Paper for Presentation next week: "Labour Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market" by Alex Mas in the Review of Economic Studies. This is a very good paper, and there is a lot of theory built into it, you just need to think about a bit. Good luck summarizing it in 20 minutes.

Intro Comments:

Today the class will do two things:

1. Methods: I will show you a bit of stuff on the bootstrap. It's easy to use and powerful, but we don't always teach it too formally. There are some ways in which it can be very useful, but also some pitfalls that I want to illustrate.

2. Difference in Difference Papers: I will discuss the use of difference in differences (henceforth DD) methods in economics. These are used so often that it is good to know what goes into them.

The Bootstrap:

The classic text is "An Introduction to the Bootstrap" by Bradley Efron and R.J. Tibshirani. If you want to know what really goes on to prove the results on the bootstrap, the best book is "The Bootstrap and Edgeworth Expansion" by Peter Hall. As well, Joel Horowitz has a chapter on the bootstrap in the most recent handbook of econometrics if you want more background.

The reason we think that the bootstrap is useful is:

- Powerful procedure that works almost all the time that standard asymptotics do.

- Replaces derivations of asymptotic distributions with computation: ultimately a much better use of your time.

- Stata makes coding these bootstraps very easy.

- The bootstrap is known to have better small sample properties than asymptotic estimators.

Let's go through the idea behind the bootstrap with a simple example. Suppose I want to compute the probability of picking Lauren to present next week. One way to do this is to do the standard derivation:

$$P[Lauren] = \frac{1}{N}$$

where N is the number of people in the room.

Another way to do this is the following:
Repeatedly pick names out of the hat. See if Lauren got picked. Repeat a bunch of time. See what the mean probabilities are of Lauren being picked. As well, we can compute how probable it would be of Lauren being picked say 5 times this semester, or picked more than 4 times using the same procedure.

Bootstrap:

- Bootstrap Sample:

  Pick N of the items in the data $x = (x_1, x_2, \cdots, x_N)$ called the bootstrap sample b. We call the empirical distribution of the data $\widehat{F}_X(\cdot)$, just the distribution of observations in the data. So the bootstrap sample is just:

$$x^{*b} \sim \prod_{i=1}^{N} \widehat{F}_X(\cdot)$$

- The bootstrap algorithm:

  For $b = 1, \cdots, B$:

  1. Draw a bootstrap sample x^{*b}.
  2. Compute the statistic $\widehat{\theta}^{*b}$ as:
  $$\widehat{\theta}^{*b} = s(x^{*b})$$
  3. Get the distribution $\vec{\theta} = (\widehat{\theta}^{*1}, \widehat{\theta}^{*2}, \cdots, \widehat{\theta}^{*B})$.
  4. I can compute the variance of θ as:
  $$\widehat{se}^B = \sum_{b=1}^{B} \left(\widehat{\theta}^{*b} - \widehat{\theta}^{*}(\cdot) \right)$$
  with:

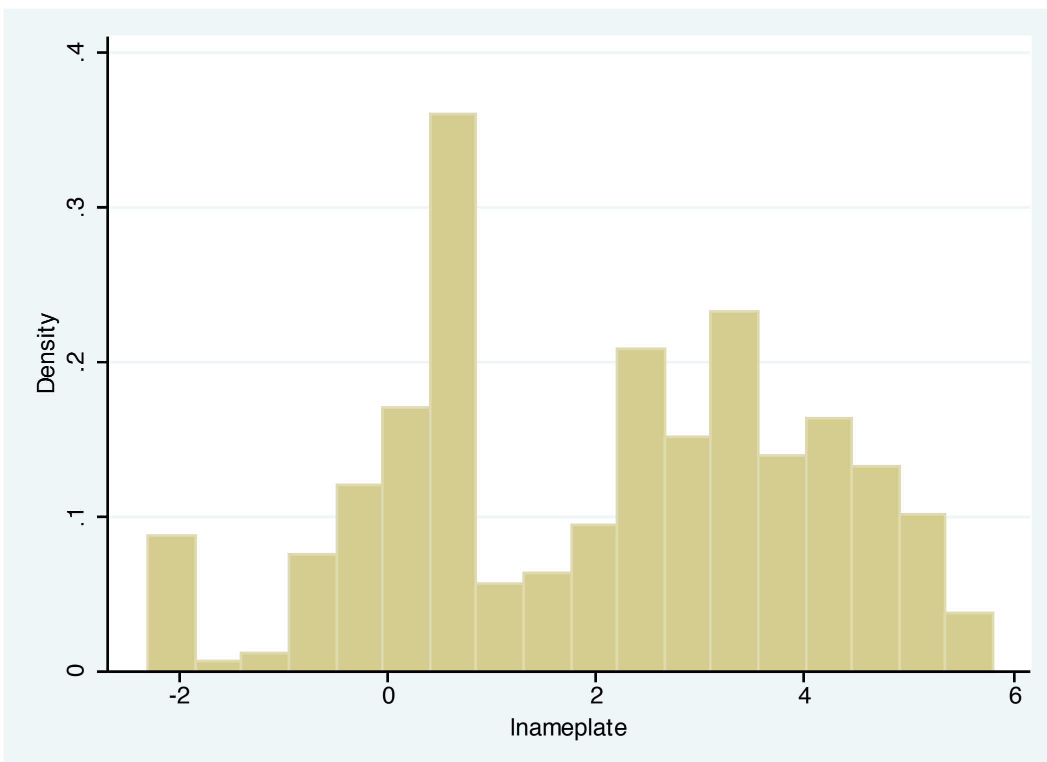
  Note that as $N \to \infty$ then we can usually have $\widehat{F}_X^N(\cdot) \to F_X^0(\cdot)$. As well, as $B \to \infty$ then $\widehat{se}^B \to$

$\widehat{se}^0$, with as $B = \infty$ this becomes the ideal bootstrap. For the standard error, $B = 200$ usually works just fine.

- The bootstrap algorithm for confidence intervals:

  For $b = 1, \cdots, B$:

  1. Draw a bootstrap sample x^{*b}.
  2. Compute the statistic $\widehat{\theta}^{*b}$ as:
  $$\widehat{\theta}^{*b} = s(x^{*b})$$
  3. Get the distribution $\vec{\theta} = (\widehat{\theta}^{*1}, \widehat{\theta}^{*2}, \cdots, \widehat{\theta}^{*B})$.
  4. I can compute the distribution of θ called $\widehat{F}_{\theta}^{B}$:



And we look at the 5th and 95th percentiles:
$\widehat{F}_{\theta}^{B,5}$ and $\widehat{F}_{\theta}^{B,95}$,

Practical Example

- So this gets really useful with more complex asymptotics. For instance there is a productivity regression:

  $$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it}$$

  but there is a censoring problem due to the fact that I only see ω_{it} of the firms that decide to stay in the market, i.e. if $\chi_{it} = 0$. Now there is a way to purging the model of this selection bias, by computing the probability that a firm exits (called the exit propensity score) and then putting it into the regression.

