

Discrimination Against Immigrants in the Criminal Justice System: Evidence from Pretrial Detentions*

Patricio Domínguez[†] Nicolás Grau[‡] Damián Vergara[§]

First version: February 22, 2022

This version: February 22, 2022

Abstract

This paper tests for discrimination against immigrant defendants in the criminal justice system in Chile using a decade of nationwide administrative records on pretrial detentions. Observational benchmark regressions show that immigrant defendants are 8.6 percentage points less likely to be released pretrial relative to Chilean defendants with similar proxies for pretrial misconduct potential. Diagnostics for omitted variable bias –including a novel test to assess the quality of the proxy vector based on comparisons of pretrial misconduct rates among released defendants– suggest that the discrimination estimates are not driven by omitted variable bias and that, if anything, failing to fully account for differences in misconduct potential leads to an underestimation of discrimination. Our estimates suggest that discrimination stems from an informational problem because judges do not observe criminal records in origin countries, with stereotypes and taste-based discrimination playing a role in the problem’s resolution. We find that discrimination is especially large for drug offenses and that discrimination increased after a recent immigration wave.

*We thank Peter Hull, Patrick Kline, and seminar participants at the IADB-Migration group, the Online Economics of Discrimination Seminar, and UC Berkeley for very helpful discussions and suggestions. Guillermo Palacios provided outstanding research assistance. Grau acknowledges financial support from the Centre for Social Conflict and Cohesion Studies (ANID/FONDAP/15130009). Vergara acknowledges financial support from the Law, Economics, and Politics Center (LEAP) at UC Berkeley. Usual disclaimers apply.

[†]Department of Industrial and Systems Engineering, Pontificia Universidad Católica de Chile, pdom-ingr@ing.puc.cl

[‡]Department of Economics, Faculty of Economics and Business, Universidad de Chile, ngrau@fen.uchile.cl.

[§]Department of Economics, UC Berkeley, damianvergara@berkeley.edu.

“When Mexico sends its people, they’re not sending their best. (...). They’re bringing drugs. They’re bringing crime. They’re rapists. And some, I assume, are good people.”

Donald Trump, presidential announcement speech, June 16, 2015.

1 Introduction

Increasing flows of international migrants have induced several policy challenges for destination countries in recent years (UN, 2013, 2019) including, among others, the mitigation of concerns about discrimination (Heath et al., 2013). Labor market, crime, and welfare concerns can trigger negative attitudes towards immigrants, particularly when cultural prejudice and misperceptions regarding foreign born and minority populations are commonplace (Dustmann and Preston, 2007; Hangartner et al., 2019; Grigorieff et al., 2020; Ajzenman et al., 2021; Alesina et al., 2021, 2022; Bursztyrn et al., 2021a). Characterizing discrimination patterns against immigrants is therefore a policy priority to improve the possibilities to integrate immigrants in destination countries.

This paper conducts an empirical analysis of discrimination against immigrant defendants in the criminal justice system, in particular, in the dictation of pretrial detention. Pretrial detention decisions are particularly relevant for two main reasons. First, one of the prevalent narratives that underpins negative attitudes towards immigration is its potential effect on crime (Fasani et al., 2019; Bursztyrn et al., 2021b). Thus, if there is discrimination against immigrants, it should be especially salient in institutions that are concerned with that rubric. Second, the pretrial detention process has been shown to be permeated by discriminatory practices (Arnold et al., 2018, 2021; Grau and Vergara, 2021) and to have a significant impact on the labor market opportunities, conviction probabilities, and the participation in public assistance programs for defendants (Heaton et al., 2017; Leslie and Pope, 2017; Dobbie et al., 2018; Dobbie and Yang, 2021a,b; Grau et al., 2021; Stevenson and Mayson, 2021). These costs are particularly harmful for immigrant populations given the challenges they face relating to integration in general and entry into the labor market in particular (Åslund et al., 2014; Dustmann et al., 2017; Brell et al., 2020; Egger et al., 2022).

We study this scenario in Chile where the pretrial detention process functions in a similar manner to the US—that is, judges decide on the pretrial detention status of defendants using limited information and in a limited time period with a mandate to not detain defendants unless misconduct potential (pretrial recidivism or failure to appear in court, or both) is considerable. Unlike in the US, however, the Chilean system is uniform across all localities and does not permit monetary bail. Since approximately 2013 Chile has become an important destination for immigrants

from other Latin American countries. And, as we discuss in Section 2, this immigration flow has generated hostile discourses in the public sphere and misperceptions regarding immigrants and immigration in the general population.

We conduct the empirical analysis using nationwide administrative records for the period 2008—2017. Our database covers 95% of Chile’s criminal prosecutions and includes detailed information on defendants and prosecutions. Despite having cleaner criminal records upon prosecution and lower pretrial misconduct rates when released, immigrant defendants are more likely to be detained pretrial relative to defendants who are Chilean nationals. We formally explore whether this descriptive fact constitutes some form of aggregate discrimination by estimating observational benchmark regressions—that is, by testing if defendants with similar proxies for misconduct potential face, on average, different pretrial release rates because of their immigration status, an exercise similar in spirit to [Gelman et al. \(2007\)](#), [Abrams et al. \(2012\)](#), and [Rehavi and Starr \(2014\)](#). The main result shows that, conditional on type of crime, criminal records, judge leniency, attorney quality, and court-by-time fixed effects, immigrant defendants are 8.6 percentage points (pp) more likely to be detained pretrial relative to Chilean defendants, a difference that represents 53% of the unconditional baseline pretrial detention rate.

We perform two sets of diagnostics that suggest that our main estimates are not driven by systematic differences in unobserved misconduct potential (i.e., omitted variable bias, OVB). First, we implement the method set out in [Oster \(2019\)](#) for assessing the degree of selection on unobservables using observed covariates. We conclude that, if anything, our observational benchmark regression underestimates discrimination against immigrants. Specifically, we show that the discrimination estimate is attenuated (although not reversed) only if we impose the nonstandard assumption that unobservables are negatively correlated with the control variables—that is, if defendants with observables that induce lower release rates are also endowed with unobservables that work in the opposite direction. This is consistent with the fact that immigrants have, on average, cleaner criminal records. In fact, and contrary to the common finding in observational benchmark regressions, we find that the release disparity between Chilean defendants and immigrant defendants is 30% smaller (6.6 pp) when control variables are omitted.

The second diagnostic is a novel test that uses observed pretrial misconduct rates of released defendants to assess the degree of OVB in the release equation. The intuition is as follows: OVB arises if there are unobservables that correlate with immigration status and predict misconduct potential. Then, we can use observed pretrial misconduct of released defendants to test whether the included control variables are sufficient to account for systematic differences in pretrial misconduct

rates between immigrant defendants and Chilean defendants. Because the sample of released defendants is selected, we estimate different selection models using assigned judges and attorneys as the excluded variables and show that, conditional on type of crime, criminal records, and time-varying court characteristics, immigration status has no predictive power on observed pretrial misconduct. This confirms that the set of controls included in the benchmark regression is sufficient to account for systematic risk differences between Chilean defendants and immigrant defendants.

The second test also suggests that, while time-varying court characteristics matter for the discrimination point estimate, they do little to correct for OVB and, therefore, they may induce included variable bias (IVB) in the main estimations. In other words, rather than controlling for differences in unobserved misconduct potential, court-by-time variation seems to mediate discrimination through the systematic assignment of immigrant defendants to less lenient courts. Under this interpretation, the baseline discrimination estimate of 8.6 pp constitutes a conservative bound given that excluding the court-by-time fixed effects increases the release disparity gap to 10.3 pp.

After establishing the robustness of the reduced-form discrimination estimate, we perform different heterogeneity analyses to obtain additional insights about the aggregate pattern. We first explore whether discrimination against immigrants stems from an informational problem. The criminal records of immigrant defendants are likely to be censored because judges cannot observe the records kept in the country of origin. Then, judges may process the information contained in criminal records differently depending on whether the defendant is immigrant or not. To test this hypothesis, we estimate our main specification separately for samples of defendants with and samples of defendants without previous prosecutions in Chile. We find that the conditional release disparities are 1.9 pp and 13.2 pp for defendants with and defendants without previous prosecutions, respectively. This implies that, unlike Chilean defendants, immigrant defendants are not rewarded for having clean criminal records. In fact, comparing the discrimination estimates with the unconditional differences in release rates suggest that judges systematically impute nonzero criminal histories for all immigrant defendants, regardless of their criminal records in Chile. We complement this analysis by showing that the discrimination estimate is 36% smaller for experienced judges. Both pieces of evidence strongly suggest that the informational problem seems to be the main driver behind the discrimination patterns.

The fact that discrimination stems from an informational problem does not necessarily mean that the aggregate estimate can be rationalized by accurate statistical discrimination. On one hand, the informational problem can be circumvented based on stereotypes (biased beliefs). On the other hand, the informational problem may provide the room for judges to exercise preferences

against immigrants (taste-based discrimination). To obtain further insights, we estimate observational outcome tests (Knowles et al., 2001; Grau and Vergara, 2021) that, when rejected, provide evidence of a nonseparately identified combination of stereotypes and taste-based discrimination (Hull, 2021). The outcome test is rejected suggesting that the discrimination pattern, although stemming from an informational problem, cannot be rationalized by accurate statistical discrimination. The combination of stereotypes and taste-based discrimination seems to play an important role in explaining the aggregate estimates of discrimination against immigrants.

We also explore whether or not the patterns of discrimination vary according to crime category. We find that discrimination is particularly severe for drug offenses: conditional release disparities are around four times larger than the baseline benchmark estimate. Statistically and economically significant differences are also found for thefts and robberies and other property crimes. Other violent crimes like homicides and sexual offenses show no significant differences, although the number of immigrants that are charged with these types of crimes is too small to confidently detect an effect. Crime categories that are associated with less hostile discourses in the public sphere (such as white collar crimes or crimes against privacy) also display no significant release disparities. This heterogeneity is consistent with the discourse that links immigration flows to drug trafficking and other crimes that tend to be prominent in surveys that measure citizens perceived fears of crime playing a role in affecting preferences and stereotypes.

