insurgent Violence

	Treatment only *All Villages* 1	Treatment with covariates *All Villages* 2	Treatment only *Groznyy Dropped* 3	Treatment with covariates *Groznyy Dropped* 4
Treatment	-0.516**	-0.506***	-0.444**	-0.579***
	(0.214)	(0.168)	(0.188)	(0.185)
Constant	-0.101	-0.645	-0.062	-1.112
	(0.093)	(0.785)	(0.835)	(0.893)
F	(1, 121) = 5.80	(11, 121) = 3.45	(1, 118) = 5.58	(11, 118) = 7.90
Prob > F	0.02	0.00	0.02	0.00
N (Clusters)	318 (123)	318 (123)	298 (119)	298 (119)

Note: Robust cluster-adjusted (on village) standard errors in parentheses. *Significant at 10% **Significant at 5% ***Significant at 1%

Table 1: Village Level "As-If" Randomization Tests and Post-Matching Statistics

Covariates	Mean Treated	Mean Control	Mean Difference	Std. Bias	Rank Sum	K-S Test
"As If" Randomization						
POPULATION	7.364	7.020	0.344	0.209	0.248	0.133
POVERTY	2.425	2.284	0.141	0.245	0.163	0.802
TARIQA	0.027	0.068	-0.041	-0.244	0.255	-
ELEVATION	5.933	5.756	0.177	0.225	0.202	0.169
ISOLATION	4.424	3.959	0.465	0.171	0.641	0.990
NEIGHBOR	0.742	0.899	-0.157	-0.213	0.321	0.542
GARRISON	0.178	0.122	0.056	0.145	0.339	-
REBEL	0.548	0.446	0.102	0.204	0.218	-
Post-Matching						
POPULATION	7.830	7.759	0.071	0.046	0.951	0.516
POVERTY	2.321	2.239	-0.082	-0.137	0.300	0.983
TARIQA	0.050	0.057	-0.007	-0.030	0.803	-
ELEVATION	5.834	5.766	0.068	0.095	0.650	0.219
ISOLATION	3.767	3.836	-0.069	-0.028	0.655	0.516
NEIGHBOR	0.896	0.882	0.014	0.021	0.839	0.516
GARRISON	0.258	0.283	-0.025	-0.057	0.614	-
REBEL	0.585	0.522	0.063	0.128	0.260	-
ATTACKS	2.113	2.151	-0.038	-0.001	0.871	0.713
SWEEPS	0.478	0.447	0.031	0.031	0.987	1.000

Mariel Boatlift Paper:

The Mariel Boatlift paper by Card is a very famous paper on natural experiments. The general question of this paper is a long-standing question in labor economics, does letting in more immigrants create pressure on domestic wages i.e. does the country have an incentive to let immigrants in, or do these immigrants and up lowering the wages for people inside the country. Depending on what you think about the answer to this question, immigration is either costly, or perhaps beneficial on the receiving country.

Some comments:

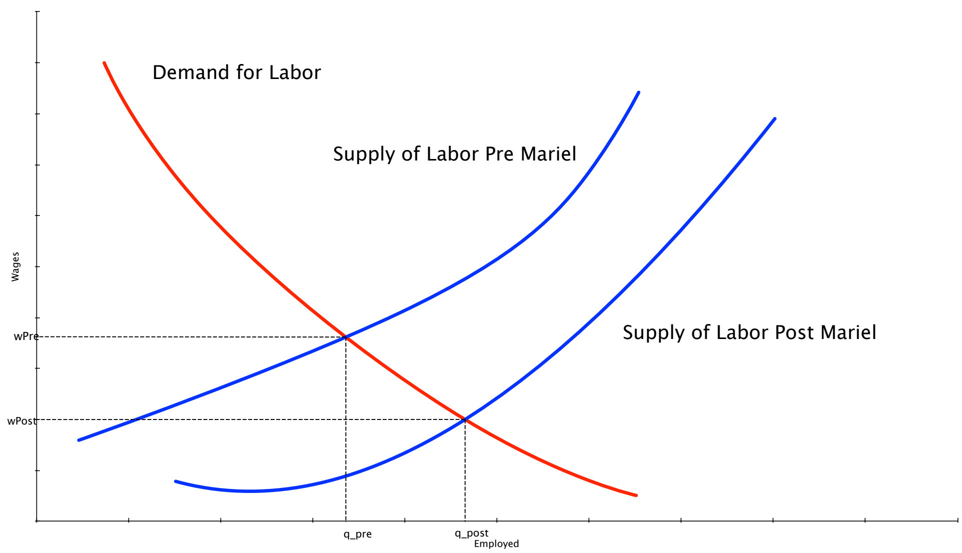
- The Mariel Boatlift refers to the arrival of $60,000$ Cubans in 1980, mainly by boat, arriving in South Florida.

