An Observational Implementation of the Outcome Test with an Application to Ethnic Prejudice in Pretrial Detentions

Appendix for Online Publication

Nicolás Grau
Universidad de Chile

Damián Vergara
UC Berkeley

A A Simple Model of Pretrial Detention Decisions

Preliminaries Judges are indexed by j and defendants by i. Judges are assigned to defendants according to the mapping $j(i)$. Judges use all available information to compute defendant-specific probabilities of pretrial misconduct and release defendants whenever the probability is smaller than or equal to a judge-specific threshold. Let G_i be an indicator variable that takes the value 1 if defendant i belongs to group G. The question we address is whether judges are prejudiced against defendants in group G in the release decision. Judges also observe other characteristics of the individual, Z_i.

Pretrial misconduct Let PM_i be an indicator variable that takes the value 1 if defendant i is engaged in pretrial misconduct. Let PM_{i0} and PM_{i1} denote pretrial misconduct if detained and released, respectively, and $Release_i$ be an indicator variable that takes the value 1 if defendant i is released. Then, $PM_i = Release_i PM_{i1} + (1 - Release_i) PM_{i0}$. Note that $PM_{i0} = 0$ for all i, given that detained defendants cannot be engaged in pretrial misconduct. We assume PM_{i1} is given by

$$PM_{i1} = 1\{PM_i^* \geq 0\} = 1\{m(Z_i, \nu_i) \geq 0\}, \tag{A.I}$$

where ν_i are variables that affect pretrial misconduct and are not observed by the judge, and m is some function. We assume that the information set is the same for all judges. Z_i may contain a defendant's criminal record and demographics, as well as the case characteristics (e.g., type of crime). On the other hand, ν_i may include both variables that the judge does not observe (e.g., defendant's informal networks) and shocks that affect the probability of misconduct that are realized after the release decision. Note that we assume $j(i)$ does not affect PM_i^*. This is for notational purposes only, we do not need that exclusion restriction for our analysis.

Selection process To make the release decision, judges use all the available information to predict PM_{i1} and compare their prediction to a threshold. Formally,

$$Release_i = 1\left\{\hat{p}(G_i, Z_i, j(i)) \leq t(G_i, Z_i, j(i))\right\}, \tag{A.II}$$

where $\hat{p}$ is a function that computes the prediction of PM_{i1}, and t is the release threshold that the judges set depending on G_i and Z_i. Note that $j(i)$ is included in both functions because judges are allowed to be heterogeneous in the way they make predictions and set thresholds.

The judge-specific prediction can written as a deviation from the true conditional probability:

$$\hat{p}(G_i, Z_i, j(i)) = \mathbb{E}_{\nu}[PM_{i1}|G_i, Z_i] + b(G_i, Z_i, j(i)), \tag{A.III}$$

where b is a function that accounts for the judge-specific bias in the risk prediction. Putting (A.II) and (A.III) together, we can write

$$\begin{aligned} Release_i &= 1\{\mathbb{E}_{\nu}[PM_{i1}|G_i, Z_i] \leq t(G_i, Z_i, j(i)) - b(G_i, Z_i, j(i))\}, \\ &\equiv 1\{\mathbb{E}_{\nu}[PM_{i1}|G_i, Z_i] \leq h(G_i, Z_i, j(i))\}. \end{aligned} \tag{A.IV}$$

We denote the function $h(G_i, Z_i, j(i)) = t(G_i, Z_i, j(i)) - b(G_i, Z_i, j(i))$ as the *effective threshold*. Defining $\mathbb{E}_{\nu}[PM_{i1}|G_i, Z_i] = p(G_i, Z_i)$ yields equation (1).

B Proofs

Let PM_i be the observed pretrial misconduct of defendant i. Then

$$\mathbb{E}[PM_i|G_i = g, Release_i^* = 0] \quad = \quad \bar{h}(g).$$

Proof. Let PM_{i1} be pretrial misconduct if released, with $\mathbb{E}[PM_{i1}|G_i, Z_i, j(i)] = \mathbb{E}[PM_{i1}|G_i, Z_i] = p(G_i, Z_i)$. $Release_i^* = 0$ implies that $p(G_i, Z_i) = h(G_i, Z_i, j(i))$. Then

$$\begin{aligned}
\mathbb{E}[PM_i|G_i = g, Release_i^* = 0] &= \mathbb{E}[PM_{i1}|G_i = g, Release_i^* = 0], \\
&= \mathbb{E}[PM_{i1}|G_i = g, p(G_i, Z_i) = h(G_i, Z_i, j(i))], \\
&= \mathbb{E}[\mathbb{E}[PM_{i1}|G_i, Z_i, j(i), p(G_i, Z_i) = h(G_i, Z_i, j(i))]|G_i = g, p(G_i, Z_i) = h(G_i, Z_i, j(i))], \\
&= \mathbb{E}[p(G_i, Z_i)|G_i = g, p(G_i, Z_i) = h(G_i, Z_i, j(i))], \\
&= \mathbb{E}[h(G_i, Z_i, j(i))|G_i = g], \\
&= \bar{h}(g).
\end{aligned}$$

Note that the argument can be replicated by allowing $p(G_i, Z_i)$ to depend on $j(i)$, being the exclusion restriction without loss of generality. □

PROPOSITION II. *Let x_1 and x_2 be two possible realizations of X_i and $\varepsilon > 0$ be a small distance from the margin of release. Under A1 and A2,*

$$\begin{aligned}
\Pr\left(Release_i^* \leq \varepsilon | X_i = x_1, Release_i = 1\right) \quad &> \quad \Pr\left(Release_i^* \leq \varepsilon | X_i = x_2, Release_i = 1\right) \\
\iff \qquad \mathbb{E}\left[Release_i | X_i = x_1\right] \quad &< \quad \mathbb{E}\left[Release_i | X_i = x_2\right].
\end{aligned}$$

Proof. Consider A1 and A2. Then, we can write the selection rule as $Release_i = 1\{\Lambda(X_i) \geq \zeta_i\}$, with $\Lambda(X_i) = \frac{d(X_i) - r_1(X_i)}{r_2(X_i)}$. Let Θ be the cdf of ζ_i. Then

$$\begin{aligned}
\Pr(Release_i^* \leq \varepsilon | X_i, Release_i = 1) &= \Pr(d(X_i) - W_i \leq \varepsilon | X_i, W_i \leq d(X_i)) \\
&= \Pr\left(\frac{d(X_i) - r_1(X_i) - \varepsilon}{r_2(X_i)} \leq \zeta_i | X_i, \zeta_i \leq \Lambda(X_i)\right) \\
&= \frac{\Pr\left(\frac{d(X_i) - r_1(X_i) - \varepsilon}{r_2(X_i)} \leq \zeta_i \leq \Lambda(X_i) | X_i\right)}{\Pr(\zeta_i \leq \Lambda(X_i))} \\
&= \frac{\Theta(\Lambda(X_i)) - \Theta\left(\frac{d(X_i) - r_1(X_i) - \varepsilon}{r_2(X_i)}\right)}{\Theta(\Lambda(X_i))}.
\end{aligned}$$

A Taylor approximation around $\varepsilon = 0$ gives

$$\Theta\left(\frac{d(X_i) - r_1(X_i) - \varepsilon}{r_2(X_i)}\right) \approx \Theta(\Lambda(X_i)) - \varepsilon \cdot (r_2(X_i))^{-1} \cdot \theta(\Lambda(X_i)),$$

where θ is the pdf of ζ_i. Then $\Pr(Release_i^* \leq \varepsilon | X_i = x_1, Release_i = 1) > \Pr(Release_i^* \leq \varepsilon | X_i = x_2, Release_i = 1)$ implies that

$$\frac{\theta(\Lambda(x_1)) \cdot (r_2(x_1))^{-1}}{\Theta(\Lambda(x_1))} > \frac{\theta(\Lambda(x_2)) \cdot (r_2(x_2))^{-1}}{\Theta(\Lambda(x_2))} \quad \Leftrightarrow \quad \frac{\Theta(\Lambda(x_1))}{\theta(\Lambda(x_1))} < \frac{\Theta(\Lambda(x_2))}{\theta(\Lambda(x_2))} \frac{r_2(x_2)}{r_2(x_1)}.$$

Note that A2 implies that $r_2(x_2)/r_2(x_1) \leq 1$. Then

$$\frac{\Theta(\Lambda(x_1))}{\theta(\Lambda(x_1))} < \frac{\Theta(\Lambda(x_2))}{\theta(\Lambda(x_2))} \frac{r_2(x_2)}{r_2(x_1)} \leq \frac{\Theta(\Lambda(x_2))}{\theta(\Lambda(x_2))}.$$

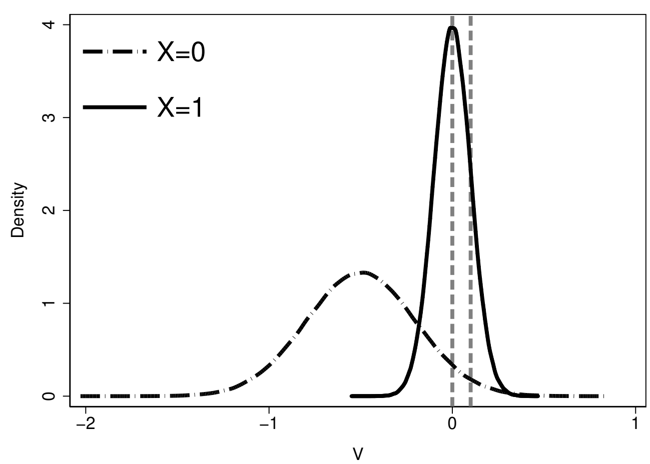
Log-concavity implies that the ratio is increasing in $\Lambda(X_i)$. Recalling that $\mathbb{E}[Release_i | X_i] = \mathbb{E}[1\{\Lambda(X_i) \geq \varepsilon_i\} | X_i] = \Theta(\Lambda(X_i))$ is also increasing in $\Lambda(X_i)$, we conclude the argument. $\square$

C Understanding A2

A2 is a sufficient but not necessary condition. In this appendix we provide examples of distributions that violate A2 and illustrate that the deviations from A2 have to be large to invalidate our identification argument.[1] The examples are only illustrative so we do not claim that their conclusions extend to more general settings.

Example 1: A specific case against P-BOT Let X_i and V_i be scalar, with $X_i \in \{0,1\}$ and $V_i|X_i = x \sim \mathcal{N}(\mu_x, \sigma_x^2)$. Assume $Release_i = 1\{V_i \geq 0\}$, so A1 is trivially satisfied. We set $\mu_0 = -0.5$, $\mu_1 = 0$, $\sigma_0 = 0.3$, and $\sigma_1 = 0.1$. We define marginally released individuals as individuals with $V_i \in [0, 0.1]$. Figure C.I displays the conditional densities for simulated data. The share of marginal among released is larger for $X_i = 1$ (68.2%) than for $X_i = 0$ (52.1%). Then, $\sigma_0 > \sigma_1$ violates A2.