  So the probability a firm exits is:

  $$\chi_{it} = \phi(x_{it}\beta)$$

  So we can compute the predicted exit probability $\widehat{\chi_{it}} = \phi(x_{it}\widehat{\beta})$. And then put it into the production function:

  $$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \gamma\widehat{\chi}_{it} + \epsilon_{it}$$

  Now the problem for asymptotics here is that $\widehat{\chi}_{it}$ is *estimated*, so it affects the standard errors for α_l and α_k need to take account of this. While there are many papers that tell you how to do this (say Murphy and Topel (1985) for instance), it's not exactly easy stuff to do. Instead one could solve this problem with a bootstrap estimator.

For $b = 1, \cdots, B$:

1. Draw a bootstrap sample x^{*b}.
2. Run the probit on exit on the bootstrap data x^{*b}.
$$\chi_{it}^{b} = \phi(x_{it}^{*b}\widehat{\beta}^{b})$$
3. Get the propensity score:
$$\widehat{\chi}_{it}^{*b} = \phi(x_{it}^{*b}\widehat{\beta}^{b})$$
4. Run the production function estimation:
$$y_{it}^{*b} = \alpha_0^{*b} + \alpha_l^{*b} l_{it}^{*b} + \alpha_k^{*b} k_{it}^{*b} + \gamma^{*b}\widehat{\chi}^{*b} + \epsilon_{it}$$
and we get α_l^{*b} and α_k^{*b}.
5. You can get confidence intervals from the distribution of α_l^{*b} and α_k^{*b} over $b = 1, \cdots, B$.

Here's a quick STATA example:

```
program acfestimates2, rclass
local exitlist="ltae comp morecomp mu lebd"
cap drop exitprob
probit jmdeath `exitlist'
predict exitprob
reg lva lsw ltae  exitprob
return scalar betalsw=b[1,1]
return scalar betaltae=b[1,2]
end

bootstrap betalsw=r(betalsw) betaltae=r(betaltae),
reps(100): acfestimates2
```

Properties of the Bootstrap

- The bootstrap almost always works when standard asymptotics do.

- The theory behind this is derived from Edgeworth Expansions.

- The cases where the bootstrap fails:

  1. Parameters on the boundary, such as extremal statistics like a max or a min.
  2. Discontinuous statistics.
  3. Cases where the variances or some other moment is infinite.

- The bootstrap typically has better small sample properties that asymptotics results, especially when these statistics are what are called asymptotically pivotal.

Bootstrap and Correlation among observations:

Note that we've assumed that the x_i 's are i.i.d.. This is often false, such as:

- Serial Correlation among observations: say if I look at the decision to open a store every year versus every day, I don't exactly have 365 times more data to identify the parameters of interest. Or let's go further and do this every second, the variance won't fall by $\sqrt{N}$.

- Some times there is correlation among groups. If I sample people in several different towns, I might do worse that if I could sample them independently.

We will discuss fixes for this, 1) clustering standard errors (which is all the rage in reduced form work), and 2) how to do block bootstrap.

Block Bootstrap:

Idea: Instead of sampling the bootstrap sample independently, we should try to replicate the correlation of the data.

Panel Block Bootstrap: Let's start with a panel y_{it}, x_{it}, with $i = 1, \cdots, M$ and $t = 1, \cdots, T$. Pick M of the markets items in the data randomly $\mathcal{M}^{*b} = \{m_1, m_2, m_3, m_2, m_3\}$, and then pull all the observations for each of the markets $t = 1, \cdots, T$ to get the bootstrap sample b.

Pure Time Series Block Bootstrap: Let's turn to a pure time series, y_t, x_t. The block bootstrap samples blocks of length K in the data:

$$x^{*b} = \{x_t, x_{t+1}, \cdots, x_{t+K}\}$$

where t is chosen randomly. The block bootstrap does not work that well if t is not really big. As well, there is an issue with "overlapping blocks" typically, and how to choose the block length K (say 10 or whatever).

Look the cluster command in bootstrap to learn more on this.

Difference in Differences:

The typical treatment effect looks at two individuals, one who get the treatment and one who gets the control.

- Suppose we have the specification for the outcome (from our model):

  $$y_{it} = \delta_t + \alpha d_{it} + \beta x_{it} + \epsilon_{it}$$

  where δ_t are time controls, d_{it} is the treatment assignment and x_{it} are observed covariates, while ϵ_{it} are unobserved covariates.

  Now the typical issue with estimating the treatment effect α is that:

  $$E[d_{it}\epsilon_{it}] \neq 0$$

- Note that if don't have any x_{it} variables this is very likely to be a problem, since certain workers (for instance) are more likely to take the treatment if they are say young, versus old. So having a lot of x's could in principle solve the problem of selection (i.e. better data is a perfectly fine, and probably better solution to the problem of selection on unobservables than the econometric techniques we will present here).

- An alternative assumption is to assume that the unobservables can be decomposed into:

  $$\epsilon_{it} = \mu_i + \eta_{it}$$

an individual specific component and a time variant component at the individual level. For instance, we might think that what we care about for the worker training program is the ability unobservable a_{it}, but it may be reasonable to assume that this unobservable is persistent, i.e. $a_{it} = a_i \forall t$.

Thus it may be more reasonable to assume the following on selection:

$$E[d_{it}\eta_{it}] = 0$$

which is strictly weaker than $E[d_{it}\epsilon_{it}] = 0$.

Let's rewrite the outcome equation in first-differences as:

$$\begin{aligned}
y_{it} - y_{it-1} =& \delta_t + \alpha d_{it} + \beta x_{it} + \epsilon_{it} \\
& - \delta_{t-1} - \alpha d_{it-1} - \beta x_{it-1} - \epsilon_{it-1} \\
=& (\delta_t - \delta_{t-1}) + \alpha(d_{it} - d_{it-1}) + \beta(x_{it} - x_{it-1}) \\
& + (\eta_{it} - \eta_{it-1})
\end{aligned}$$

Now we can estimate the treatment effect by assuming:

$$E[(d_{it} - d_{it-1})(\eta_{it} - \eta_{it-1})] = 0$$

- Difference in differences:

  Now the idea of diff in diff is that the may be some other variables that change over time that may be correlated with the assignment of the treatment. For instance Steven Ryan (2009) has a nice paper on the effect of the Clean Air Act on entry costs in the cement industry. The issue is that maybe there are other things that are correlated with the introduction of the Clean Air Act in (1991), like a spike in demand for cement that was not expected by the industry.

