Lecture Note 17 — Signaling and Statistical Discrimination: An Application to 'Ban the Box' Legislation

David Autor, MIT and NBER

December 13, 2022

Introduction

It's possible to develop the misimpression that information economics is mostly designed to explain cute and quirky phenomena: why GED holders earn more than similarly-skilled non-GED holders, why people are reluctant to buy used cars, why insurance companies charge such high prices for low-deductible auto policies, etc. This lecture will correct that misimpression. Information—or the lack thereof—matters in markets.

One of the places where information matters most is in employment: specifically, who gets hired and who does not. Hiring a worker is somewhat akin to buying a jet ski or a used car. You don't really know what 'quality' of worker you're going to get until after you've worked with an individual for a while. If it turns out that you have made a bad hiring decision, you can ultimately dismiss (AKA fire) the worker. But firing is unpleasant, disruptive, and sometimes legally risky (AKA expensive). Employers therefore have an incentive to gather information about applicants prior to hiring to assess their quality and fit.

You might surmise that job applicants would therefore have an incentive to *conceal* information so employers don't draw any negative inferences. But that's not necessarily true. Job applicants who have *positive* information to reveal will have an incentive to reveal that information. And as we saw in the signaling model, one group's decision to reveal positive information about itself can have adverse consequences for other groups that don't have similarly positive information. In the Spence model, workers who reveal themselves to be of High ability by obtaining high levels of schooling implicitly reveal workers who did not obtain high levels of schooling to be of Low ability.

Noting the harm done to Low ability workers by this signaling process, policymakers

might decide to pass a law that banned employers from asking applicants about how much schooling they had obtained. As a 14.03/003 student, you might hypothesize that this rule would be bad for High ability workers but good for Low ability workers (effectively a transfer between these two groups). In reality, it could be worse than that. Depending on employers' aversion to Low ability workers, they may choose to not hire anybody rather than risk hiring a Low ability worker.

1 Criminal History, Employment Screening, and the Ban-the-Box Movement

Aside from the tiny Republic of Seychelles, the U.S. has the highest fraction of its population incarcerated in the world: 693 prisoners per 100K residents.[1] Of course, many more people have been to prison than are currently in prison. A recent paper estimates that in 2010, 6% of non-African-American adults and 23% of African-American adults were either current felons or former felons.[2] Given that most (not all) felonies are committed by men, these statistics would suggest that the fraction of non-African-American and African-American *men* who are felons or former felons is on the order of 10% and (perhaps) 35% respectively. It is hard *not* to find these numbers startling.

Individuals convicted of a felony face substantial barriers to reintegration and economic self-sufficiency after their felony convictions. Many employers ask job applicants whether they have a felony conviction. Employers may frequently choose not to consider applicants who answer yes. This type of applicant screening is not unlawful. In the U.S., it is legal to discriminate on any characteristic *other* than race, sex, disability, union membership, and age (over 40), so long as the characteristic in question can be considered job relevant (e.g., it would not be legal to discriminate on physical strength for a telephone operator position, but it would be legal to discriminate on physical attractiveness for a maï¿œtre d' position at a restaurant).

As reported in a famous 2018 *Quarterly Journal of Economics* paper by Agan and Starr on your syllabus, in an effort to reduce barriers to employment for people with criminal records, more than 100 U.S. jurisdictions (cities and towns) and 23 states have passed "Ban-the-Box" (BTB) policies. Although the details vary, these policies all prohibit employers from asking about criminal history on the initial job application and in job interviews. Employers may

[1] `https://en.wikipedia.org/wiki/List_of_countries_by_incarceration_rate`

[2] Shannon, Sarah K.S., Christopher Uggen, Jason Schnittker, Melissa Thompson, Sara Wakefield, and Michael Massoglia. "The Growth, Scope, and Spatial Distribution of People with Felony Records in the United States, 1948 to 2010." *Demography*, 2017.

still conduct criminal background checks, but only at or near the end of the employment process. Most BTB policies apply to public employers only, but seven states (including New Jersey) and a number of cities (including New York City) have now also extended these restrictions to private employers.