Figure C.I: Example 1



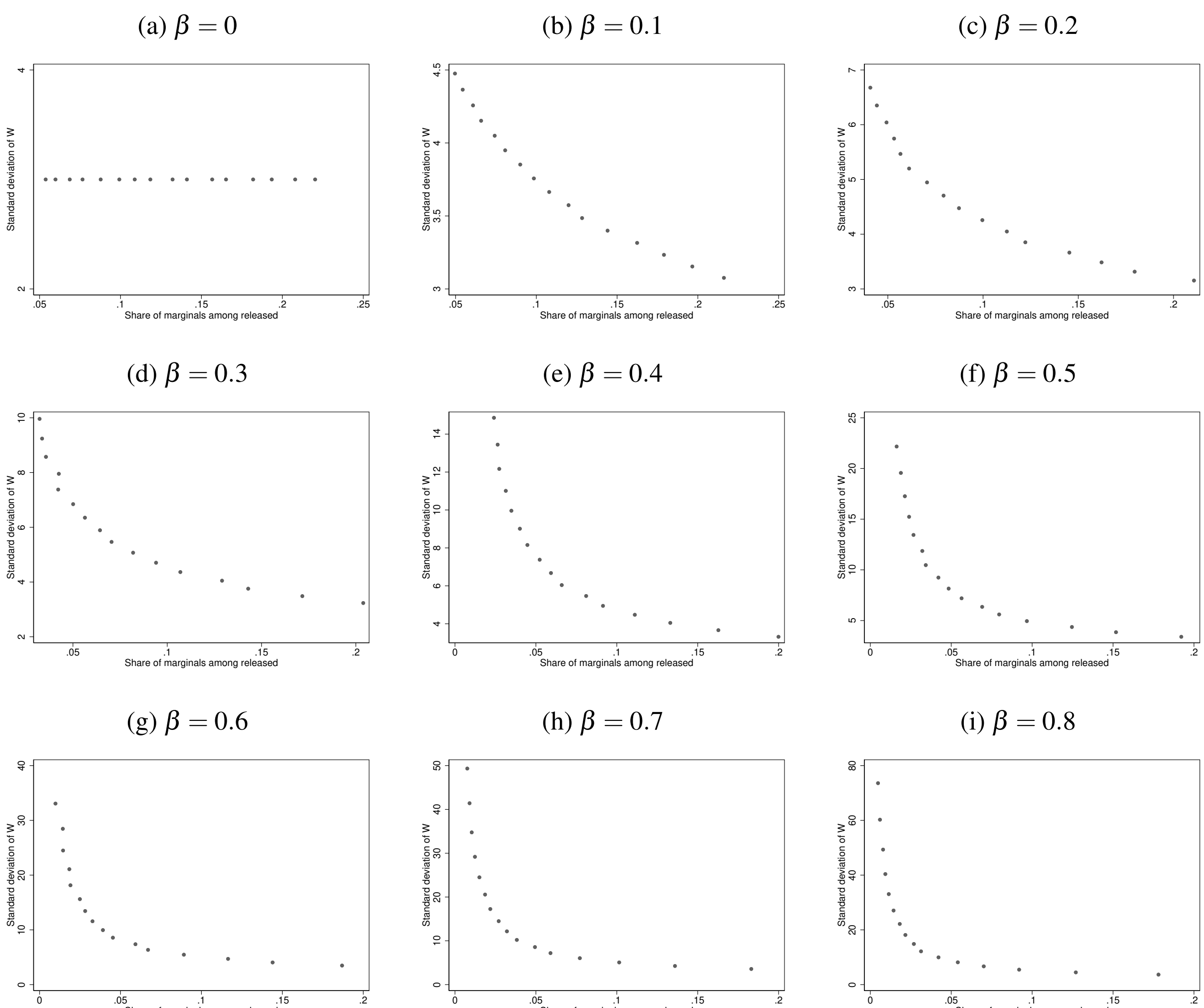
While $Release_i^*$ does not explicitly depends on X_i, it is still the case that X_i is a good predictor given the high correlation with V_i. Since $\mu_0 < \mu_1$, $\mathbb{E}[Release_i|X_i = 0] < \mathbb{E}[Release_i|X_i = 1]$. Then, this specific pattern of heterokedasticity invalidates the identification argument: the X_i realization that induces a lower propensity score also has a smaller share of marginals among released. To shed some light on the intuition of this result and its connection to the violation of A2 (in particular, to the violation of $r_2(X_i)$ monotonicity restriction), notice that while $\mathbb{E}[Release_i|X_i = 1]$ is equal to 0.5 for any value of σ_1, it is possible to find a small enough value of σ_1 such that the share of marginal among released for $X_i = 1$ is arbitrarily close to 1.

[1] We thank Chris Walters for suggesting these examples.

This example is very specific and unfavorable for our case. In fact, simulations show that setting, for example, $\sigma_1 = 0.2$ reestablishes A2. To get a slightly more general intuition on how to think about A2, below we provide a more complex example.

Example 2: A2 violations have to be large Let X_i be scalar, with $X_i \in \{0.25, 0.5, ..., 3.75, 4\}$. Let $W_i = 1.5(X_i - 2) + 3\exp(\beta X_i)\zeta_i$, with $\zeta_i \sim \mathcal{N}(0,1)$. Let $Release_i = 1\{W_i \geq 0\}$ and define as marginals individuals with $W_i \in [0, 0.5]$. We simulate the model for $\beta \in \{0, 0.1, 0.2, ..., 0.6, 0.7\}$. Figure C.II shows that this model (i) violates A2 since $r_2(X_i) = 3\exp(\beta X_i)$ decreases with the share of marginals given X_i (except for $\beta = 0$), and (ii) higher values of β induce steeper functions, i.e., the larger the β, the stronger the deviation from A2.

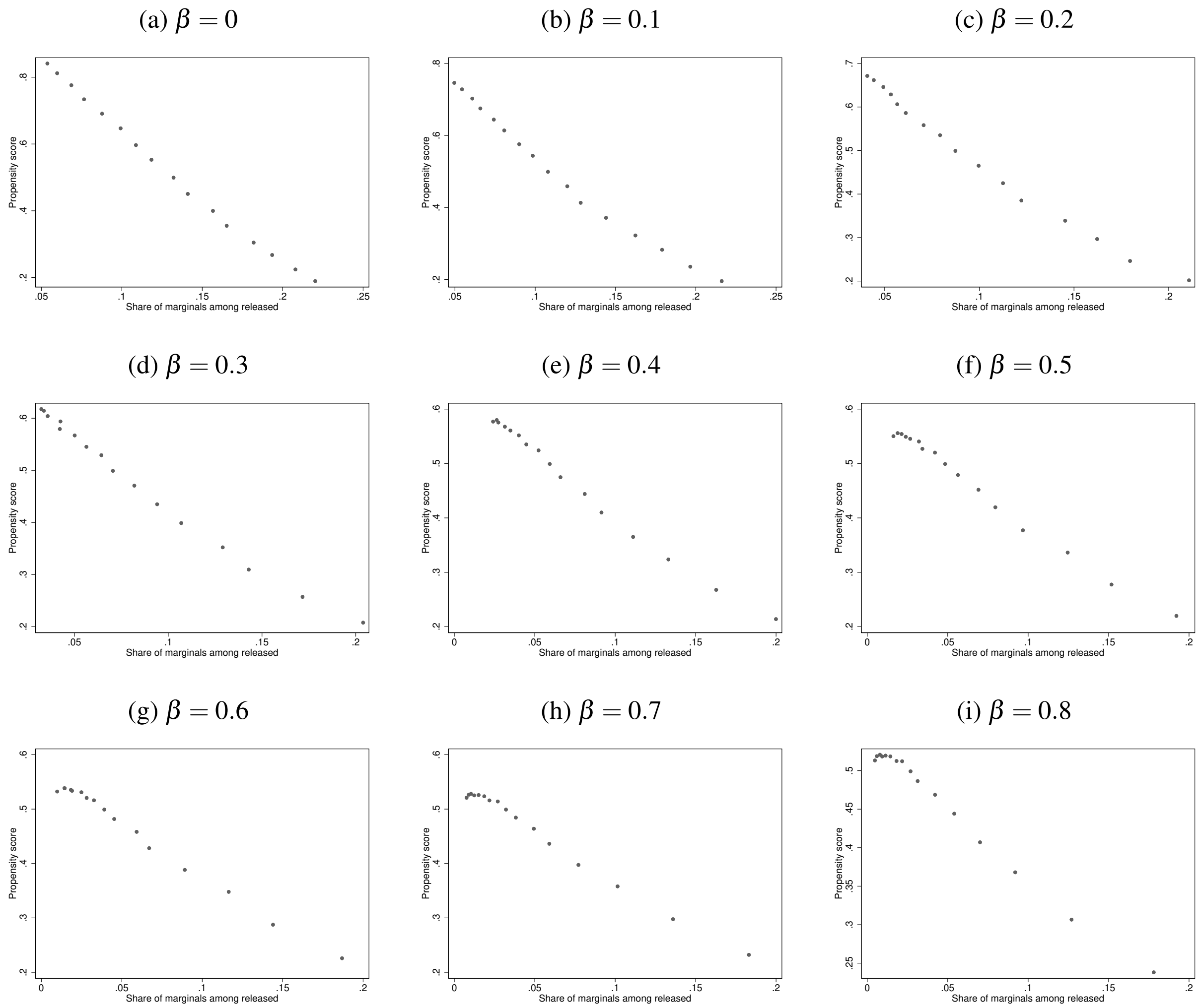
Figure C.II: $r_2(X_i)$ as a Function of the Share of Marginal Defendants Among Released



Notes: These figures plot the relationship between $r_2(X_i)$ and the share of marginal defendants among released, for different levels of β.

Figure C.III shows, for each value of β, the simulated relationship between the share of marginals among released and the propensity score. Figures overall suggest that the identification argument holds in this setting. We only see problems in the ranking (i) when β is large, and (ii) for observations that are far from the margin in expectation. Then, this example suggest that regular deviations from A2 should not be problematic for the P-BOT application.

Figure C.III: Propensity Score as a Function of the Share of Marginal Defendants Among Released



Notes: These figures plot the relationship between the propensity score and the share of marginal defendants among released, for different levels of β.

D Monte Carlo Simulations

This section presents the results of different Monte Carlo simulations. The objective of this exercise is twofold. First, it shows that when both assumptions (A1 and A2) are met, the presence of correlation between the observed and unobserved variables does not affect identification. Indeed, the P-BOT increases its precision when the correlation is large. Second, it shows that the P-BOT converges to a less precise version of the full sample outcome test (Knowles et al., 2001) when the conditional variance goes to infinity. Intuitively, when the observable component has no predictive power, the P-BOT is essentially an OLS regression of a random subsample of released defendants.

To these purposes, we simulate the model using the following equations:

$$
\begin{aligned}
Release_i &= 1\{(\alpha_X+\delta_X)X_i+(\alpha_W+\delta_W)W_i+\delta_R R_i \leq \beta_0-\beta_R R_i\}, \\
PM_i &= 1\{\alpha_X X_i+\alpha_W W_i+\varepsilon_i > \gamma_0\}\cdot Release_i, \\
\varepsilon_i &= \delta_X X_i+\delta_W W_i+\delta_R R_i+\nu_i, \\
W_i &= (X_i+R_i)\beta_W+\zeta_i,
\end{aligned}
$$

where X_i and W_i are defendant i's characteristics other than race, R_i is defendant i's race, $\nu_i \sim \mathcal{N}(0,\sigma_\nu^2)$, and $\zeta_i \sim \mathcal{N}(0,\sigma_\zeta^2)$. We assume that the judge observes both X_i and W_i, but the econometrician only observes X_i. As argued in the paper, β_0 is a leniency measure, β_R measures racial taste-based discrimination, and δ_R measures racial statistical discrimination. For simplicity, we assume that judges are homogeneous, i.e., β_0 and β_R are not function on $j(i)$. Note that both A1 and A2 hold in this setting. The parameter of interest is β_R.

The two set of simulations have the following random structure:

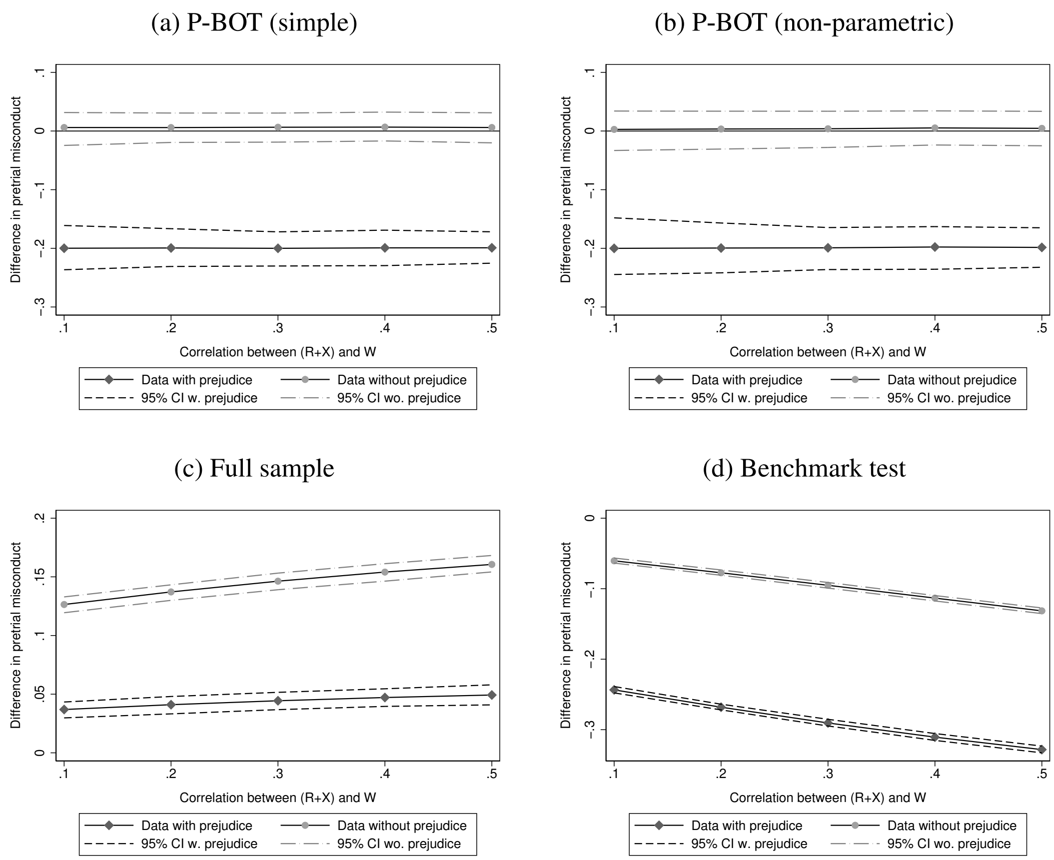
$$
\begin{pmatrix} X_i \\ R_i^* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ -0.5 \end{pmatrix}, \begin{pmatrix} 1 & 0.15 \\ 0.15 & 1 \end{pmatrix}\right),
$$

where R_i^* is a latent variable such that $R_i = 1\{R_i^* \geq 0\}$. We simulate the model using $\alpha_X = \alpha_W = \delta_X = \delta_W = \delta_R = 0.1$, $\beta_0 = 0.5$, $\gamma_0 = 0.1$, and $\sigma_\nu = 0.1$. We provide simulations for a model *without discrimination* (i.e., $\beta_R = 0$) and *with discrimination*, with $\beta_R = 0.2$.

