

## Pre-Analysis Plan: Labor Demand for Workers with a Criminal Conviction

**Date:** 3/20/2020

**Investigators:** Zoë B. Cullen (Harvard Business School), Will Dobbie (Harvard Kennedy School), Mitchell Hoffman (University of Toronto Rotman School of Management)<sup>1</sup>

### Background

Employers are significantly less likely to hire workers with a past criminal conviction (WCCs) compared to otherwise identical workers without a conviction. Despite the high unemployment rate among workers with a past conviction, there is little evidence on why firms are unwilling to hire these workers or what policies might increase their employment rates. Survey data suggest that wage subsidies may increase the demand for these workers, but the hypothetical (and hence, low-stakes) nature of these studies makes it difficult to know whether employers are expressing their true preferences or just their aspirations.

We employ a randomized discrete choice experiment on a large on-demand staffing platform to estimate the labor demand for WCCs under different policies and market conditions. The platform currently screens out candidates with a past criminal conviction. Employers using the platform will have the option to have WCCs assigned to their vacancies at a randomly chosen wage subsidy, both with and without the option of crime and safety insurance and increased performance screening, and under different hypothetical labor market conditions. Employer choices at different wage subsidies will yield willingness to pay (WTP) estimates for this population of jobseekers under different counterfactual policies. These estimates will allow us to estimate a labor demand function.

### Study Timeline

TASKS	START DATE	DURATION
Finalization of details of insurance policy with a third-party provider.	Nov. 1, 2019	2 weeks
Staggered delivery of survey experiment, week-by-week.	Mar. 5, 2020	1 month
Analysis of survey responses and write-up.	Apr. 15, 2020	6 months

---

<sup>1</sup> The investigators have no conflicts of interest to declare.

## Experimental Design

The experimental design is as follows:

### A. Recruitment

Our sampling frame includes firms in the United States who hire through the online platform. We will recruit participants through the platform's account managers. Each of approximately 9,000 active clients in their database is assigned to an account manager. Typical personal correspondence on the platform occurs via email and phone. Our recruitment email emphasizes that the survey may have implications for the pool of workers who will receive that employer's job postings but does not explicitly mention the topic of past convictions. Survey completion will be incentivized by a \$35 to \$100 Amazon gift card and by the chance to win up to an additional \$2 if respondents answer factual questions correctly.

### B. Survey Design

In the survey experiment, we will ask firms about their interest in hiring individuals with a criminal record. The survey is incentive-compatible because subjects are informed that their choices will affect their future applicant pool. For each question, subjects can respond that they would be interested in hiring such individuals, that they would only be interested if it was otherwise hard to fill their jobs, or that they would not be interested.

We will cross randomize values of each of the following dimensions using a computer random number generator:

- Wage discount: 0%; 5%; 10%; 25%; 50%; 100%.
- Level of crime and safety insurance: \$1,000; \$5,000; \$100,000; \$5,000,000.
- Number of positive past performance scores on the platform: 1; 5; 25.
- Hypothetical local unemployment rate: 2%; 6%; 10%.
- Years since the most recent criminal act: 1; 3; 7.

We will elicit interest in hiring individuals based on their specific conviction (e.g., drug-related crimes, property-related crimes, violent crimes, financial crimes). We will also elicit their beliefs about what percentage of WCCs would perform well.

In 2017, the online platform transitioned to a new background checking procedure that added identity verification, and thus improved screening. Before this transition, some WCCs who would later have been excluded were permitted to accept jobs. If we receive information on the performance of those WCCs in a reasonable timeframe, then we will add an information intervention near the end of the survey (after the key questions described above). We will present information about the true proportion of WCCs who perform well to a subset of respondents. We will subsequently repeat the questions eliciting beliefs about WCC performance and willingness to hire WCCs under the randomized subsidy level to obtain posterior beliefs from all participants (including those not shown the information treatment).

Data will be collected through an online Qualtrics survey. After the survey response window is closed, the survey data will be linked to administrative data from the online labor platform, including firms' past work orders and firm characteristics.

### Hypotheses Tested

#### A. Primary

H1.To determine whether the demand for WCCs is low with no wage subsidy.

H2.To determine whether the demand for WCCs is higher among employers randomized to higher wage subsidies over the range [5%, 100%].

H3.To determine whether the demand for WCCs is higher and/or more elastic among employers randomized to higher levels of crime and safety insurance over the range [\$1,000, \$5 million].

H4.To determine whether the demand for WCCs is higher and/or more elastic among employers randomized to a greater number of positive past performance scores over the range [1, 25] positive reviews.

H5.To determine whether the demand for WCCs is higher and/or more elastic among employers randomized to a tighter labor market as characterized by an unemployment rate over the range [2%, 10%].

#### B. Secondary

H6.To determine whether the demand for WCCs changes based on the number of years since the crime was committed, randomized over the range [1,7] years.

H7.To determine whether the demand for WCCs is different for (i) property/financial felonies, (ii) violent felonies, (iii) substance-abuse felonies, (iv) property/financial misdemeanors, (v) violent misdemeanors, and (vi) substance-abuse misdemeanors.