Agan and Starr succinctly describe the rationale for these policies. "These laws seek to increase employment opportunities for people with criminal records. They are often also presented as a strategy for reducing unemployment among black men, who in recent years have faced unemployment rates approximately double the national average (Bureau of Labor Statistics 2015). The theory underlying this strategy is straightforward: black men are more likely to have criminal convictions than other groups (Shannon et al. 2017), and having a criminal record is a substantial barrier to employment (Pager 2003; Holzer, Raphael, and Stoll 2006; Holzer 2007; Pager, Western, & Bonikowski 2009). Thus, a policy that increases the employment of people with records should disproportionately help minority men."

They also succinctly describe the potential concern with these policies. "This effort could have unintended consequences, however. In the absence of individual information about which applicants have criminal convictions, employers might statistically discriminate against applicants with characteristics correlated with criminal records, such as race. In this scenario, applicants with no criminal records who belong to groups with higher conviction rates, such as young black males, would be adversely affected by BTB policies."

2 A Simple Model of Ban-the-Box

Let's formalize a model of how Ban-the-Box might help felons or hurt non-felon minorities. Consider a firm looking to hire its next worker. The firm pays a fixed wage for a given position, and prefers to hire a worker from Group 1 (e.g., a minority worker) but could also hire a worker from Group 0 (e.g., a non-minority worker). The firm can instantly determine whether an applicant is from Group 1 or Group 0.

Unfortunately, a fraction λ of Group 1 workers are felons (F), which the firm is averse to hiring. Non-felons (N) comprise a fraction $(1 - \lambda)$ of Group 1, and there is a negligible proportion of felons in Group 2. (We could add a small fraction of felons into Group 2 without changing the takeaway of the model.)

Denoting the firm's payoffs by π, the firm has

$$\pi_1^N > \pi_0 > \pi_1^F$$

When firms are able to distinguish between felons and non-felons, they will hire a non-felon.

If a non-felon from Group 1 is unavailable, they will hire someone from Group 0. Felons will very rarely be hired.

2.1 Hiring when no screening is available

Next, consider what happens when firms cannot distinguish between felons and non-felons in Group.

1. Employers will hire from Group 1 if this is more profitable than hiring from Group 0. Specifically, if
$$(1-\lambda)\pi_1^N + \lambda\pi_1^F > \pi_0,$$
then it's worth hiring from group 1, despite the uncertainty. Examining this inequality, you can see that felons will be hired more often when (1) non-felons from Group 1 are very profitable (π_1^N large); (2) felons from Group 1 are not very *un*profitable (π_1^F reasonably close to π_0); and (3) felons are a modest proportion of Group 1 (λ small).

2. Employers will *not* hire from Group 1 if this is less profitable than hiring from Group 0. Specifically, if
$$(1-\lambda)\pi_1^N + \lambda\pi_1^F < \pi_0,$$
then employers will 'statistically discriminate' against group 1 members, not out of animus but out of the concern that too many have low productivity (and given that the employer cannot distinguish high productivity π_1^N from low productivity π_1^F.

2.2 Adding a low-cost screening technology, AKA 'the box'

Now imagine that a low cost screening technology is available. Suppose firms had to pay c units in profit to screen non-felons from felons. Call this cheap screen "the box." It is a checkbox on the application form that applicants are obliged to check if they have been convinced of a felony. That is, for each applicant they screened, they would pay cost c. If the applicant turned out to be a felon, they would then have to screen a new applicant.[3] The expected number of times a firm would need to screen an applicant to get one non-felon hire is $1/(1-\lambda)$. To see why, note that if there no felons, they you'd only expect to screen one person. If 50% of applicants are felons, you'd expect to screen two applicants, and so on.

[3]Applicants can lie, of course, but most employers who use the box will do a criminal background check before hiring, meaning it doesn't make much sense for the applicant to lie on the application. But the box does save the employer money. If the employer has to pay for a criminal background check for *all* applicants, this is more expensive than only paying for the applicants who have reported (likely truthfully) that they do not have a felony conviction.

