Lecture Note 1 - Introduction and a First Application: The Minimum Wage Debate and Causal Inference in Economics

David Autor, MIT Department of Economics and NBER*

14.03/14.003 Microeconomic Theory and Public Policy, Fall 2022
(Compiled on September 7, 2022)

*Thanks to Professor Tobias Salz for his improvements to this lecture note.

1 Introduction to 14.03/14.003

This is an intermediate course in microeconomic theory and its application to real world policy problems. The class assumes you are familiar with economic theory at the 14.01 level. It's also helpful if you've taken some statistics or econometrics, though not required. For those who have not, the Teaching Assistants will provide a primer on some key statistical and mathematical concepts during the initial recitations. In addition, the handout Math Tools for 14.03/003 (on the class website), offers a primer on some mathematical concepts that you'll be using in lectures, problem sets, exams.

The class is organized around three themes:

1. Economic theory: Where does it come from, what does it predict, and in what ways is it useful?

2. Causality: What is it, and how do we measure or estimate causal effects?

3. Empirical applications: Economic theory is a way of organizing facts and interpreting and patterns in the world. This class will use data to test theory and use theory to interpret data. We will analyze a number of randomized experiments and quasi-experiments in the light of economic theory.

Definition 1. Randomized experiment (also frequently called a Randomized Controlled Trial or RCT). In the statistical theory of design of experiments, randomization involves randomly allocating the experimental units across the treatment groups. (Source: Wikipedia, `http://en.wikipedia.org/wiki/Randomized_experiment`).

Example: If an experiment compares a new drug against a proven existing drug, patients would be allocated to either the new drug or to the existing drug using randomization. A comparison of outcomes among patients allocated the new drug and those allocated the existing drug would provide an estimate of the causal effect of the new drug relative to the existing drug. (Note, it would not provide an estimate of the causal effect of the new drug relative to no treatment unless there was also a placebo group in the experiment.)

Definition 2. Quasi-experiment. An event that unintentionally creates conditions similar to a randomized experiment.

Example: One million people buy one lottery ticket each and one hundred of them win. This quasi-experiment could be used to evaluate the effect of unanticipated wealth increases on happiness, health, marital dissolution, obesity.

There is an impressive diversity of experiments and quasi-experiments that economists have applied to analyze important questions in social science. You may ask, why would a researcher use a quasi-experiment instead of an RCT? In reality, RCTs are the gold standard of evidence, and economists use them wherever possible. Indeed, we will study multiple RCTs during the semester.

However, many key economic questions center either around major life choices and outcomes such as health, wealth, education, and risk, or 'macro-scale' treatments such as international trade, civil war, or epidemic disease. For both ethical and practical reasons, these outcomes are often not suitable for randomized experimentation. In these cases, we look for chance events in the real world that approximate the experiment we would conduct if it were ethically or practically feasible.

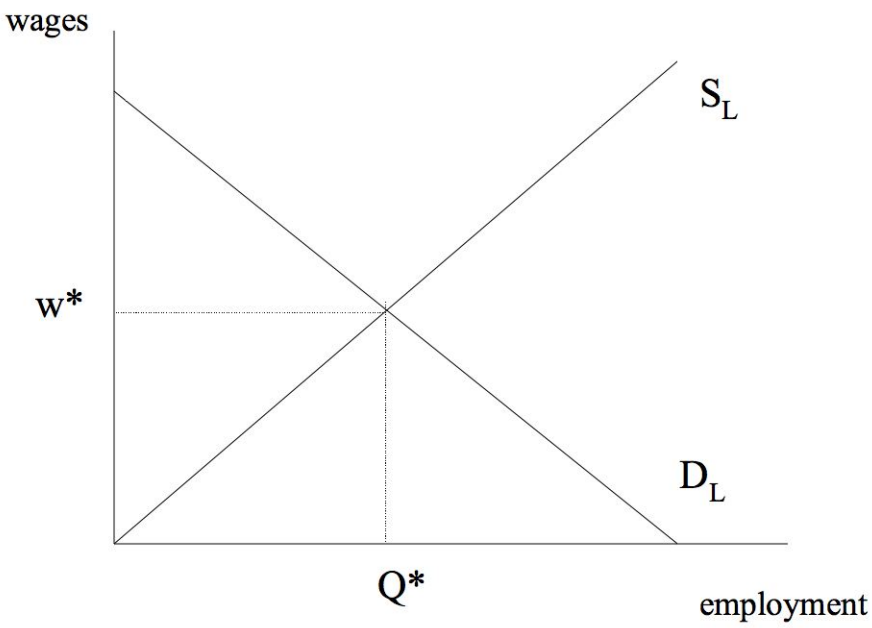
2 A first example: The minimum wage and employment

Rather than start the class with a discussion of economic methodology, we'll start with an application and return to the big picture when that is complete (sometime in the second lecture). The application we'll consider is the impact of the minimum wage on wages and employment. The costs and benefits of legislated minimum wages are a venerable topic in economics—and area of ongoing controversy and active policymaking (e.g., President Obama tried to raise the national minimum wage from $7.25 to $10.10 per hour, a substantial increase. President Trump's top economic advisor called a higher federal minimum wage "a terrible idea" that would "damage" small businesses.[1] In general, the Republican party vehemently opposes minimum wage regulations. Arkansas and Missouri had successful ballot measures in November 2018 to raise the minimum wage, and in recent years, several cities have enacted minimum wage regulations). Here's a mini-outline of what we'll cover on this topic in the first two lectures:

1. Textbook model of competitive labor market
2. Impact of minimum wage on employment in the textbook model
3. Assumptions behind this model
4. What happens when we relax a key assumption: price-taking by firms
5. Impact of minimum wage on employment when employers have market power
6. Testing the textbook model and alternatives
7. Natural experiments in economics
8. The Fundamental Problem of Causal Inference
9. Workarounds for the Fundamental Problem
10. Notation for causal inference
11. Estimating causal effects using "Differences-in-Differences" (DD)
12. The Card and Krueger minimum wage study

[1] Source: Washington Post, November 1, 2018. "President Trump's top economic adviser: A federal minimum wage is a terrible idea. A terrible idea."