The idea of DD methods is to filter these problems out by comparing two groups, one that gets the treatment, the other which is a control group. We use the following to estimate the effect:

$$\begin{aligned}
&(y_{it} - y_{it-1} | d_{it} = 1) - (y_{jt} - y_{jt-1} | d_{it} = 0) \\
&= \alpha(d_{it}) + \beta[(x_{it} - x_{it-1}) - (x_{jt} - x_{jt-1})] \\
&\qquad\qquad + (\eta_{it} - \eta_{it-1}) - (\eta_{jt} - \eta_{jt-1})
\end{aligned}$$

Note that the time effects fall out here, and we only have 2 parameters to estimate (1 if the x_{it} don't matter, or are all time invariant).

- First-Differences don't always work

  One of the real conceptual issues with doing estimates in D, DD or DDD is that taking difference changes the identification of the model radically. I am only using variation over time to identify the coefficients. If I want to look at the effect of intelligence on wages, do I want to look at:

1. Intelligence regressed on wages in the cross-section.
2. Changes in Intelligence regressed on changes in wages.

Well, what would generate changes in intelligence? Would having a larger measurement issues with them (like I take a test twice and get slightly difference readings on my IQ), or with wages I might be unemployed at some point in time, which would generate large changes in wages that are not entirely “real”.

There is a nice paper by Griliches and Mairaisse on the search for identification in production function analysis, that goes through the formulas for the signal to noise ratio when you have mismeasured variables. It is a good paper to read if you’ve got the time.

- Quasi Differences:

  Suppose instead of assuming the decomposition $\epsilon_{it} = \mu_i + \eta_{it}$, we assume that:

  $$\epsilon_{it} = \rho\epsilon_{it} + \eta_{it}$$

  which is just an AR(1) assumption on the evolution of the error.

  The the quasi-difference we are looking at is:

  $$y_{it} - \rho y_{it-1} = (\delta_t - \rho\delta_{t-1}) + \alpha(d_{it} - \rho d_{it-1}) + \beta(x_{it} - \rho x_{it-1}) + \eta_{it}$$

I'm looking for this type of idea to be used more often in empirical work in the future.

Attenuation Bias

When we first-difference we change the attenuation bias inherent in the problem, and the paper "Errors in Variables in Econometrics" by Griliches and Hausman (1986) in the Journal of Econometrics does a good job with this. Let's do the classic errors-in-variable problem.

Suppose the relationship we are looking at is:

$$y_{it} = \alpha x^*_{it} + \epsilon_{it}$$

and there is error-in-variables for x^*_{it}, i.e. we get to see x_{it} given by:

$$x_{it} = x^*_{it} - v_{it}$$

where v_{it} is measurement error with variance σ_v. What this means is that $v_{it} \perp x^*_{it}$, i.e. it's not endogenous, and it is mean 0.

But this means that $cov(x_{it}, v_{it}) = cov(v_{it}, v_{it}) = \sigma^2_v$. Thus our estimate will be *attenuated*:

$$\widehat{\alpha} \to \alpha \left(\frac{\sigma^2_X}{\sigma^2_X + \sigma^2_v} \right)$$

which is know as the noise to signal problem, i.e. how much noise is there compared to the information in x.

Attenuation Bias in Panel Data

Suppose we have the following relationship:

$$y_{it} = \alpha x^*_{it} + \mu_i + \epsilon_{it}$$

which we first difference to get rid of the μ_i:

$$\begin{aligned} y_{it} - y_{it-1} &= \alpha(x^*_{it} - x^*_{it-1}) + (\epsilon_{it} - \epsilon_{it-1}) \\ &= \alpha(x_{it} - x_{it-1} + v_{it} - v_{it-1}) + (\epsilon_{it} - \epsilon_{it-1}) \end{aligned}$$

so our new variables are:

$$y_{it} - y_{it-1} = \alpha(\Delta x_{it} + \Delta v_{it}) + \Delta \epsilon_{it}$$

and the associated attenuation bias is:

$$\widehat{\alpha} \to \alpha \left(\frac{\sigma^2_{\Delta X}}{\sigma^2_{\Delta X} + \sigma^2_{\Delta v}} \right)$$

Is this better or worse than before?

Helena Smoking Experiment:

- Question: What is the effect of second hand smoke on the incidence of heart disease.

- There is some medical experiments that show that this may be the case, but the effects "in the field" so to speak are ambiguous.

- Policies under consideration:

  Banning smoking in bars and restaurants. This policy was implemented in many cities and countries over the last 10 years, mainly because of the effects of second hand smoke on health outcomes.

- Policy Shift:

  > Helena, Montana, USA, is a geographically isolated community that imposed such a law from 5 June 2002. Opponents won a court order suspending enforcement of the law on 3 December 2002. This allowed us to examine the association of the ordinance with admissions for myocardial infarction from within Helena (intervention) and from outside Helena, where the ordinance did not apply (control).

- The statistical model they have in mind is a diff in diff:

$$y_{it} = \alpha_i + \delta_t + \beta d_{it} + \epsilon_{it}$$

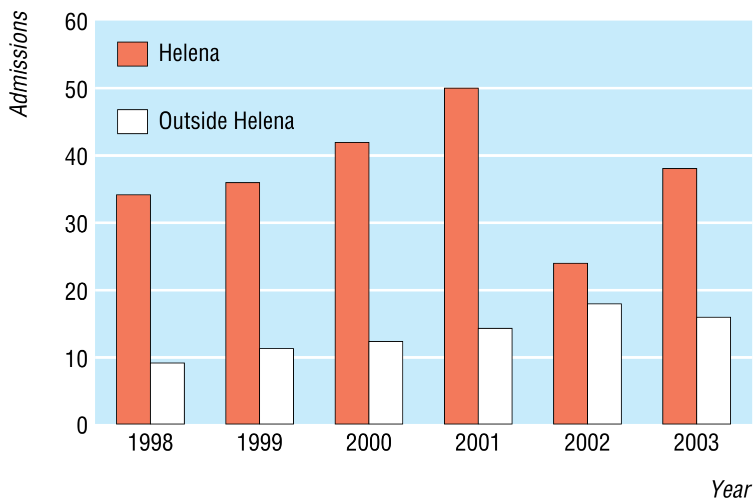
where α_i is a market fixed effect, δ_t is a time fixed effect and $d_{it} \in \{0, 1\}$ indicates whether the treatment of banning smoking was in effect or not. Again the implicit assumption to identify β is unconfoundness, i.e.

$$E[d_{it}\epsilon_{it}|i, t] = 0$$

hence the real assumption here is both: 1) linearity of the effects we are looking at and 2) there isn't something else changing in one of these markets that is contemporaneous with the treatment being turned on. For example, perhaps there was a health campaign that made people exercise more in Helena and also generated the decision to ban smoking in Helena. Again, the issue is the policies get enacted for a reason, and these reasons may be correlated with other changes we don't observe.