Finally, we explore the evolution of discrimination patterns during the recent wave of immigration to Chile. Immigration inflows in Chile grew considerably and became more racially diverse after 2012. We find that benchmark estimates are 55% larger in the 2013–2017 period relative to the 2008–2012 period. We also find that discrimination estimates are remarkably similar across different groups of immigrants, suggesting that the increase over time is driven by an intensification of discrimination rather than a composition effect. Although we do not establish a causal relationship, we conjecture that this trend is related to the crime-related narrative response to the immigration wave lead by political authorities and the media (Ajzenman et al., 2021).

2 Setting and Data

This section describes the Chilean context and institutional setting. We start by characterizing immigration in Chile, discussing its magnitude, trends, composition, and the attitudes to immigration among Chilean citizens. We then describe the pretrial detention system. Finally, we describe the data used in the empirical analysis and provide descriptive statistics.

2.1 Immigration in Chile

Immigration in Chile has changed dramatically since 2013. According to the national census, immigrants represented around 1.5% of the Chilean population in 2002 and in 2012 this figure was 2%. In 2017, however, the share of immigrants in the population increased to 4.5%, and projections from the National Institute of Statistics suggest that the number reached 7.8% in 2019. In terms of gross inflows, both the volume and composition has changed considerably. Between 2008 and 2012 the inflow of immigrants was fairly stable at around 100,000 individuals per year. Peruvians accounted for the largest share of immigrants arriving each year (roughly 50% of new arrivals) and Argentinians, Bolivians, and Colombians collectively represented around 25% of inflows. The annual inflow of immigrants grew to around 350,000 individuals in 2017. In addition, the composition changed substantially. For the 2008–2012 period, Haitian and Venezuelan immigrants accounted for around 2% of the arrivals; however, for the 2013–2017 period they represented almost 20% of the arrivals.¹ Inflows of Colombian nationals also increased significantly after 2013. Their share among new arrivals to Chile grew from 10% in the period 2008–2012 to 19% in the period 2013–2017.

The recent immigration wave has not gone unnoticed by Chile’s population, politicians, and media. Similar to the inhabitants of other countries (Hopkins et al., 2019; Jørgensen and Osmundsen, 2020; Alesina et al., 2022), Chileans overestimate the size of the immigrant population. As summarized in Ajzenman et al. (2021), surveys show that in 2019 Chileans thought that immigrants represented 33% of the total population, which compares unfavorably to the administrative estimate of 7.8%. Additional survey evidence shows that *increase in crime* is the most prominent concern related to immigration among Chile’s population. Ajzenman et al. (2021) report that these perceptions are not related to changes in victimization rates, even though the concerns were severe enough to induce increases in preventive behavioral responses, such as investments in home security. In addition, Ajzenman et al. (2021) show that the media plays an important role in the divergence between perceptions and actual crime rates, for example, by increasing the number of crime-related news segments with an alleged perpetrator that is an immigrant.

Public attitudes toward immigrants can also be inferred from recent political campaigns. Immigration, which was not a public issue before the recent increase in immigration inflows, has become a prominent topic in the national political debate, with an increasing number of politicians framing their public discourse in the context of immigration regulation. For example, inspired by

¹To illustrate the prominence of Chile as a destination of choice within the region, in March 2021 Chile was the country with the second highest number of Venezuelan immigrants equating to 600,000 individuals with regular status, just below Colombia with 720,000. See <https://data2.unhcr.org/en/situations/vensit> for details.

Donald Trump, the far-right candidate in the last presidential election, José Kast, proposed to build a ditch on Chile’s border with Bolivia. In most cases, a link between crime and immigration is proffered as the most important reason for stopping or reducing the arrival of new immigrants. For example, when the National Congress of Chile passed President Piñera’s migration project in 2019, he described the initiative as a “(...) great step forward to continue putting our house in order regarding matters of migration, and thus better combat illegal immigration and the evils such as delinquency, drug trafficking, and organized crime entering Chile” (Twitter, January 16, 2019). These types of statements, that display either implicit or explicit anti-immigrant sentiment, have been accompanied by instances of misinformation that appear to be designed to stigmatize immigrants or immigration. For example, in 2018, a widely shared social media post claimed that the former President Bachelet received USD 3,000 from the UN for each Haitian immigrant that entered Chile. The increase in negative attitudes towards immigrants has contributed towards Chile’s political climate becoming more polarized and an increase in the relative popularity of far-right nationalist groups with conservative-leaning voters.²

2.2 Pretrial Detentions in Chile

The procedure to define pretrial detention for people arrested in Chile works as follows: During the 24 hours after the initial detention, there is an arraignment hearing in which a detention judge must decide whether the defendant will be incarcerated during the criminal investigation. Monetary bail is not an option in the Chilean system. Following the legal principle of the presumption of innocence, judges should not incarcerate defendants unless there is clear danger of escape (i.e., a high probability of failing to appear in court), the defendant represents a significant threat to society (i.e., a high probability of committing another crime during the investigation), or the imprisonment of the defendant aids the investigation of the criminal case. As described in [Grau et al. \(2021\)](#), pretrial detention became more frequent between 2007 (17,891 cases) and 2018 (34,815 cases). In the same period, the proportion of cases where pretrial detention is imposed has increased from

²There is a contrast between the hostile expressions against immigrants, such as President’s Piñera’s tweet, with what is usually presented as the traditional view on Chilean attitudes towards immigration as expressed in Chito Faró’s famous 1942 folk song “Si vas para Chile” (“If you go to Chile”). This popular and traditional song describes Chileans as particularly receptive to foreigners stating: “Y verás como quieren en Chile al amigo cuando es forastero” (“and you will see how they love a friend in Chile when is a foreigner”). Other music artists have offered a different perspective on attitudes towards immigration. For example, a 2012 track by Chilean rapper Portavoz called “Poblador del mundo” (“World citizen”) tells a story in the voice of an immigrant that feels constant discrimination living in Chile. In one verse, he raps “Llegué con fe para doblarle la mano al destino y me encontré con que aquí soy un enemigo. Y es que con mi llegada los dramas también han venido, ya me han hecho entender que no soy para nada bienvenido” (“I arrived with the belief that I could turn the hand of destiny but found that here I am an enemy. With my arrival the dramas have also come, they have already made me understand that I am not welcome at all”).

7.3% to 9.6%, and pretrial detainees as a share of total prisoners rose from 21.9% to 36%.

Pretrial detention hearings are brief (lasting about 15 minutes) and judges have to make a decision whether or not to detain the defendant at the end of the hearing. Judges are assigned in a quasi-random fashion at the court-by-time level: in every court at the beginning of each month judges are assigned to different time slots to lead arraignment hearings for no reason other than splitting this duty among the judges in the particular court. The assignment of public attorneys to each case is also quasi-random at the court-by-time level.

2.3 Data

We use administrative records from the Public Defender’s Office (PDO). The PDO is a centralized public service under the oversight of the Ministry of Justice. It offers criminal defense services to all individuals accused of or charged with a crime; as such, it ensures the right to a defense by a lawyer and due process in criminal trials. Our estimation sample covers more than 95% of the criminal cases for the period between 2008 and 2017, and contains detailed case and defendant characteristics, including the nationalities of the defendants. We define immigrant defendants as defendants that are nationals of any other country than Chile. In addition, we can identify the judges assigned to each case at the beginning of the criminal process (i.e., when pretrial detention decisions take place), as well as the public attorney assigned to the arraignment hearing.

To build the estimation sample, we consider all detention hearings for adult defendants who were arrested between 2008 and 2017 and defended by the PDO. We exclude hearings due to legal summons because the information set available to the judge may be different in those cases. To be able to focus on arraignment hearings in which pretrial detention is a plausible outcome, we only consider types of crimes that have at least a 5% probability of pretrial detention. For the same reason, when defendants are accused of more than one crime during the same arraignment hearing, we only retain the information related to the most severe crime (with severity measured as the probability of pretrial detention). A more detailed description of the data, the sample restrictions, and the variables is presented in Appendix B.

Table 1 shows descriptive statistics for the estimation sample. This table provides three stylized facts that motivate our research question. First, relative to Chilean defendants, immigrants are more likely to be detained before the trial. The pretrial release rate for immigrant defendants and Chilean defendants is 77% and 84%, respectively. Second, immigrant defendants have cleaner criminal records. Chilean defendants are more likely to have previous prosecutions, convictions,

Table 1: Descriptive Statistics

	Chilean	Immigrant
Released	0.84	0.77
Outcomes (only for released)		
Nonappearance in court	0.17	0.16
Pretrial recidivism	0.19	0.13
Pretrial misconduct	0.29	0.24
Individual Characteristics		
Male	0.88	0.87
At least one previous case	0.68	0.41
At least one previous pretrial misconduct	0.40	0.22
At least one previous conviction	0.65	0.39
No. of previous cases	4.58	2.48
Severity previous case	0.09	0.05
Severity current case	0.18	0.18
Court Characteristics		
Average severity (year/Court)	0.09	0.11
No. of cases (year/Court)	3,023	3,557
No. of judges (year/Court)	46	58
Observations (released)	580,400	4,900
Observations (nonreleased)	112,606	1,462

Notes: This table presents descriptive statistics for our estimation sample. The sample considers all arraignment hearings for adult defendants who were arrested between 2008 and 2017. We do not include hearings due to legal summons and only consider types of crimes with at least a 5% probability of pretrial detention. When defendants are accused of more than one crime, we retain the information related to the most severe crime (with severity measured as the probability of pretrial detention). More details about the data and the variables can be found in Appendix B.

and instances of pretrial misconduct. The criminal records of immigrants may be censored because judges cannot observe the criminal histories of immigrant defendants before they arrive to Chile. Therefore, how judges interpret immigrants' (lack of) criminal records could play a role in explaining the release disparities. Third, once released, immigrants are less likely to be engaged in any kind of pretrial misconduct; pretrial recidivism rates are 6 pp larger for Chilean defendants.

3 Theoretical Framework

To motivate a formal test for discrimination in pretrial detention decisions, this section develops a simple theoretical framework to guide the empirical analysis and formally defines the notion of discrimination we use throughout the paper.