The first set of simulations sets $\sigma_\zeta = 1$ and tests the performance of the P-BOT for different values of β_W, which measures the correlation between the observed and unobserved variables. In particular, we consider $\beta_W \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. For estimating the P-BOT, we compute conditional predicted release probabilities and run OLS regressions of PM_i on R_i on samples of released defendants with predicted probabilities up to the 5th percentile. We also perform the non-parametric estimation evaluated at the 1st percentile. The predicted probability is estimated using a probit model of $Release_i$ on X_i and R_i. We compare the P-BOT to other models that are likely to be affected by the magnitude of β_W. Specifically, we run OLS regressions of PM_i on R_i using the complete sample of released defendants (outcome test with full sample). We also estimate probit regressions of PM_i on R_i and X_i and report the coefficient on R_i (benchmark test).

Figure D.I shows the results. The point estimates are the mean estimate across 200 Monte Carlo simulations, and the confidence intervals are formed using the 2.5 and 97.5 percentiles of the estimated models. The figure shows that the P-BOT correctly identifies β_R regardless of the value of β_W. Moreover, as β_R increases, the precision of the estimation increases. This is consistent with the discussion in the main text. Importantly, these correlation values are large enough to make both the model subject to substantial inframarginality bias (and, therefore, strongly affecting the performance of the outcome test using the full sample) and to omitted variable bias in the release equation (and, therefore, strongly affecting the performance of the benchmark test).

Figure D.I: Tests' Performance for Different Values of β_W

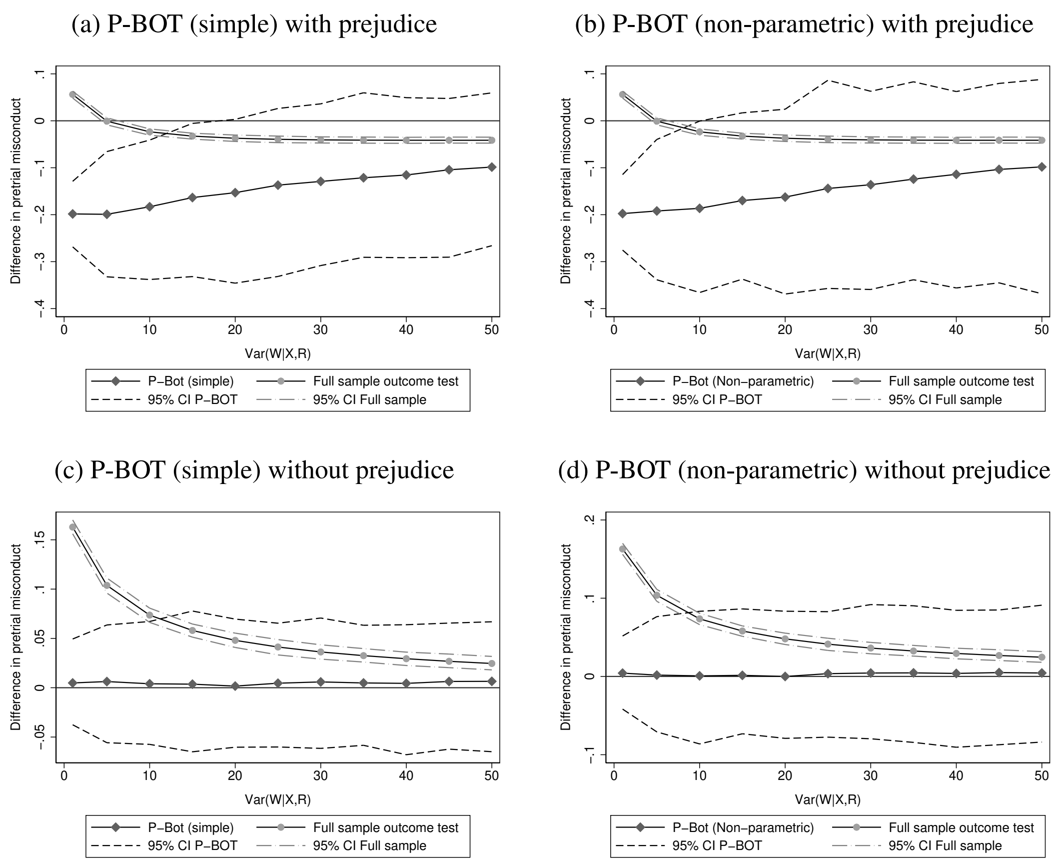


Notes: These figures plot the performances of P-BOT and alternative tests to test discrimination for different levels of correlation between observables and unobservables. The P-BOT is implemented using both approaches explained in Section 4 of the paper using the 5th percentile of the release probability as the threshold to define marginal defendants. *Full sample* is the outcome test considering the full sample of released defendants. *Benchmark test* reports the R coefficient of an OLS regression of release on X and R.

The second set of simulations sets $\beta_w = 0.5$ and tests the performance of the P-BOT for different values of σ_ζ, which measures the conditional variance of the unobserved component. In particular, we consider $\sigma_\zeta^2 \in \{1,5,10,15,20,25,30,35,40,45,50\}$. To assess the magnitude of these variances, recall that $\sigma_X = 1$ and $\sigma_{R^*} = 0.1$.

Figure D.II shows the results. As before, the point estimates are the mean estimate across 200 Monte Carlo simulations, and the confidence intervals are formed using the 2.5 and 97.5 percentiles of the estimated models. The figures show that when σ_ζ increases, the P-BOT converges to a less precise version of Knowles et al. (2001) test. The reason is that large conditional variances decrease the performance of the prediction model. Then, as the relative predictive power of X_i and R_i decreases, the P-BOT ends essentially selecting a random set of the sample of released individuals.

Figure D.II: P-BOT versus Full Sample Outcome Test, for Different Values of σ_ζ^2



Notes: These figures plot the performances of the P-BOT and the full sample outcome test for different levels of σ_ζ^2. The P-BOT is implemented using both approaches explained in Section 4 of the paper using the 5th percentile of the release probability as the threshold to define marginal defendants. *Full sample* is the outcome test considering the full sample of released defendants.

E Data Appendix

This appendix gives a more detailed description of the data, the sample restrictions, and the construction of the variables.

E.1 Sources

We merge three different sources of data to build our database.

PDO administrative records We use administrative records from the Public Defender Office (PDO, see http://www.dpp.cl/). The PDO is a centralized public service under the oversight of the Ministry of Justice that provides criminal defense services to all individuals accused of or charged with a crime who lack an attorney. The centralized nature of the PDO ensures that the administrative records contain information for all the cases handled only by the PDO or in tandem with a private attorney (as opposed to by only private attorneys), which covers more than 95% of the universe of criminal cases of Chile. The unit of analysis is a criminal case and contains: defendants characteristics (ID, name, gender, self-reported ethnicity, and place of residence, among other characteristics) and case characteristics (case ID, court, public attorney assigned, initial and end dates, different categories for type of crime, pretrial detention status and length, and outcome of the case, among other administrative characteristics). We consider cases whose arraignment hearings occurred between 2008 and 2017.

Registry of judges In addition, we have access to detention judges and their assigned cases, for hearings that occurred between 2008 and 2017. We merge this registry with the administrative records using the cases' IDs. We do not observe other characteristics of the judges in addition to their names and IDs. This data was shared by the Department of Studies of the Chilean Supreme Court (https://www.pjud.cl/corte-suprema).

Mapuche surnames The registry of Mapuche surnames was provided by the Mapuche Data Project (http://mapuchedataproject.cl/). The Mapuche Data Project is an interdisciplinary project that seeks to identify, digitalize, compile, process, and harmonize quantitative information of the Mapuche people for research and policy purposes. The surnames registry, one of the several datasets publicly available in their website, contains 8,627 different Mapuche surnames. The identification is based on the works of Amigo and Bustos (2008) and Painemal (2011). Since we observe names and surnames in the PDO records, we can directly identify defendants with Mapuche surnames.

E.2 Estimation sample

The initial sample contains $3,571,230$ cases and covers all the cases recorded by the PDO whose arraignment hearing occurred between 2008 and 2017. To create our estimation sample, we make the following adjustments.

Basic data cleaning Due to potential miscoding, we drop observations where the initial date of the case is later than the end date, and observations where the length of pretrial detention is larger than the length of the case. After these adjustments, the sample size reduces to $3,559,019$ (i.e, we drop $12,211$ cases).

Sample restrictions We then make the following sample restrictions:

- We exclude hearings due to legal summons ($1,233,909$ observations). We do this because the information set of the judges is likely to be different.
- We drop juvenile defendants ($254,243$ observations). We do this because the juvenile criminal system works differently, so the mandated selection rule and the preventive measures differ between systems (see Cortés, Grau, and Rivera, 2019 for details).
- We drop cases where the defendant hires a private attorney as his exclusive defender ($103,092$ observations). We do this because we do not observe the result of the arraignment hearing (and what happens after in the prosecution) in these cases.
- We drop cases whose length is larger than two years ($55,495$ observations).
- For defendants that are accused of more than one crime in a given case and, therefore, the records provide multiple observations, we consider the most severe crime (see below the severity definition). In this step we drop $193,720$ observations. To be clear, in this step we do not drop defendants, but only cases. We do this to have only at most one case/defendants pair per day of arraignment hearing.
- We drop cases where the detention judge is missing ($67,440$ observations).

- We drop types of crime whose likelihood of pretrial detention is less than 5% ($945,753$ observations). We do this because we want to study judges' decisions in cases where pretrial detention is a plausible outcome.
- We drop cases handled by judges that see less than 10 cases in the whole time-period ($2,846$ observations). We also only consider cases whose public attorney defended at least 10 cases, but we do not drop any data because of this restriction.
- We drop defendants from ethnic groups different than Mapuche ($2,789$ observations).

After all these adjustments the sample size is $699,732$. That matches the numbers of Table 1.

E.3 Variables

Many of the variables used in our empirical application are directly contained in the administrative records. Here we describe how we construct the other variables.

- *Mapuche*: we build four indicators of Mapuche combining self-reporting and surnames information. See Section 4 for details.
- *Severity*: we proxy crime severity by computing the share of cases within the type of crime that use pretrial detention.
- *Criminal record*: we can track all arrests of a given defendant using their IDs. Then, the variables *Previous prosecution*, *Number of previous prosecutions*, *Previous pretrial misconduct*, *Previous conviction*, and *Severity of previous prosecution* are constructed by looking at the characteristics of the cases associated to the defendant's ID that were initiated before the current one. For individuals with no previous prosecutions, these variables are set to zero. For building these variables, we can track cases up to 2005.
- *Pretrial misconduct*: pretrial misconduct is an indicator variable that takes value 1 if the defendant do not return to a scheduled hearing and/or is engaged in pretrial recidivism. Non-appearance in court is recorded in the administrative data. Pretrial recidivism is built by looking at arrests associated to the same defendant's ID whose initial date is between the initial and end dates of the current prosecution.
- *Attorney quality and judge leniency*: as in Dobbie, Goldin, and Yang (2018), we use the residualized (against court-by-time fixed effects) leave-out mean release rate.
- *Year of prosecution fixed effects*: we consider the initial date to set the fixed effects.