- This a very large influx given that most of the Mariel Cubans settled in one city: Miami.

- However, Miami was one of the nation's most multiethnic cities at the time, with over 30% of it's population born abroad, and much of that population was hispanic, and much of those hispanics were cubans. LA in comparison only had 22% of it's population born abroad.

- The Mariel Cubans were a difficult immigrant group, since many of them were less educated that the Cubans who had arrived before, and there is some suspicion that Fidel Castro used the occasion to send Cuban inmates into the U.S..

- Theory would tell us that we need to think about: Supply and Demand for Labor.



- Who would be the most affected group: other unskilled workers, cubans, blacks.

- What are the distributional effects of the Mariel Boatlift? Which wages get affected and by how much?

Table 2. Characteristics of Mariel Immigrants and Other Cubans: Tabulations from March 1985 CPS.

Characteristic	*Mariel Immigrants*	*All other Cubans*
Educational Attainment (Percent of Population in Each Category):		
No High School	56.5	25.4
Some High School	9.1	13.3
Completed High School	9.5	33.4
Some College	6.8	12.0
Completed College	18.1	15.8
Percent Male	55.6	50.7
Percent Under 30 in 1980	38.7	29.6
Mean Age in 1980 (Years)	34.9	38.0
Percent in Miami in 1985	53.9	52.4
Percent Worked in 1984	60.6	73.4
Mean Log Hourly Earnings	1.37	1.71
Occupation Distribution (Percent Employed in Each Category):		
Professional/Managers	19.3	21.0
Technical	0.0	1.5
Sales	4.5	11.2
Clerical	2.5	13.5
Craftsmen	9.5	19.9
Operatives	19.1	13.8
Transportation Ops.	3.8	4.3
Laborers	10.8	3.3
Farm Workers	0.0	1.8
Less-Skilled Service	26.0	7.4
More-Skilled Service	4.6	2.3
Sample Size	50	528
Weighted Count	42,300	476,900

Note: The sample consists of all Cubans in the March 1985 Current Population Survey age 21–66 (i.e., age 16–61 in 1980). Mariel immigrants are identified as those Cubans who stated that they lived outside the United States 5 years previously.

Table 3. Logarithms of Real Hourly Earnings of Workers Age 16–61 in Miami and Four Comparison Cities, 1979–85.

Group	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Miami:							
Whites	1.85	1.83	1.85	1.82	1.82	1.82	1.82
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)	(.05)
Blacks	1.59	1.55	1.61	1.48	1.48	1.57	1.60
	(.03)	(.02)	(.03)	(.03)	(.03)	(.03)	(.04)
Cubans	1.58	1.54	1.51	1.49	1.49	1.53	1.49
	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)	(.04)
Hispanics	1.52	1.54	1.54	1.53	1.48	1.59	1.54
	(.04)	(.04)	(.05)	(.05)	(.04)	(.04)	(.06)
Comparison Cities:							
Whites	1.93	1.90	1.91	1.91	1.90	1.91	1.92
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Blacks	1.74	1.70	1.72	1.71	1.69	1.67	1.65
	(.01)	(.02)	(.02)	(.01)	(.02)	(.02)	(.03)
Hispanics	1.65	1.63	1.61	1.61	1.58	1.60	1.58
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)

Note: Entries represent means of log hourly earnings (deflated by the Consumer Price Index—1980 = 100) for workers age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. See note to Table 1 for definitions of groups.