3.1 Model

Let i index defendants and j index judges, with judges assigned to defendants according to the mapping $j(i)$. Let $Y_i^* \in \{0, 1\}$ be pretrial misconduct if released, with Y_i being the realized value, so $Y_i = Y_i^*$ for pretrial released defendants and $Y_i = 0$ for pretrial detained defendants. Judges are mandated to detain defendants with $Y_i^* = 1$ and release defendants with $Y_i^* = 0$ but they do not observe Y_i^* at the time of the decision.³ Instead, judges base the release decision on their predictions on the probability that $Y_i^* = 1$. The prediction is made using all the available information, which includes the nationality of the defendant, $I_i = 1\{i \text{ is immigrant}\}$, and a vector of other characteristics, Z_i .

The principle of the presumption of innocence and the documented impact of pretrial detention imply that only defendants with a high predicted probability of pretrial misconduct should be detained. Because the decision is discretionary and judges potentially have heterogeneous preferences, release thresholds are idiosyncratic. Then, the release equation for defendant i can be written as

$$Release_i = 1 \{p_{j(i)}(I_i, Z_i) \leq t_{j(i)}(I_i, Z_i)\}, \quad (1)$$

where $p_{j(i)}(I_i, Z_i) \in [0, 1]$ is the prediction of the pretrial misconduct probability of defendant i made by the assigned judge $j(i)$, and $t_{j(i)}(I_i, Z_i) \in [0, 1]$ is the judge-specific release threshold.

Let $p(I_i, Z_i) = \Pr(Y_i^* = 1 | I_i, Z_i)$. Because the conditional distribution is possibly unknown to judges, their predictions regarding pretrial misconduct are potentially biased. Let $b_{j(i)}(I_i, Z_i)$ be the judge-specific bias in predictions, such that $p_{j(i)}(I_i, Z_i) = p(I_i, Z_i) + b_{j(i)}(I_i, Z_i)$. Then,

$$Release_i = 1 \{p(I_i, Z_i) \leq t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)\} \equiv 1 \{f_{j(i)}(I_i, Z_i) \geq 0\} = R(I_i, Z_i, j(i)). \quad (2)$$

As [Hull \(2021\)](#) stresses, this analysis does not require additional assumptions on judge behavior, relating to the way they set thresholds or make predictions, for example. The analysis only needs the judge decision to be binary and to follow a threshold-crossing model. As discussed in [Section 2](#), both assumptions represent a good approximation of the Chilean setting. In this simple model, the decision is deterministic from the judge's perspective. We denote the function that determines whether a defendant with immigration status I_i and characteristics Z_i is released or not when assigned to judge $j(i)$ by $R(I_i, Z_i, j(i)) \in \{0, 1\}$.

³Alternatively, Y_i^* could be defined as the probability of defendants engaging in some form of pretrial misconduct if released, with $Y_i^* \in [0, 1]$ and the judges' mandate to detain defendants pretrial when Y_i^* exceeds some threshold.

3.2 Sources of Release Disparities and Normative Definitions

Equation (2) suggests that a defendant’s nationality may affect the release decision through different channels. First, judges may use I_i to predict (or interpret signals of) variables they do not observe that affect Y_i^* if distributions vary with I_i . This is defined as statistical discrimination (Aigner and Cain, 1977) and directly affects the conditional outcome probabilities, $p(I_i, Z_i)$, and the bias, $b_{j(i)}(I_i, Z_i)$, if these parameters impact how priors are updated. Second, if statistical discrimination is inaccurate and the degree of inaccuracy varies with I_i , this induces a differential prediction bias to the decision. We refer to this as biased beliefs or stereotypes, and its effect is captured by $b_{j(i)}(I_i, Z_i)$ (Bordalo et al., 2016; Bohren et al., 2021). Third, release thresholds may depend on I_i because of animus. This affects $t_{j(i)}(I_i, Z_i)$ and gives form to the standard notion of taste-based discrimination or discriminatory preferences (Becker, 1957, 1993).

In this paper, we focus on an aggregate notion of discrimination that incorporates the three sources of release disparities mentioned: statistical discrimination, biased beliefs or stereotypes, and taste-based discrimination. A possibly complex combination of these drivers of release disparities may yield a reduced-form effect of I_i on $Release_i$, captured by the function $f_{j(i)}(I_i, Z_i)$ even within defendants with equal Y_i^* . As emphasized in Yang and Dobbie (2020) and Arnold et al. (2021), all sources of disparities are problematic through the lens of legal principles similar to the Equal Protection Clause in the US Constitution, which suggests that this reduced form effect has normative content regardless of the specific sources. Intuitively, the composite effect of all sources may lead defendants with equal misconduct potential but different immigration status to face different pretrial detention rates. This motivates the following definition.

DEFINITION (DISCRIMINATION): *discrimination, \mathbf{D} , is defined as*

$$\mathbf{D} = \mathbb{E}[\mathbf{d}(Y_i^*)] = \mathbb{E}[\mathbb{E}[d(Z_i, j(i))|Y_i^*]] = \mathbb{E}[\mathbb{E}[R(1, Z_i, j(i)) - R(0, Z_i, j(i))|Y_i^*]], \quad (3)$$

where the inner expectation integrates over $(Z_i, j(i))$, and the outer expectation integrates over Y_i^* .

Under this discrimination definition, there is system-wide discrimination against immigrants if $\mathbf{D} < 0$. The inner object, $d(Z_i, j(i))$, is the judge-specific gap in release rates between immigrant and native defendants endowed with Z_i . $\mathbf{d}(Y_i^*)$ is the expectation of the judge-specific measure over $(Z_i, j(i))$ conditional on misconduct potential Y_i^* and, therefore, does not impose restrictions on the group-specific distributions. Finally, the aggregate measure of discrimination averages over the distribution of misconduct potential. Similar to the definition outlined in Arnold et al. (2021), this definition compares release rates between immigrant and native defendants with equal pretrial

misconduct potential, with potential differences arising because of statistical discrimination, biased beliefs, or taste-based discrimination, or a combination of these sources of discrimination.

Empirical test One fundamental challenge when testing for discrimination following (3) is that Y_i^* is most likely not observable; therefore, a regression of $Release_i$ on I_i controlling for Y_i^* is potentially not implementable. Arnold et al. (2021) propose a weighting scheme that equalizes group-specific distributions of Y_i^* under which omissions of Y_i^* from the (weighted) regression of $Release_i$ on I_i do not induce OVB and, therefore, identifies \mathbf{D} . Unfortunately, the immigrant population in Chile is small, which prevents us from implementing this methodology.⁴ As an alternative, we implement observational benchmark regressions similar in spirit to Gelman et al. (2007), Abrams et al. (2012), and Rehavi and Starr (2014). Formally, we assume that Y_i^* can be expressed as a function of observed variables, X_i , so the discrimination definition can be written as

$$\mathbf{D} = \mathbb{E}[\mathbf{d}(g(X_i))] = \mathbb{E}[\mathbb{E}[d(Z_i, j(i))|g(X_i)]] = \mathbb{E}[\mathbb{E}[R(1, Z_i, j(i)) - R(0, Z_i, j(i))|g(X_i)]], \quad (4)$$

where $Y_i^* = g(X_i)$. Although X_i and Z_i possibly overlap, X_i does not have to be equal to Z_i . Under this assumption, regressing $Release_i$ on I_i controlling for $g(X_i)$ identifies \mathbf{D} .

In Section 4 we discuss our choice of X_i and provide diagnostics to assess its performance in approximating Y_i^* . This is central to the analysis because the extent to which (4) constitutes a good approximation of (3) fundamentally depends on the vector X_i . More specifically, as noted by Arnold et al. (2021), observational benchmark regressions can suffer both from IVB and OVB. IVB arises when some of the included variables in X_i mediate discrimination against, in this case, immigrant defendants in a disparate impact sense. For example, if there is discrimination based on place of living and immigrants are overrepresented in a discriminated locality, then controlling for place of living will attenuate the estimated disparity. OVB arises when the vector X_i does not fully account for group differences in misconduct potential and, therefore, the estimated coefficient reflects differences in unobservables rather than discriminatory forces. This empirical challenge generates a tension between over and under controlling.

We deal with this tension by choosing the smallest set of variables that successfully eliminates

⁴The Arnold et al. (2021) method combines quasi-random assignment of judges, observational misconduct data on released defendants behavior, and extrapolation techniques. The extrapolation step requires several judges handling several cases that involve defendants that are immigrants. Our limitation is, therefore, statistical power. Specifically, the extrapolation technique used in Arnold et al. (2021) is based on regressing average pretrial misconduct rates of released immigrant defendants against judge leniency to infer the average pretrial misconduct rate of a supremely lenient judge using an identification at infinity argument. For this technique to work properly several judges are required in order to have enough support, in addition to several cases per judge so that pretrial misconduct rates can be precisely estimated. Both requirements are not met in our immigrant subsample.

OVB and therefore minimizes the likelihood of IVB. Section 4 presents several tests that reject OVB given our choice of X_i and develops sensitivity analyses that flag potential sources of IVB in our setting. In particular, variables that do little to attenuate OVB concerns and significantly affect the discrimination estimate are more likely to contaminate the result with IVB.

On the normative interpretation Related to the discussion on IVB and OVB, our definition of discrimination seeks to identify a disparate impact (rather than disparate treatment) notion of discrimination (Arnold et al., 2021; Rose, 2022). The disparate impact doctrine focuses on release disparities among defendants with equal misconduct potential but different group membership. Under this perspective, different structural models that rationalize similar release gaps conditional on misconduct potential are normatively equivalent. On the other hand, the disparate treatment doctrine focus on the actual release decision, understanding as discrimination explicit judge-level practices against a particular group. This implies that the normative implication of release disparities among defendants with equal misconduct potential is different if they follow explicit discriminatory practices against group membership or if, for example, are driven by discriminatory practices against other characteristics that correlate with group membership.