F Assessing Assumptions' Validity

In this appendix, we provide suggestive evidence that the identification strategy is valid in our setting. It has to be kept in mind that these assumptions are not directly testable and, therefore, these tests, while reassuring, are only suggestive. We first study the common support assumption. Then, we assess the separability (monotonicity) assumption. Finally, we propose a diagnostic that assesses, in more general terms, the validity of the propensity score-based ranking argument.[2]

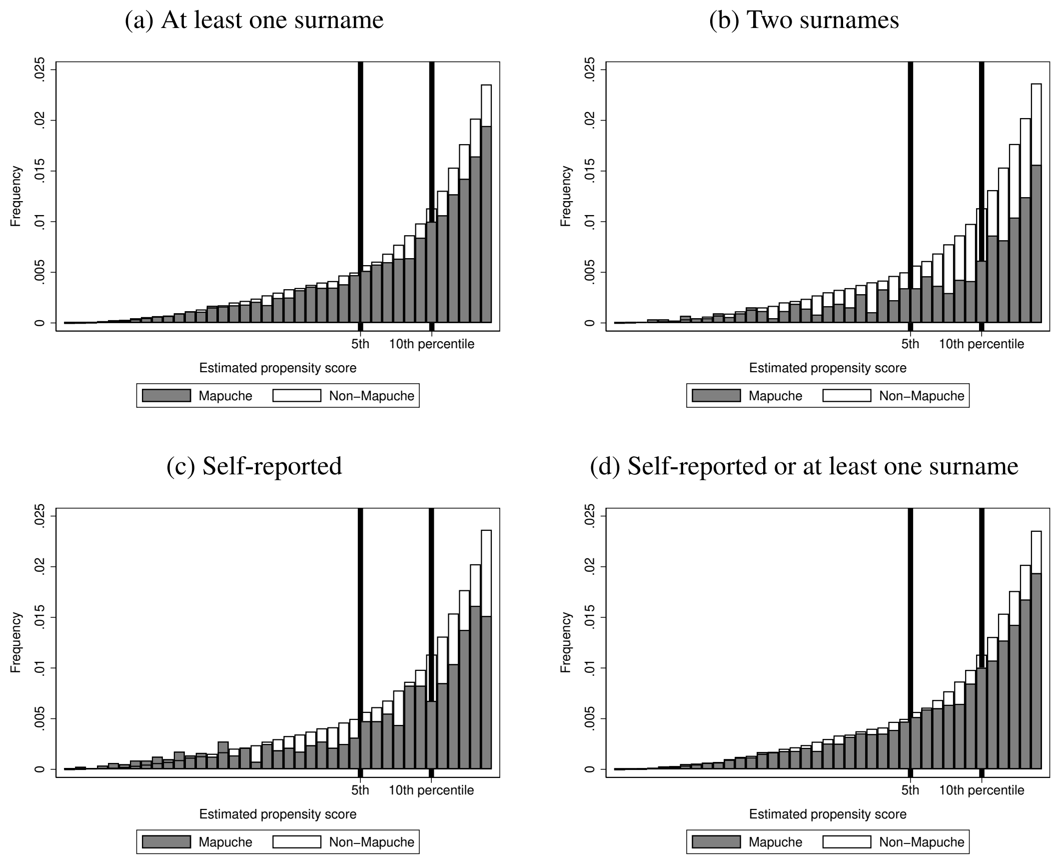
Assumption 0 Figure F.I shows the (estimated) propensity score distributions for released defendants, separating Mapuche and non-Mapuche defendants. The figures suggest that the continuity and full-support assumptions are met in our setting, especially for the more comprehensive Mapuche definitions.

Assumption 1 Recall that A1 says that there are functions d and g such that $1\{f(X_i, V_i) \geq 0\} = 1\{d(X_i) - g(V_i) \geq 0\}$. This implies that the direction in which X_i affects the likelihood of being released is not affected by the value of V_i. One way to assess this assumption is to check whether the coefficients of a regression of $Release_i$ on X_i are stable (in terms of sign) when considering subsamples with (probably) different unobservables. Likewise, recall that, through the lens of the model, A1 implies monotonicity on observables in the expected risk equation. Then, a similar exercise can be done with the coefficients of a regression of PM_i on X_i among different subsamples of released defendants with (probably) different unobservables. This test is similar to the monotonicity tests performed by Arnold, Dobbie, and Yang (2018) and Bald et al. (2019).

Tables F.I and F.II show the results using $Release_i$ and PM_i as dependent variables, respectively. Each cell reports the estimated coefficient of the regressor specified in the column, using the sample specified in the first column. Each row represents a different estimation. The first row reports the coefficients using the whole sample, and then rows are paired by mutually exclusive sample categories that are (probably) characterized by different unobservables. For example, row 2 shows results for the Mapuche subsample, while row 3 shows results for the non-Mapuche subsample. Then, rows 4 and 5 split the sample by gender, and so on. Results strongly support the monotonicity assumption. In all but two cases (i.e., 96% of cases) the sign of the coefficient is consistent across samples. Moreover, the magnitudes are also similar. This suggests that the direction of the effect of observables is unlikely to be affected by the unobserved variables.

[2] This can be interpreted as a joint test for A1 and A2. However, since A2 is not necessary, this test could be well-behaved without necessary meeting A2.

Figure F.I: Propensity Score Histograms (up to the 20th percentile)



Note: These plots show the propensity score histograms for Mapuche and Non-Mapuche released defendants. The two vertical lines represent the 5th and 10th percentile of the distribution. For presentation purposes, we only show each histogram up to the 20th percentile. However, histograms are calculated considering the entire population of released defendants.

Ranking validity This test builds on the intuition of Altonji, Elder, and Taber (2005) and Oster (2019).[3] Recall that V_i are variables that the judges observe, so X_i can be interpreted as elements of V_i that the econometrician happened to see. Then, we can use observed variables to simulate unobservables and assess the validity of the identification argument.

We perform the following exercise. Assume that our set of observed variables, X_i, is a good approximation (up to some small well-behaved noise) of the judges' (complete) information set. Under that assumption, the identification of marginally released defendants using the ranking based on the propensity score is accurate. We fit the propensity score and label as marginal the bottom 5% of the predicted probability distribution (among released defendants). Then, we omit one observable (label it as V_i) and (i) estimate the propensity score with the restricted set of observables

[3]Their methodologies are not exactly suitable to our setting since (i) we allow for standard omitted variable bias, and (ii) we do not require the estimated coefficients of the selection equation to have causal interpretation.

Table F.I: Testing for Monotonicity in Observables (Dep. Variable: Release Status)

Estimation sample	Previous case	Previous pretrial misconduct	Previous conviction	Severity previous case	Severity current case
All	-0.029	-0.027	-0.015	-0.110	-0.753
	[-0.034, -0.024]	[-0.029, -0.025]	[-0.020, -0.010]	[-0.116, -0.104]	[-0.758, -0.749]
Mapuche	-0.028	-0.025	-0.014	-0.091	-0.742
	[-0.046, -0.010]	[-0.032, -0.017]	[-0.031, 0.004]	[-0.112, -0.069]	[-0.756, -0.727]
Non-Mapuche	-0.029	-0.027	-0.015	-0.112	-0.754
	[-0.035, -0.024]	[-0.029, -0.025]	[-0.020, -0.010]	[-0.118, -0.106]	[-0.758, -0.749]
Male	-0.031	-0.030	-0.015	-0.098	-0.763
	[-0.036, -0.025]	[-0.032, -0.027]	[-0.020, -0.009]	[-0.105, -0.092]	[-0.768, -0.758]
Female	-0.013	-0.005	-0.022	-0.242	-0.682
	[-0.025, 0.000]	[-0.011, 0.001]	[-0.034, -0.010]	[-0.262, -0.223]	[-0.695, -0.668]
Low income	-0.029	-0.024	-0.016	-0.106	-0.761
	[-0.037, -0.021]	[-0.027, -0.021]	[-0.024, -0.008]	[-0.116, -0.097]	[-0.768, -0.754]
High income	-0.030	-0.029	-0.015	-0.114	-0.748
	[-0.036, -0.023]	[-0.032, -0.026]	[-0.021, -0.008]	[-0.122, -0.106]	[-0.754, -0.743]
Low judge leniency	-0.030	-0.028	-0.017	-0.112	-0.778
	[-0.037, -0.022]	[-0.031, -0.025]	[-0.024, -0.010]	[-0.121, -0.103]	[-0.784, -0.772]
High judge leniency	-0.029	-0.025	-0.013	-0.109	-0.728
	[-0.036, -0.021]	[-0.028, -0.022]	[-0.020, -0.006]	[-0.117, -0.100]	[-0.734, -0.722]
Low attorney quality	-0.028	-0.027	-0.018	-0.122	-0.820
	[-0.035, -0.020]	[-0.031, -0.024]	[-0.025, -0.011]	[-0.130, -0.113]	[-0.826, -0.814]
High attorney quality	-0.030	-0.026	-0.012	-0.099	-0.687
	[-0.037, -0.024]	[-0.029, -0.023]	[-0.019, -0.006]	[-0.108, -0.091]	[-0.693, -0.681]
Small Court (No. of cases)	-0.019	-0.025	-0.019	-0.127	-0.789
	[-0.026, -0.011]	[-0.028, -0.022]	[-0.026, -0.012]	[-0.136, -0.118]	[-0.795, -0.783]
Big Court (No. of cases)	-0.039	-0.028	-0.012	-0.099	-0.721
	[-0.046, -0.032]	[-0.031, -0.025]	[-0.019, -0.006]	[-0.107, -0.091]	[-0.727, -0.715]
Small Court (No. of judges)	-0.020	-0.027	-0.017	-0.128	-0.795
	[-0.027, -0.012]	[-0.030, -0.024]	[-0.024, -0.009]	[-0.137, -0.119]	[-0.801, -0.789]
Big Court (No. of judges)	-0.038	-0.026	-0.015	-0.096	-0.713
	[-0.045, -0.031]	[-0.029, -0.023]	[-0.021, -0.008]	[-0.104, -0.088]	[-0.719, -0.707]
Low severity court	-0.032	-0.023	-0.011	-0.086	-0.634
	[-0.039, -0.026]	[-0.025, -0.020]	[-0.017, -0.005]	[-0.093, -0.078]	[-0.640, -0.628]
High severity court	-0.025	-0.031	-0.020	-0.137	-0.876
	[-0.033, -0.018]	[-0.034, -0.027]	[-0.027, -0.013]	[-0.146, -0.127]	[-0.882, -0.869]

Note: This table presents the results of the test for monotonicity in observables. Each reported value is the marginal effect of the variable of the column on the probability of release, estimated using a different sample in each row. The continuous variables were discretized using the respective median as the threshold. The values in parenthesis are 95% confident intervals, estimated using bootstrap with 500 repetitions.

and identify marginals using the ranking strategy, and (ii) compute the conditional probabilities of being marginal, namely the shares of marginals identified in the first step for different combinations of the observables used in the restricted estimation. We then compute the rank correlation between (i) the share of marginals using the restricted propensity-score ranking and the conditional probabilities, and (ii) the estimated propensity score using the restricted set of observables and the conditional probabilities of being marginal. In case (i), the correlation is expected to be positive. In case (ii), the correlation is expected to be negative. If the identification argument holds, we should expect these rank correlations to be large.