H8.To determine whether the demand for WCCs is different among employers who post jobs that (i) involve interactions with customers, (ii) handling cash/high-value inventory, and (iii) allow for more discretion in hiring decisions.

H9.To determine whether the demand for WCCs is increasing in employers' perceptions of WCCs' on-the-job performance.

H10.To determine whether demand for WCCs is higher and/or more elastic among employers in a tighter labor market as characterized by (i) the share of employers only willing to hire when the vacancy would otherwise go unfilled and (ii) whether the employer is located in

a region that has an unemployment rate above or below the median unemployment level in our sample.

Because we are testing multiple hypotheses, we will use techniques that limit the false discovery rate such as correcting the p-values following the standard approach (e.g., Benjamini, Krieger, and Yekutieli, 2006; Anderson, 2008).

### Outcomes

Our primary outcome variable is employers' choices over whether to have WCCs assigned to their vacancies at randomly chosen wage subsidies. For the main questions, respondents are presented with a scenario and asked if, under the given condition, they would be interested in hiring a WCC. They have the option of responding "Yes," "Only if it's hard to fill my jobs," or "No." For all primary hypotheses, we will use this response as our main dependent variable. We will estimate demand assuming the outside option for the employer is a replacement worker at the market wage, and so will code a binary variable  $Y_i = 1$  if the response is "Yes" and  $Y_i = 0$  otherwise.

A secondary outcome variable is employers' perceptions of WCCs' performance. This is measured as a continuous variable from [0%,100%] based on respondents' beliefs about what percentage of jobs completed by WCCs would result in a positive outcome (high rating or absence of low rating).

To help understand how representative our sample of respondents will be, we will also include questions from a national human resources survey on attitudes about hiring WCCs. We will compare summary statistics based on responses from these questions in our survey to those from the national survey.

### Econometric Specifications

To test H1, we will calculate the proportion of respondents who respond that they are willing to hire a WCC with a randomly assigned subsidy of 0%.

To test H2, we will use the following specification:

$$(1) Y_i = \beta_1 + \sum_{k=2}^6 \beta_k \text{Subsidy}_{ki} + X_i' \mathbf{B} + \varepsilon_i$$

where  $i$  indexes firms (i.e. survey respondents).  $Y_i$  is an indicator for whether the firm would be interested in hiring an individual with a criminal conviction at the randomized subsidy level, defined in the Outcomes section. *Subsidy* is an indicator for the randomly assigned subsidy level that the firm sees. The key  $\beta_k$  coefficients represent the differential impact of offering various subsidies relative to the lowest level of incentive.  $X_i$ , included in some specifications, will be a set of background characteristics about the respondent and firm that may influence their baseline demand for WCCs. Including these variables in the analysis should not affect our estimates of the treatment effect but could increase their precision. In some specifications, we will treat *Subsidy* as a continuous variable to help interpret the responses to later questions in terms of their subsidy-equivalent effect.

To test H3, we will use Equation (2), where *Insurance* represents indicators for the randomized levels of crime and safety insurance. We include interactions between the subsidy and treatment levels in this and all following regressions to estimate changes in elasticity.

$$(2) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{24} \beta_k \text{Subs}_{ki} * \text{Insurance}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

To test H4, we will use Equation (3), where *Performance* represents indicators for the randomized numbers of positive past performance scores.

$$(3) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{18} \beta_k \text{Subs}_{ki} * \text{Performance}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

To test H5, we will use Equation (4), where *Unemployment* represents indicators for the randomized unemployment rates.

$$(4) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{18} \beta_k \text{Subs}_{ki} * \text{Unemployment}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

To test H6, we will use Equation (5), where *Years* represents indicators for the number of years since the most recent criminal act.

$$(5) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{18} \beta_k \text{Subs}_{ki} * \text{Years}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

To test H7, we will use Equation (6), where *CrimeType* represents the type of conviction (property/financial felony, property/financial misdemeanor, violent felony, violent misdemeanor, substance-related felony, substance-related misdemeanor). We will run a separate regression for each type of conviction.

$$(6) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{18} \beta_k \text{Subs}_{ki} * \text{CrimeType}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

To test H8, we will use Equation (7), where *JobType* represents whether the employer posts jobs that involve interactions with customers, handling cash/high-value inventory, or if the employer has a company-wide policy on hiring WCCs.

$$(7) Y_i = \sum_{k=1}^6 \beta_k \text{Subsidy}_{ki} + \left( \sum_{k=1}^{18} \beta_k \text{Subs}_{ki} * \text{JobType}_{ji} \right) + X_i' \mathbf{B} + \varepsilon_i$$

We will calculate the elasticity of the estimated labor demand curves using a linear or quadratic model to smooth the raw points. We will then test the robustness of these estimates by calculating elasticities from the raw points and from different smoothing models, including quartic and spline smoothers. We will also estimate mean effects of each of the treatment levels described in H3 to

H8 as compared to the subsidy-only baseline by averaging the treatment effects over each of the subsidy levels.