As before, the extent to which our definition relates to each normative perspective depends on the variables included in X_i . More specifically, it is dependent on how accurately these variables proxy misconduct potential. If X_i exactly matches the information set of judges (Z_i), the definition will probably approximate a disparate treatment notion of discrimination because judges observe variables through which discrimination towards immigrants manifests itself and that are unlikely to predict misconduct potential on top of other controls. That is why we purposely exclude (and why we are not concerned that we do not observe) all demographics other than immigration status that are potentially observed by judges (age, place of living, physical aspects, or income, for example). By contrast, we include the minimum set of variables that, we argue, are sufficient to eliminate OVB because they constitute a reasonable proxy of misconduct potential—namely, criminal records and case characteristics.⁵

⁵Arnold et al. (2021) show that discrimination estimates are biased when criminal record and type of crime variables are included in regressions that already control for misconduct potential through an IVB mechanism, which implies that these variables are imperfect proxies for misconduct potential. This can be explained if, for example, the distributions of types of crime differ by immigration status and judges are engaged in crime-specific discrimination patterns, or if there is discrimination against immigrants in previous stages of the judicial process that is manifest in differences in the criminal records. Note, however, that Arnold et al. (2021) estimated observational release disparities that control for type of crime and criminal records are close in magnitude to the weighted benchmark regressions that recover the unbiased estimate, which suggests that this set of controls do a reasonably good job when approximating potential misconduct. They document unconditional gaps of 7.2 pp that are reduced to 5.2 pp when including controls and fixed effects. The bias-corrected weighted regressions with no controls report discrimination estimates that range from 4.2 pp to 5.4 pp, depending on the specific extrapolation technique.

4 Empirical Analysis

This section presents the empirical strategy and the main results. We estimate substantial discrimination against immigrant defendants in pretrial detention decisions. We discuss and implement different tests for assessing the degree of OVB and conclude that unobservables are unlikely to explain the results. If anything, unobserved differences in misconduct potential are likely to exert downward bias on the discrimination estimates.

4.1 Benchmark Regressions

Empirical strategy Discrimination, as defined by equation (4), can be potentially identified by *benchmark regressions*—that is, by reduced-form regressions of $Release_i$ on I_i and $g(X_i)$. We assume that g is linear in X_i —that is, $g(X_i) = X_i' \alpha_X$. And, importantly for the discussion on OVB, we further define $X_i = [X_i^o \ X_i^u]$, where o accounts for variables that are observed by the econometrician and u accounts for variables that are unobserved by the econometrician. The observational benchmark test therefore consists of regressions of the following type:

$$Release_i = \alpha_0 + \alpha_D I_i + X_i^{o'} \alpha_{X_o} + \varepsilon_i, \tag{5}$$

where $\varepsilon_i = X_i^{u'} \alpha_{X_u} + \xi_i$ is the error term. We distinguish between two sources of unobserved heterogeneity that affect the release decision. First, $X_i^{u'} \alpha_{X_u}$ accounts for unobservables that predict misconduct potential. The vector X_i that accurately predicts Y_i^* may be extensive and may therefore contain information that is not observed by the econometrician. Second, ξ_i accounts for unobservables that do not predict misconduct potential but may affect the release decision. For example, ξ_i may contain variables that judges use to discriminate on top of immigration status but that are not correlated with Y_i^* after controlling for X_i .

Let $\hat{\alpha}_D$ be the ordinary least squares (OLS) estimator of (5). The main identification assumption for $\mathbb{E}[\hat{\alpha}_D | I_i, X_i^o] = \mathbf{D}$ is $\mathbb{C}(I_i, X_i^u) = 0$. That is to say, the unobservables that predict misconduct potential, X_i^u , are not correlated with immigration status, I_i . This follows from equation (4) because discrimination is defined as average release disparities conditional on misconduct potential. If I_i and X_i^u are correlated, then $\hat{\alpha}_D$ will capture both discrimination and unobserved differences in misconduct potential and will therefore be contaminated by standard OVB. Note, however, that correlation between I_i and ξ_i is allowed by our definition of discrimination given our focus on a disparate impact notion of discrimination so its presence is, therefore, not problematic for the

Table 2: Main Results–Benchmark Regressions: $\hat{\alpha}_D$

	(1)	(2)	(3)	(4)	(5)
Immigrant	-0.067 (0.014)	-0.103 (0.013)	-0.063 (0.013)	-0.030 (0.012)	-0.087 (0.012)
Mean dep. variable	0.84	0.84	0.84	0.84	0.84
Individual controls	No	Yes	No	No	Yes
Attorney and judge controls	No	No	Yes	No	Yes
Court-by-year fixed effects	No	No	No	Yes	Yes
No. of Immigrants	6,362	6,362	6,362	6,362	6,362
No. of Chileans	693,006	693,006	693,006	693,006	693,006
R-squared	0.000	0.167	0.011	0.023	0.191

Notes: This table presents the estimated benchmark coefficient, $\hat{\alpha}_D$, of equation (5) and its corresponding standard error in parentheses (clustered at the court-by-year level). Each column represents a different regression that includes different sets of controls. *Individual controls* include an indicator for whether the individual has previous prosecutions, the number of previous prosecutions, the severity of the last prosecution (measured as the average pretrial detention rate of the type of crime), an indicator for whether the individual was engaged in pretrial misconduct during a previous prosecution, an indicator for whether the individual has been convicted of a crime in the past, and type of crime fixed effects for the current prosecution (consisting of nine mutually exclusive categories). *Attorney and judge controls* include judge leniency and public attorney quality and their squares, measured as residualized (against court-by-year fixed effects) leave out release rates, as in [Dobbie et al. \(2018\)](#).

identification of \mathbf{D} . After presenting the main results, we discuss and implement diagnostics to assess the pervasiveness of potential OVB in our main benchmark estimations.

We estimate equation (5) using OLS and include the following sets of controls in the vector X_i^o . First, we include individual-level characteristics related to the criminal history and the type of crime. These include an indicator for whether the individual has previous prosecutions, the number of previous prosecutions, the severity of the last prosecution (measured as the average pretrial detention rate of the type of crime), an indicator for whether the individual was engaged in pretrial misconduct during a previous prosecution, an indicator for whether the individual has been convicted of a crime in the past, and type of crime fixed effects for the current prosecution. These variables are expected to proxy individual-level misconduct potential. Second, to improve precision, we include measures of judge leniency and public attorney quality and their squares, measured as residualized (against court-by-year fixed effects) leave out release rates, as in [Dobbie et al. \(2018\)](#). Third, we include court-by-year fixed effects that can potentially account for unobserved shocks that correlate with misconduct potential. Standard errors are clustered at the court-by-year level.

Results Table 2 presents the results. Each column represents a different regression and reports the estimated benchmark coefficient, $\hat{\alpha}_D$, with its corresponding standard error. In the absence

of OVB, $\hat{\alpha}_D$ is a valid measure of discrimination, with negative values accounting for discrimination against immigrant defendants. Column 1 shows that the raw (uncontrolled) release rate for immigrants is 6.7 pp lower than the release rate for defendants that are Chilean nationals. The difference is quantitatively important: it represents 8% of the unconditional average release rate or, put differently, almost 42% of the unconditional average detention rate.

If there are differences in misconduct potential between immigrant defendants and Chilean defendants, this result could be explained by OVB. Columns 2, 3, and 4 explore how the gap changes when adding the specific sets of controls previously discussed. When adding the individual controls (Column 2), the release rate gap increases to 10.3 pp. Under the assumption that criminal records and type of crime are positively correlated with misconduct potential, this suggests that the raw gap cannot be explained by differences in latent risk. In fact, the result conjectures that ignoring differences in misconduct potential may lead to an underestimation of the measure of discrimination. When adding the (residualized) controls for the assigned judge and attorney (Column 3), the release disparity remains unchanged. This is consistent with judges and attorneys being quasi-randomly assigned at the court-by-year level. When adding court-by-year fixed effects (Column 4), the release rate gap decreases to 3 pp. There are two potential explanations for this: The first explanation is that court-by-time fixed effects could mediate discrimination against immigrant defendants if they are systematically assigned to more severe courts (i.e., IVB). The second explanation is that the 3 pp decrease captures unobserved heterogeneity in misconduct potential that varies at the court-by-year level. Below we present suggestive evidence in favor of the first interpretation. Finally, when adding the three sets of controls (Column (5)), the gap in release rates between immigrant defendants and Chilean defendants is 8.7 pp. This represents 10.3% of the unconditional average release rate or, put differently, 54.3% of the unconditional average detention rate. We interpret these results as strong evidence of discrimination against defendants that are immigrants.

4.2 Assessing Omitted Variable Bias

The main threat to identification in benchmark regressions is OVB. If there are omitted variables that predict misconduct potential and that are correlated with immigration status, then the discrimination estimate will be biased. We conduct two sets of diagnostics to assess the extent to which OVB affects our main estimations. Both exercises suggest that unobservables that predict misconduct potential are unlikely to explain the results and, if anything, they are likely to exert downward bias on our discrimination estimates. The second diagnostic also provides elements for assessing the degree of IVB in our benchmark regressions.

Diagnostic 1: Bounds on selection on unobservables One way of assessing the role of unobservables is to follow the extension proposed by [Oster \(2019\)](#) to the method proposed in [Altonji et al. \(2005\)](#). This approach interprets observables as unobservables that the econometrician happened to observe, establishing an implicit relationship between them. Under this interpretation, and assuming some degree of correlation between X_i^o and X_i^u , the sensitivity of $\hat{\alpha}_D$ to the inclusion of observed control variables is informative about potential OVB because the bias arising from omitting observables is related to the bias induced by the unobserved variables. This allows for the computation of bounds on $\hat{\alpha}_D$ given different assumptions on the informativeness of observables about unobservables.