Table F.II: Testing for Monotonicity in Observables (Dep. Variable: Pretrial Misconduct)

Estimation sample	Previous case	Previous pretrial misconduct	Previous conviction	Severity previous case	Severity current case
All	0.073 [0.066, 0.080]	0.090 [0.087, 0.093]	0.035 [0.029, 0.042]	0.039 [0.029, 0.049]	0.034 [0.025, 0.044]
Mapuche	0.055 [0.030, 0.081]	0.082 [0.071, 0.093]	0.039 [0.014, 0.063]	0.059 [0.022, 0.096]	0.044 [0.011, 0.078]
Non-Mapuche	0.075 [0.068, 0.082]	0.091 [0.088, 0.094]	0.035 [0.028, 0.042]	0.037 [0.026, 0.048]	0.033 [0.024, 0.043]
Male	0.076 [0.068, 0.083]	0.092 [0.089, 0.095]	0.034 [0.027, 0.041]	0.044 [0.033, 0.055]	0.040 [0.030, 0.050]
Female	0.064 [0.044, 0.083]	0.076 [0.066, 0.085]	0.041 [0.022, 0.060]	-0.024 [-0.061, 0.012]	-0.012 [-0.039, 0.016]
Low income	0.069 [0.058, 0.080]	0.083 [0.078, 0.088]	0.038 [0.027, 0.049]	0.038 [0.022, 0.054]	0.076 [0.062, 0.090]
High income	0.075 [0.066, 0.084]	0.093 [0.089, 0.097]	0.034 [0.026, 0.043]	0.040 [0.026, 0.053]	0.001 [-0.012, 0.013]
Low judge leniency	0.064 [0.054, 0.074]	0.086 [0.082, 0.090]	0.044 [0.035, 0.054]	0.042 [0.027, 0.057]	0.033 [0.019, 0.046]
High judge leniency	0.083 [0.073, 0.093]	0.094 [0.090, 0.098]	0.027 [0.017, 0.036]	0.036 [0.022, 0.051]	0.036 [0.023, 0.049]
Low attorney quality	0.070 [0.060, 0.080]	0.094 [0.089, 0.098]	0.042 [0.032, 0.051]	0.052 [0.037, 0.067]	0.029 [0.016, 0.043]
High attorney quality	0.077 [0.067, 0.087]	0.087 [0.082, 0.091]	0.029 [0.019, 0.038]	0.026 [0.012, 0.041]	0.039 [0.026, 0.051]
Small Court (No. of cases)	0.062 [0.052, 0.072]	0.087 [0.083, 0.092]	0.036 [0.026, 0.046]	0.051 [0.036, 0.066]	0.092 [0.079, 0.106]
Big Court (No. of cases)	0.083 [0.074, 0.093]	0.090 [0.086, 0.094]	0.036 [0.027, 0.045]	0.031 [0.017, 0.045]	-0.013 [-0.025, 0.000]
Small Court (No. of judges)	0.075 [0.064, 0.085]	0.090 [0.086, 0.095]	0.029 [0.019, 0.039]	0.049 [0.034, 0.065]	0.059 [0.045, 0.072]
Big Court (No. of judges)	0.073 [0.064, 0.083]	0.086 [0.082, 0.090]	0.041 [0.032, 0.050]	0.030 [0.016, 0.044]	0.010 [-0.002, 0.023]
Low severity court	0.074 [0.064, 0.083]	0.084 [0.080, 0.088]	0.038 [0.028, 0.047]	0.038 [0.025, 0.052]	0.047 [0.035, 0.059]
High severity court	0.072 [0.061, 0.082]	0.095 [0.090, 0.099]	0.034 [0.024, 0.044]	0.038 [0.023, 0.054]	0.019 [0.005, 0.033]

Note: This table presents the results of the test for monotonicity in observables. Each reported value is the marginal effect of the variable of the column on pretrial misconduct, estimated using a different sample of released defendants in each row. The continuous variables were discretized using the respective median as the threshold. The values in parenthesis are 95% confident intervals, estimated using bootstrap with 500 repetitions.

We perform this exercise by using each of the 15 observables used in the estimation as V_i.[4] To compute the rank-correlations, we discretize the non-discrete regressors (using the median) to define $2^{(15-1)} = 16,384$ categories of observables. For each of these categories, we compute the average restricted estimated propensity score, the average share of marginals using the restricted propensity score, and the conditional probability of being marginal using the base estimation as the true share of marginals. Table F.III presents the results. We report both the Spearman's-ρ and

[4]Number of previous cases, severity of previous case, severity of current case, average severity by year-court, number of cases by year-court, judge leniency, jugde leniency squared, attorney quality, attorney quality squared, Mapuche indicator, gender, previous case indicator, previous pretrial misconduct indicator, previous conviction indicator.

Kendall's-τ statistics for rank correlation. It can be seen that in all variables by one (severity of current case), the correlations are very large. We interpret this as strong suggestive evidence of the validity of our identification argument.

Table F.III: Rank Correlations

	Corr. btw. $\Pr(Marg\|X=x, Release=1)$ and $\mathbb{E}[Marg\|X=x]$ using restricted p-score		Corr. btw. $\Pr(Marg\|X=x, Release=1)$ and $\mathbb{E}[Release\|X=x]$ using restricted p-score	
Excluded predictor	Spearman	Kendall	Spearman	Kendall
No of previous cases	0.966	0.946	-0.676	-0.553
Severity previous case	0.952	0.934	-0.691	-0.567
Severity current case	0.499	0.439	-0.491	-0.368
Average severity (year/court)	0.930	0.896	-0.709	-0.582
No of cases (year/court)	0.993	0.986	-0.707	-0.581
No of judges (year/court)	0.980	0.967	-0.707	-0.582
Judge leniency	0.976	0.964	-0.707	-0.581
Judge leniency square	1.000	0.999	-0.703	-0.578
Attorney quality	0.959	0.938	-0.704	-0.579
Attorney quality square	0.993	0.988	-0.705	-0.579
Mapuche	0.998	0.997	-0.725	-0.595
Male	0.997	0.996	-0.725	-0.596
Previous case	0.975	0.967	-0.702	-0.575
Previous pretrial misconduct	0.985	0.973	-0.688	-0.565
Previous conviction	0.996	0.993	-0.717	-0.588

Note: This table presents the rank-correlations between the ranking of the conditional probabilities of being marginal and (i) the ranking of the conditional share of marginals using the restricted propensity score estimation, and (ii) the ranking of the predicted propensity score using the restricted estimation. We report the Spearman's-ρ and the Kendall's-τ_b rank correlation statistics. The excluded predictor is specified in the first column. All regressions include year fixed effects. The unit of analysis to build the ranking is the combination of all possible values of the predictors, without considering the excluded category (i.e., 14 predictors), where the continuous predictors were transformed into binary variables by using the median among released as threshold. Then, each combination of predictors defines a cell, where the maximum number of cells is $2^{14} = 16,384$. Since there are cells without released defendants, in practice this number is between $4,449$ and $7,609$, depending on the excluded predictor.

G Prediction Models

Table G.I: Determinants of Release Probability Using a Probit Model (Marginal Effects)

	At least one Surname	Two Surnames	Self-Reported	Self-Reported or at least one surname
Mapuche	-0.004	-0.008	-0.013	-0.004
	(0.002)	(0.004)	(0.004)	(0.002)
Male	0.003	0.002	0.002	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Previous prosecution	-0.028	-0.029	-0.028	-0.028
	(0.003)	(0.003)	(0.003)	(0.003)
Previous pretrial misconduct	-0.028	-0.027	-0.028	-0.027
	(0.001)	(0.001)	(0.001)	(0.001)
Previous conviction	-0.015	-0.015	-0.015	-0.015
	(0.003)	(0.003)	(0.003)	(0.003)
No. of previous Prosecution	-0.007	-0.007	-0.007	-0.007
	(0.000)	(0.000)	(0.000)	(0.000)
Severity (previous prosecution)	-0.111	-0.112	-0.113	-0.111
	(0.004)	(0.004)	(0.004)	(0.004)
Severity (current prosecution)	-0.757	-0.756	-0.758	-0.757
	(0.008)	(0.008)	(0.008)	(0.008)
Average severity of the cases (court/year)	-1.021	-1.028	-1.030	-1.020
	(0.032)	(0.033)	(0.033)	(0.032)
No. of cases per court/year	-0.0000029	-0.0000030	-0.0000029	-0.0000028
	(0.0000007)	(0.0000007)	(0.0000007)	(0.0000007)
No. of judges per court/year	0.00030	0.00030	0.00030	0.00030
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Judge leniency	0.541	0.541	0.537	0.540
	(0.028)	(0.028)	(0.028)	(0.028)
Judge leniency squared	0.789	0.705	0.744	0.780
	(0.363)	(0.369)	(0.371)	(0.364)
Attorney quality	0.531	0.531	0.528	0.530
	(0.027)	(0.027)	(0.027)	(0.027)
Attorney quality squared	0.613	0.601	0.590	0.611
	(0.118)	(0.117)	(0.119)	(0.118)
Year of Prosecution fixed effects	YES	YES	YES	YES
Court fixed effects	NO	NO	NO	NO
No. of Mapuche	50,818	9,710	9,423	52,002
No. of Non-Mapuche	647,730	647,730	647,730	647,730
pseudo-R-squared	0.23	0.23	0.23	0.23
Correctly classified (0.5 prob as threshold)	0.85	0.85	0.85	0.85
Correctly classified (prediction: Non-Released)	0.60	0.60	0.59	0.60
Correctly classified (prediction: Released)	0.87	0.87	0.87	0.87

Note: This table presents the marginal effects of a probit model for the determinants of the release status using the data described in Table 1. The standard errors (in parenthesis) are clustered at the year/court level. The four models correspond to the four definitions of Mapuche considered in this paper.

Table G.II: Determinants of Release Probability Using a Linear Probability Model

	At least one Surname	Two Surnames	Self-Reported	Self-Reported or at least one surname
Mapuche	-0.002	-0.007	-0.006	-0.002
	(0.002)	(0.003)	(0.004)	(0.002)
Male	-0.002	-0.003	-0.003	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Previous prosecution	-0.002	-0.002	-0.002	-0.002
	(0.002)	(0.003)	(0.003)	(0.002)
Previous pretrial misconduct	-0.024	-0.024	-0.025	-0.024
	(0.001)	(0.001)	(0.001)	(0.001)
Previous conviction	-0.014	-0.014	-0.014	-0.014
	(0.002)	(0.002)	(0.002)	(0.002)
No. of previous prosecution	-0.009	-0.009	-0.009	-0.009
	(0.000)	(0.000)	(0.000)	(0.000)
Severity (previous prosecution)	-0.160	-0.162	-0.162	-0.160
	(0.005)	(0.005)	(0.005)	(0.005)
Severity (current prosecution)	-1.012	-1.008	-1.009	-1.012
	(0.011)	(0.011)	(0.011)	(0.011)
Average severity of the cases (court/year)	-1.060	-1.066	-1.069	-1.061
	(0.028)	(0.029)	(0.029)	(0.028)
No. of cases per court/year	-0.0000048	-0.0000050	-0.0000051	-0.0000048
	(0.0000018)	(0.0000018)	(0.0000018)	(0.0000018)
No. of judges per court/year	-0.00013	-0.00013	-0.00013	-0.00013
	(0.00007)	(0.00008)	(0.00008)	(0.00007)
Judge leniency	0.558	0.558	0.553	0.557
	(0.026)	(0.027)	(0.027)	(0.026)
Judge leniency squared	0.552	0.482	0.535	0.548
	(0.340)	(0.352)	(0.352)	(0.340)
Attorney quality	0.527	0.528	0.525	0.527
	(0.020)	(0.020)	(0.020)	(0.020)
Attorney quality squared	-0.069	-0.078	-0.087	-0.070
	(0.087)	(0.089)	(0.089)	(0.087)
Year of Prosecution fixed effects	YES	YES	YES	YES
Court fixed effects	YES	YES	YES	YES
No. of Mapuche	50,818	9,710	9,423	52,002
No. of Non-Mapuche	647,730	647,730	647,730	647,730
R-squared	0.22	0.22	0.21	0.22
Correctly classified (0.5 prob as threshold)	0.85	0.85	0.85	0.85
Correctly classified (prediction: Non-Released)	0.65	0.65	0.64	0.65
Correctly classified (prediction: Released)	0.86	0.86	0.86	0.86

Note: This table presents the point estimates of a linear probability model for the determinants of the release status using the data described in Table 1. The standard errors (in parenthesis) are clustered at the year/court level. The four models correspond to the four definitions of Mapuche considered in this paper.