For H9, we will first test whether employers update their beliefs in response to information about WCC performance. We will estimate the learning rate,  $\gamma_0$ , from Equation (8) below, which assumes respondents are Bayesian in their updating.

$$(8) M_i^{post} - M_i^{prior} = \gamma_0(M_i^{signal} - M_i^{prior})T_i^M + \gamma_1(M_i^{signal} - M_i^{prior}) + \varepsilon_i$$

$M_i^{post}$  represents raw posterior beliefs about the relative performance of WCCs and  $M_i^{prior}$  represents raw prior beliefs about the relative performance of WCCs. All respondents, regardless of whether they are randomly assigned to receive feedback about historical performance (the signal of true performance), will be asked for their posterior beliefs about performance at the end of the survey experiment.  $M_i^{signal}$  represents the true historical relative performance of WCCs (shown to the subset of respondents as a signal of expected performance). To calculate  $M_i^{signal}$  we will exploit the 2017 policy change described earlier. We will re-run background checks and compare the outcomes for workers who “slipped through the cracks” before the policy change with those who had clean records.  $T_i^M$  is an indicator for whether the respondent received the information intervention.

We will then estimate the elasticity of willingness to hire with respect to beliefs about WCC performance using an instrumental variables (IV) approach. In the first stage, Equation (9), we will regress posterior beliefs on prior beliefs and the gap between the signal and prior, interacted with whether the person was shown the signal:

$$(9) M_i^{post} = \mu_0 + \mu_1(M_i^{signal} - M_i^{prior}) \cdot T_i^M + \mu_2(M_i^{prior}) + \mu_3X_i + \psi_i$$

In the second stage, Equation (10), we will regress willingness to hire on the estimated posterior beliefs from the first stage, controlling for the gap between the signal and the prior:

$$(10) \log(Y_i^{post}) = \pi_0 + \eta M_i^{post} + \pi_1(M_i^{signal} - M_i^{prior}) + \pi_2(M_i^{prior}) + \pi_3X_i + \varepsilon_i$$

Where  $Y_i^{post}$  represents posterior willingness to hire WCCs and  $X_i$  represents controls that help with the precision of these estimates, including the hiring experience of the employer and their authority in making hiring decisions.

Finally, we will estimate heterogeneity in the effect on hiring for firms with different baseline perceptions. Assuming there is enough density in each of the following groups, we will split the sample into respondents whose perceptions were (i) incorrect and negative and (ii) incorrect and positive.

To test H10, we will first determine the employment rate in the employer’s region. Using this information, we will create a subsample split (above/below the median unemployment rate), and re-test hypotheses H1-H6 in these two subsamples.

We will also test H10 by calculating the share of employers only willing to hire when the vacancy would otherwise go unfilled, holding the subsidy level constant. In other words, we will also estimate demand when a vacancy is the outside option, and redefine our main variable (“Yes” “Only if it's hard to fill my jobs” or “No”) to = 0 if the response is “No” and = 1 otherwise, and use this to re-examine our main hypotheses.

### Robustness & Additional Data Quality Checks

For the specifications used to examine H2 to H8, we will test robustness to using a logit specification. We will also test the robustness of the results by reweighting these regressions using observable characteristics to match a sample of firms from a national human resources survey. Further, we will examine the subsample of respondents who self-report that they would have the authority to hire a WCC.

Finally, we will perform a series of standard data quality checks to ensure that the randomization was successful and that the experimental design was not compromised. For example, we will perform a balance test by regressing respondent/firm characteristics on indicators for the main randomized treatments from H2-H5 (e.g., subsidy level, insurance level). We will also use JPAL’s replication service before publication and pending IRB approval for the replication service to access the data.

### Registration Timeline

The study was registered with the American Economic Association (AEA) on December 23, 2019. IRB clearance was obtained from Harvard University (#IRB19-1079) and the University of Toronto. On December 23, 2019, the investigators uploaded the first copy of the pre-analysis plan (PAP) that outlines the main hypotheses.

On Jan 10, 2020, approximately two months before the experiment began, we made the following changes to the PAP:

- Updated the timeline to reflect minor delays in the roll-out;
- Changed the survey incentive range from \$17-\$50 with the possibility of an additional \$10 to \$35-\$100 with the possibility of an additional \$2;
- Changed the 5% randomized wage discount to 10%;
- Changed the \$1,000 randomized insurance cap to \$5,000;
- Included interaction terms in the main regression specifications;
- Added details about elasticity calculations; and
- Added details about robustness checks.

On March 7, 2020, two days after the experiment began, we learned that the intended switch from 5% to 10% for the wage discount and \$1,000 to \$5,000 for the randomized insurance cap was not included in the first batch of surveys sent to respondents. Fortunately, we had randomly split our initial sample into two batches for delivery and both sets of values are economically interesting. Thus, batch #1 had a 5% wage discount and \$1,000 insurance cap, and batch #2 had

the 10% wage discount and \$5,000 insurance cap. In subsequent batches of newer potential respondents, subjects in the low level of wage discount were evenly split randomly between 5% and 10%, and subjects in the low level of insurance cap were evenly split between \$1,000 and \$5,000. As such, the values referenced in this PAP now include both sets of values.