Based on (5), let $W_{1i} = X_i^o \alpha_{X_o}$ and $W_{2i} = X_i^u \alpha_{X_u}$, and define $\sigma_{jI} = \mathbb{C}(W_{ji}, I_i)$ and $\sigma_j^2 = \mathbb{V}(W_{ji})$. The intuition behind the method proposed in [Altonji et al. \(2005\)](#) and extended by [Oster \(2019\)](#) is to assume that the correlation between the vectors W_{ji} and I_i is proportional—that is,

$$\delta \frac{\sigma_{1I}}{\sigma_1^2} = \frac{\sigma_{2I}}{\sigma_2^2}, \quad (6)$$

with δ being the coefficient of proportionality. This can be rationalized by assuming a correlation structure, in a projection sense, between the vectors X_i^o and X_i^u . [Oster \(2019\)](#) assumes $\delta = 1$. That is to say, biases from omitting observables and unobservables are of similar magnitudes. Bounds, however, can be computed under different assumptions on δ to accommodate different a priori correlation structures.⁶

With an argument based on auxiliary regressions, [Oster \(2019\)](#) characterizes OVB as a function of δ that naturally suggests bounds on the estimated coefficient. [Oster \(2019\)](#) also notes that the degree to which the sensitivity analysis informs us about OVB depends on the variance of ξ_i or on the contribution of X_i to the variance of $Release_i$, conditional on I_i . Intuitively, coefficient stability may also arise from lack of explanatory power, so the implied OVB has to be scaled by the explanatory power of the observed and unobserved covariates. This implies that the econometrician also has to make an assumption about R_{max} —that is, the hypothetical R^2 that would be obtained after estimating (5) by OLS including both X_i^o and X_i^u . Bias is less likely to be problematic if coefficient stability after the inclusion of X_i^o is accompanied by important increases in goodness of fit relative to R_{max} .

Table 3 presents the estimated bounds for the coefficient $\hat{\alpha}_D$ under different assumptions on δ and R_{max} . Each cell shows the estimated $\hat{\alpha}_D$ parameter adjusted for OVB under the assumption

⁶Assuming $\delta = 1$ mimics a setting in which X_i^o is a random subsample of X_i . If one assumes that the most important elements of X_i are included in X_i^o , then δ is likely to be smaller than one.

Table 3: Oster (2019) Test for Coefficient Stability: $\hat{\alpha}_D$

R_{max}	δ					
	-1	-0.5	-0.25	0.25	0.5	1
0.25	-0.081	-0.084	-0.085	-0.088	-0.090	-0.093
0.5	-0.056	-0.071	-0.079	-0.095	-0.103	-0.121
0.75	-0.033	-0.059	-0.072	-0.102	-0.117	-0.151
1	-0.012	-0.047	-0.066	-0.109	-0.132	-0.185

Notes: This table presents bounds for the estimated coefficient $\hat{\alpha}_D$ that correct for potential OVB by implementing the method developed by Oster (2019), under different assumptions on δ and R_{max} . δ is the coefficient of proportionality that accounts for the correlation between observables and unobservables. $\delta = 0$ is not included because it rules out OVB by construction. R_{max} is the R^2 of a regression that includes observables and unobservables.

that selection on observables informs us about selection on unobservables. Note that larger values for δ and R_{max} imply more conservative bounds on $\hat{\alpha}_D$ because both impose a stronger role for unobservables in the release equation.

Imposing $\delta > 0$, which is the standard assumption, suggests that discrimination is underestimated: bounds are larger (in absolute value) than the fully controlled benchmark regressions displayed in Column 5 of Table 2. This follows from the fact that, on average, immigrant defendants have cleaner criminal records than Chilean defendants. It also relates to the discussion on Column 2 in Table 2, which posited that, if anything, unobservables exert downward bias on the discrimination estimates. The results show that the only way in which OVB works against finding discrimination is by assuming that X_i^o and X_i^u are negatively correlated—that is, that defendants with observables that induce lower misconduct potential also have unobservables that work in the opposite direction. Although it is counterintuitive, this scenario can be explored by assuming negative values for δ . Note that, even in the most conservative (and presumably implausible) scenario in which the negative correlation is large ($\delta = -1$) and the explanatory power of omitted variables is sizable ($R_{max} = 1$), the point estimate still suggests (mild) discrimination against immigrant defendants. These results suggest that our main findings are unlikely to be driven by unobservables that explain misconduct potential and correlate with immigration status. By contrast, we provide suggestive evidence that, if anything, the results in Table 2 underestimate the level of discrimination against immigrant defendants in pretrial release decisions.

Diagnostic 2: Using the outcome equation to test for omitted variable bias We develop an alternative test to assess the role of unobservables in explaining the estimated release disparities. This approach is based on our theoretical framework and uses the estimation of the outcome

equation (controlling for selection bias) to validate the benchmark equation. Intuitively, relevant omitted variables predict misconduct potential and correlate with immigration status. A regression of observed pretrial misconduct against immigration status controlling for our proxies of misconduct potential should therefore inform about the partial correlation between being an immigrant and pretrial misconduct.

If we assume linear models for simplifying the exposition, then the release and outcome equations can be written as

$$Release_i = \alpha_0 + \alpha_D I_i + X_i^{o'} \alpha_{X_o} + X_i^{u'} \alpha_{X_u} + \xi_i, \quad (7)$$

$$PM_i = \beta_0 + \beta_D I_i + X_i^{o'} \beta_{X_o} + X_i^{u'} \beta_{X_u} + \epsilon_i. \quad (8)$$

In what follows, the estimated coefficients from regressions that include (I_i, X_i^o, X_i^u) are denoted by $(\hat{\alpha}_D^{ou}, \hat{\beta}_D^{ou})$, and the estimated coefficients from regressions that only include (I_i, X_i^o) are denoted by $(\hat{\alpha}_D^o, \hat{\beta}_D^o)$. Concerns about OVB in observational benchmark regressions (possibly) imply that $\mathbb{E}[\hat{\alpha}_D^{ou}|I_i, X_i] \neq \mathbb{E}[\hat{\alpha}_D^o|I_i, X_i^o]$. We argue that failing to reject $\beta_D = 0$ in an outcome regression that omits X_i^u implies that $\mathbb{E}[\hat{\alpha}_D^{ou}|I_i, X_i] = \mathbb{E}[\hat{\alpha}_D^o|I_i, X_i^o]$; thus, the outcome regression can be used as an indirect test for OVB in the release equation.

To see why, recall that, by assumption, $X_i = [X_i^o \ X_i^u]$ is an accurate proxy of misconduct potential, implying that $\beta_D = 0$: immigration status cannot predict pretrial misconduct after controlling for misconduct potential. The hypothetical (unselected) regression of PM_i versus I_i and X_i yields unbiased estimates. Specifically, $\mathbb{E}[\hat{\beta}_D^{ou}|I_i, X_i] = \beta_D = 0$.⁷ This logic does not hold in a regression that omits X_i^u because $\mathbb{E}[\hat{\beta}_D^o|I_i, X_i^o] = \beta_D + bias$, where the bias depends on β_{X_u} and the correlation between I_i and X_i^u through a standard OVB formula. Assuming $\beta_{X_u} \neq 0$ (the contrary would imply that OVB is ruled-out by assumption), nonzero estimates of $\hat{\beta}_D^o$ indicate that the correlation between I_i and X_i^u is nonzero. Intuitively, I_i becomes a good predictor for pretrial misconduct only if it correlates with elements of X_i that are not observed by the econometrician.

The nonzero correlation between I_i and X_i^u is exactly what induces bias when estimating discrimination in the release equation without observing X_i^u . In particular, $\mathbb{E}[\hat{\alpha}_D^{ou}|I_i, X_i] = \mathbf{D}$ because controlling by X_i implies that $\hat{\alpha}_D^{ou}$ identifies the direct effect of I_i , as well as the effects of other unobservables that are correlated with I_i but that are not related with misconduct potential (con-

⁷This requires us to assume that ϵ_i is uncorrelated with I_i , so $\mathbb{E}[\epsilon_i|I_i, X_i] = \mathbb{E}[\epsilon_i|X_i] = 0$. If we think of ϵ_i as the projection error, $Y_i^* = g(X_i) + \epsilon_i$, then the second equality is true by construction, and the first equality requires us to assume that the projection error does not vary with immigration status, which is a plausible assumption given that X_i^u is unrestricted (eventually making ϵ_i negligible) and allowed to correlate with I_i .

tained in ξ_i) and, therefore, constitute valid sources of discrimination under a disparate impact perspective. By contrast, $\mathbb{E}[\hat{\alpha}_D^o | I_i, X_i^o] = \mathbf{D} + \text{bias}$, where the bias depends on α_{X_u} and, again, the correlation between I_i and X_i^u . Note that the correlation that induces OVB in the release equation is the same that induces nonzero $\hat{\beta}_D^o$ coefficients in the outcome equation.

Putting the two arguments together, we note that failing to reject $\beta_D = 0$ in a regression that omits X_i^u (i.e., when the point estimate is $\hat{\beta}_D^o$) implies $\mathbf{D} = \mathbb{E}[\hat{\alpha}_D^{ou} | I_i, X_i] = \mathbb{E}[\hat{\alpha}_D^o | I_i, X_i^o]$. Put simply, not rejecting that immigration status does not predict pretrial misconduct on top of X_i^o confirms that the vector X_i^o successfully controls for group differences in misconduct potential and, therefore, the discrimination estimates from equation (7) are not affected by standard OVB. One challenge in implementing this test is that the sample of released defendants is selected because pretrial misconduct if released is not observed for detained individuals. Accordingly, we estimate selection models that use the judge and attorney controls as the excluded variables in the outcome equation based on the assumption that judges' leniency and attorneys' quality affect selection into treatment but do not affect misconduct potential.^{8,9} Our main specification implements a standard two-step parametric Heckit selection correction (Heckman, 1974). We also provide results using the semiparametric correction proposed by Newey (2009).¹⁰

Table 4 presents the results using the parametric selection correction. Table A.I of Appendix A presents the results using the semiparametric selection correction. Each column represents a different regression and reports the estimated $\hat{\beta}_D^o$ coefficient from equation (8) with its corresponding standard error. In all regressions, the selection equation includes the complete set of controls, but the included covariates in the outcome equation vary between columns. Note that the F-test for judge leniency and attorney quality (and their squares) suggests they effectively help with the

⁸To avoid saturating the nonlinear first-stage model with the court-by-time fixed effects, we replace them with time-varying variables at the court level: number of judges, average pretrial release rate, and number of prosecutions (within a court in a given year). Estimating the fully controlled benchmark regression (Column 5 in Table 2) and replacing the court-by-time fixed effects by these time-varying court-level controls yields a point estimate of -0.084, suggesting that the time-varying variables do a good job of approximating the court-by-time fixed effects.