Source: Based on samples of employed workers in the outgoing rotation groups of the Current Population Survey in 1979–85. Due to a change in SMSA coding procedures in 1985, the 1985 sample is based on individuals in outgoing rotation groups for January–June of 1985 only.

Table 4. Unemployment Rates of Individuals Age 16–61 in Miami and Four Comparison Cities, 1979–85. (Standard Errors in Parentheses)

Group	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Miami:							
Whites	5.1	2.5	3.9	5.2	6.7	3.6	4.9
	(1.1)	(0.8)	(0.9)	(1.1)	(1.1)	(0.9)	(1.4)
Blacks	8.3	5.6	9.6	16.0	18.4	14.2	7.8
	(1.7)	(1.3)	(1.8)	(2.3)	(2.5)	(2.3)	(2.3)
Cubans	5.3	7.2	10.1	10.8	13.1	7.7	5.5
	(1.2)	(1.3)	(1.5)	(1.5)	(1.6)	(1.4)	(1.7)
Hispanics	6.5	7.7	11.8	9.1	7.5	12.1	3.7
	(2.3)	(2.2)	(3.0)	(2.5)	(2.1)	(2.4)	(1.9)
Comparison Cities:							
Whites	4.4	4.4	4.3	6.8	6.9	5.4	4.9
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)
Blacks	10.3	12.6	12.6	12.7	18.4	12.1	13.3
	(0.8)	(0.9)	(0.9)	(0.9)	(1.1)	(0.9)	(1.3)
Hispanics	6.3	8.7	8.3	12.1	11.8	9.8	9.3
	(0.6)	(0.6)	(0.6)	(0.7)	(0.7)	(0.6)	(0.8)

Note: Entries represent means of unemployment indicator variable for individuals age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. Samples are based on individuals in the labor force. See notes to Table 3 for definitions of groups and data sources.

Table 5. Means of Log Wages of Non-Cubans in Miami by Quartile of Predicted Wages, 1979–85. (Standard Errors in Parentheses)

	Mean of Log Wage by Quartile of Predicted Wage				
Year	*1st Quart.*	*2nd Quart.*	*3rd Quart.*	*4th Quart.*	*Difference of Means: 4th − 1st*
1979	1.31	1.61	1.71	2.15	.84
	(.03)	(.03)	(.03)	(.04)	(.05)
1980	1.31	1.52	1.74	2.09	.77
	(.03)	(.03)	(.03)	(.04)	(.05)
1981	1.40	1.57	1.79	2.06	.66
	(.03)	(.03)	(.03)	(.04)	(.05)
1982	1.24	1.57	1.77	2.04	.80
	(.03)	(.03)	(.03)	(.04)	(.05)
1983	1.27	1.53	1.76	2.11	.84
	(.03)	(.04)	(.03)	(.05)	(.06)
1984	1.33	1.59	1.80	2.12	.79
	(.03)	(.04)	(.04)	(.04)	(.05)
1985	1.27	1.57	1.81	2.14	.87
	(.04)	(.04)	(.04)	(.05)	(.06)

Note: Predicted wage is based on a linear prediction equation for the log wage fitted to individuals in four comparison cities; see text. The sample consists of non-Cubans (male and female, white, black, and Hispanic) between the ages of 16 and 61 with valid wage data in the earnings supplement of the Current Population Survey. Wages are deflated by the Consumer Price Index (1980 = 100).