- Notice that the market fixed effects suck out any of the variables $x_{it} = x_i$ that don't change.

- Notice that the time fixed effects purge the model of issues of common shocks, $x_{it} = x_t$ that don't change.



Admissions for acute myocardial infarction during six month periods June-November before, during (2002), and after the smoke-free ordinance (ordinance did not apply outside Helena). The law was implemented on 5 June 2002

Admissions for acute myocardial infarction during six month period (June to November) when smoking ban was enforced and equivalent months in years before and after ban, according to areas with (Helena) and without enforcement*

	Helena	**Not Helena**
Ordinance year (2002)	24	18
Other years†	40	12.4
Difference (95% CI)	–16 (–31.7 to –0.3)	5.6 (–5.2 to 16.4)
Helena difference–not Helena difference (95% CI)	–21.6 (–40.6 to –2.6)	

*All comparisons done assuming Poisson distribution.
†Average number of admissions during six month period for years other than 2002.

Card and Krueger on Minimum Wages:

- Question: What is the impact of minimum wages on employment and business activity.

- The Model: Standard price theory says that the number of workers hired is chosen so that:

$$\frac{\partial R}{\partial q}\frac{\partial F(\theta)}{\partial L} = w(\theta)$$

  So raising the wage for a worker of type θ (say where the workers are ranked by θ according to their marginal product of labor, should by diminuishing returns reduce the number of hired workers (for the worker types θ for which the minimum wage laws are binding).

- Question: Is the minimum wage effect large or not?

- Natural Experiment:

  New Jersey Raised it's minimum wage from \$4.25 to \$5.05 on April 1st 1992. The control group is Pennsylvania, where this did not happen.

- Data Collection: Fast food restaurants which tend to employ people at wages below the minimum wage. So Card and Krueger collect data from these restaurants before and after policy shift.

- There are several effects that will be measured:
  1. Effect on Employment.
  2. Effect on Employment at restaurants with previously low wages, i.e. for those restaurants where the minimum wage binds.
  3. Effect on Fringe Benifits.
  4. Pass-through of increases in wages to consumers.

Table 1—Sample Design and Response Rates

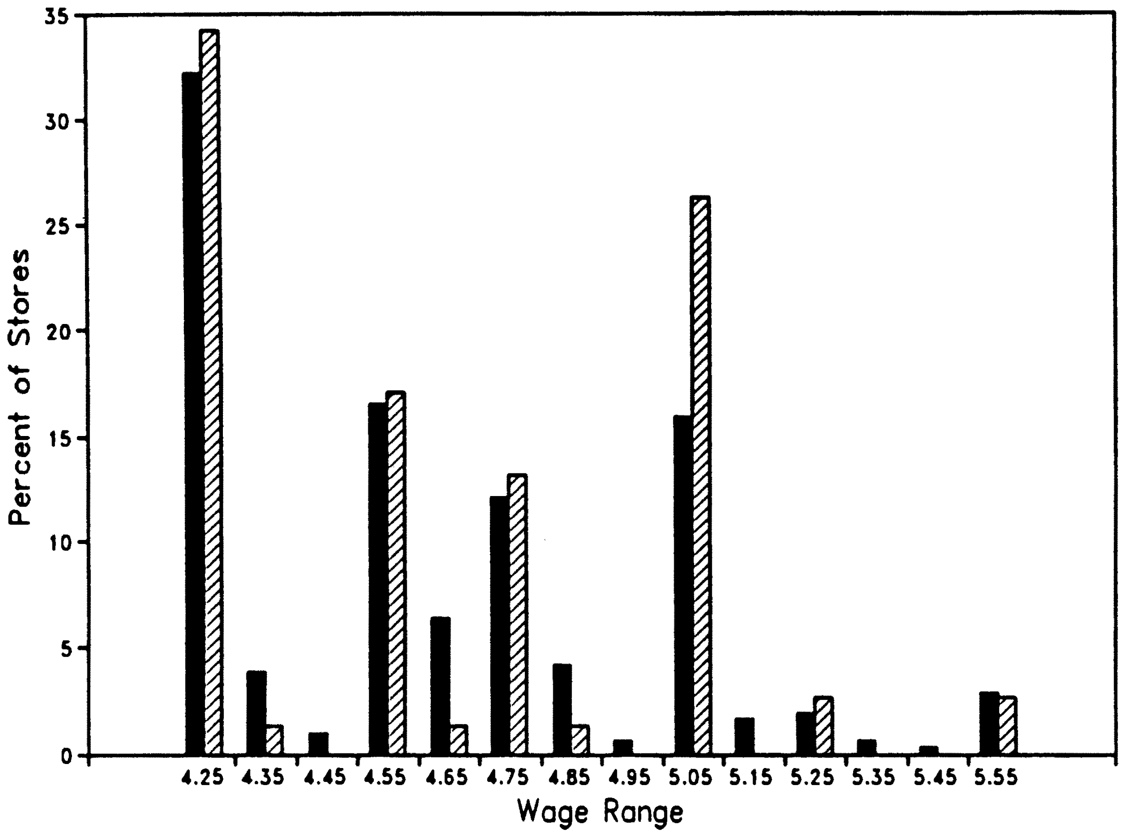
	All	Stores in: NJ	Stores in: PA
Wave 1, February 15–March 4, 1992:			
Number of stores in sample frame:[a]	473	364	109
Number of refusals:	63	33	30
Number interviewed:	410	331	79
Response rate (percentage):	86.7	90.9	72.5
Wave 2, November 5–December 31, 1992:			
Number of stores in sample frame:	410	331	79
Number closed:	6	5	1
Number under rennovation:	2	2	0
Number temporarily closed:[b]	2	2	0
Number of refusals:	1	1	0
Number interviewed:[c]	399	321	78

[a] Stores with working phone numbers only; 29 stores in original sample frame had disconnected phone numbers.

[b] Includes one store closed because of highway construction and one store closed because of a fire.

[c] Includes 371 phone interviews and 28 personal interviews of stores that refused an initial request for a phone interview.