⁹Although the test proposed here is different from the weighting correction made in Arnold et al. (2021), it is worth noting that their proposal is also based on the intuition that the unselected outcome equation can be used to assess the degree of OVB in the selection equation.

¹⁰The Heckit selection correction assumes that the error terms of the first and second steps are jointly normal. The semiparametric correction of Newey (2009) uses series approximations to compute control function corrections. We implement the semiparametric correction following Low and Pistaferri (2015) where the first step uses Gallant and Nychka (1987) estimator to approximate the unknown density by third degree Hermite polynomial expansions and the second step controls for non-linear transformations of the density prediction. As in Low and Pistaferri (2015), we consider three models. Let \hat{f} denote the predicted density. The control function used in Model I is \hat{f} and its square, in Model II is $\Phi(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{f})$ and its square—where Φ is the normal cumulative distribution function and $(\hat{\alpha}_0, \hat{\alpha}_1)$ are the estimated coefficients of a Probit model of *Release* on a constant and \hat{f} —, and in Model III is $\lambda(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{f})$ and its square—where $\lambda(x) = \phi(x)/\Phi(x)$ is the inverse Mills ratio and ϕ the normal density.

Table 4: Parametric Selection Model for Assessing OVB: $\widehat{\beta}_D^o$

	(1)	(2)	(3)	(4)
Immigrant	-0.078 (0.007)	-0.012 (0.006)	-0.090 (0.007)	-0.008 (0.006)
Mean dep. variable	0.29	0.29	0.29	0.29
Individual controls	No	Yes	No	Yes
Court-by-year controls	No	No	Yes	Yes
No. of Immigrants	4,900	4,900	4,900	4,900
No. of Chileans	580,406	580,406	580,406	580,406
F-test for excluded variables	969.6	969.6	969.6	969.6

Notes: This table presents the results from the two-step parametric sample selection model (Heckman, 1974) with different sets of controls, where the dependent variable is pretrial misconduct. We report the point estimate for the immigrant indicator (i.e., the coefficient $\widehat{\beta}_D^o$) of equation (8) and its standard error. Both sets of controls (*individual controls* and *court-by-year controls*) are always included in the selection equation, but the columns vary in their inclusion in the outcome equation. Judge and attorney controls are defined as in Table 2 and are excluded from the outcome equation. The last row presents the joint F-test for judge and attorney (excluded) controls in the selection equation. *Individual controls* are defined as in Table 2. To avoid saturating the nonlinear first-stage with court-by-year fixed effects they are replaced in the regressions by court-by-year time varying covariates—namely, the average number of judges, the average pretrial release rate, and the number of prosecutions (within a court in a given year).

selection correction.

Column 1 in Table 4 shows that, after controlling for selection, pretrial misconduct rates are 7.8 pp smaller for immigrant defendants. This suggests that the benchmark regression with no controls (Column 1 in Table 2) is affected by OVB. The implied OVB, however, suggests that abstracting from controls leads to an underestimation of the discrimination against immigrant defendants. Column 2 shows that including individual controls (criminal records and type of crime) reduces the pretrial misconduct rate differences to 1.2 pp. That is to say, conditioning on the individual controls reduces group variation in Y_i^* by 85%. This suggests that, although imperfect, criminal records and type of crime are good predictors of pretrial misconduct potential. Adding time-varying court controls do little to correct for unobserved differences in misconduct potential (Column 3); however, adding both individual and time-varying court controls reduces the misconduct rate disparities to a nonsignificant gap of less than a percentage point. Taken together, we interpret this as support for our choice of X_i^o . The vector X_i^o seems to do a good job in accounting for group differences in misconduct potential, which implies that the discrimination estimates of Column 5 of Table 2 are not driven by unobserved variation in latent risk. The model with the semiparametric selection correction proposed by Newey (2009) yield the same conclusions.¹¹

¹¹The three models reproduce the pattern displayed in Table 4. After accounting for selection, immigrants have lower pretrial misconduct rates. These gaps are substantially reduced after including individual controls in the

A similar test can be implemented by comparing the role of unobservables in the selection and outcome equations using Kitagawa–Oaxaca–Blinder (KOB) decompositions (Kitagawa, 1955; Oaxaca, 1973; Blinder, 1973). Intuitively, unobservables can affect the interpretation of the estimated benchmark coefficients if they correlate with immigration status and explain misconduct potential. If unobservables are important for predicting group-specific misconduct potential when making release decisions, they should also explain differences in observed pretrial misconduct rates.¹² Appendix C shows KOB decompositions based on the vectors of individual covariates, using both raw data and residualized data against court-by-time fixed effects. The share of variation that is not explained by observables is large in the release equation but it is essentially zero in the outcome equation. This suggests that the immigrant indicator in the release equation is effectively capturing discrimination and not differences in unobservables.

On included variable bias Our diagnostics do not explicitly deal with IVB concerns, but the second OVB diagnostic provides some insights to assess the pervasiveness of IVB. Individual controls are much more important than court-by-time variation in terms of accounting for group differences in pretrial misconduct rates. This indicates that the inclusion of court-by-time fixed effects in the benchmark regression may induce IVB in the sense that these particular fixed effects impact the estimated discrimination coefficient but do little to attenuate OVB. In other words, regressions that include court-by-year fixed effects could be over controlling and therefore inducing bias to our discrimination estimates.

IVB affects the interpretation of the sensitivity analysis. As shown in Table 2, including court-by-year fixed effects significantly reduces the discrimination estimate. The results presented in Table 4 suggest that the sensitivity of $\hat{\alpha}_D$ to the inclusion of court-by-year fixed effects is not explained by unobserved variation in misconduct potential but rather by an alternative mechanism that mediates discrimination against immigrant defendants. For example, the fact that discrimination decreases when controlling by court-by-time fixed effects may indicate that part of the discrimination against immigrant defendants is driven by their systematic allocation to courts that are more severe on

outcome equation. In Model I, individual controls reduce the gap by two thirds, although the gap remains significant at 3 pp after including the complete set of controls. In Models II and III the individual controls completely close the gap in pretrial misconduct to a nonsignificant estimate of less than a percentage point, even in the specification that excludes time-varying court characteristics.

¹²One caveat of this analysis is that the sample used in the outcome equation is selected, presumably based on unobservables. To the extent that the selected sample has variation in unobservables, however, the exercise should still capture their effects. This would not be the case if, for example, the release decision is based on a binary unobservable such that all released defendants have the same value. In this case, the lack of variation in the selected sample would prevent the decomposition from capturing the effect of unobservables—that is, the KOB exercise could be noninformative under specific patterns of selection on unobservables.

average. Therefore, from a disparate impact perspective, the relevant discrimination estimate could be the one that only includes individual controls (Column 2 of Table 2), which suggests larger levels of discrimination against immigrant defendants than the fully controlled specification. Because this conjectured IVB attenuates the discrimination estimate, we take a conservative position by presenting the fully controlled benchmark regression as our preferred specification. It is important to keep in mind, however, that the impact of IVB may mean that the actual discrimination against immigrant defendants is even larger.

5 Dissecting the Discrimination Estimates

Section 4 documents a robust pretrial release rate disparity between immigrant defendants and Chilean defendants after controlling for variables that proxy for misconduct potential that we interpret as evidence of discrimination against immigrant defendants. This section dissects this reduced-form result to get a more comprehensive characterization of the discrimination patterns against immigrant defendants. All results presented in this section consider the full set of controls and should therefore be compared to Column 5 in Table 5.

First, we explore some behavioral drivers of the discrimination estimate. We show that a sizable portion of the effect stems from an informational problem: immigrant criminal records are censored and this leads judges to punish immigrants with no previous prosecutions. Outcome test estimations show that the result cannot be rationalized by accurate statistical discrimination, suggesting that biased beliefs and taste-based discrimination play a role in the resolution of the informational problem. Second, we explore heterogeneities by type of crime and find that discrimination is especially large for drug-related offenses. This is consistent with a hostile discourse in the public sphere against immigrants that links immigration to drug trafficking. Finally, we explore the time trend of discrimination patterns considering the recent immigration wave discussed in Section 2. We find that discrimination has dramatically increased in recent years, and that this increase is not explained by the change in the composition of the immigrant population.

5.1 Behavioral Drivers of Discrimination

The criminal records of immigrant defendants are censored (i.e., judges cannot observe the criminal histories of immigrant defendants in their origin countries). This means that judges may interpret the information available to them differently. In Bayesian jargon, the quality of the signal may differ between Chilean defendants and immigrant defendants. This is likely to be more detrimental

Table 5: Behavioral Drivers of Discrimination and the Role of Information

	Benchmark Test			Outcome Test	
	Previous Prosecution			KPT (4)	GV (5)
	Yes (1)	No (2)	(3)		
Immigrant	-0.019 (0.007)	-0.132 (0.015)	-0.135 (0.023)	-0.050 (0.007)	-0.074 (0.016)
Immigrant x experienced judge			0.049 (0.022)		
Mean dep. variable	0.80	0.92	0.82	0.29	0.37
No. of Immigrants	2,630	3,732	4,187	4,900	840
No. of Chileans	472,636	220,370	357,520	580,406	57,690
R-squared	0.18	0.17	0.21		

Notes: Columns 1 to 3 present the estimated benchmark coefficient, $\hat{\alpha}_D$, of the extensions of equation (5) and its corresponding standard error (clustered at the court-by-year level). Each column represents a different regression and includes all sets of controls analogous to Column 5 in Table 2 (see Table 2 notes for details). Columns 1 and 2 present $\hat{\alpha}_D$ separately for defendants with and for defendants without previous prosecutions. Column 3 presents $\hat{\alpha}_D$ for the period 2013–2017 and includes an interaction between I_i and an indicator that takes the value one if the case was handled by an experienced judge (i.e., judges that handled 200 or more cases in the 2008–2012 period). Columns 4 and 5 present the outcome test estimations using Knowles et al. (2001) (KPT) and Grau and Vergara (2021) (GV) approaches, respectively. KPT approach consists on an OLS regression of pretrial misconduct on the immigrant indicator for the complete sample of released defendants. GV performs a sample selection procedure that limits the extent of inframarginality bias in the KPT implementation. Details of the outcome test implementation of GV can be found in Appendix D.

to immigrant defendants with no criminal history in Chile because the main information contained in the criminal records is at the extensive margin. We explore whether this informational gap is in fact driving discrimination patterns against immigrant defendants.