Table G.III: Determinants of Release Probability Using a Heteroskedastic Probit Model

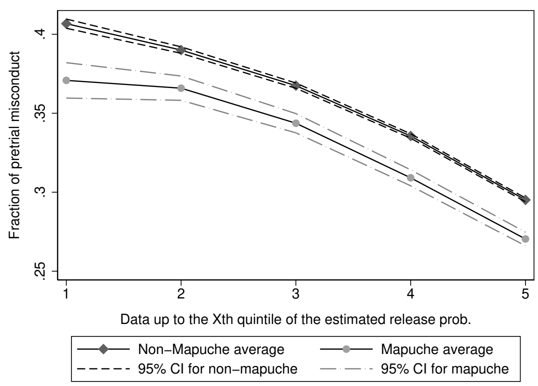
	At least one Surname	Two Surnames	Self-Reported	Self-Reported or at least one surname
Mapuche	-0.042	-0.115	-0.146	-0.043
	(0.014)	(0.030)	(0.031)	(0.014)
Male	0.092	0.081	0.082	0.091
	(0.028)	(0.029)	(0.029)	(0.028)
Previous prosecution	-0.142	-0.150	-0.139	-0.143
	(0.061)	(0.064)	(0.064)	(0.061)
Previous pretrial misconduct	-0.161	-0.160	-0.166	-0.160
	(0.022)	(0.023)	(0.023)	(0.022)
Previous conviction	-0.340	-0.331	-0.343	-0.339
	(0.056)	(0.056)	(0.057)	(0.055)
No. of previous Prosecution	-0.045	-0.045	-0.044	-0.045
	(0.002)	(0.002)	(0.002)	(0.002)
Severity (previous prosecution)	-0.769	-0.771	-0.767	-0.771
	(0.063)	(0.065)	(0.065)	(0.063)
Severity (current prosecution)	-5.640	-5.590	-5.599	-5.647
	(0.226)	(0.231)	(0.233)	(0.226)
Average severity of the cases (court/year)	-7.965	-7.959	-7.958	-7.956
	(0.315)	(0.321)	(0.324)	(0.317)
No. of cases per court/year	-0.000035	-0.000035	-0.000034	-0.000034
	(0.000013)	(0.000014)	(0.000014)	(0.000014)
No. of judges per court/year	0.003434	0.003296	0.003260	0.003436
	(0.000801)	(0.000803)	(0.000804)	(0.000802)
Judge leniency	6.220	6.269	6.190	6.232
	(0.532)	(0.539)	(0.536)	(0.533)
Judge leniency squared	13.879	13.494	13.518	13.876
	(3.474)	(3.520)	(3.517)	(3.474)
Attorney quality	5.568	5.555	5.525	5.575
	(0.365)	(0.374)	(0.371)	(0.365)
Attorney quality squared	6.543	6.473	6.368	6.546
	(0.916)	(0.928)	(0.931)	(0.917)
Conditional variance:				
Male	0.062	0.058	0.057	0.062
	(0.015)	(0.015)	(0.015)	(0.015)
Previous prosecution	0.055	0.052	0.057	0.055
	(0.031)	(0.033)	(0.033)	(0.031)
Previous pretrial misconduct	0.026	0.024	0.022	0.026
	(0.013)	(0.014)	(0.014)	(0.013)
Previous conviction	-0.148	-0.145	-0.150	-0.147
	(0.028)	(0.029)	(0.029)	(0.028)
No. of previous Prosecution	0.017	0.017	0.017	0.017
	(0.001)	(0.001)	(0.001)	(0.001)
Severity (previous prosecution)	0.221	0.226	0.228	0.220
	(0.035)	(0.036)	(0.036)	(0.035)
Severity (current prosecution)	1.335	1.338	1.344	1.337
	(0.054)	(0.055)	(0.055)	(0.054)
Average severity of the cases (court/year)	0.566	0.588	0.598	0.576
	(0.181)	(0.186)	(0.186)	(0.182)
No. of cases per court/year	-0.000002	-0.000002	-0.000001	-0.000002
	(0.000007)	(0.000007)	(0.000007)	(0.000007)
No. of judges per court/year	0.000643	0.000585	0.000581	0.000640
	(0.000334)	(0.000340)	(0.000339)	(0.000334)
Judge leniency	1.126	1.189	1.163	1.137
	(0.248)	(0.253)	(0.252)	(0.248)
Attorney quality	0.779	0.798	0.804	0.784
	(0.113)	(0.114)	(0.114)	(0.114)

Note: This table presents the point estimates of a probit model for the determinants of the release status using the data described in Table 1 and the point estimates for the relationship between covariates and the variance of the unobservable component (modeled as $\exp(X\beta)$). The standard errors (in parenthesis) are clustered at the year/court level. The four models correspond to the four definitions of Mapuche considered in this paper.

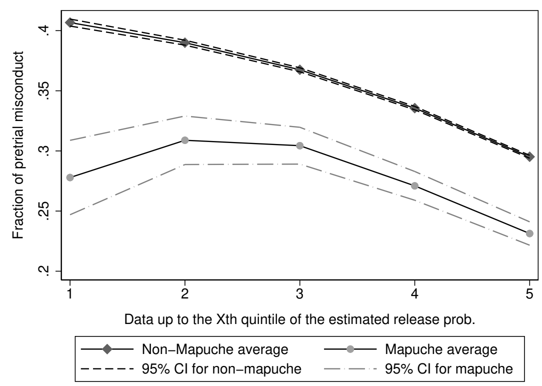
H Main results using alternative Mapuche definitions

Figure H.I: Pretrial Misconduct Rates for Different Quintiles of the Predicted Release Probability

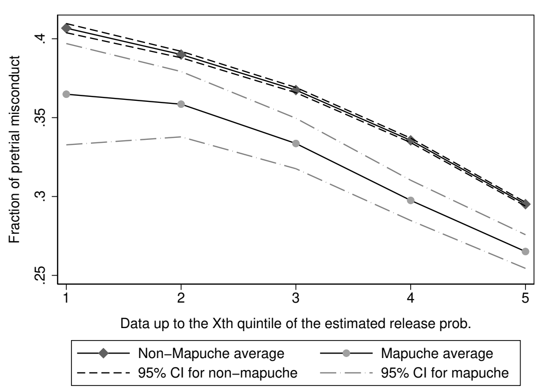
(a) At least one surname



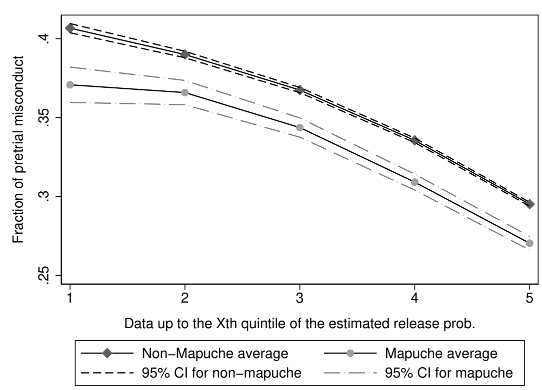
(b) Two surnames



(c) Self-reported



(d) Self-reported or at least one surname



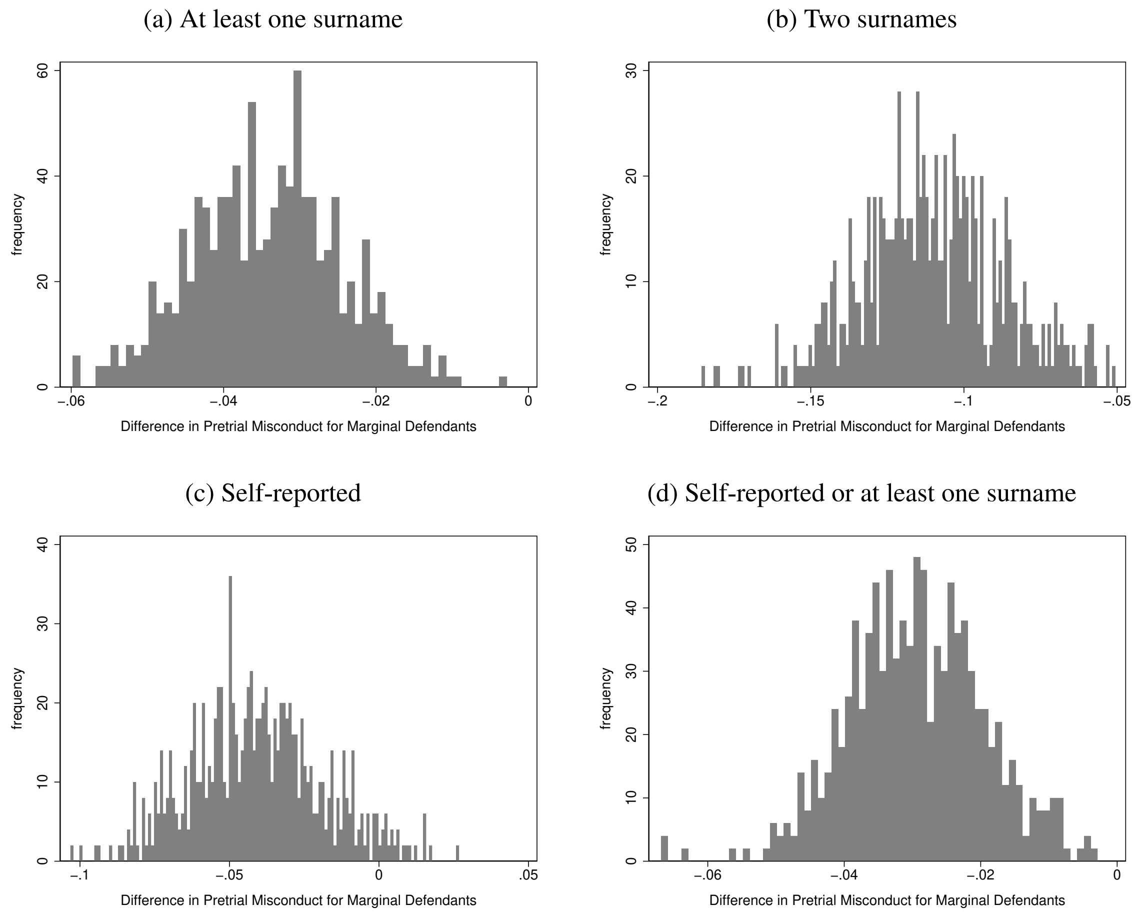
Note: These plots present the Mapuche and non-Mapuche pretrial misconduct rates for different groups of predicted release probability quintiles (1: quintile 1; 2: quintiles 1-2; 3: quintiles 1-3; 4: quintiles 1-4; 5: full sample). Predictions are estimated using a probit model. Each plot presents the results for one of the four definitions of Mapuche. Confidence intervals are analytically calculated assuming that quintiles are given. Pretrial misconduct accounts for non-appearance in court and/or pretrial recidivism.

Table H.I: Prediction-Based Outcome Test, Using Probit to Estimate the Release Probability (Outcome: Pretrial Misconduct)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.045	-0.145	0.069	-0.040
C.I. (95%)	[-0.070, -0.025]	[-0.197,-0.080]	[-0.119,-0.017]	[-0.064, -0.020]
(a) Mapuche expectation	0.363	0.264	0.340	0.368
(b) Non-Mapuche expectation	0.408	0.408	0.408	0.408
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.035	-0.138	-0.071	-0.031
C.I. (95%)	[-0.061,-0.009]	[-.212,-0.068]	[-0.128,-0.013]	[-0.058, -0.007]
(a) Mapuche expectation	0.390	0.287	0.354	0.393
(b) Non-Mapuche expectation	0.425	0.425	0.425	0.425
No. of Mapuche ($\leq$ 5th pctl.)	1,916	269	321	1,986
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,231	27,241	27,166	27,299
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.043	-0.160	-0.040	-0.040
C.I. (95%)	[-0.057, -0.026]	[-0.200, -0.116]	[-0.079, -0.001]	[-0.054, -0.024]
(a) Mapuche expectation	0.361	0.243	0.363	0.363
(b) Non-Mapuche expectation	0.403	0.404	0.403	0.403
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.040	-0.149	-0.064	-0.036
C.I. (95%)	[-0.060, -0.020]	[-0.197, -0.097]	[-0.116, -0.016]	[-0.056, -0.017]
(a) Mapuche expectation	0.371	0.262	0.347	0.375
(b) Non-Mapuche expectation	0.411	0.411	0.411	0.411
No. of Mapuche ($\leq$ 10th pctl.)	3,774	497	636	3,901
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,699	54,523	54,338	54,669

Note: This table presents the results from the P-BOT using the data described in Table 1, considering two approaches to estimate the outcome equation and two criteria to determine who is the margin. Release probabilities are predicted using a probit model. The outcome is any pretrial misconduct. Panel A shows the estimates using the simple approach, considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using the non-parametric approach. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. Details of the covariates included in the prediction model can be found in Appendix G. The confidence intervals are calculated using bootstrap with 500 repetitions.