3 Textbook model of wages and employment



Definition 3. Labor supply curve. All potential workers in the labor market, arranged according to their "reservation wage," which is the lowest wage at which they will accept to take a job (from low to high)

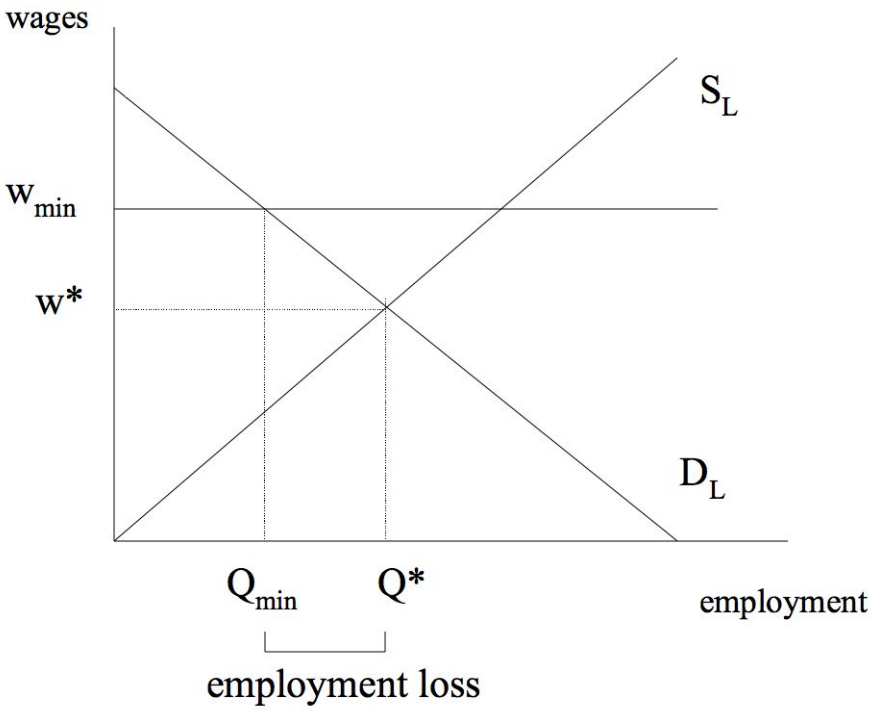
Definition 4. Labor demand curve. All potential employers in labor market, arranged according to their willingness to pay for a worker (from high to low)

- What is the key outcome variable in this model: the wage or the number of employed workers? *Neither*. They are simultaneously determined. Another way to say this is that these outcomes are *endogenous*.
- In the example above, the demand and supply curves are *exogenous*. The equilibrium wage and employment levels are *endogenous*.

Definition 5. Endogenous. Internally determined. An effect rather than a cause.

Definition 6. Exogenous. Externally determined. A causing or forcing variable.

What happens when we impose a minimum wage in this labor market?



- Wages:

$$w_{\min} > w^*$$

  Employment:

$$Q_{\min} < Q^*$$

- If this model is right why would a policymaker ever want to impose a minimum wage?
- One possible answer: Total earnings

$$w_{\min} Q_{\min} \gtrless w^* Q^*$$

  Total worker earnings may increase even if employment falls.

- What does this depend on? The *elasticity* of demand:

$$\eta = \frac{\partial Q}{Q} \frac{w}{\partial w} \gtrless -1$$

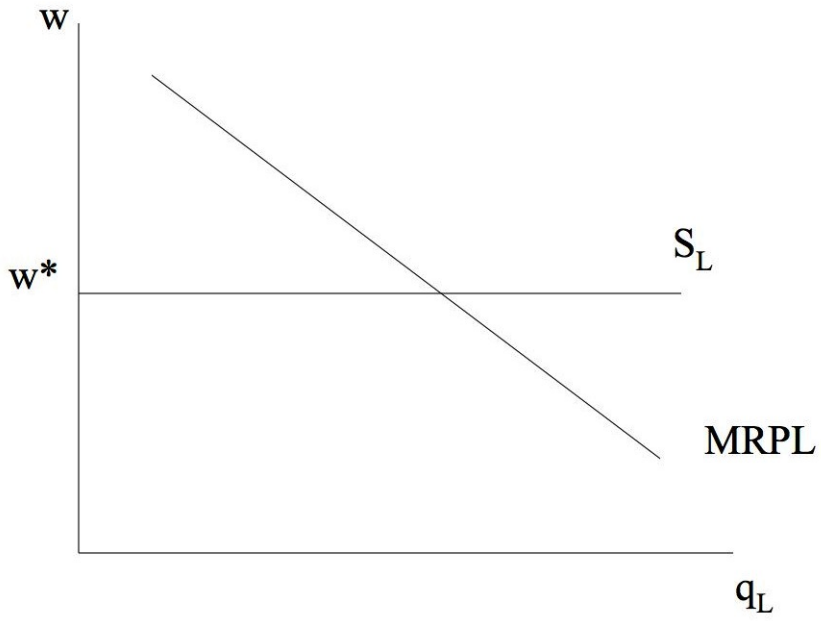
  If the proportional increase in wages is larger than the (induced) proportional decline in employment $\Longrightarrow$ wage-bill increases. Specifically, if $\sigma > -1$ (i.e., $|\sigma| < 1$), then a 1% rise in wages reduces employment by less than 1%, so total wages paid (wages $\times$ workers) rises.

Definition 7. Elasticity. The ratio of the proportional change in a variable Y caused by a given proportional change in a variable X. So for example, if the elasticity of Y with respect to X is 2, then a 1% increase in X causes a 2% increase in Y.

3.1 Why do minimum wages reduce employment? Revisiting the theory

- What is the primary assumption in the textbook model that yields the prediction that (binding) minimum wages always and everywhere reduce employment?
- The answer is *price-taking* behavior, both in labor and product markets. That is, the price of the good the firm is producing does not fall if the firm makes a few more, and the prevailing wage the firm faces does not rise if it hires a few more workers. Formally, product demand and labor supply are both *perfectly elastic* as the far as the firm is concerned.

Individual "price-taking" firm



- MRPL = Marginal Revenue Product of Labor $\Longrightarrow$ "What the marginal worker produces".
- We normally assume that at any given firm, MRPL is decreasing in employment due to decreasing returns in the production function. All else equal, the next worker produces marginally less than the prior hire. This could be because the most important tasks are always done first; hence, adding more workers means that some less important tasks are also accomplished.
- You learned in 14.01 that the firm equates the Marginal Revenue Product with the wage:

$$MRPL = w^*.$$

  Where did that come from?

- Recall the firm's profit maximization problem, which is to maximize the difference between revenues and costs (i.e, profits). Assume that the firm's only input is labor. Denote the first

derivative of a function $f(\cdot)$ by $f'(\cdot)$ and the second by $f''(\cdot)$. The firm's problem is:

$$\max \ \pi = p \cdot f(L) - w(L) \cdot L,$$

where p is the product price, $w(L)$ is the wage necessary to "call forth" L workers, and $f(L)$ is the amount of output produced. We assume that $f'(\cdot) > 0$ and $f''(\cdot) < 0$, so an additional worker always raises output, but marginal productivity declines as we add workers. Note that p is not a function of L, meaning we assume that the price of output is taken as given (it's exogenous).

- Differentiate this expression with respect to L and set it equal to zero. (Why zero? At the optimum, this derivative must equal zero. If not, the firm would want to adjust L further. If the marginal profit were positive, the firm would want to hire more labor. If the marginal profit were negative, the firm would want to hire less labor.)

$$\frac{\partial \pi}{\partial L} = p \cdot \frac{\partial f(L)}{\partial L} - w(L) - \frac{\partial w(L)}{\partial L} \cdot L = 0$$

Rearranging:

$$pf'(L) = w(L) + w'(L)L$$

where:

  - $pf'(L)$ is the marginal revenue product of labor ($MRPL$)
  - $w(L)$ is the equilibrium wage
  - $w'(L)L$ is an additional change in *total labor costs* caused by hiring an extra worker. It is equal to the product of the firm's *entire* work force and the marginal wage increase

- This third term is potentially important. It says that each additional worker hired (each "marginal" worker) could potentially raise the cost of all of the previous workers hired ("infra-marginal" workers). Why? If all workers are paid a single wage ($w(L)$), and calling forth an additional worker raises that wage, then the cost of the additional worker is not simply w but $w + w'(L)L$.

