Graduate Labor Economics
ECN 250A, Winter 2018

Professor Brendan Price*
University of California, Davis

This is a closed-book exam. Answer all questions briefly but clearly. The exam is graded out of 75 points; each subquestion is worth 5 points.

I Short-answer questions (30 points)

(a) Producer theory predicts that—if the labor market is perfectly competitive—there will be a "law of one price" for labor. In practice, of course, different workers receive vastly different hourly wages. Why doesn't the mere existence of wage dispersion prove that the labor market is not competitive?

Theory predicts that all workers receive the same hourly wage for each *efficiency unit* of labor they supply. The existence of wage dispersion is consistent with a law of one wage provided that wage differences are solely driven by variation in workers' efficiency units (i.e., their productivity or skill level). Another explanation is that wage differences reflect differences in amenities across jobs, since in this case the wage must adjust to clear the market for on-the-job amenities as well as the market for labor output.

(b) Retail pharmacies like CVS rely on a mixture of human cashiers and automated self-checkout kiosks. Suppose that the price of an automated kiosk falls by 50% in the next 5 years. In a two-factor model, how will this price change affect employment in the retail pharmacy sector?

Declines in the price of a kiosk have theoretically ambiguous effects on cashier employment. The substitution effect says that pharmacies will substitute away from human cashiers and towards kiosks, but the scale effect says that it will be cheaper to operate a pharmacy establishment, so companies may open up new pharmacy locations. The net effect is ambiguous and it depends on the elasticity of substitution between cashiers and kiosks.

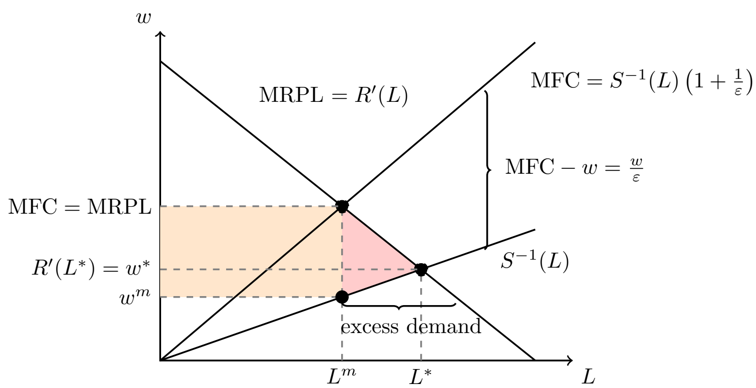
(c) In class, we discussed numerous explanations for the rise in wage inequality observed in the United States since 1980 (as measured, for example, in the March CPS). List three such explanations.

Possible answers include skill-biased technical change (together with a slowdown in rising educational attainment); increased exposure to trade with low-income countries (which in traditional trade theory should increase the skill premium in high-income countries); superstar effects stemming from globalization; the declining real value of the minimum wage, especially during the 1980s; private-sector deunionization; growing dispersion of firm-level pay practices; and "domestic outsourcing" of low-skill workers to outside contractors.

 Given my subsequent affiliation at the Federal Reserve Board, I note that any views presented here are my own and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff.

(d) Explain why, in a static monopsony model, increasing the minimum wage can lead to either increases or decreases in employment. (A graph may be helpful here.)

In the absence of a minimum wage (or when the minimum wage doesn't bind for a particular employer), a monopsonist is constrained by the supply curve she faces: at her profit-maximizing wage, she would love to hire additional workers but she can't attract them without raising wages. Increasing the minimum wage forces the monopsonist to raise wages, which attracts more workers and can therefore raise employment. But once the minimum wage rises above the wage that would exist under perfect competition, employment is determined by the demand curve, so that increasing the minimum wage even further leads to a decrease in employment. All of this is readily seen in the figure we discussed in class:



(e) DiNardo and Lee (2004) use a regression-discontinuity design to estimate the economic effects of plant unionization, using union elections conducted during 1984–1999. As they explain, "outcomes for employers where unions *barely* won the election (e.g., by one vote) are compared with those where the unions *barely* lost." Even if DiNardo and Lee's results are internally valid, however, we might worry about their external validity. Why we might hesitate to generalize from DiNardo and Lee's results?

DiNardo and Lee's RD design identifies the local average treatment effect of unionization *among firms close to the 50% vote-share threshold*. The causal effects might be very different at other vote shares: for example, unions might have a dramatic impact on firm outcomes when they win 90% of the vote, since a high vote share indicates a "strong" union with a powerful mandate from the workers. DiNardo and Lee also point out that unions organized during their sample period might be "marginal" unions that are a lot weaker than the inframarginal unions organized in earlier periods. And of course results on US unions might not generalize to other countries.

(f) In their study of the earnings losses experienced by displaced workers, Jacobson, Lalonde, and Sullivan (1993) compare changes in these workers' earnings to changes in earnings among two comparison groups: (i) non-displaced workers employed by healthy firms and (ii) non-displaced workers employed by the displaced workers' former employers—that is, the treatment group's former coworkers. List one pro and one con of using (ii) as the control group, rather than (i).