February 1992



November 1992

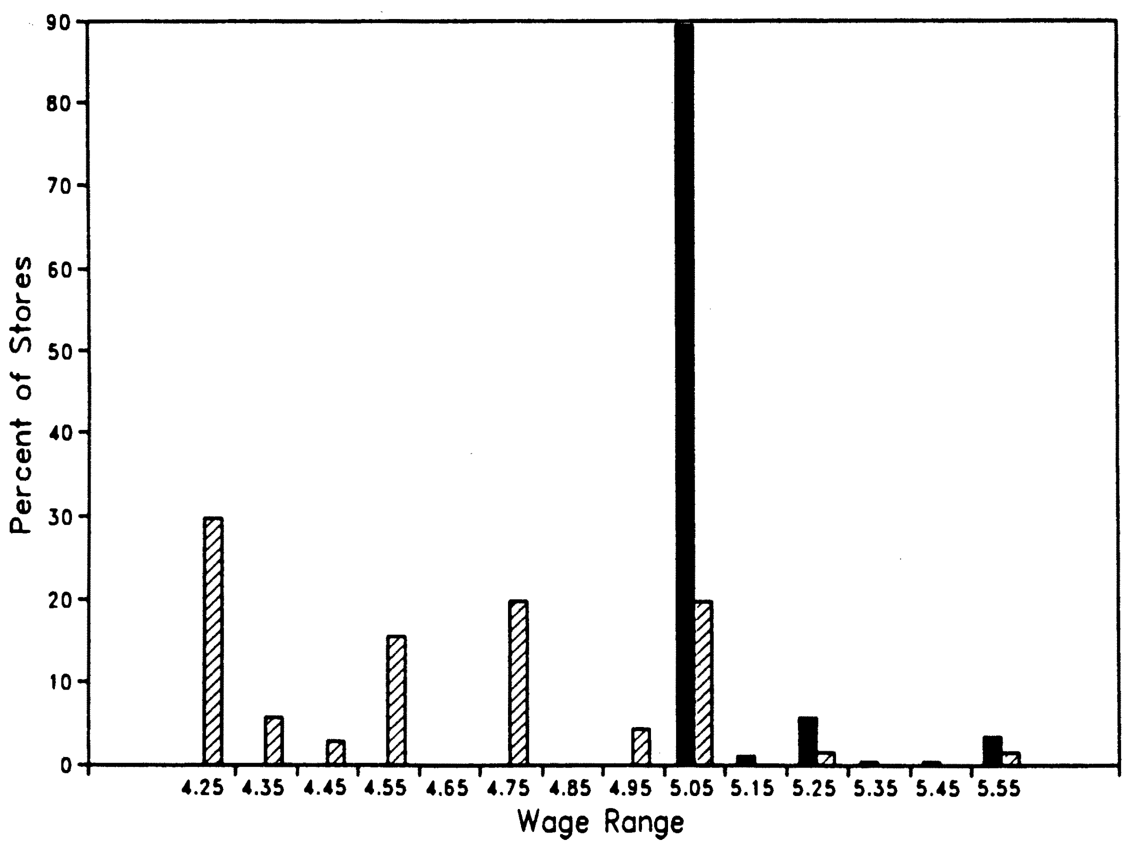


TABLE 3—AVERAGE EMPLOYMENT PER STORE BEFORE AND AFTER THE RISE IN NEW JERSEY MINIMUM WAGE

	Stores by state			Stores in New Jersey[a]			Differences within NJ[b]	
Variable	PA (i)	NJ (ii)	Difference, NJ − PA (iii)	Wage = \$4.25 (iv)	Wage = \$4.26–\$4.99 (v)	Wage ≥ \$5.00 (vi)	Low–high (vii)	Midrange–high (viii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	−2.89 (1.44)	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)	−2.69 (1.37)	−2.17 (1.41)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	−0.14 (1.07)	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)	0.67 (1.44)	0.75 (1.27)
3. Change in mean FTE employment	−2.16 (1.25)	0.59 (0.54)	2.76 (1.36)	1.32 (0.95)	0.87 (0.84)	−2.04 (1.14)	3.36 (1.48)	2.91 (1.41)
4. Change in mean FTE employment, balanced sample of stores[c]	−2.28 (1.25)	0.47 (0.48)	2.75 (1.34)	1.21 (0.82)	0.71 (0.69)	−2.16 (1.01)	3.36 (1.30)	2.87 (1.22)
5. Change in mean FTE employment, setting FTE at temporarily closed stores to 0[d]	−2.28 (1.25)	0.23 (0.49)	2.51 (1.35)	0.90 (0.87)	0.49 (0.69)	−2.39 (1.02)	3.29 (1.34)	2.88 (1.23)

TABLE 5—SPECIFICATION TESTS OF REDUCED-FORM EMPLOYMENT MODELS

	Change in employment		Proportional change in employment	
Specification	NJ dummy (i)	Gap measure (ii)	NJ dummy (iii)	Gap measure (iv)
1. Base specification	2.30 (1.19)	14.92 (6.21)	0.05 (0.05)	0.34 (0.26)
2. Treat four temporarily closed stores as permanently closed[a]	2.20 (1.21)	14.42 (6.31)	0.04 (0.05)	0.34 (0.27)
3. Exclude managers in employment count[b]	2.34 (1.17)	14.69 (6.05)	0.05 (0.07)	0.28 (0.34)
4. Weight part-time as 0.4 × full-time[c]	2.34 (1.20)	15.23 (6.23)	0.06 (0.06)	0.30 (0.33)
5. Weight part-time as 0.6 × full-time[d]	2.27 (1.21)	14.60 (6.26)	0.04 (0.06)	0.17 (0.29)
6. Exclude stores in NJ shore area[e]	2.58 (1.19)	16.88 (6.36)	0.06 (0.05)	0.42 (0.27)
7. Add controls for wave-2 interview date[f]	2.27 (1.20)	15.79 (6.24)	0.05 (0.05)	0.40 (0.26)
8. Exclude stores called more than twice in wave 1[g]	2.41 (1.28)	14.08 (7.11)	0.05 (0.05)	0.31 (0.29)
9. Weight by initial employment[h]	—	—	0.13 (0.05)	0.81 (0.26)
10. Stores in towns around Newark[i]	—	33.75 (16.75)	—	0.90 (0.74)
11. Stores in towns around Camden[j]	—	10.91 (14.09)	—	0.21 (0.70)
12. Pennsylvania stores only[k]	—	−0.30 (22.00)	—	−0.33 (0.74)