- So the key assumption of the *competitive* model is:

$$w'(L) = 0 \Longleftrightarrow \text{Price taking firms}$$

No firm is large enough to raise the market wage simply by hiring a few more workers.

- If the firm is a price taker in the labor market, it chooses employment so that:

$$pf'(L) = w^*,$$

where w^* is the market wage, which the firm takes as given.

- How does firm choose employment when it is *not* price taker? According to the FOC above:

$$pf'(L) = w(L) + w'(L)L$$

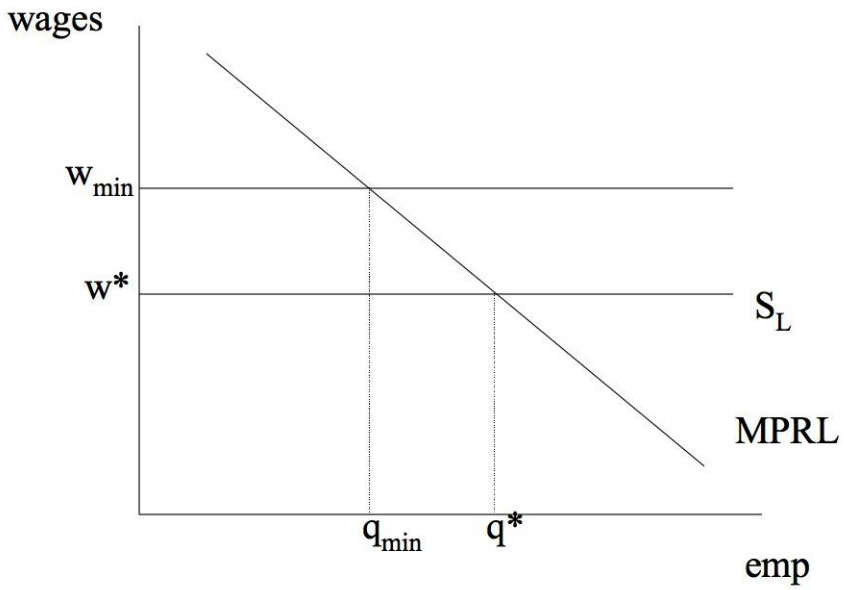
If $w'(L) \neq 0$ then firms must pay *all* of its workers higher wages with each additional worker it hires.

- Here's one convenient way to express this result

$$\begin{aligned} pf'(L) &= w(L) + w'(L)L \\ MRPL &= w + \frac{\partial w}{\partial L} L \\ &= w\left(1 + \frac{\partial w}{\partial L}\frac{L}{w}\right) \\ &= w\left(1 + \frac{1}{\eta}\right) \end{aligned}$$

where η is the elasticity of labor supply (the percent change in labor supply for a 1 percent change in the wage) as experienced by the single firm. For a price-taking firm, $\eta \to \infty$, meaning that $1/\eta \to 0$. So, if a firm is a price taker, the wage is exactly equal to MRPL (since the denominator of the above expression is equal to one). If the firm is not a price taker in the labor market, then the wage it pays is *strictly less* than MRPL.

3.2 Conventional case: Individual Price Taking Firm



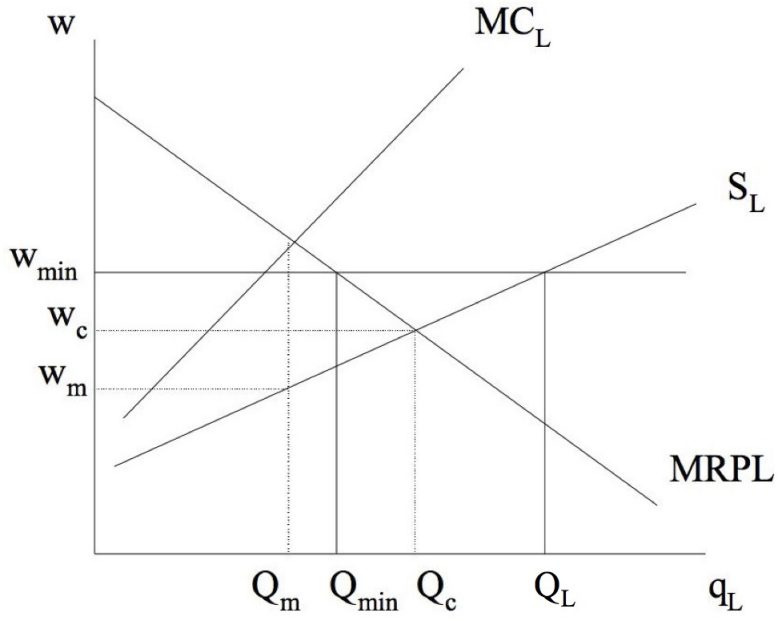
- Notice that the labor supply curve is upward-sloping at the market level, but it is flat as perceived by the single firm.

- If we imposed a binding minimum wage in this market ($w_{\min} > w^*$), each firm in this market would reduce its quantity of workers demanded.

3.3 Monopsonistic employer

Definition 8. Monopsony. "One buyer, many sellers." More generally, monopsony is a case where an agent (firm or consumer) is not a price taker in a market in which it is a buyer. Its own demand affects the price it pays. (Conversely, monopoly is a case where a firm is not a price taker in a market in which it is a seller. Its own supply affects the price it commands in the market.)

- The labor supply curve for a monopsonist is upward sloping. To obtain one more worker, the monopsonist must raise the wage by a small amount.
- Assuming that all workers receive the same pay (i.e., the late-comers don't get paid more), the marginal cost of the next worker is not simply her wage but also the wage increase given to all of the other inframarginal workers.
- Thus, the marginal cost of labor curve for a monopsonistic firm is *even more* upward sloping than its labor supply curve. The additional cost for each worker is given by the higher wage of that worker and by the increase in wage given to the entire pool of workers.



- What happens if we impose a binding minimum wage on a monopsonistic employer?
- One case is illustrated above. In this example, implementation of a binding minimum wage raises wages *and* employment:

$$w_{\min} > w_m$$
$$Q_{\min} > Q_m$$

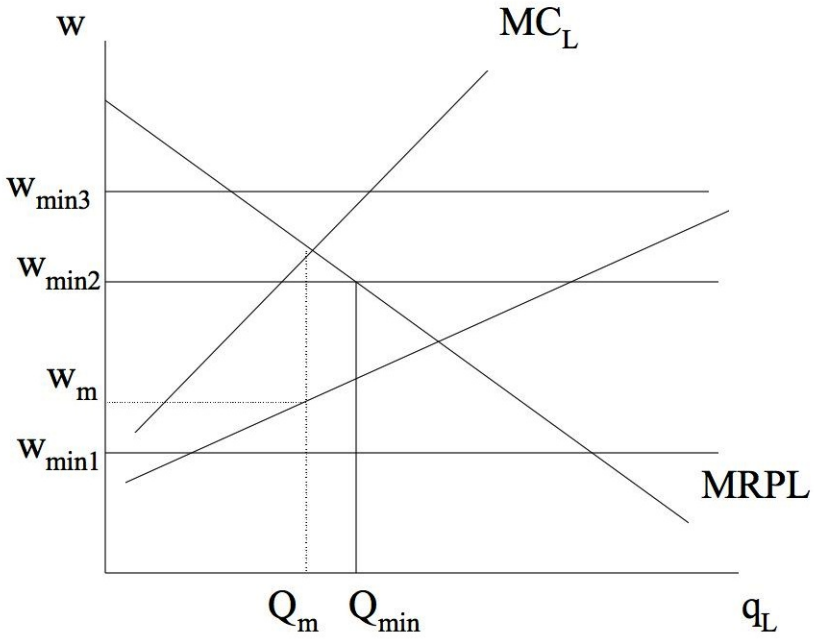
- Why does that happen? The firm is now a *price-taker* for labor at $w_{\min}$. That is, it is required by law to pay at least $w_{\min}$, although it would like to set a lower wage. At this wage, the firms only want to hire $Q_{\min}$ workers since if it hires more, then the MRPL (the amount a marginal worker produces) exceeds the wage. Therefore the firms chooses $Q_{\min}$ so that:

$$w = MRPL.$$

  Note that the number of workers that would like to be employed is larger than $Q_{\min}$ – it is given by Q_L. So, although more workers try to get employed, the monopsonist only gives a job to fewer.

- Thus, paradoxically, raising the minimum wage can raise both wages and employment in a monopsonistic labor market.

- Does raising minimum wage to monopsonists *always* increase wages and employment? Definitely not.



1. $w_{\min 1}$ - Introduction of minimum wage $w_{\min 1}$ has no effect because the minimum wage is below w_m and hence doesn't bind
2. $w_{\min 2}$ - Introduction of minimum wage $w_{\min 2}$ raises wages and employment

3. $w_{\min 3}$ - Introduction of minimum wage $w_{\min 3}$ raises wages but reduces employment

- So, this bit of simple theory presents an interesting possibility. It is conceivable—though not necessarily likely—that a mandated minimum wage could raise both earnings and employment. If so, this is a policy that many policymakers would support (though not most businesses; this is primarily, though not exclusively, a redistribution of wealth from firms to workers).

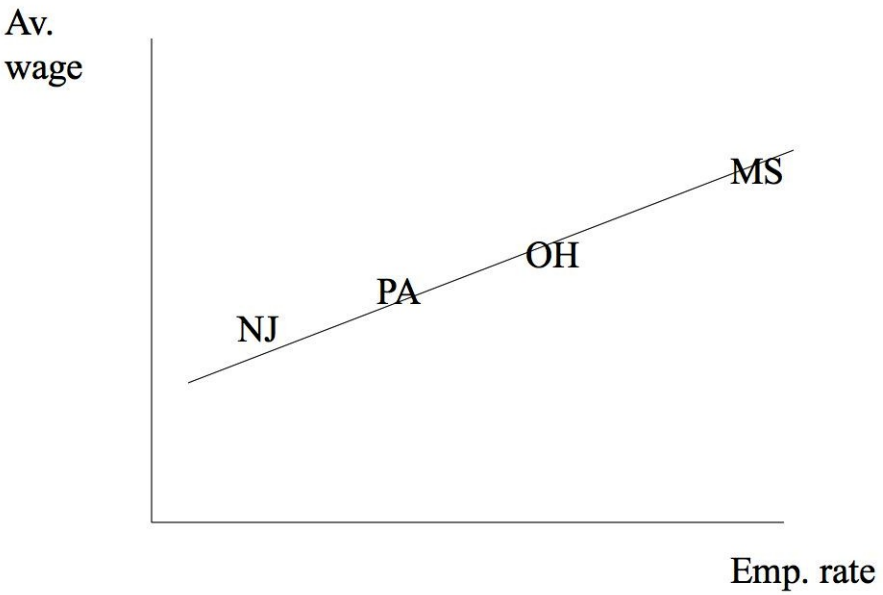
4 Testing for monopsony in the labor market

If monopsony were present in the labor market, where would you expect to find it? (Remember the criterion: the firm's own labor demand changes the market wage.)

- "Company towns" such as old mining towns, where the mining company was the only employer
- Cases where skills are very specific, e.g. Tesla automobile technicians
- "Captive" labor markets, e.g., spouses of soldiers at rural military bases or in remote island locations.
- How about Fast food restaurants located in nearby towns in New Jersey and Pennsylvania?

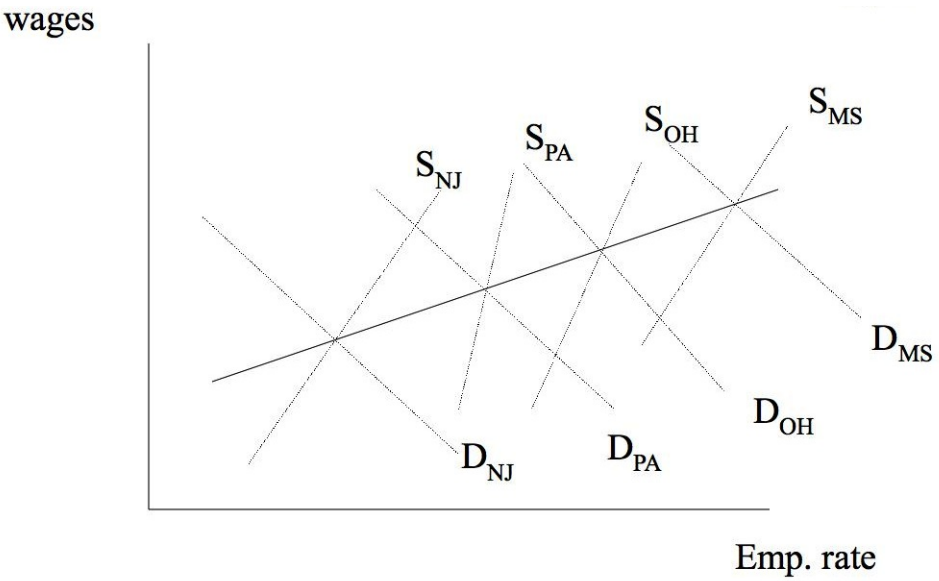
4.1 Testing for monopsony in the labor market

- How do we go about testing the monopsony vs competitive model of the labor market?
- *Focus on the key empirical implication that distinguishes these models:*
  - In the competitive model, an increase in the minimum wage always reduces employment: $W \uparrow \rightarrow L \downarrow$
  - In the monopsonistic model, an increase in the minimum wage (may) raise employment: $W \uparrow \rightarrow L \uparrow$
- How do you test this implication?
- We could look across different states and ask ourselves the following question: Is employment higher in states where wages are higher?
- Let's suppose you find the following pattern:



Would this convince you that higher wage levels *caused* higher employment? I hope not!

- What's the problem with the wage here? We don't know why the wage differs across states. For example, there might be different demand and supply schedules in each state (after all, we don't think the labor market for fast food workers is a nationwide one; presumably, people won't move cross-country for a McDonalds job).



Since both employment and wages are endogenous outcomes—determined by both supply and demand—this picture tells us essentially nothing about the impact of minimum wages on employment.