To test the information hypothesis, we estimate the main benchmark regressions separately for defendants with and for defendants without previous prosecutions. These estimates include all sets of controls analogous to Column 5 in Table 2. Results are displayed in Columns 1 and 2 in Table 5. Immigrant defendants with previous prosecutions are 1.9 pp less likely to be released compared to Chilean defendants with previous prosecutions; compared to a difference of 13.2 pp for defendants without previous prosecutions. To assess the magnitude of the difference between the estimations, note that the (unconditional) pretrial release rate is 12 pp larger for defendants with clean criminal records, which is roughly similar in magnitude to the difference in the discrimination estimate between samples. This implies that not having previous prosecutions is, essentially, irrelevant for immigrant defendants: they are not “rewarded” for having clean criminal records. This result is consistent with judges systematically imputing nonzero criminal histories for all immigrant defendants regardless of their criminal record in Chile.

To further explore this intuition, we test whether judge experience matters for the aggregate discrimination patterns. If judges face an informational problem whose imperfect solution leads to discriminatory practices, they should gradually implement nondiscriminatory solutions as they gain experience. Consequently, we should observe that more experienced judges exhibit lower levels of discrimination. To test this, we define experienced judges as those who handled more than 200 cases in the period 2008–2012 and estimate the benchmark regressions for the period 2013–2017, including the experience indicator and its interaction with the immigrant indicator. Column 3 shows that release disparities are around 35% smaller for experienced judges.

As discussed in Section 3, aggregate patterns of discrimination represent a possibly complex combination of different behavioral sources—these being, statistical discrimination, biased beliefs, and taste-based discrimination, which is not possible to identify separately without additional structural assumptions (Arnold et al., 2021; Bohren et al., 2021; Hull, 2021). The fact that discrimination in pretrial detention decisions stems from an informational problem does not necessarily mean that the aggregate estimate can be rationalized by accurate statistical discrimination. On the one hand, judges can circumvent the informational problem based on stereotypes. This may, for example, lead judges to “overcorrect” for censored criminal records for immigrants. On the other hand, the informational problem may provide judges with the room to exercise taste-based discrimination. The fact that experienced judges still display significant discrimination against immigrants suggests that these drivers are likely to play a role in the analysis.

To get additional insights, we implement the outcome test (Becker, 1957, 1993) using the observational approaches of Knowles et al. (2001) and Grau and Vergara (2021).¹³ The outcome test is a diagnostic for differences in effective selection thresholds between groups. Using the notation set out in Section 3 (equation (2)), the test explores whether $\mathbb{E}[t_{j(i)}(1, Z_i) - b_{j(i)}(1, Z_i)]$ is different from $\mathbb{E}[t_{j(i)}(0, Z_i) - b_{j(i)}(0, Z_i)]$, where the expectation is taken across Z_i and $j(i)$. Differences in effective thresholds can be identified by differences in observed misconduct among marginally released defendants (Arnold et al., 2018; Hull, 2021). Knowles et al. (2001) propose using the average behavior of released defendants. Grau and Vergara (2021) refine Knowles et al. (2001) approach by performing a sample selection procedure that limits the extent of inframarginality bias.¹⁴

¹³The small share of the immigrant population prevents us from implementing the instrument-based approach proposed by Arnold et al. (2018).

¹⁴When risk distributions vary by group, the behavior of the average released defendant is potentially different from the behavior of the marginally released defendants. This is called inframarginality bias. Grau and Vergara (2021) deal with this problem in two stages. First, it uses observational variation to identify samples of marginal defendants. The identification argument works under sufficient conditions which are supported by suggestive evidence presented in Appendix D. Second, it performs differences in means of pretrial misconduct rates between immigrant defendants and Chilean defendants using the samples of marginal defendants. The sample of marginal defendants is defined as

Intuitively, discrimination in pretrial detention decisions may originate from judges setting stricter release thresholds for immigrant defendants, in a disparate impact sense, that imply lower release rates at equivalent true pretrial misconduct probabilities. These differences in the treatment of immigrant defendants could be driven by stereotypes or animus or both. This source of discrimination does not incorporate sources of release disparities driven by differences in group-specific misconduct potential. Differences in true pretrial misconduct potential affect how often a given defendant crosses the release threshold, but the threshold itself is not informative about them. Then, rejecting the outcome test implies that the observed behavior of marginally released defendants is inconsistent with a selection process that lacks stereotypes and animus (Hull, 2021).

Columns 4 and 5 in Table 5 show the results of the outcome test using Knowles et al. (2001) and Grau and Vergara (2021) approaches, respectively. Estimates indicate that immigrant defendants face effective release thresholds that are, on average, between 5 and 7.2 pp stricter. That is to say, immigrants are required to have, on average, smaller pretrial misconduct probabilities to qualify for pretrial release relative to Chilean defendants. Although the magnitudes between benchmark and outcome regressions are not directly comparable (Domínguez et al., 2022), this result rejects accurate statistical discrimination as the only driver of the discrimination estimates, implying that a combination of stereotypes and animus plays an important role in explaining aggregate patterns of discrimination against immigrants.¹⁵

5.2 Heterogeneity by Type of Crime

If statistical structures, stereotypes, or judges preferences vary with type of crime, then discrimination patterns should do as well. Exploring this particular pattern of heterogeneity is instructive because a hostile facet of the public discourse regarding immigrants links immigration with drug trafficking and violent crimes, so discrimination should be particularly salient in pretrial detention decisions concerning these crimes.

To investigate this, we split prosecutions between the following nine mutually exclusive categories: homicides, sexual offenses, thefts and robberies, other property crimes, drug offenses, white collar crimes and tax crimes, crimes against public trust, crimes against the freedom and privacy

the 10% of released defendants with larger conditional probabilities of being marginal. That is why the number of observations is smaller when using this approach. Additional details about the implementation of the outcome test proposed by Grau and Vergara (2021) are presented in Appendix D.

¹⁵The average pretrial misconduct rate is larger in Column 5 (37%) relative to Column 4 (29%). This suggests that Grau and Vergara (2021) implementation of the outcome test successfully attenuates inframarginality concerns by focusing on a sample of released defendants that are, on average, riskier and, therefore, presumably more likely to be at the margin of pretrial detention.

Table 6: Discrimination by Type of Crime: $\hat{\alpha}_D^k$

	Coeff.	SE
Homicide	-0.030	(0.067)
Sexual offense	0.001	(0.033)
Theft or robbery	-0.038	(0.018)
Other property crime	-0.052	(0.014)
Drug offense	-0.351	(0.034)
White collar or tax crime	-0.020	(0.019)
Crime against public trust	0.021	(0.015)
Crime against people’s freedom and privacy	-0.007	(0.006)
Other crimes	0.019	(0.021)
Mean dep. variable	0.84	
No. of Immigrants	6,362	
No. of Chileans	693,006	
R-squared	0.192	

Notes: This table presents the estimated benchmark coefficients by type of crime, $\hat{\alpha}_D^k$, of equation (9) and its corresponding standard error in parentheses (clustered at the court-by-year level). Each column represents a different regression and includes all sets of controls analogous to Column 5 in Table 2 (see Table 2 notes for details).

of people, and other crimes.¹⁶ Using these categories, we estimate the following regression:

$$Release_i = \alpha_0 + \sum_{k=1}^9 \alpha_D^k I_i \cdot C_i^k + X_i^{o'} \alpha_{X_o} + \varepsilon_i, \quad (9)$$

where C_i^k is an indicator that takes the value one if the imputed crime is in category k . All other variables are defined as in equation (5); therefore, the noninteracted crime indicators are included in X_i^o . In this specification, $\hat{\alpha}_D^k$ identifies the controlled pretrial release disparities between Chilean defendants and immigrant defendants (our measure of discrimination) in crime category k .

Table 6 shows the results. Discrimination against immigrants is especially large for drug offenses—around four times larger than the baseline benchmark estimate. Significant effects are also found for thefts and robberies and property crimes. Other violent crimes like homicides and sexual offenses show no significant differences; however, as shown in Table A.II of Appendix A, the number of immigrants imputed for sexual offences and homicides in the period considered is too low, therefore potentially being underpowered to precisely estimate disparities in those particular categories. Crime categories that are associated with less hostile discourses in the public sphere (such as white collar crimes and tax crimes or crimes against privacy, and for which the number of

¹⁶For examples of the crimes included in each category, see Appendix B.

Table 7: Discrimination in Different Periods and Group Heterogeneity: $\hat{\alpha}_D$

	2008-2012		2013-2017	
	(1)	(2)	(3)	(4)
Immigrant	-0.066 (0.016)		-0.101 (0.015)	
Immigrant (old group)		-0.064 (0.017)		-0.099 (0.015)
Immigrant (new group)				-0.108 (0.022)
Mean dep. variable	0.85	0.85	0.82	0.82
No. of Immigrants (old group)	1,956	1,956	2,873	2,873
No. of Immigrants (new group)	219	0	1,314	1,314
No. of Chileans	335,486	335,486	357,520	357,520
R-squared	0.172	0.172	0.207	0.207

Notes: This table presents the estimated benchmark coefficient, $\hat{\alpha}_D$, of extensions for equation (5) and its corresponding standard error (clustered at the court-by-year level). Each column represents a different regression and includes all sets of controls analogous to Column 5 in Table 2 (see Table 2 notes for details). Columns 1 and 2 use data from the period 2008–2012, and Columns 3 and 4 use data from the period 2013–2017. In columns 2 and 4 the immigrant indicator is defined separately depending on the nationality of the immigrant. *New group* includes immigrants from Colombia, Haiti, and Venezuela. *Old group* considers immigrants from all the other countries.

prosecuted immigrants is in fact larger) also show no significant release disparities. Overall, these results suggest that discrimination patterns are mainly driven by drug offenses, property crimes, and thefts and robberies. This is consistent with idea that public discourses regarding immigrants play a role in the formation of stereotypes and hostile preferences.