Table 6. Comparison of Wages, Unemployment Rates, and Employment Rates for Blacks in Miami and Comparison Cities.
(Standard Errors in Parentheses)

	All Backs				*Low-Education Blacks*			
	Difference in Log Wages, Miami − Comparison		*Difference in Emp./Unemp., Miami − Comparison*		*Difference in Log Wages, Miami − Comparison*		*Difference in Emp./Unemp., Miami − Comparison*	
Year	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*
1979	−.15	−.12	.00	−2.0	−.13	−.15	.03	−.8
	(.03)	(.03)	(.03)	(1.9)	(.05)	(.05)	(.04)	(3.8)
1980	−.16	−.12	.05	−7.1	−.07	−.07	.03	−8.2
	(.03)	(.03)	(.03)	(1.6)	(.05)	(.05)	(.04)	(3.5)
1981	−.11	−.10	.02	−3.0	−.05	−.11	.04	−7.7
	(.03)	(.03)	(.03)	(2.0)	(.05)	(.05)	(.04)	(4.2)
1982	−.24	−.20	−.06	3.3	−.17	−.20	−.04	.6
	(.03)	(.03)	(.03)	(2.4)	(.05)	(.05)	(.04)	(4.7)
1983	−.21	−.15	−.02	.1	−.13	−.11	.04	−3.3
	(.03)	(.03)	(.03)	(2.7)	(.06)	(.05)	(.04)	(4.7)
1984	−.10	−.05	−.04	2.1	−.04	−.03	.05	.1
	(.03)	(.03)	(.03)	(2.4)	(.06)	(.05)	(.04)	(4.7)
1985	−.05	−.01	−.06	−5.5	.18	.09	.00	−4.7
	(.04)	(.04)	(.04)	(2.6)	(.07)	(.07)	(.06)	(5.6)

Notes: Low-education blacks are those with less than 12 years of completed education. Adjusted differences in log wages between blacks in Miami and comparison cities are obtained from a linear regression model that includes education, potential experience, and other control variables; see text. Wages are deflated by the Consumer Price Index (1980 = 100). "Emp.-Pop. Rate" refers to the employment:population ratio. "Unemp. Rate" refers to the unemployment rate among those in the labor force.

Results: What does Card find:

- No effect on wages.
- No effect on unemployment rates.
- No effect of relatively less skilled workers (blacks and hispanics).
- No effect on the distribution of wages.
- No effect on unemployment rates.
- Note: The control cities really matter! There is a recession in the early 80's which is doing far more work than the Boatlift in Miami to move around employment rates and wages.
- Perhaps natives responded to the Mariel Boatlift by moving to other cities than Miami (response to immigration is in internal migration choices).
- Perhaps Miami and Cubans in Miami in particular were particularly well suited for immigration?
- Effect of Mariel on the social welfare system, criminality, schools and so on is not discussed in detail.

Greenstone-Moretti on Million Dollar Plants:

Greenstone and Moretti study the decision of plants of were to locate (in which county in particular), and what the effect of these plants locating a specific county has on the county itself. The reason that they're interested in this problem, is that there's this real question in public economics about giving tax breaks so that large industries will locate in a specific area. In particular there's a question of while politically it may be beneficial to be seen as having attracted a large plants. However, economically these plants may have negative effects on local residents, especially when these plants or firms are offered very generous tax rebates, or tax incentives. (i.e. see the Montreal Olympics as an example of this problem)

Of course the problem of selection is that plants might be more likely to locate in certain types of counties, so we can't just compare the counties that had a plant locate versus plants that did not have the plant locate there. The ingenious idea in this paper is what we should do is use as a control group the counties that were in the running for this plant to locate in their area, but ended up losing at the very end. While these counties might also be different in some way from the winning counties, there's a better chance that there closely matched to the winning counties on covariates.

- Note the treatment and control groups being used in this paper.

- Which treatment effect is being estimated? Effect conditional on being in the running. This is not representative of all counties in the country.

- The Welfare Effect here is really not discussed. We are subsidizing firms to locate in a certain location using a dollar of tax money t which costs $1+\lambda$ dollars of social welfare to raise (note λ is the deadweight loss of taxation, estimated to be about 0.3). Essentially, at best firms will locate in the best location, and may even locate in second best location if the tax incentives are misaligned in different areas.

- The best one can do if figure out how the incentives for locating million dollar plants looks.