An advantage of comparing displaced workers to their former coworkers is that these groups may be especially similar. A disadvantage is that these coworkers might also be adversely affected by the same firm-level shocks that caused the displaced workers to be displaced. In that case, comparing displaced workers to their former coworkers may lead us to underestimate the adverse earnings consequences of displacement, relative to a counterfactual in which workers are not subjected to economic distress.

II Between-firm wage dispersion (20 points)

The table below, taken from Krueger and Summers (1988), reports the estimated coefficients on 1-digit industry dummies in cross-sectional Mincer regressions of individual log wages on demographics (sex, education, experience, etc.), occupation dummies, and industry dummies, estimated separately in 1974, 1979, and 1984. The coefficients indicate that some industries (e.g., construction) systematically pay above-average wages to observationally similar workers.

Industry	(1) 1974	(2) 1979	(3) 1984
Construction	.195	.126	.108
	(.021)	(.031)	(.034)
Manufacturing	.055	.044	.091
	(.020)	(.029)	(.032)
Transportation & Public Utilities	.111	.081	.145
	(.021)	(.031)	(.034)
Wholesale & Retail Trade	−.128	−.082	−.111
	(.020)	(.030)	(.033)
Finance, Insurance and Real Estate	.047	−.010	.055
	(.022)	(.035)	(.034)
Services	−.070	−.055	−.078
	(.021)	(.030)	(.032)
Mining	.179	.229	.222
	(.035)	(.058)	(.075)
Weighted Adjusted Standard Deviation of Differentials[b]	.097**	.069**	.094**
Sample Size	29,945	8,978	11,512

(a) Explain why these coefficients do not necessarily reveal the causal effect of industry membership on a worker's wages.

High-paying industries may simply employ higher-skilled workers. For example, consulting firms employ lots of highly educated workers who attended prestigious universities. These workers would probably earn high salaries even if they worked in other industries.

(b) Provide two explanations for why some industries may systematically pay a given worker more than other industries do.

(i) Industries might pay above-average wages to compensate workers for poor working conditions, such as an elevated risk of injury (e.g., as in mining) or grueling hours (e.g., as in investment banking). (ii) Industries might pay "efficiency wages": that is, in some industries paying above-market wages might increase productivity by attracting better workers, encouraging workers to exert greater effort, or decreasing turnover. The benefits of offering higher wages depends on industry-level factors like the importance of unobserved worker skill, the cost of monitoring effort, and the cost of turnover.

Since Abowd, Kramarz, and Margolis (1999), many papers have used linked worker-firm data to estimate models of the form

$$\log w_{it} = \alpha_i + \gamma_{J(i,t)} + \varepsilon_{it},$$

where α_i is a worker fixed effect and $\gamma_{J(i,t)}$ is a fixed effect for the firm that employs worker i in year t.

(c) Card, Heining, and Kline (2013) find that a large share of the rise in West German wage inequality can be attributed to increases in $\text{Var}(\gamma_{J(i,t)})$ and to increases in $\text{Cov}(\alpha_i, \gamma_{J(i,t)})$. Explain in words what these two statistical statements represent.

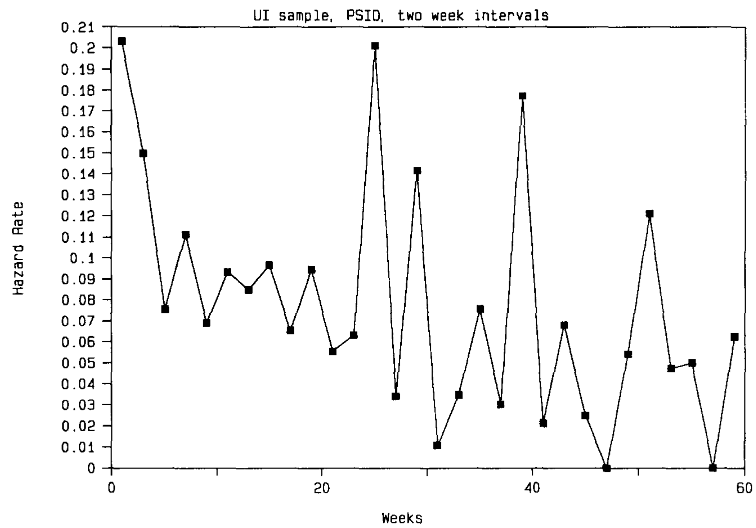
The increase in $\text{Var}(\gamma_{J(i,t)})$ indicates that there is growing dispersion in firm-level pay practices: firms are increasingly likely to pay "above-average" or "below-average" wages to a worker of given ability. Put differently, there are more high-paying ("good") firms and also more low-paying ("bad") firms than there used to be. The increase in $\text{Cov}(\alpha_i, \gamma_{J(i,t)})$ indicates an increase in assortative matching: high-earning workers are increasingly likely to be employed by high-paying firms.