TABLE 6—EFFECTS OF MINIMUM-WAGE INCREASE ON OTHER OUTCOMES

	Mean change in outcome			Regression of change in outcome variable on:		
Outcome measure	NJ (i)	PA (ii)	NJ − PA (iii)	NJ dummy (iv)	Wage gap[a] (v)	Wage gap[b] (vi)
Store Characteristics:						
1. Fraction full-time workers[c] (percentage)	2.64	−4.65	7.29	7.30	33.64	20.28
	(1.71)	(3.80)	(4.17)	(3.96)	(20.95)	(24.34)
2. Number of hours open per weekday	−0.00	0.11	−0.11	−0.11	−0.24	0.04
	(0.06)	(0.08)	(0.10)	(0.12)	(0.65)	(0.76)
3. Number of cash registers	−0.04	0.13	−0.17	−0.18	−0.31	0.29
	(0.04)	(0.10)	(0.11)	(0.10)	(0.53)	(0.62)
4. Number of cash registers open at 11:00 A.M.	−0.03	−0.20	0.17	0.17	0.15	−0.47
	(0.05)	(0.08)	(0.10)	(0.12)	(0.62)	(0.74)
Employee Meal Programs:						
5. Low-price meal program (percentage)	−4.67	−1.28	−3.39	−2.01	−30.31	−33.15
	(2.65)	(3.86)	(4.68)	(5.63)	(29.80)	(35.04)
6. Free meal program (percentage)	8.41	6.41	2.00	0.49	29.90	36.91
	(2.17)	(3.33)	(3.97)	(4.50)	(23.75)	(27.90)
7. Combination of low-price and free meals (percentage)	−4.04	−5.13	1.09	1.20	−11.87	−19.19
	(1.98)	(3.11)	(3.69)	(4.32)	(22.87)	(26.81)
Wage Profile:						
8. Time to first raise (weeks)	3.77	1.26	2.51	2.21	4.02	−5.10
	(0.89)	(1.97)	(2.16)	(2.03)	(10.81)	(12.74)
9. Usual amount of first raise (cents)	−0.01	−0.02	0.01	0.01	0.03	0.03
	(0.01)	(0.02)	(0.02)	(0.02)	(0.11)	(0.11)
10. Slope of wage profile (percent per week)	−0.10	−0.11	0.01	0.01	−0.09	−0.08
	(0.04)	(0.09)	(0.10)	(0.10)	(0.56)	(0.57)

TABLE 7—REDUCED-FORM MODELS FOR CHANGE IN THE PRICE OF A FULL MEAL

Independent variable	Dependent variable: change in the log price of a full meal (i)	(ii)	(iii)	(iv)	(v)
1. New Jersey dummy	0.033 (0.014)	0.037 (0.014)	—	—	—
2. Initial wage gap[a]	—	—	0.077 (0.075)	0.146 (0.074)	0.063 (0.089)
3. Controls for chain and[b] ownership	no	yes	no	yes	yes
4. Controls for region[c]	no	no	no	no	yes
5. Standard error of regression	0.101	0.097	0.102	0.098	0.097

TABLE 8—ESTIMATED EFFECT OF MINIMUM WAGES ON NUMBERS OF MCDONALD'S RESTAURANTS, 1986–1991

Independent variable	Dependent variable: proportional increase in number of stores (i)	(ii)	(iii)	(iv)	Dependent variable: (number of newly opened stores) ÷ (number in 1986) (v)	(vi)	(vii)	(viii)
Minimum-Wage Variable:								
1. Fraction of retail workers in affected wage range 1986[a]	0.33 (0.20)	—	0.13 (0.19)	—	0.37 (0.22)	—	0.16 (0.21)	—
2. (State minimum wage in 1991) ÷ (average retail wage in 1986)[b]	—	0.38 (0.22)	—	0.47 (0.22)	—	0.47 (0.23)	—	0.56 (0.24)
Other Control Variables:								
3. Proportional growth in population, 1986–1991	—	—	0.88 (0.23)	1.03 (0.23)	—	—	0.86 (0.25)	1.04 (0.25)
4. Change in unemployment rates, 1986–1991	—	—	−1.78 (0.62)	−1.40 (0.61)	—	—	−1.85 (0.68)	−1.40 (0.65)
5. Standard error of regression	0.083	0.083	0.071	0.068	0.088	0.088	0.077	0.073

Leslie and Jin

Jin and Leslie (2003) and Milyo and Waldfogel (1999) are fundamentally about the role of information in consumers decision making:

1. Prices.

2. Product Characteristics: hygiene and mortality of surgery.

Advertising is just one way in which firms transmit information to firms. There are several questions about what type of information will be send to consumers, i.e. will this be truthful revelation or not, and the impact of information transmission on consumer's behavior. In particular, consumers may already have a strong incentive to learn about the products that firms are offering so it is not clear why a firm message would be credible.

Jin and Leslie (2003) exploit a "natural experiment" that occurred in Los Angeles in 1997. The city government introduced rules forcing restaurants to post the results of their hygiene report, including a very visible colored grade, on the front of their restaurant door. Figure **??** show one such example.

TABLE II
TIMING OF MANDATORY GRADE CARD ORDINANCE

Panel A: Number of restaurants located in cities with or without the ordinance

Quarter	Total restaurants	Ordinance not adopted: No. restaurants	Ordinance not adopted: % of restaurants	Ordinance adopted: No. restaurants	Ordinance adopted: % of restaurants
1998 Q1	10126	9626	95.06	500	4.94
Q2	10238	3806	37.18	6432	62.82
Q3	10222	2972	29.07	7250	71.93
Q4	9883	2009	20.33	7874	79.67

Panel B: Disclosure status of restaurants

Quarter	% of restaurant days under Regime I: *voluntary disclosure without standard format*	% of restaurant days under Regime II: *voluntary disclosure with standard format*	% of restaurant days under Regime III: *mandatory disclosure*
1998 Q1	84.64	15.04	0.32
Q2	40.78	43.20	16.02
Q3	15.59	41.28	43.13
Q4	4.04	34.27	61.69

Every restaurant receives an official grade card following inspections conducted after January 16, 1998. However, the restaurant is only required to post the grade card if it is located in a city which has adopted the ordinance. Restaurants not yet inspected after January 16, 1998, fall under Regime I.

The introduction of report cards was staggered in different cities within Los Angeles County. This allows Jin and Leslie (2003) to use a "difference in differences" strategy to identify the effect of report cards on both hygiene scores and food related disease. A "difference in differences" estimation identifies the effect of a policy from the differences between changes for the treatment and control group.