Table 1: "The Million Dollar Plant" Sample

	(1)
Number of Observations	
Winners	82
Losers	129
Distribution of the Number of Loser Counties per Winner	
1	57
2	14
3	7
4	2
7	1
8	1
Distribution of Cases Across Industries	
Manufacturing	63
Transportation and Public Utilities	8
Trade	4
Finance, Insurance and Real Estate	1
Services	6
Distribution of Year of Announcement of Plant Location	
1982	5
1983	3
1984	3
1985	6
1986	6
1987	8
1988	9
1989	7
1990	6
1991	12
1992	10
1993	7

Notes: The "Million Dollar Plant" sample is derived from various issues of *Site Selection*. See the text for more details.

Table 2: County Characteristics by Winner Status

	Winning Counties	Losing Counties	(1)-(2) t-stat [p-value]	All US Counties
	(1)	(2)	(3)	(4)
(1) Levels of Outcome Variables				
Total Wage Bill in New Plant's Industry ($millions)	1127	1145	0.17	175
	(1480)	(1455)	[0.87]	(760)
Employment in New Plant's Industry	40635	41568	0.13	6846
	(54143)	(49986)	[0.88]	(26990)
Aggregate Property Values ($millions)	17084	19099	-1.00	--
	(8773)	(10630)	[0.30]	
(2) Trends in Outcome Variables				
Percent Growth in Total Wage Bill in New Plant's Industry	.023	.016	1.04	.018
	(.089)	(.084)	[.30]	(.282)
Percent Growth in Employment in New Plant's Industry	.022	.011	1.57	.006
	(.092)	(.089)	[.12]	(.283)
Percent Growth in Property Values	.050	.055	-0.25	--
	(.128)	(.092)	[0.80]	
(3) Socio-Economic and Demographic Characteristics				
Per Capita Income	13660	15223	-2.70	11416
	(3211)	(4250)	[0.008]	(2636)
Percent Growth in Per capita Income	0.014	0.011	-1.02	0.011
	(0.010)	(0.019)	[0.31]	(0.057)
Per Capita Total earnings	8993	10102	-1.78	7236
	(4359)	(3730)	[0.08]	(2253)
Percent Growth in Per Capita Total Earnings	0.013	0.015	-0.89	0.011
	(0.028)	(0.025)	[0.37]	(0.124)
Employment-Population Ratio	0.541	0.573	-1.40	0.468
	(0.170)	(0.132)	[0.17]	(0.126)
Percent Growth in Employment-Population Ratio	0.010	0.011	-0.51	0.009
	(0.017)	(0.015)	[0.60]	(0.042)
Per Capita Transfers	1770	1930	-1.76	2333
	(628)	(554)	[0.080]	(1480)
Population	342876	449280	-1.79	90139
	(424939)	(346988)	[0.076]	(400341)
Fraction of Population with High- School Degree	0.729	0.762	-1.63	0.695
	(0.088)	(0.092)	[0.10]	(0.103)
Fraction of Population with College Degree	0.197	0.238	-2.22	0.134
	(0.074)	(0.089)	[0.02]	(0.063)
Fraction of Pop 17 or Younger	0.257	0.246	1.41	0.269
	(0.037)	(0.027)	[0.16]	(0.035)
Fraction of Pop 65 or Older	0.125	0.123	0.13	0.149
	(0.051)	(0.029)	[0.89]	(0.043)
(4) Geographic Distribution				
Northeast	0.093	0.250	-2.86	0.069
	(0.292)	(0.392)	[0.005]	(0.254)
Midwest	0.129	0.203	-1.43	0.340
	(0.320)	(0.375)	[0.16]	(0.473)