Figure H.II: Perturbation Test



Note: These plots present the results of the perturbation test described in Section 3. They are produced in the following steps. First, we estimate the probit model. Then, for each released individual in the sample, we simulate 500 realizations from a standardized normal distribution to simulate $Release_i^*$ and redefine the samples of marginal individuals. Within each sample, we estimate the outcome test and plot its distribution across simulations.

Table H.II: Alternative Tests for Prejudice

	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Outcome test (full sample):				
Coeff.	-0.023	-0.059	-0.026	-0.023
Robust SE	(0.003)	(0.006)	(0.010)	(0.003)
Observations	698,548	657,440	657,153	699,732
IV-Outcome test:				
Mapuche coeff.	0.418	-0.153	12.688	0.240
Mapuche robust SE	(0.527)	(0.288)	(141.0)	(0.478)
Non-Mapuche coeff.	0.363	0.363	0.363	0.363
Non-Mapuche robust SE	(0.059)	(0.059)	(0.059)	(0.059)
No. of Mapuche	49,570	8,055	7,853	50,802
No. of non-Mapuche	647,701	647,701	647,701	647,701

Note: This table presents the results from alternative tests for prejudice using the data described in Table 1. The outcome is any pretrial misconduct. The outcome test using the full sample reports the estimated coefficient of an OLS regression of pretrial misconduct on a Mapuche indicator. Following Arnold, Dobbie, and Yang (2018), the IV-outcome test reports the coefficient of a 2SLS regression of pretrial misconduct on release, instrumenting release with the residualized leave-out mean release rate of the assigned judge. In the IV estimation, standard errors are clustered at the year/court level.

I Robustness Checks

Table I.I: Prediction-Based Outcome Test, Using OLS to Estimate the Release Probability (Outcome: Pretrial Misconduct)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.037	-0.130	-0.052	-0.031
C.I. (95%)	[-0.058, -0.015]	[-0.176, -0.077]	[-0.102, -0.000]	[-0.053, -0.010]
(a) Mapuche expectation	0.349	0.256	0.334	0.354
(b) Non-Mapuche expectation	0.385	0.386	0.386	0.385
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.033	-0.140	-0.068	-0.030
C.I. (95%)	[-0.060, -0.007]	[-0.198, -0.075]	[-0.131, -0.014]	[-0.056, -0.005]
(a) Mapuche expectation	0.374	0.267	0.339	0.378
(b) Non-Mapuche expectation	0.407	0.407	0.408	0.407
No. of Mapuche ($\leq$ 5th pctl.)	1,990	297	341	2,061
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,247	27,213	27,146	27,224
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.045	-0.158	-0.036	-0.042
C.I. (95%)	[-0.058, -0.028]	[-0.187, -0.119]	[-0.077, 0.002]	[-0.054, -0.026]
(a) Mapuche expectation	0.332	0.219	0.340	0.334
(b) Non-Mapuche expectation	0.376	0.377	0.376	0.376
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.038	-0.144	-0.057	-0.034
C.I. (95%)	[-0.056, -0.019]	[-0.185, -0.098]	[-0.101, -0.012]	[-0.051, -0.014]
(a) Mapuche expectation	0.348	0.242	0.329	0.352
(b) Non-Mapuche expectation	0.386	0.387	0.386	0.386
No. of Mapuche ($\leq$ 10th pctl.)	3,900	575	629	4,023
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,573	54,445	54,345	54,547

Note: This table presents the results from the P-BOT using the data described in Table 1, considering two approaches to estimate the outcome equation and two criteria to determine who is the margin. Release probabilities are predicted using a linear probability model. The outcome is any pretrial misconduct. Panel A shows the estimates using a simple difference between the Mapuche and non-Mapuche averages in pretrial misconduct, only considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using a non-parametric local estimation for the conditional expectation of pretrial misconduct at the margin of release, for Mapuche and non-Mapuche defendants. The point estimate is calculated by subtracting these two estimations. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. Details of the covariates included in the prediction model can be found in Appendix G. The confidence intervals are calculated using bootstrap with 500 repetitions.

Table I.II: Prediction-Based Outcome Test, Using OLS to Estimate the Release Probability and Lasso to Select Predictors (Outcome: Pretrial Misconduct)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.052	-0.143	-0.085	-0.044
C.I. (95%)	[-0.073, -0.028]	[-0.193, -0.084]	[-0.125, -0.030]	[-0.067, -0.022]
(a) Mapuche expectation	0.367	0.275	0.335	0.375
(b) Non-Mapuche expectation	0.418	0.419	0.420	0.419
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.040	-0.139	-0.068	-0.035
C.I. (95%)	[-0.068, -0.012]	[-0.204, -0.077]	[-0.125, -0.005]	[-0.063, -0.005]
(a) Mapuche expectation	0.393	0.293	0.365	0.398
(b) Non-Mapuche expectation	0.432	0.432	0.433	0.433
No. of Mapuche ($\leq$ 5th pctl.)	2,065	316	394	2,137
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,172	27,194	27,093	27,148
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.046	-0.139	-0.048	-0.045
C.I. (95%)	[-0.061, -0.028]	[-0.178, -0.096]	[-0.089, -0.005]	[-0.060, -0.026]
(a) Mapuche expectation	0.372	0.280	0.372	0.374
(b) Non-Mapuche expectation	0.419	0.419	0.420	0.419
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.045	-0.139	-0.064	-0.041
C.I. (95%)	[-0.068, -0.024]	[-0.185, -0.097]	[-0.110, -0.015]	[-0.064, -0.019]
(a) Mapuche expectation	0.379	0.285	0.360	0.383
(b) Non-Mapuche expectation	0.424	0.424	0.425	0.425
No. of Mapuche ($\leq$ 10th pctl.)	3,912	574	667	4,032
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,561	54,446	54,307	54,538

Notes: This table presents the results from the P-BOT with the release probabilities predicted using a linear model. The predictors were selected using Lasso. The original set of covariates included 1,568 variables to be chosen: the predictors considered in Table I.I, their squared terms, their interactions, and judge fixed effects. When Mapuche is defined as *at least one surname*, lasso selected 880 predictors, 878 when it is defined as *two surnames*, 871 when it is defined as *self-reported*, and 877 when it is defined as *self-reported or at least one surname*. In all these models, 85% of the cases are correctly classified by the prediction model. Specifically, those who are predicted as released and detained are correctly classified in 87% and 64% of the cases, respectively. The other characteristics of this table replicates Table I.I. Panel A shows the estimates using a simple difference between the Mapuche and non-Mapuche averages in pretrial misconduct, only considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using a non-parametric local estimation for the conditional expectation of pretrial misconduct at the margin of release, for Mapuche and non-Mapuche defendants. The point estimate is calculated by subtracting these two estimations. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. The confidence intervals are calculated using bootstrap with 500 repetitions.

Table I.III: Prediction-Based Outcome Test, Using Heteroskedastic Probit to Estimate the Release Probability (Outcome: Pretrial Misconduct)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.049	-0.148	-0.081	-0.043
C.I. (95%)	[-0.070, -0.025]	[-0.191, -0.095]	[-0.126, -0.031]	[-0.066, -0.021]
(a) Mapuche expectation	0.357	0.258	0.325	0.362
(b) Non-Mapuche expectation	0.405	0.406	0.406	0.405
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.033	-0.144	-0.073	-0.029
C.I. (95%)	[-0.062, -0.007]	[-0.200, -0.084]	[-0.132, -0.017]	[-0.057, -0.003]
(a) Mapuche expectation	0.389	0.278	0.350	0.393
(b) Non-Mapuche expectation	0.422	0.422	0.423	0.422
No. of Mapuche ($\leq$ 5th pctl.)	1,965	299	354	2,036
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,272	27,211	27,133	27,249
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.040	-0.154	-0.028	-0.038
C.I. (95%)	[-0.057, -0.025]	[-0.196, -0.123]	[-0.064, 0.008]	[-0.053, -0.022]
(a) Mapuche expectation	0.357	0.244	0.369	0.360
(b) Non-Mapuche expectation	0.398	0.398	0.398	0.398
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.041	-0.151	-0.059	-0.037
C.I. (95%)	[-0.058, -0.023]	[-0.196, -0.108]	[-0.099, -0.015]	[-0.054, -0.019]
(a) Mapuche expectation	0.364	0.254	0.347	0.369
(b) Non-Mapuche expectation	0.405	0.406	0.406	0.405
No. of Mapuche ($\leq$ 10th pctl.)	3,841	528	658	3,971
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,632	54,492	54,316	54,599

Note: This table presents the results from the P-BOT using the data described in Table 1, considering two approaches to estimate the outcome equation and two criteria to determine who is the margin. Release probabilities are predicted using a heteroscedastic probit model. The outcome is any pretrial misconduct. Panel A shows the estimates using a simple difference between the Mapuche and non-Mapuche averages in pretrial misconduct, only considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using a non-parametric local estimation for the conditional expectation of pretrial misconduct at the margin of release, for Mapuche and non-Mapuche defendants. The point estimate is calculated by subtracting these two estimations. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. Details of the covariates included in the prediction model can be found in Appendix G. The confidence intervals are calculated using bootstrap with 500 repetitions.

Table I.IV: Prediction-Based Outcome Test, Using Probit to Estimate the Release Probability (Outcome: Non-Appearance in Court)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.020	-0.057	-0.033	-0.019
C.I. (95%)	[-0.037, -0.004]	[-0.095, -0.014]	[-0.069, 0.002]	[-0.036, -0.004]
(a) Mapuche expectation	0.157	0.119	0.143	0.158
(b) Non-Mapuche expectation	0.176	0.176	0.176	0.176
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.010	-0.059	-0.023	-0.010
C.I. (95%)	[-0.030, 0.010]	[-0.098, -0.011]	[-0.066, 0.026]	[-0.029, 0.010]
(a) Mapuche expectation	0.165	0.117	0.153	0.165
(b) Non-Mapuche expectation	0.176	0.176	0.176	0.176
No. of Mapuche ($\leq$ 5th pctl.)	1,916	269	321	1,986
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,321	27,241	27,166	27,299
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.027	-0.083	-0.029	-0.026
C.I. (95%)	[-0.040, -0.015]	[-0.111, -0.057]	[-0.058, 0.002]	[-0.039, -0.014]
(a) Mapuche expectation	0.166	0.111	0.165	0.167
(b) Non-Mapuche expectation	0.194	0.194	0.194	0.194
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.020	-0.069	-0.029	-0.019
C.I. (95%)	[-0.037, -0.005]	[-0.102, -0.037]	[-0.063, 0.008]	[-0.035, -0.003]
(a) Mapuche expectation	0.161	0.112	0.151	0.162
(b) Non-Mapuche expectation	0.181	0.181	0.181	0.181
No. of Mapuche ($\leq$ 10th pctl.)	3,774	497	636	3,901
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,699	54,523	54,338	54,669

Note: This table presents the results from the P-BOT using the data described in Table 1, considering two approaches to estimate the outcome equation and two criteria to determine who is the margin. Release probabilities are predicted using a probit model. The outcome is non-appearance in court. Panel A shows the estimates using a simple difference between the Mapuche and non-Mapuche averages in non-appearance in court, only considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using a non-parametric local estimation for the conditional expectation of non-appearance in court at the margin of release, for Mapuche and non-Mapuche defendants. The point estimate is calculated by subtracting these two estimations. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. Details of the covariates included in the prediction model can be found in Appendix G. The confidence intervals are calculated using bootstrap with 500 repetitions.