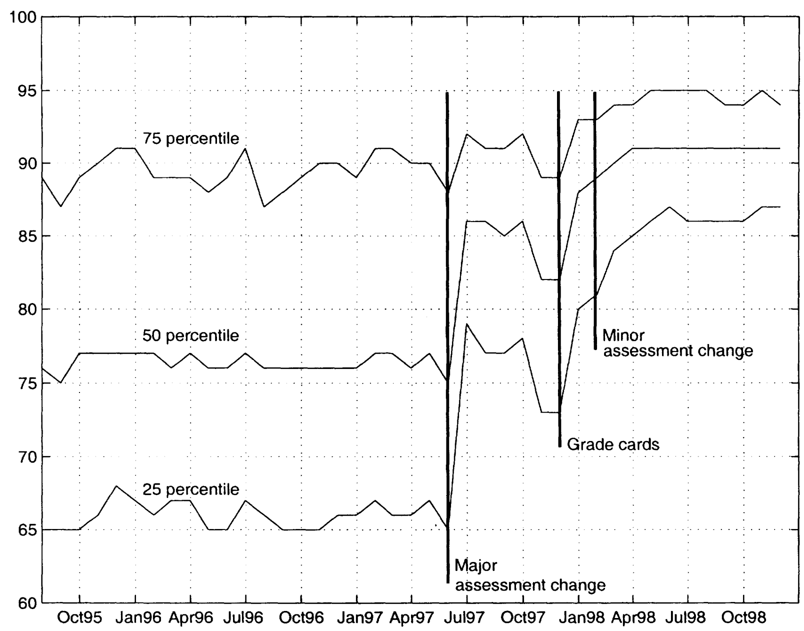
$$h_i^t = \alpha_i + \beta_1 m_i^t + \beta_2 v_i^t + \gamma_1 c_1^t + \gamma_2 c_2^t + \epsilon_i^t \qquad (1)$$

Suppose instead that Jin and Leslie (2003) could only use a "differences" approach, comparing the treatment group before and after the policy was introduced.

$$h_i^t = \alpha_i + \beta_1 m^t + \beta_2 v_i^t + \gamma_1 c_1^t + \gamma_2 c_2^t + \epsilon_i^t \qquad (2)$$

In this case, it would be impossible to separate the effect of the policy change m^t from changes in the grading scheme used by the DHS (the c_1^t and c_2^t terms). Without the use of difference in differences strategy it could be quite easy to look for changes in behaviour that coincide with policy changes and attribute these to the change in the policy. Moreover, the actual policy which is being implemented is often unclear, even if the text of the regulation is precise. Wolfram and Bushnell (2006) look at the effect of New Source Review for power plants on emissions. Unfortunately, was is the case that many plants increased their generating capacity without investing into pollution control technology.

One nice characteristic of Jin and Leslie (2003) is that they investigate the potential that the grading criteria for the hygiene score also changed during the period. Most notably, after grading cards are introduced, there is a large spike at 90% (the threshold for a A grade) which did not exist before the introduction of grade cards.



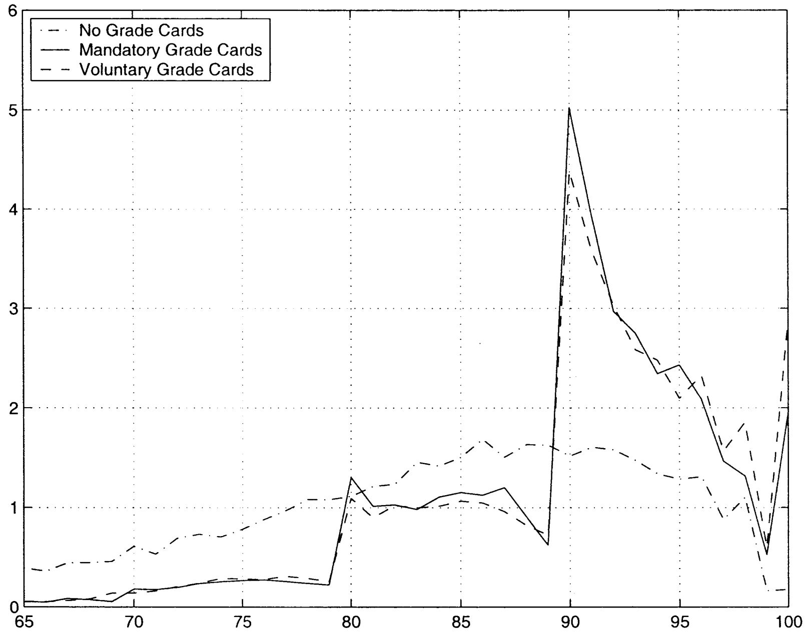


FIGURE II
Distributions of Hygiene Scores under Different Disclosure Regimes
The figure is no different from a histogram (or an unsmoothed nonparametric density). Units on the vertical axis are meaningless.

One important question is why didn't these restaurants decide to post their health inspection reports before mandatory report cards were introduced. However, while a restaurant may want to distinguish itself from rivals by drawing attention to it's hygiene score, this may also alert consumers to the dangers of eating at restaurants in general. The best example of these "product related risks" can be found in airline advertising. If United showed an ad that said that United is a safer airline than American since it had fewer crashes, consumer would probably react by not flying any airline at all instead of substituting towards United.

TABLE III
THE EFFECTS OF GRADE CARDS AND DISCLOSURE REGULATION ON HYGIENE SCORES

	Without fixed effects		With fixed effects	
	Coefficient	Std. error	Coefficient	Std. error
Mandatory disclosure	4.9432	1.1384***	4.3958	1.4046***
Voluntary disclosure	4.0585	0.3199***	3.2528	0.3550***
Inspection Criteria II	7.7192	0.9181***	8.0886	0.9907***
Inspection Criteria III	9.9838	1.2233***	10.4158	1.3542***
Observations	69,991			
No. restaurants	13,544			
R^2	0.3574		0.5874	

TABLE IV
EFFECTS OF GRADE CARDS AND DISCLOSURE REGULATION ON
ln(QUARTERLY RESTAURANT REVENUE)

	Coefficient	Std. error
Mandatory disclosure	0.0569	0.0153***
Voluntary disclosure	0.0326	0.0149**
B-grade	−0.0074	0.0084
C-grade	0.0039	0.0074
D-grade	−0.0023	0.0057
Mandatory × B-grade	−0.0497	0.0151***
Mandatory × C-grade	−0.0670	0.0304**
Mandatory × D-grade	−0.0565	0.0437
Voluntary × B-grade	−0.0029	0.0128
Voluntary × C-grade	−0.0238	0.0216
Voluntary × D-grade	−0.0758	0.0469
Missing grade	−0.0001	0.0096
Observations	74,321	
R^2	0.9506	

The regression also includes a restaurant fixed effects, a full set of quarterly dummies and city-level random effects (i.e., we cluster the standard errors by city with Huber-White standard errors).

D-Grade is equivalent to any score below 70 (i.e., less than a C-grade). Missing Grade is for restaurants that have opened but have not yet been inspected.

Excluded dummy is for voluntary disclosure without a standard format. Interactions with A-grade are also excluded.

The sample size is slightly reduced because we discard (i) observations for the first and last quarter when a restaurant is a new entrant or exitor, since we do not know the date of entry or exit; (ii) observations with negative tax, and hence negative revenue (due to overpayment of tax in a prior quarter); and (iii) restaurants with merged tax accounts (see text for a detailed explanation).

Stars denote significance levels: 99 percent confidence level (***), 95 percent confidence level (**), and 90 percent confidence level (*).

Looking at the effect of mandatory disclosure on hygiene score does not tell us much about their effect on things we care about.