South	0.660	0.391	3.88	0.449
	(0.474)	(0.460)	[0.000]	(0.097)
West	0.127	0.152	-0.77	0.140
	(0.323)	(0.348)	[0.44]	(0.347)
<u>(5) Industry Distribution of the Labor Force</u>				
Construction	0.067	0.059	1.79	0.050
	(0.036)	(0.019)	[0.075]	(0.043)
Manufacturing	0.268	0.222	2.30	0.236
	(0.156)	(0.107)	[0.02]	(0.181)
Transportation, Utilities	0.052	0.059	-1.56	0.053
	(0.031)	(0.022)	[0.12]	(0.044)
Wholesale	0.068	0.068	-0.09	0.065
	(0.046)	(0.022)	[0.92]	(0.054)
Retail	0.217	0.216	0.17	0.263
	(0.053)	(0.045)	[0.87]	(0.104)
Finance, Insurance, Real Estate	0.059	0.074	-3.54	0.051
	(0.026)	(0.030)	[0.001]	(0.035)
Services	0.248	0.284	-2.99	0.246
	(0.088)	(0.070)	[0.003]	(0.101)

Notes: The figure in columns 1 and 2 are averages for the three years before the plant opening. The figures in the top panel of column 4 are a weighted average for years 1982 to 1993, with weights proportional to the number of Million Dollar cases in each year and industry. The figures in panels 2 to 5 of column 4 are a weighted average for years 1982 to 1993, with weights proportional to the number of Million Dollar cases in each year (see bottom of Table 1 for the distribution of cases across years). All monetary values are in 1983 dollars.

Figure 1: The Effect of Plant Opening on 1-Digit Industry Wage Bill in Winner and Loser Counties

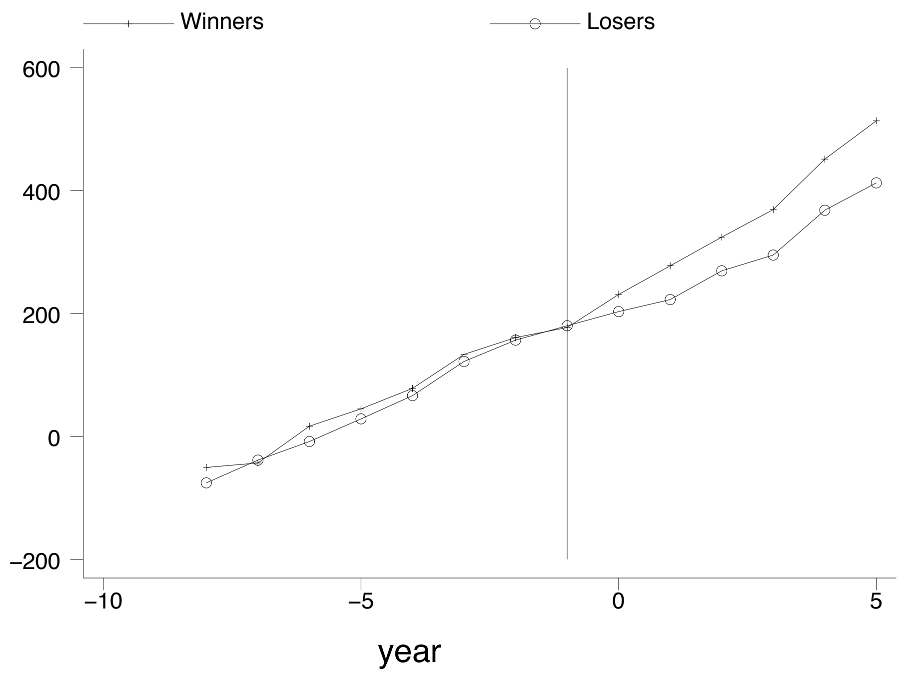


Figure 2: The Effect of Plant Opening on 1-Digit Industry Wage Bill in Winner and Loser Counties - Winner and Loser Sample