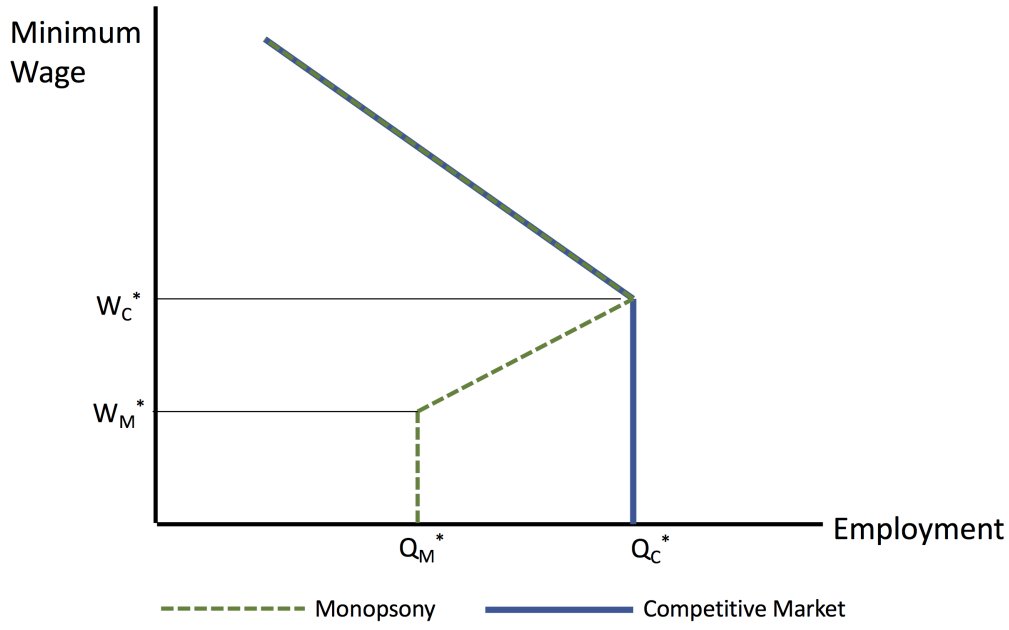
- A further problem: While we may believe in the existence of supply and demand curves as an outcome of market processes, we do not ever see these curves. What we *observe* is the equilibrium wage level and the quantity employed. Thus, we cannot directly see whether or not individual firms face upward sloping labor supply (as would occur if they had monopsony power).

- How do we overcome this problem? We need an experiment, specifically, one in which wages are raised exogenously. (Why not an experiment that shifts supply inward so that fewer workers are willing to work at any given wage? Consider the empirical implications in both the monopsonistic and competitive cases.)

- If we could exogenously manipulate the minimum wage, we could study its impact on employment to infer the slope of the relationship between wages and employment (downward sloping $\rightarrow$ competitive market, upward sloping $\rightarrow$ monopsony)

4.2 An idea: New Jersey's 1993 minimum wage change

- The discussion above shows that the relationship between the minimum wage and employment may be different for the monopsonistic market vs the perfect competition market. The figure below illustrates the different relationships.



- Notice that the conditions under which the introduction of a minimum wage raises employment in a monopsonistic market are only locally satisfied—that is, raising the minimum wage by "too much" will reduce employment even in the monopsonistic setting. Thus, if we find that an increase in the minimum wage raises employment, this is *sufficient* but not *necessary* to establish the existence of monopsony.

- By studying the pre/post change in employment following the adoption of the minimum wage, we can explore whether employment rises (monopsony) or falls (competitive market).
- Before we explore this relationship empirically, we need to take a moment (okay, a half hour) to discuss causal inference.

5 Causal inference in social science

In addition to this lecture note, I encourage you to read the 1986 JASA article "Statistics and Causal Inference" by Paul Holland to gain a deeper understanding of this topic.

- Much of social science (psychology, anthropology, sociology, political science, epidemiology and some subset of economics) concerns analyzing correlations among variables, e.g., the correlation between education and income, the correlation between obesity and heart disease, the correlation between happiness and longevity.
- Correlation describes the statistical relationship between two observed variables. Correlation has no necessary relationship with cause and effect. You can measure the correlation between happiness and longevity with great precision and yet know nothing about whether making someone happier will improve their longevity. (Let's say that happier people live longer. Perhaps these people are happy precisely because they feel fortunate to have lived so long.)
- In 14.03/14.003, we are not primarily interested in these correlations. Science advances primarily through analyzing cause and effect relationships, not by documenting correlations (though correlations are not irrelevant).
- Causal questions:
  - What is the causal effect of education on income?
  - What is the causal effect of obesity on heart disease?
  - What is the causal effect of happiness on longevity?
  - What is the causal effect of the minimum wage on employment?
- Causal questions are much harder to answer than the correlational questions. Here's why: *Correlations are readily measured from observational data. Causal effects can never be measured directly (to be explained).*
- The reason that causal questions are difficult to answer is that they intrinsically concern a *counterfactual* state—that is, something that by definition has not or did not occur.
- To be concrete, let's examine a simple case where we have a binary causal variable $X \in \{0, 1\}$ and a response variable Y (which may be continuous).

- If X has a causal effect on Y, this implies that the value of Y we observe depends on the value of X.
- We can define Y_0 and Y_1 as the values that Y would take if X were equal to 0 and 1 respectively.
- Y_0 and Y_1 are counterfactuals of one another. If $X = 0$ then we will have $Y = Y_0$. The counterfactual question is what value of Y would we have observed if instead we had $X = 1$. Obviously, the answer is Y_1. But Y_1 is a notional concept. We can posit that both Y_0 and Y_1 have well defined values, but we will never see both of them.

5.1 The fundamental problem of causal inference

- Let Y_i be the outcome of interest for unit i, where i could be a person, a cell, a drop of water, a sovereign country. We'll suppress the i subscript where possible.
- We want to consider two possible outcomes for i. Let Y_0 be the value of Y for $X = 0$ and Y_1 be the value of Y for $X = 1$, as above. Thus, for every unit i, we can imagine two potential outcomes $\{Y_0, Y_1\}$ that we would observe if the unit were untreated $(X = 0)$ or treated $(X = 1)$.
- We observe either Y_0 or Y_1, but we assume that both are well defined. That is, there is a precise alternative state of the world that would have occurred had we chosen $X = 1$ instead of $X = 0$ or vice versa.
- In this framework, the causal effect of X on Y is $T = Y_1 - Y_0$, where T stands for Treatment Effect.
- The problem that this immediately reveals is that *we never observe* $Y_1 - Y_0$ for an individual i. Instead, we observe

$$Y_i = Y_{1i}X_i + Y_{0i}\left(1 - X_i\right).$$

  That is, we observe Y_1 or Y_0 but not both.

Definition 9. Fundamental Problem of Causal Inference. It is not possible to observe the value Y_{1i} and Y_{0i} for the same unit i, so we cannot measure the causal effect of X on Y for unit i.