- Effect on Revenue. It may be the case that consumers do not react to report cards at all. However, Jin and Leslie (2003) find that restaurants which increased their report card score also had an increase in revenues.

- Effect on Food Poisoning. The main reason that we might care about hygiene report cards is because consumers get sick because of unsanitary conditions at restaurants. Jin and Leslie (2003) find a very large effect of report cards on food poisoning (in the order of 20% less food poisoning due to the introduction of report cards).

TABLE V
NUMBER OF HOSPITAL ADMISSIONS IN CALIFORNIA FOR DIGESTIVE DISORDERS

	Los Angeles County				California, except Los Angeles County			
	Food-related		Nonfood-related		Food-related		Nonfood-related	
Year	Number	% Change	Number	% Change	Number	% Change	Number	% Change
1995	401		54,412		607		128,949	
1996	431	7.5%	56,692	4.2%	675	11.2%	131,623	2.1%
1997	405	−6.0%	59,585	5.1%	634	−6.1%	139,645	6.1%
1998	351	−13.3%	61,305	2.9%	654	3.2%	145,261	4.0%
1999	309	−12.0%	60,915	−0.6%	601	−8.1%	148,338	2.1%

TABLE VI
THE EFFECTS OF GRADE CARDS ON ln (NO. HOSPITALIZATIONS FOR DIGESTIVE DISORDERS)

	Coefficient	Std. error
Mandatory disclosure	0.0271	0.0246
Voluntary disclosure	0.0716	0.0238***
Food-related × mandatory disclosure	−0.2243	0.0426***
Food-related × voluntary disclosure	−0.2055	0.0350***
Observations	6,840	
R^2	0.9809	

Covariates not shown include fixed effects for food-related illnesses in each three-digit zip code, fixed effects for nonfood-related illnesses in each three-digit zip code, and year and month dummies. We also include three-digit zip code illness-type random-effects (i.e., we cluster the standard errors by three-digit zip code and illness-type with Huber-White standard errors).

Stars denote significance levels: 99 percent confidence level (***), 95 percent confidence level (**), and 90 percent confidence level (*).

Milyo and Walfogel

Milyo and Waldfogel (1999) exploit another "natural experiment" that occurred in Rhode Island. Before May 1996 Rhode Island had laws which prohibited advertising liquor prices, either in newspapers, television or shop storefronts. In May 1996 the Supreme Court struck down this law.

Previous work on the relationship between advertising and prices had focused on the cross-sectional relationships between laws permitting advertising and prices. For instance, Kwoka (1984)'s work on eyeglasses and advertising legislation showed that both prices for eyeglasses and price dispersion were significantly higher in states that banned price advertising. While this is consistent with a model of search and price dispersion, it is not clear if this pattern is caused by the lack of advertising or state characteristics which are correlated with higher prices. For instance, smaller and more rural states such as Wyoming or Kansas will have higher search costs due to the greater distance between eyeglass retailer, and may also be more likely to have banned advertising. So it is not clear what is causing this correlation.

Data Collection

- Milyo and Waldfogel (1999) looked at the Supreme Court docket (the list of cases which might be heard by the court) and found that the legislation banning advertising in Rhode Island might be overturned. This allowed them to collect data *before* and after the decision.

- Data is collected for prices in liquor stores in both Rhode Island and Massachussetts (which is used a control group since liquor advertising was permitted prior to 1996).

- Since it would be quite time consuming to collect data on all products in each liquor store, Milyo and Waldfogel select a sample of products for which to collect data. They take a limited sample of products since retailers typically do not allow them write down prices inside the store.

- Milyo and Waldfogel also collect data on advertising by liquor store, both in newspapers and in store fronts.

Results

Waldfogel and Milyo find little effect of lifting advertising restrictions on prices:

- There is no impact of permitting advertising on the average price of liquor. It is not clear to me, if this effect is not significant economically, or just not significant statistically.

- However the price of items that are advertised falls both at the store that advertised it and at stores which are located nearby.

TABLE 3—TEST OF CONTROLS

Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island	Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island
0.65 (0.52)	0.47 (0.63)	0.45 (0.64)	0.77 (0.46)
State-product and store fixed effects		Store-product fixed effects	
$F_{(2,6319)}$		$F_{(2,4047)}$	

Notes: These are test statistics of the hypotheses that, prior to the change in the law, Massachusetts and Rhode Island prices and markups move together. Regressions in columns 1 and 2 include separate product effects for each state, as well as store fixed effects. Regressions in columns 3 and 4 include store-product effects All regressions include 6,480 observations. Coefficients are in percentages. Probability values appear in parentheses.

TABLE 4—OVERALL EFFECT OF ADVERTISING ON PRICES

Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island	Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island
−0.51 (−1.15)	−0.73 (−1.58)	−0.39 (−1.02)	−0.80 (−1.94)
State-product and store fixed effects		Store-product fixed effects	

Notes: Coefficients are in percentages. *T*-statistics are in parentheses. Regressions in columns 1 and 2 include separate product effects for each state, time effects, and store fixed effects. Regressions in columns 3 and 4 include time effects and store-product effects. All regressions are based on 6,480 observations.

TABLE 5—EFFECT OF ADVERTISING ON PRICES, BY STORE TYPE

	State-product and store fixed effects		Store-product fixed effects	
	Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island	Log price in Massachusetts and Rhode Island	Markup in Massachusetts and Rhode Island
Nonadvertising Rhode Island store [1,328]	−0.15 (−0.38)	−0.56 (−1.37)	−0.26 (−0.58)	−0.48 (−1.03)
Nonadvertised product at an advertising Rhode Island store [124]	−0.19 (−0.23)	−0.41 (−0.48)	−0.13 (−0.14)	−0.28 (−0.29)
Own-advertised product at an advertising Rhode Island store [22]	−21.43 (−11.83)	−22.14 (−11.41)	−24.16 (−13.14)	−24.84 (−12.94)
H_0: Same coefficient for all nonadvertised products (Probability value)	0.00 (0.86)	0.00 (0.96)	0.02 (0.88)	0.05 (0.83)

Notes: Coefficients are in percentages. *T*-statistics are in parentheses. Number of price observations by category in brackets reported in heading column. Regressions in columns 1 and 2 include separate product effects for each state, time effects, and store fixed effects. Regressions in columns 3 and 4 include time effects and store-product effects. All regressions are based on 6,480 observations.

*References

Jin, G. Z., and P. Leslie (2003): “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *Quarterly Journal of Economics*, 118(2), 409–51.

Kwoka, John E., J. (1984): “Advertising and the Price and Quality of Optometric Services,” *American Economic Review*, 74(1), 6.

Milyo, J., and J. Waldfogel (1999): “The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart,” *American Economic Review*, 89(5), 1081–96.