Firms would use this screening technology if

$$\begin{aligned}\pi_1^N - \frac{c}{(1-\lambda)} &> (1-\lambda)\pi_1^N + \lambda\pi_1^F \\ \lambda\pi_1^N &> \lambda\pi_1^F + \frac{c}{(1-\lambda)} \\ \pi_1^N &> \pi_1^F + \frac{c}{\lambda\,(1-\lambda)}\end{aligned}$$

That is, the productivity difference between non-felons and felons is large enough to justify the screening cost, adjusting for the odds that applicants are felons. If $\lambda = 0$, then it's never worth it to screen since there are no felons. If $\lambda = 1$, then it's never worth it to screen because all applicants are felons. If $\lambda \in (0, 1)$, it may be worthwhile to screen.

Two additional subtleties

- If $\pi_1^N > \pi_1^F + \frac{c}{\lambda(1-\lambda)}$, then firms will prefer using 'the box' then not using it, even if they would be willing to hire from group 1 without the box (that is, $(1-\lambda)\pi_1^N + \lambda\pi_1^F > \pi_0$)
- It may be that the availability of 'the box' actually causes firms to hire group 1 workers whom they otherwise would not. Say $(1-\lambda)\pi_1^N + \lambda\pi_1^F < \pi_0$, meaning that group 1 hires are too risky to hire without a screen. It may still be the case that

  $$\pi_1^N - \frac{c}{(1-\lambda)} > \pi_0 > (1-\lambda)\pi_1^N + \lambda\pi_1^F,$$

  that is, firms will wish to hire from group 1 if they can use 'the box' but won't do so if they cannot.

2.3 Banning the box

Now let's consider a 'ban the box' policy that makes makes it illegal for firms to ask applicants if they have a felony conviction. But firms can still do criminal background checks after they've interviewed and decided to hire an applicant. Let's say that the cost of a criminal background check is $c' > c$.[4]

What will imposing this higher cost do to hiring from group 1? There are three possible scenarios

1. The cost is not *too* high so firms pay c' instead of c:

   $$\pi_1^N - \frac{c'}{(1-\lambda)} > \pi_0 > (1-\lambda)\pi_1^N + \lambda\pi_1^F.$$

[4]This assumption is realistic. The box is virtually free. Criminal background checks cost $10 to $20 per applicant according to current price offers on the Internet as of December 2018.

Here, 'ban the box' has no effect. Firms drop the box but use criminal background checks instead. They hire precisely the same people, albeit at greater cost.

2. Firms stop screening for criminal backgrounds and simply hire from both groups. This would occur if
$$\pi_1^N < \pi_1^F + \frac{c'}{\lambda(1-\lambda)} \text{ and } (1-\lambda)\pi_1^N + \lambda\pi_1^F > \pi_0$$

3. The cost of criminal background checks for all applicants is too high *but* non-screened applicants are too risky. This would occur if
$$\pi_1^N < \pi_1^F + \frac{c'}{\lambda(1-\lambda)} \text{ and } (1-\lambda)\pi_1^N + \lambda\pi_1^F < \pi_0.$$

   In this case, employers will simply hire group 0 applicants instead of group 1 applicants.

This reasoning makes clear that 'banning the box' *will help* group 1 workers if it induces a pooling equilibrium where employers wish to hire both felons and non-felons from group 1 when it becomes too expensive to use a screen to distinguish them.

However, 'banning the box' will *harm* group 1 workers if it induces a separating equilibrium where employers choose to hire only group 0 workers because it's too expensive to screen group 1 workers but too risky to hire them without screening. Notice that in this scenario, the workers who are harmed are group 1 workers who are *not* felons. Felons were not going to be hired regardless (since $(1-\lambda)\pi_1^N + \lambda\pi_1^F < \pi_0$). Banning the box prevents non-felons from group 1 from being considered. This is a perverse outcome—the opposite of what ban the box advocates would ever intend—but it is nevertheless a realistic (though by no means certain) possibility.

3 Evidence

So is ban-the-box *good* for former felons who could use a leg up? Under some conditions, ban-the-box legislation is likely to help felons by pooling them with desirable non-felons. However, in other cases, ban-the-box legislation could hurt minorities as a whole, since firms could shy away from hiring minorities that might be felons when they can't screen for this characteristic at a low cost. Unfortunately, the evidence that economists have accumulated so far suggests that the second case is more likely to hold in the modern U.S. labor market.

In class, we will discuss Agan and Starr (2018), which provides the best evidence available to date.