- Natural question: Why can't we just switch X from 0 to 1 and back to observe both Y_1 and Y_0? In fact, this procedure is not informative about $Y_1 - Y_0$ without further assumptions (discussed below).
- One useful observation to build intuition: many causal relationships are irreversible. If X corresponds to attending MIT vs. another college and Y corresponds to post-college earnings, we can either observe your post-MIT earnings or your post-non-MIT earnings, not both.

5.2 Solving the fundamental problem of causal inference

Since the problem is fundamental, there is no solution. But there are several "work-arounds."

Work-around I: Postulate stability and reversibility (AKA 'causal transience')

- One work around is to assume stability and reversibility (what Holland calls *temporal stability* and *causal transience*).
- If the causal effect of X on Y is the same at every point in time (now and the future) *and* the causal effect of X on Y is reversible (so having once been exposed X doesn't permanently change the effect of X on Y), then we can observe $Y_{1i} - Y_{0i}$ simply by repeatedly changing X from 0 to 1.
- Formally, these assumptions are: $Y_{1it} = Y_{1i}$, $Y_{0it} = Y_{0i}$ for all i and t where t indexes time.
- Of course, temporal stability and causal transience are *postulates*. They cannot be tested
- Example: You can turn water from ice to steam and back repeatedly to analyze the causal effect of temperature change on water molecules. But what allows you to make the *causal inference* that steam is the counterfactual for ice when the treatment is 100 degrees versus 0 degrees Celsius are the postulates that (1) water molecules are not fundamentally altered by heating and cooling; and (2) that the relationship between temperature and the behavior of water is stable (e.g., does not depend on the phase of the moon).
- Counter-example: It would probably not be valid to assess the effectiveness of a treatment for high cholesterol for patient i by repeatedly administering the cholesterol reducing treatment, testing the patient's cholesterol level, then withdrawing the treatment, testing the patient's cholesterol level, etc. Cholesterol levels are sluggish state variables. And they might be permanently affected by even a one-time treatment.

Work-around II: Postulate homogeneity

- We may alternatively assume *unit homogeneity*. If Y_{1i} and Y_{0i} are identical for all i, we can measure the causal effect of X on Y simply by taking the difference $Y_{1i} - Y_{0j}$ for $i \neq j$.
- Of course, unit homogeneity is also a postulate; one cannot know that two things are identical in all respects.
- But under certain laboratory conditions, unit homogeneity seems quite reasonable (e.g., experimenting with two molecules of water). This assumption would clearly be invalid for two cholesterol patients. Or for any two people more generally.

Work-around III: Estimate causal effects for populations rather than individuals

- For human subjects, neither (1) temporal stability and causal transience nor (2) unit homogeneity can plausibly be expected to hold in any setting. No two people are alike. And no one person is identical to him or herself at a different point in time.

- *We should therefore acknowledge that we will never be able to credibly estimate* $T_i = Y_{1i} - Y_{0i}$ for a person i.

- We might, however, be satisfied to settle for some kind of population average treatment effect instead:

$$T^* = E\left[Y_1 - Y_0 | X = 1\right],$$

  where $E[\cdot]$ is the *expectations* operator, denoting the mean of a random variable. This expression above defines the Average Treatment Effect for the Treated (ATT), that is the causal effect of the treatment on the people who received the treatment (i.e., for whom $X = 1$).

- The ATT should be distinguished from the Average Treatment Effect (ATE), defined as

$$T^\dagger = E\left[Y_1 - Y_0\right].$$

  The difference between T^* and $T^\dagger$ is this: ATT measures the causal effect only for those who receive treatment whereas the ATE is the causal effect one would notionally obtain if *everyone* were treated. These can be quite different. The ATT for a cholesterol lowering drug given to morbidly obese patients is probably not comparable to the ATE for a cholesterol lowering drug given to the entire population of adults.

- Returning to our discussion of ATT, how do we estimate this quantity? Since we cannot directly observe T for any given individual i, how do we measure $E[Y_1 - Y_0|X = 1]$ for some population of $i's$?

- One idea: We could compare $E[Y|X = 1]$ and $E[Y|X = 0]$ to form $\tilde{T} = E[Y|X = 1] - E[Y|X = 0]$. For example, let X be the cholesterol treatment and Y be a measure of serum cholesterol level. We could compare cholesterol levels among those taking the treatment ($E[Y|X = 1]$) versus those not taking the treatment ($E[Y|X = 0]$) to estimate the causal effect of the treatment on cholesterol levels. Is this a good idea?

- A moment's thought should suggest that $\tilde{T}$ is *not* a good estimator for T^*. The problem is that people who take the cholesterol treatment are likely to have abnormally high cholesterol whereas those who do not take the treatment are likely to have normal cholesterol levels. Thus, *even if* the treatment lowered cholesterol, we might erroneously conclude the opposite because our comparison group ($X = 0$) had low cholesterol to begin with whereas our treatment group ($X = 1$) had abnormally high cholesterol—and may still have above average cholesterol, even if the treatment lowered it somewhat.

- So, if $\tilde{T}$ is not a good measure of T^*, what would a good comparison look like? We need to find treatment and control populations that have the same expected levels of cholesterol *but for* the treatment. Formally, we want to identify a set of people for whom the *counterfactual* outcomes

are comparable between the treatment and comparison (AKA control) groups. Specifically:

$$E[Y_1|X=1] = E[Y_1|X=0] \tag{1}$$
$$E[Y_0|X=1] = E[Y_0|X=0].$$

These equalities imply that the treatment and control groups are 'exchangeable.' If we swapped treatment and control groups prior to the experiment, we'd estimate the same treatment effect as we'd get from the initial assignment.

- If these conditions are satisfied, then it's straightforward to see that a contrast of the outcomes of the treatment and control groups will provide a valid estimate of the causal effect of treatment for the treated group. Specifically,

$$\begin{aligned} E[Y_1|X=1] - E[Y_0|X=0] &= E[Y_1|X=1] - E[Y_0|X=1] \\ &= T^* \end{aligned}$$

Notice that our substitution above of $E[Y_0|X=1]$ for $E[Y_0|X=0]$ is justified by the assumption of treatment-control balance in equation (1). If the subjects who didn't receive the treatment are just like those who did *but for* not having received the treatment, then the contrast between the treated and untreated groups provides an unbiased estimate of the causal effect of the treatment on the treated group (ATT).

- Returning to our *invalid* estimator $\tilde{T}$, let's ask how likely is it that the counterfactual outcomes would be balanced among a set of people selected from the population according to whether or not they are currently receiving the treatment. It does not take a great leap of logic to hypothesize that the patients receiving the cholesterol drug are more likely to be suffering from high cholesterol than those who are not taking the drug. This would imply that:

$$E[Y_1|X=1] > E[Y_1|X=0]$$
$$E[Y_0|X=1] > E[Y_0|X=0].$$

In words, patients receiving the drug are more likely to suffer from high cholesterol *whether or not* they are receiving the drug (presumably, they are receiving the drug specifically because they were diagnosed with high cholesterol).