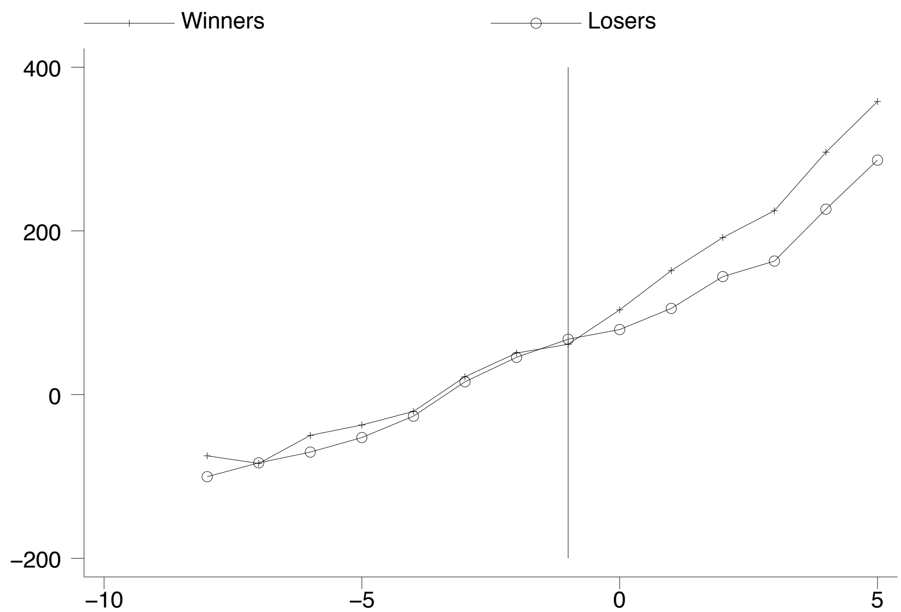


Figure 4: The Effect of Plant Opening on Wage Bill - Same Industry, Neighboring Counties - Winner and Loser Sample

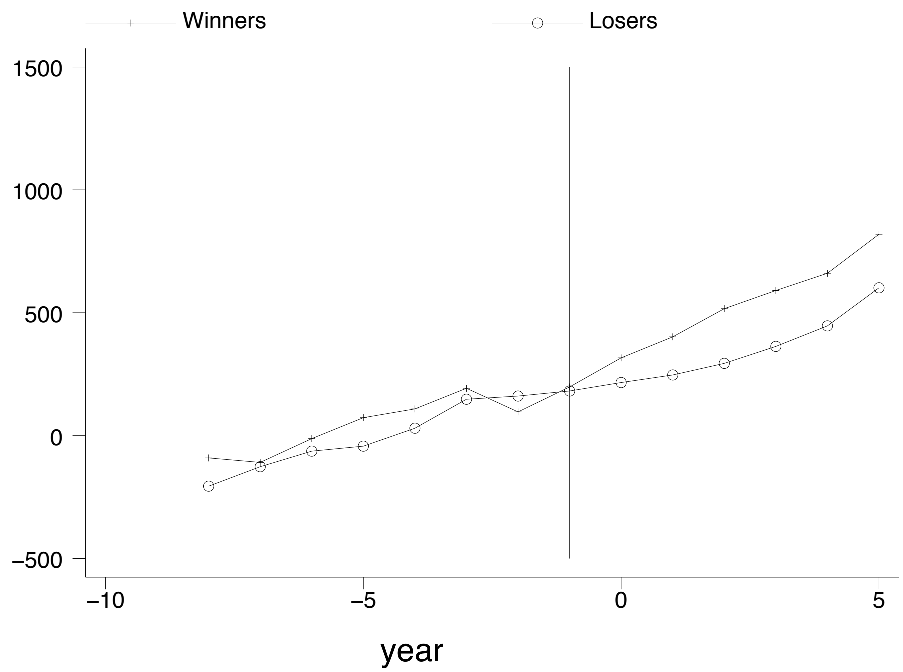


Table 3: The Effect of Plant Openings on the Wage Bill, by Year and Winner Status Relative to the Date of the Plant Location Announcement