(d) While AKM models were originally developed to decompose wages into worker and firm components, they can be adapted to other settings where two sets of agents or institutions match with each other and where this matching changes over time. Give an example of a non-wage application of the AKM model (either one I mentioned in class or one you come up with yourself).

For full credit, you needed to specify both sides of the market (in a setting where there is a meaningful amount of switching) and also specify the outcome variable to be decomposed. (i) Test scores can be decomposed into student and teacher fixed effects. (ii) Health outcomes can be decomposed into patient and hospital fixed effects. (iii) Physician prescribing behavior can be decomposed into physician and hospital fixed effects. (iv) Teaching evaluations can be decomposed into professor and course fixed effects.

III Duration dependence (10 points)

The figure below, taken from Katz and Meyer (1990), plots empirical job-finding hazards among US unemployment insurance claimants, as a function of time since entry into unemployment.



(a) Why does the hazard rate spike around 26 weeks and again around 39 weeks?

In the United States, most workers receive about 26 weeks of UI benefits during normal times, with a 13-week extension during recessions. As a large theoretical and empirical literature has found, there is a spike in job-finding around the time when UI benefits are exhausted.

(b) Aside from these spikes, the hazard rate mostly exhibits negative duration dependence (i.e., the longer a worker has been unemployed, the less likely that worker is to become reemployed next period). Provide two explanations for negative duration dependence.

On the supply side, negative duration dependence arises if unemployed workers become discouraged over time and don't search as hard for a job, or if workers quickly exhaust the stock of good matches and hence find it optimal to spend less time searching. On the demand side, negative duration dependence arises if employers statistically discriminate against the long-term unemployed (a "scarring" effect), since being long-term unemployed might signal that a worker is not very productive. Negative duration dependence can also reflect the declining rate of recall from temporary layoff. Finally, negative duration dependence can arise at the aggregate level even if each worker has a constant hazard rate, since over time the risk set of still-unemployed workers is increasingly comprised of workers with low job-finding hazards.

IV Incentivizing job search (15 points)

Consider a stylized continuous-time search model in which a worker can generate job offers at flow rate s by incurring a search cost $\psi(s)$. Workers receive flow utility b during unemployment, and all jobs pay a constant wage $w > b$. So far, this is a stationary setup: b, w, and $\psi(\cdot)$ are all constant over time.

Some policymakers have proposed using "reemployment bonuses" to incentivize job search. Imagine that California offers a lump sum payment of $1000 to any unemployed jobseeker who becomes reemployed within 3 months of entering unemployment. After 3 months, jobseekers are no longer eligible for the bonus.

(a) Draw a graph plotting a given worker's search effort as a function of time since entry into unemployment (i) in the absence of the bonus policy and (ii) in the presence of the bonus policy. You don't have to do any math to answer this question, but you might find it helpful to think about the underlying Bellman equation. Make sure to label your graph clearly.

In the absence of the bonus policy, search effort is constant over time. In the presence of the bonus policy, search effort is strictly higher during the first three months than in the absence of the bonus policy, since during these early months of unemployment finding a job is more valuable than in the no-bonus counterfactual. Furthermore, the hazard rate rises as workers approach the three-month mark, for the same reason that job-finding rises as workers approach benefit exhaustion in Mortensen (1977) and in Price (2018). Once the bonus option expires at three months, there is no extra incentive to search for a job, so the hazard rate immediately falls to the constant level we saw in the initial no-bonus scenario.

(b) Suppose that, to assess the efficacy of this policy, California conducts a randomized controlled trial (RCT) in which 1% of jobless workers are sampled, half are treated, and half are used as a control group. Suppose further that the RCT finds a 5% reduction in the median unemployment duration. Explain why this treatment effect may not be a good estimate of what would happen if California adopted this policy for all unemployed workers.

This RCT won't detect any general equilibrium effects of the bonus policy on the probability of finding a job. As we discussed in class (and as we saw in Lalive, Landais, and Zweimüller 2015 and Marinescu 2017), when one group of workers searches harder for a job, they might make it harder for other workers to find jobs through a "rat-race effect". But they might also incentivize firms to create additional job vacancies, which makes it easier for unemployed workers to find jobs.

(c) Explain how we could modify this RCT to provide estimates of both the individual and market-level effects of the bonus policy.

If all we care about is the market-level effect, we could randomly assign each California city (or county, or whatever geographic unit we like) to either treatment or control, then make everyone in the treatment regions (and no one in the control regions) eligible for the bonus policy. If we'd like to estimate both the individual and market-level effects of the policy, we could use a two-stage randomization in the manner of Crépon et al. (2013), which I described very briefly in class. For each county c, we could randomly draw a proportion $p_c \in [0, 1]$, say from a uniform distribution on $[0, 1]$. Then we would offer the bonus scheme to $100 \times p_c$ percent of the unemployed workers in county c. Differences in jobless durations between treated and untreated individuals in each county would identify the individual (partial equilibrium) effect of treatment. Differences in jobless durations among *untreated* workers residing in more- vs. less-intensively-treated counties would identify the market-level (general equilibrium) effect of treatment.