- So, if we calculated the contrast $\tilde{T} = E[Y|X=1] - E[Y|X=0]$ for this unbalanced group, what would we get?

$$\begin{aligned} E[Y_1|X=1] - E[Y_0|X=0] = &\underbrace{E[Y_1|X=1] - E[Y_0|X=1]}_{T^*} \\ &+ \underbrace{\{E[Y_0|X=1] - E[Y_0|X=0]\}}_{Bias}. \end{aligned} \tag{2}$$

The first term on the right-hand side of this equation is the true, causal effect of the cholesterol treatment on those who take it (the ATT). The second term is the potential bias that occurs if the counterfactual (non-treated) outcomes of the comparison group (those not taking the treatment) differ from the counterfactual (non-treated) outcomes of the treatment group (those taking treatment) if they were untreated.

- We've just argued above that $E[Y_0|X=1] > E[Y_0|X=0]$. Thus, the bias in this case is positive—that is, it goes in the direction of generating an estimate of $\tilde{T}$ that is *larger* than the true casual effect T^* (so, $E\left[\tilde{T}\right] > E[T^*]$). Even if T^* were hypothetically negative (that is, the drug reduces cholesterol), we could easily conclude through this naive comparison that the drug increases cholesterol levels.

Another example

- Let's say Y is the number of mathematical expressions you can differentiate in an hour after 4 years of college and MIT is an indicator variable for whether or not you attended MIT.
- If we administered math tests at random to students at Boston area colleges, we would certainly find that $\tilde{T} = E[Y_1|MIT=1] - E[Y_0|MIT=0] > 0$, i.e., MIT students can solve more calculus problems in an hour than non-MIT students.
- But $\tilde{T}$ a not a valid estimate of the causal effect of attending MIT on calculus skills (that is, $\tilde{T} \neq E[Y_1 - Y_0|\text{MIT}=1]$). Students who are skilled in calculus choose MIT, and they would be more skilled than the average student in calculus, regardless of whether they attended MIT. So, $E[Y_0|MIT=1] > E[Y_0|MIT=0]$.
- The substantive problem (again) is that the "treatment," MIT attendance, is endogenous. Students come to MIT in part because they are good at math. It is unwise to assume that non-MIT students are a valid comparison group for MIT students.

5.3 Implementing the statistical solution using randomization

- Let's say that we picked a large number of $i's$ at random and randomly assigned half to $MIT=1$ and half to $MIT=0$.
- This pretty much guarantees (unless we are very unlucky) that

$$E[Y_1|MIT=1] = E[Y_1|MIT=0] \text{ and } E[Y_0|MIT=1] = E[Y_0|MIT=0].$$

So, condition (1) should be satisfied.

- Plugging back into (2):

$$\hat{T} = E\left[Y_1|MIT=1\right] - E\left[Y_0|MIT=0\right] \;\; = \\ E\left[Y_1|MIT=1\right] - E\left[Y_0|MIT=1\right] + \underbrace{\left\{E\left[Y_0|MIT=1\right] - E\left(Y_0|MIT=0\right)\right\}}_{bias\,=0}.$$

- Thus, randomization has eliminated the bias term (in expectation) by balancing the counterfactual outcomes between the treatment and control groups. Specifically, the students assigned to MIT would have been expected to fare comparably to the students who were *not* assigned to MIT had these students instead been assigned to MIT. What randomization has bought us specifically is this:

$$\left\{E\left[Y_0|MIT=1\right] - E\left[Y_0|MIT=0\right]\right\} = 0.$$

- In summary, randomization overcomes the causal inference problem by making the treatment status $X = \{0, 1\}$ independent of potential outcomes: $Y_1, Y_0 \perp X$ so

$$E\left[Y_1|X=1\right] = E\left[Y_1|X=0\right] \quad \text{and} \quad E\left[Y_0|X=1\right] = E\left[Y_0|X=0\right]$$

- This observation motivates the idea of using a randomly selected control group to ensure that the group not receiving the treatment provides a valid estimate of the *counterfactual* outcome for the treated group.

- Bottom line:

  - It is rarely plausible for human behavior that either of the two alternative solutions will be plausible (temporal stability + causal transience or unit homogeneity). By contrast, so long as we can randomize, the statistical solution is likely to work.
  - Needless to say, to solve the Fundamental Problem of Causal Inference in economics, we always use the statistical solution. That is, we never assume that humans are homogeneous or temporally stable. In some cases, we use randomized experiments. Where randomized experiments are infeasible or impractical, we use quasi-experiments.

6 Difference-in-Difference Estimation

- Often, we don't simply measure the level of Y but it's change as a function of X (the treatment) *and* time. For example, if we have a treatment and control group, we can form:

	Before	After	Change	
Treatment	Y_{jb}	Y_{ja}	ΔY_j	where b stands for before and a stands for after
Control	Y_{kb}	Y_{ka}	ΔY_k	

- Why do we want to make a pre-post comparison?

- We actually do not need to do this if we have a very large population of (randomly assigned) treatment and control units to work with. In that case, we could simply calculate

$$\hat{T} = E\left[Y|X=1\right] - E\left[Y|X=0\right] = E\left[Y_1 - Y_0|X=1\right].$$

  If X is randomly assigned and the population of treated units is large, then the conditions in equation (1) should apply and hence the *cross-sectional* (as opposed to over-time) comparison should provide a valid estimate of the causal effect of interest.

- However, we often don't have very large samples of treatment and control individuals to work with.

- Let's say we are assessing the effect of a new drug treatment on cholesterol levels. We could pick 10 people each for the treatment and control groups, give the treatment group the drug treatment and the control group the placebo, and then compare the average cholesterol level between these two groups.

- There is nothing wrong with this approach. But we might be concerned that, just by chance, these two groups started out with somewhat different cholesterol levels.

- Because of this concern, we could also take baseline data (prior to treatment) to ensure that these groups were comparable.

- Let's say the baseline averages were comparable but not identical; by chance, the treatment group had a slightly lower cholesterol level than the treatment group. We'd be concerned that our experiment would be biased in favor of the finding that the treatment lowered cholesterol (since the treatment group started with a better outcome).

- It's that concern that motivates us to compare the *change* in cholesterol in the treatment group to the change in cholesterol in the control group. By studying the change in the outcome variable, we subtract off initial differences in levels that could potentially prove confounding in small samples. Thus, we focus on the *improvement* (or change) in the treatment group *relative to* the control group.

- Formally, let's say that prior to treatment, we observe:

$$Y_{jb} = \alpha_j.$$
$$Y_{kb} = \alpha_k.$$

  We would hope that $\alpha_j \simeq \alpha_k$, but this does not strictly have to be the case.

- Now, imagine that after treatment, we observe

$$Y_{ja} = \alpha_j + \delta + T,$$

where T is the causal effect and δ is any effect of time. For example, cholesterol levels may tend to rise over time as people age.

- So, if we take the first difference for Y_j, we get:

$$\Delta Y_j = Y_{ja} - Y_{jb} = (\alpha_j - \alpha_j) + \delta_j + T$$

  This does not recover T. But it does remove the "level effect" α_j.

- Similarly, $\Delta Y_k = (\alpha_k - \alpha_k) + \delta_k$. Differencing removes the level effect for group j.

- If we are willing to postulate that the time effect operates identically on the treatment and control groups, $\delta_j = \delta_k = \delta$, then we have

$$\Delta Y_j - \Delta Y_k = T + \delta - \delta = T.$$

- So, the difference-in-difference estimator allows us to potentially recover the causal effect of treatment even when the treatment and control groups are not entirely identical and when there is a potentially confounding effect of time.