Time	Winners (1)	Losers (2)	Difference (3)
τ = -8	-320.4	-75.5	-244.8
	(159.0)	(118.1)	
τ = -7	-313.1	-38.5	-274.6
	(158.9)	(118.0)	
τ = -6	-253.3	–8.2	-245.0
	(158.8)	(118.0)	
τ = -5	-225.1	28.5	-253.7
	(158.9)	(118.0)	
τ = -4	-191.7	66.3	-258.1
	(158.9)	(118.0)	
τ = -3	-136.4	121.9	-258.4
	(158.9)	(118.1)	
τ = -2	-109.0	156.7	-265.8
	(158.9)	(118.1)	
τ = -1	-92.7	180.1	-272.9
	(158.7)	(118.6)	
τ = 0	-39.0	203.1	-242.1
	(159.0)	(118.3)	
τ = 1	7.7	222.2	-215.0
	(158.7)	(118.1)	
τ = 2	54.3	269.4	-215.0
	(158.7)	(118.1)	
τ = 3	99.1.	295.1	-196.0
	(158.7)	(118.1)	
τ = 4	181.3	368.1	-186.7
	(158.6)	(118.1)	
τ = 5	243.6	412.6	-169.0
	(158.7)	(118.2)	

Notes: The Table reports the $\pi_{W\tau}$ (column 1) and $\pi_{L\tau}$ (column 2) coefficients and their standard errors from the estimation of equation (6) on the *County Business Patterns* data. Column (3) reports the difference between the column (1) and (2) entries. See the text for more details. The normalized coefficients are plotted in the top panel of Figure 1.

Table 4: The Effect of Plant Openings on 1-Digit Industry Wage Bill

	All US Counties		Winner and Loser Sample		All US Counties Naïve Estimator	
	(1)		(2)		(3)	
	(a)	(b)	(a)	(b)	(a)	(b)
Wage Bill						
Change in Time Trend in Winner Counties Relative to Loser Counties	16.8 (5.2)	16.7 (5.2)	13.2 (5.6)	12.6 (5.6)	24.7 (4.9)	24.8 (4.9)
Average Wage Bill in Winner Counties at $\tau = -1$	1127 (1498)	1127 (1498)	1127 (1498)	1127 (1498)	1127 (1498)	1127 (1498)
Ratio of Row 1 and Row 2	0.015	0.015	0.012	0.012	0.022	0.022
N	356589	356589	17580	17580	356589	356589
Region by Year Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: The entries in the first row are estimates of β and their heteroskedastic consistent standard errors from the fitting of equation (7). The second row reports the change in trend in one of the loser counties relative to the trend in other loser counties. The third row presents the average wage bill in the 1-digit industry of winner counties. Wage bill is measured in millions of dollars. N refers to the number of observations in the estimation of equation (6). The sample in columns 1 and 3 includes all US counties. The sample in column 2 includes only the 166 winner and loser counties.

Table 5: The Effect of Plant Openings on Wage Bill – Other Industries and Other Counties

	Other Industries, Same County (1)		Same Industry, Contiguous Counties (2)		Other Industries, Contiguous Counties (3)	
	(a)	(b)	(a)	(b)	(a)	(b)
Wage Bill						
Change in Time Trend in Winner Counties Relative to Loser Counties	42.0 (17.5)	32.1 (37.5)	40.4 (17.6)	39.1 (18.9)	72.8 (55.7)	33.8 (75.1)
Average Wage Bill in Relevant Industry in Winner Counties (or Winners' Neighbors) at τ =-1	3702 (5805)	3702 (5805)	3020 (4202)	3020 (4202)	10697 (13655)	10697 (13655)
Ratio of Row 1 and Row 2	0.011	0.009	0.012	0.012	0.007	0.003
N	17580	17580	17580	17580	17580	17580
Region by Year Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: The entries in the first row are estimates of β and their heteroskedastic consistent standard errors from the fitting of equation (7). In column (1), the $\pi_{W\tau}$'s and $\pi_{L\tau}$'s are derived from the sample of all industries besides the new plant's industry in the winner and loser counties. In columns (2) and (3), the $\pi_{W\tau}$'s and $\pi_{L\tau}$'s are obtained from by fitting equation (6) to observations from counties that neighbor the winner and loser counties for the same industry as the new plant and all other industries, respectively. See the previous tables and the text for more details.