Table I.V: Prediction-Based Outcome Test, Using Probit to Estimate the Release Probability (Outcome: Pretrial Recidivism)

Data up to 5th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.040	-0.121	-0.049	-0.036
C.I. (95%)	[-0.064, -0.019]	[-0.167, -0.072]	[-0.098, 0.001]	[-0.059, -0.015]
(a) Mapuche expectation	0.289	0.208	0.280	0.293
(b) Non-Mapuche expectation	0.329	0.330	0.330	0.329
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.033	-0.115	-0.056	-0.030
C.I. (95%)	[-0.057, -0.007]	[-0.175, -0.057]	[-0.108, 0.005]	[-0.054, -0.003]
(a) Mapuche expectation	0.316	0.233	0.292	0.319
(b) Non-Mapuche expectation	0.348	0.348	0.349	0.348
No. of Mapuche ($\leq$ 5th pctl.)	1,916	269	321	1,986
No. of Non-Mapuche ($\leq$ 5th pctl.)	27,321	27,241	27,166	27,299
Data up to 10th percentile	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Panel A: Simple Version				
Point estimate, (a)-(b):	-0.033	-0.132	-0.023	-0.031
C.I. (95%)	[-0.049, -0.018]	[-0.168, -0.095]	[-0.056, 0.015]	[-0.046, -0.016]
(a) Mapuche expectation	0.282	0.183	0.292	0.284
(b) Non-Mapuche expectation	0.315	0.315	0.315	0.315
Panel B: Non-Parametric				
Point estimate, (a)-(b):	-0.035	-0.124	-0.047	-0.032
C.I. (95%)	[-0.052, -0.017]	[-0.166, -0.077]	[-0.090, 0.001]	[-0.049, -0.013]
(a) Mapuche expectation	0.295	0.206	0.283	0.298
(b) Non-Mapuche expectation	0.330	0.330	0.330	0.330
No. of Mapuche ($\leq$ 10th pctl.)	3,774	497	636	3,901
No. of Non-Mapuche ($\leq$ 10th pctl.)	54,699	54,523	54,338	54,669

Note: This table presents the results from the P-BOT using the data described in Table 1, considering two approaches to estimate the outcome equation and two criteria to determine who is the margin. Release probabilities are predicted using a probit model. The outcome is pretrial recidivism. Panel A shows the estimates using a simple difference between the Mapuche and non-Mapuche averages in pretrial recidivism, only considering the individuals whose estimated release probability is lower than or equal to the 5th/10th percentile. Panel B shows the estimates using a non-parametric local estimation for the conditional expectation of pretrial recidivism at the margin of release, for Mapuche and non-Mapuche defendants. The point estimate is calculated by subtracting these two estimations. The margin of release is defined as the 1st percentile of the estimated release probability. The bandwidth is the same for both estimations (for Mapuche and non-Mapuche) and it is defined as the distance between the 1st percentile and the 5th/10th percentile of the estimated release probability. Details of the covariates included in the prediction model can be found in Appendix G. The confidence intervals are calculated using bootstrap with 500 repetitions.

J Randomization Test

Table J.I: Predicting Release Status

		Mapuche			
	Non-Mapuche	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Male	-0.007	-0.000	-0.013	0.008	-0.000
	(0.002)	(0.005)	(0.010)	(0.010)	(0.005)
Previous prosecution	-0.002	-0.005	-0.015	0.005	-0.004
	(0.003)	(0.008)	(0.020)	(0.021)	(0.008)
Previous pretrial misconduct	-0.010	-0.010	-0.017	-0.010	-0.010
	(0.000)	(0.001)	(0.002)	(0.002)	(0.001)
Previous conviction	-0.164	-0.137	-0.179	-0.161	-0.137
	(0.005)	(0.016)	(0.044)	(0.047)	(0.016)
No. of previous prosecutions	-0.028	-0.022	0.011	0.032	-0.022
	(0.005)	(0.008)	(0.023)	(0.022)	(0.008)
Severity (previous prosecution)	-0.025	-0.029	-0.001	0.125	-0.021
	(0.009)	(0.022)	(0.047)	(0.048)	(0.022)
Severity (current prosecution)	0.005	0.013	0.012	0.014	0.013
	(0.001)	(0.003)	(0.007)	(0.008)	(0.003)
Court-by-time fixed effects	YES	YES	YES	YES	YES
Observations	647,730	50,818	9,710	9,423	52,002
Joint-F-test	1286.1	473.6	110.1	108.9	470.4
p-value	0.000	0.000	0.000	0.000	0.000
Cragg-Donald F-test (first stage)	353.0	3.9	10.7	0.0	4.3

Note: This table presents the results of an OLS regression of release status on covariates using the data described in Table 1. Drug crime, homicide, and property crime are dummies for the crime types. The null hypothesis in the joint-F-test is that all coefficients are jointly zero. Standard errors are clustered at the year/court level. The Cragg-Donald F-test for the first stage is presented at the bottom of the table.

Table J.II: Predicting Judge Leniency

		Mapuche			
	Non-Mapuche	At least one surname	Two surnames	Self-reported	Self-reported or at least one surname
Male	0.000	-0.000	-0.001	0.002	-0.000
	(0.000)	(0.001)	(0.002)	(0.002)	(0.001)
Previous prosecution	0.000	0.001	0.001	0.002	0.001
	(0.000)	(0.001)	(0.003)	(0.003)	(0.001)
Previous pretrial misconduct	0.000	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Previous conviction	-0.000	0.001	-0.003	-0.002	0.001
	(0.000)	(0.002)	(0.007)	(0.005)	(0.002)
No. of previous prosecutions	0.000	-0.001	0.001	0.001	-0.001
	(0.000)	(0.001)	(0.003)	(0.003)	(0.001)
Severity (previous prosecution)	0.000	-0.001	-0.003	-0.007	-0.001
	(0.000)	(0.002)	(0.005)	(0.009)	(0.002)
Severity (current prosecution)	-0.000	-0.000	-0.001	-0.001	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Court-by-time fixed effects	YES	YES	YES	YES	YES
Observations	647,701	49,570	8,055	7,853	50,802
Joint-F-test	1.0	0.9	1.4	0.4	0.8
p-value	0.465	0.548	0.172	0.944	0.608
Cragg-Donald F-test (first stage)	353.0	3.9	10.7	0.0	4.3

Note: This table presents the results of an OLS regression of judge leniency on covariates using the data described in Table 1. Judge leniency is measured using the residualized leave-out race-specific release rate, as in Arnold, Dobbie, and Yang (2018). Drug crime, homicide, and property crime are dummies for the crime types. The null hypothesis in the joint-F-test is that all coefficients are jointly zero. Standard errors are clustered at the year/court level. The Cragg-Donald F-test for the first stage is presented at the bottom of the table.

K Comparing P-BOT and IV Marginal Defendants

This appendix compares, in terms of observed characteristics, the marginal defendants identified by the P-BOT and the instrument-based approach proposed by Arnold, Dobbie, and Yang (2018). Given that our IV model only has statistical power in the sample of non-Mapuche defendants, we limit the comparison to this group.

The P-BOT explicitly identifies marginally released defendants. Then, it is straightforward to characterize their distribution of observables. In the case of the instrument-based approach, under the standard IV assumptions, the marginal defendants are given by the compliers. Then, we characterize the compliers' observables following the method developed by Abadie (2003) and extended to the judges design framework by Dahl, Kostøl, and Mogstad (2014), Dobbie, Goldin, and Yang (2018), and Bald et al. (2019).

Let $\bar{z}$ and $\underline{z}$ denote the maximum and the minimum value for the judge leniency instrument, respectively. The fraction of compliers is identified by $\Pr(Release_i = 1|Z_i = \bar{z}) - \Pr(Release_i = 1|Z_i = \underline{z}) = \Pr(Release_i(\bar{z}) > Release_i(\underline{z}))$. This expression can be estimated using the IV first stage estimation, in particular, by multiplying the estimated coefficient on the instrument by $(\bar{z} - \underline{z})$. In practice, we assign the top and bottom percentile of the distribution of the instrument to $\bar{z}$ and $\underline{z}$, respectively.[5] By repeating the same procedure but restricting the sample to individuals with $X_i = x$, we can estimate the probability of being complier given that $X_i = x$, i.e., $\Pr(Release_i(\bar{z}) > Release_i(\underline{z})|X_i = x)$. Then, by Bayes rule

$$\Pr(X_i = x|Release_i(\bar{z}) > Release_i(\underline{z})) = \frac{\Pr(Release_i(\bar{z}) > Release_i(\underline{z}))}{\Pr(Release_i(\bar{z}) > Release_i(\underline{z})|X_i = x)} \Pr(X_i = x).$$

Using this equation we can characterize the compliers' distribution of observables.

Tables K.I presents these conditional probabilities for the marginal defendants identified by the P-BOT and the instrument-based approach, defining P-BOT marginal defendants as those released individuals whose propensity score is in the bottom 5% or 10% of the distribution, respectively. As this table shows, in all variables but one (an indicator that takes value 1 if the defendant is accused of a drug crime) when the probability of belonging to some particular group conditional on being IV-complier is higher (lower) than the unconditional one, it is also the case that the conditional probability of being a marginal defendant according to the P-BOT is higher (lower) than the unconditional probability. In the case of gender there is also a change in the direction, but the differences are small in magnitude. In other words, under both methodologies, marginally

[5] These conditional probabilities can be also estimated by local regressions. Results are similar to the linear case.

released defendants are more likely to have previous prosecutions, to have been engaged in pretrial misconduct in the past, to have been convicted in the past, and to be accused of more severe crimes. We interpret this as evidence that the non-Mapuche marginal defendants identified by the P-BOT and the instrument-based approach have similar distribution of observables. Reassuringly, around 6% of non-Mapuche defendants are compliers, while in the P-BOT the share of non-Mapuche defendants identified as marginals are 4% and 8%, when looking at the bottom 5% and 10% of the released defendants propensity score distribution, respectively.

Table K.I: Characteristics of Marginal Defendants

	$\Pr[X=x]$	$\Pr[X=x \mid \text{Marginal}]$ IV	$\Pr[X=x \mid \text{Marginal}]$ P-BOT (5%)	$\Pr[X=x \mid \text{Marginal}]$ P-BOT (10%)
Male	0.885	0.884	0.917	0.920
	(0.0003)	(0.0120)	(0.0016)	(0.0012)
Female	0.115	0.118	0.083	0.080
	(0.0003)	(0.0117)	(0.0016)	(0.0012)
At least one previous case	0.680	0.821	0.927	0.876
	(0.0006)	(0.0161)	(0.0018)	(0.0019)
No previous case	0.320	0.175	0.073	0.124
	(0.0006)	(0.0164)	(0.0018)	(0.0019)
At least one previous pretrial misconduct	0.401	0.546	0.678	0.645
	(0.0006)	(0.0195)	(0.0032)	(0.0024)
No previous pretrial misconduct	0.599	0.444	0.322	0.355
	(0.0006)	(0.0206)	(0.0032)	(0.0024)
At least one previous conviction	0.653	0.803	0.901	0.852
	(0.0006)	(0.0170)	(0.0020)	(0.0020)
No previous conviction	0.347	0.192	0.099	0.148
	(0.0006)	(0.0172)	(0.0020)	(0.0020)
High Severity (previous case)	0.591	0.711	0.798	0.751
	(0.0006)	(0.0181)	(0.0025)	(0.0022)
Low Severity (previous case)	0.409	0.292	0.202	0.249
	(0.0006)	(0.0184)	(0.0025)	(0.0022)
High Severity (current case)	0.513	0.807	0.997	0.989
	(0.0006)	(0.0144)	(0.0004)	(0.0006)
Low Severity (current case)	0.487	0.162	0.003	0.011
	(0.0006)	(0.0143)	(0.0004)	(0.0006)
Drug crime	0.124	0.177	0.021	0.066
	(0.0004)	(0.0147)	(0.0010)	(0.0013)
Non-drug crime	0.876	0.819	0.979	0.934
	(0.0004)	(0.0158)	(0.0010)	(0.0013)
Property crime	0.182	0.080	0.002	0.009
	(0.0005)	(0.0114)	(0.0003)	(0.0005)
Non-property crime	0.818	0.919	0.998	0.991
	(0.0005)	(0.0115)	(0.0003)	(0.0005)

Note: This table presents the probability of belonging to different groups of observables (which are binary or were discretized using the respective median as the threshold). The sample is restricted to non-Mapuche defendants. This probability is calculated unconditionally, conditioning on being an IV-complier, and conditioning of being identified as marginal by the P-BOT. The standard errors are calculated by bootstrap (500 repetitions).

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics 113*(2), 231 – 263.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

Amigo, H. and P. Bustos (2008). Apellidos mapuche historia y significado. *Maigret Ltda*.

Arnold, D., W. Dobbie, and C. Yang (2018). Racial bias in bail decisions. *Quarterly Journal of Economics 133*(4), 1885–1932.

Bald, A., E. Chyn, J. S. Hastings, and M. Machelett (2019). The causal impact of removing children from abusive and neglectful homes. *Working Paper*.

Cortés, T., N. Grau, and J. Rivera (2019). Juvenile incarceration and adult recidivism. *Working Paper*.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family welfare cultures. *Quarterly Journal of Economics 129*(4), 1711–1752.

Dobbie, W., J. Goldin, and C. S. Yang (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review 108*(2), 201–240.

Knowles, J., N. Persico, and P. Todd (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy 109*(1), 203–229.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Painemal, N. (2011). Apellidos mapuche vinculados a títulos de merced. *Santiago: Corporación Nacional de Desarrollo Indígena (CONADI)*.