7 Back to Jersey

- Let $Y_{n,t}$ be the level of employment in New Jersey before ($t = 1992$) or after ($t = 1993$) the introduction of the minimum wage. The minimum wage hike occurs in 1993, so we will be comparing outcomes in 1992 to those in 1993/94.

- If we want to estimate the causal effect of the minimum wage hike on New Jersey employment, we could calculate:

$$\hat{T} = Y_{n,1992} - Y_{n,1993},$$

  which is simply the before/after change in New Jersey employment.

- What are the potential weaknesses of this estimate of the causal effect? One is that it requires assumption of temporal stability: were it not for the minimum wage hike, New Jersey employment would have remained unchanged.

- Is this plausible? Probably not. In our previous example, $Y_{j,1993} - Y_{j,1992} = T + \delta$. Our causal estimate so far conflates the true treatment effect with any incidental, contemporaneously occurring "time effects," such as a change in fast food demand in New Jersey.

- So, what do we need to improve on this experiment? We could select a group of states at random and assign the minimum wage increase to half of them and not to the other half. Then, we could compare employment in each group of states. A problem here is that this experiment is not available to us. But it's a good idea!

- Another possibility is to select a single state that we think is closely comparable to NJ and use it as our "control group." Here, that state is Pennsylvania.
- In this case, we could take baseline data in both states and then compare the change in NJ to the change in PA. This is our difference-in-difference estimator.

7.1 Card & Krueger (1994)

- The 1994 Card and Krueger article is a widely cited study of the impact of the minimum wage on employment levels. It created huge controversy in both policy circles and among academic economists, and arguably caused millions of workers to get a legislated raise from the Clinton administration in 1995.
- April 1, 1992: the New Jersey minimum wage rose from \$4.25 to \$5.05 per hour (a sizable increase)
- Eastern Pennsylvania (bordering NJ) didn't raise the minimum wage. It maintained the Federal minimum wage of \$4.25 per hour.
- Card & Krueger surveyed 410 fast food restaurants.
- For purposes of the analysis, the pre-period is Feb-Mar 1992, and the post-period is period is Nov-Dec 1992.
- The setup:

	Before	After	Δ
NJ	$Y_{n,1992}$	$Y_{n,1993}$	ΔY_n
PA	$Y_{p,1992}$	$Y_{p,1993}$	ΔY_p

$$\hat{T} = \Delta Y_n - \Delta Y_p$$

- Table 3 in the paper shows "Per store employment"

	Before	After	Δ
NJ	20.44	21.03	$\Delta Y_n = +0.59$
PA	23.33	21.37	$\Delta Y_p = -2.16$

- $\hat{T} = 0.59 - (-2.16) = 2.76$ with a standard error of 1.36 (so, it is statistically significant at the 5 percent since the t-ratio is ≈ 2.0).
- The paper contains many more tests, but this is the basic result: $2.76 \approx 13.5\%$ increase in employment in NJ relative to PA.

7.1.1 Interpretations?

1. Monopsony

Other interpretations?

2. Hungry teens
3. Motivational effects
4. Confounding variables (shocks to PA that are not accounted for in the test)
5. Wrong venue (why did they study fast food?)

We will have much more to say in class about the interpretation of the Card and Krueger findings.

8 Brief discussion: Methodology of Economics – *or* Why Economic Theory?

Stepping back, I want to provide a brief overview of the methodology of economic analysis.

Definition 10. *Positive Economics*

(1) The study of "what is." A descriptive endeavor free of value judgements; (2) Building models to make sense of, and generalize, the phenomena we observe; (3) Making predictions based on those models.

Definition 11. *Normative Economics*

Assessing "what ought to be done;" Making economic policy prescriptions. Note: Sometimes positive economics gives us all the tools we need to say that one policy is preferable to another. For example, when one policy is Pareto superior to another. (Not too many of these)

Definition 12. *Pareto Improvement*

A choice/policy/outcome that can make at least one person better off without making anyone else worse off. This is a pretty timid moral criterion (though it's not entirely uncontroversial).

- In reality, Pareto improving policy options are very rare. (We tend to expect that people would already have made those types of improvements without any policy interventions!)
- Most policy choices involve value judgements, ethical preferences, trade-offs among competing goals (e.g., employment and inflation; equity and efficiency).
- Economic theory rarely tells you what policies to choose. But it often makes the trade-offs clear.

8.1 Strength of economic approach to social science

- *Rigorous*: assumptions are stated, methods are formal, conclusions are *internally* consistent.
- *Cohesive:* built on a foundation of first principles and theory.
- *Refutable*: makes strong, testable (refutable) predictions, many (far from all!) of which appear correct.
- *Practical*: will help you to better understand how the world works.

8.2 Weaknesses of the economic approach

- "Economics is marked by a startling crudeness in the way it thinks about individuals and their motivations..."— Paul Krugman
- Strong, simplifying assumptions that are often unpalatable and cannot be completely right (e.g., people act rationally to pursue self-interested—distinct from selfish—objectives...)

8.3 But there is some strength in this weakness

- We have a model of the world; it is called "the world"—and it's generally too complicated to analyze in its totality, considering all factors at once.
- Economic theory typically presents a very simplified, highly stylized depiction of the world. But this can be quite helpful.
- "The test of the validity of a model is the accuracy of its predictions about real economic phenomena, *not* the realism of its assumptions"—Milton Friedman
- "A hypothesis is important if it explains much by little"—Milton Friedman
- Our approach: simple models, significant insights.

8.4 Three significant insights of economic approach

1. Economics is about "people doing the best with what they have." This observation gets you a long way in understanding human activity, both positively and normatively. You can understand much by starting from the premise that people are *trying* to make the best choices for themselves. Many alternative assumption—for example, people are largely irrational and guided by forces they do not perceive or comprehend—appear much less attractive.
2. Equilibrium—The market 'aggregates' individual choices to produce collective outcomes that are sometimes *spectacularly different* from individual decisions. (An example from 14.01: if the firms in a fully competitive marketplace are each trying to maximize profits, none makes economic profits in equilibrium, yet the market equilibrium maximizes the sum of consumer and producer surplus.)

3. Properties of equilibrium can be evaluated using the criterion of efficiency:

   - Given: Individuals are trying to make best choices for themselves. Does market equilibrium produce an outcome that cannot be improved upon without making at least one person worse off (i.e., is it Pareto efficient)?
   - There is no obvious reason to assume that markets would lead to desirable outcomes, i.e., that we couldn't do *much* better by engaging in central planning than in relying on the haphazard result of everyone selfishly making independent choices.
   - Yet, one of the stunning insights of economics is that under some key conditions, the market will produce Pareto efficient outcomes. (Not to say that all Pareto efficient outcomes are desirable outcomes. But a Pareto efficient allocation is often a good starting point!)
   - And, where the market does not produce Pareto efficient outcomes, theory provides insight into why this occurs, and may provide guidance on how to get to a better outcome.
   - The question of when market outcomes are first-best efficient comes from the study of General Equilibrium. We will build towards this topic from the start of the class and develop it rigorously towards mid-semester.