5.3 Time Trend and the Immigration Wave

Recall in Section 2 it was premised that immigration flows have substantially increased in recent years, generating a hostile discourse against immigrants in the public sphere. It was also argued that the composition of the immigrant population by country of origin has changed considerably. We now examine whether these changes have been accompanied by a change in the aggregate discrimination patterns across time.

Following the changes in immigration flows described in Section 2, Table 7 shows the benchmark estimates separately for the periods 2008–2012 and 2013–2017. The fully controlled discrimination estimate increased from 6.6 to 10.1 pp, a 53% increase. A potential explanation for this increase is the increase in the hostility of the public discourse that may have affected stereotypes and preferences. Alternatively, this result could be driven by a composition effect. Although Peru

was the principal source of immigration flows to Chile for many years, beginning in 2013 the relative share of immigrants from Colombia, Haiti, and Venezuela increased considerably. If there is heterogeneity in discrimination patterns—that is, discrimination is more intense against the “new” group of immigrants (for example, because of more salient racial differences)—then the increase of discrimination between the two periods could be explained by a composition effect and not by an increase in the average intensity of discrimination for a given nationality.¹⁷ Column 4 presents the discrimination estimate for the 2013–2017 period separately for the “new” and the “old” groups of immigrants.¹⁸ We cannot reject that the discrimination estimates for both groups are equal, which suggests that the increase in discrimination is unlikely to be driven by the change in the composition of the immigrant population.

6 Conclusion

This paper leverages rich administrative data to test for discrimination in pretrial detention decisions against defendants who are immigrants in Chile. We find that immigrant defendants are released pretrial at substantially lower rates relative to Chilean defendants with similar criminal records, case characteristics, and court-by-time fixed effects. Several diagnostics suggest that these estimated disparities are not driven by unobserved differences in pretrial misconduct potential between Chilean defendants and immigrant defendants. We therefore interpret the results as robust evidence of discrimination against immigrant defendants in pretrial detention decisions.

We show that release disparities are an order of magnitude larger for defendants with no previous criminal prosecutions in Chile. We argue this implies that discrimination is originated in an informational problem driven by censored criminal records (or, put differently, weaker signal quality). However, we show that accurate statistical discrimination cannot rationalize the release disparities, suggesting that biased beliefs and taste-based discrimination play an important role. Consistent with the hostile discourses against immigrants in the public sphere, we also find that discrimination in pretrial detention decisions is particularly severe for drug offenses, and that the intensity of discrimination has increased in recent years during which the magnitude and the composition of the immigrant population has changed considerably.

Overall, our findings suggest that discrimination against immigrant defendants in pretrial de-

¹⁷Potential differences in risk between the two groups are not a concern for the dynamic results because benchmark regressions include the full set of controls.

¹⁸This exercise cannot be computed for the 2008–2012 period because the “new” group, by definition, is essentially nonexistent before 2013, as the number of observations suggest.

tention decisions is substantial in the Chilean context and highlight the importance of implementing effective anti-discrimination policies to better assimilate the integration of immigrants in destination countries. Several countries are enacting anti-discrimination policies (or strengthening existing ones) to foster the integration of immigrant populations (OECD, 2019). This policy challenge appears to be particularly urgent in countries where immigration is a relatively recent phenomenon and institutions are not adequately equipped to deal with unexpected shocks in this regard.

Bibliography

- Abrams, D. S., Bertrand, M., and Mullainathan, S. (2012). Do judges vary in their treatment of race? *The Journal of Legal Studies*, 41(2):347–383.
- Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2):175–187.
- Ajzenman, N., Dominguez, P., and Undurraga, R. (2021). Immigration, crime, and crime (mis) perceptions.
- Alesina, A., Harnoss, J., and Rapoport, H. (2021). Immigration and the future of the welfare state in Europe. *The Annals of the American Academy of Political and Social Science*, 697(1):120–147.
- Alesina, A., Miano, A., and Stantcheva, S. (2022). Immigration and redistribution. *Review of Economic Studies*.
- Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184.
- Arnold, D., Dobbie, W., and Hull, P. (2021). Measuring racial discrimination in bail decisions. *Working Paper*.
- Arnold, D., Dobbie, W., and Yang, C. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, 133(4):1885–1932.
- Åslund, O., Hensvik, L., and Skans, O. N. (2014). Seeking similarity: How immigrants and natives manage in the labor market. *Journal of Labor Economics*, 32(3):405–441.
- Becker, G. (1957). *The Economics of Discrimination*. University of Chicago Press.
- Becker, G. (1993). Nobel Lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101:385–409.

- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human resources*, 8(4):436–455.
- Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2021). Inaccurate statistical discrimination: An identification problem. *Working Paper*.
- Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.
- Brell, C., Dustmann, C., and Preston, I. (2020). The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives*, 34(1):94–121.
- Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2021a). The immigrant next door: Exposure, prejudice, and altruism. *Working Paper*.
- Bursztyn, L., Haaland, I. K., Rao, A., and Roth, C. P. (2021b). Disguising prejudice: Popular rationales as excuses for intolerant expression. *Working Paper*.
- Cortés, T., Grau, N., and Rivera, J. (2019). Juvenile incarceration and adult recidivism. *Working Paper*.
- Dobbie, W., Goldin, J., and Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2):201–240.
- Dobbie, W. and Yang, C. (2021a). The economic costs of pretrial detention. *Brookings Papers on Economic Activity*.
- Dobbie, W. and Yang, C. (2021b). The US pretrial system: Balancing individual rights and public interests. *Journal of Economic Perspectives*, 35(4):49–70.
- Domínguez, P., Grau, N., and Vergara, D. (2022). Combining discrimination diagnostics to identify sources of statistical discrimination. *Economics Letters*, 212, 110294.
- Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic policy*, 32(91):497–550.
- Dustmann, C. and Preston, I. (2007). Racial and economic factors in attitudes to immigration. *The BE Journal of Economic Analysis & Policy*, 7(1).
- Egger, D., Auer, D., and Kunz, J. (2022). Effects of migrant networks on labor market integration, local firms and employees. *Working Paper*.

- Fasani, F., Mastrobuoni, G., Owens, E. G., and Pinotti, P. (2019). *Does Immigration Increase Crime?* Cambridge University Press.
- Gallant, A. R. and Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, 55(2):363–390.
- Gelman, A., Fagan, J., and Kiss, A. (2007). An analysis of the New York City police department’s “stop-and-frisk” policy in the context of claims of racial bias. *Journal of the American Statistical Association*, 102(479):813–823.
- Grau, N., Marivil, G., and Rivera, J. (2021). The effect of pretrial detention on labor market outcomes. *Journal of Quantitative Criminology*.
- Grau, N. and Vergara, D. (2021). An observational implementation of the outcome test with an application to ethnic prejudice in pretrial detentions. *Working Paper*.
- Grigorieff, A., Roth, C., and Ubfal, D. (2020). Does information change attitudes toward immigrants? *Demography*, 57(3):1117–1143.
- Hangartner, D., Dinas, E., Marbach, M., Matakos, K., and Xefteris, D. (2019). Does exposure to the refugee crisis make natives more hostile? *American Political Science Review*, 113(2):442–455.
- Heath, A., Liebig, T., and Simon, P. (2013). Discrimination against immigrants - Measurement, incidence and policy instruments. *International Migration Outlook - OECD*, pages 191–230.
- Heaton, P., Mayson, S., and Stevenson, M. (2017). The downstream consequences of misdemeanor pretrial detention. *Stanford Law Review*, 69:711.
- Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42(4):679–694.
- Hopkins, D. J., Sides, J., and Citrin, J. (2019). The muted consequences of correct information about immigration. *The Journal of Politics*, 81(1):315–320.
- Hull, P. (2021). What marginal outcome tests can tell us about racially biased decision-making. *Working Paper*.
- Jørgensen, F. J. and Osmundsen, M. (2020). Correcting citizens’ misperceptions about non-western immigrants: Corrective information, interpretations, and policy opinions. *Journal of Experimental Political Science*, pages 1–10.
- Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, 50(272):1168–1194.

- Knowles, J., Persico, N., and Todd, P. (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, 109(1):203–229.
- Leslie, E. and Pope, N. G. (2017). The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments. *The Journal of Law and Economics*, 60(3):529–557.
- Low, H. and Pistaferri, L. (2015). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review*, 105(10):2986–3029.
- Newey, W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, 12:S217–S229.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.
- OECD (2019). International migration outlook. *Organisation for Economic Co-operation and Development*.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.
- Rehavi, M. M. and Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, 122(6):1320–1354.
- Rose, E. K. (2022). A constructivist perspective on empirical discrimination research. *Working Paper*.
- Stevenson, M. T. and Mayson, S. G. (2021). Pretrial detention and the value of liberty. *Working Paper*.
- UN (2013). International migration policies: Government views and priorities. *United Nations - Population Division*.
- UN (2019). International migration 2019 report. *United Nations - Department of Economic and Social Affairs*.
- Yang, C. and Dobbie, W. (2020). Equal protection under algorithms: A new statistical and legal framework. *Michigan Law Review*.

Discrimination Against Immigrants in the Criminal Justice System: Evidence from Pretrial Detentions

Online Appendix

Patricio Domínguez, Nicolás Grau, Damián Vergara

A	Additional Figures and Tables	i
B	Data appendix	iii
B.1	Sources	iii
B.2	Estimation sample	iii
B.3	Variables	v
C	Kitawaga-Oaxaca-Blinder decompositions	vii
D	Outcome test	ix

A Additional Figures and Tables

Table A.I: Semiparametric Selection Model for Assessing OVB: $\hat{\beta}_D^o$

	(1)	(2)	(3)	(4)
Model I: Immigrant	-0.091 (0.0064)	-0.102 (0.0064)	-0.029 (0.0061)	-0.030 (0.0061)
Model II: Immigrant	-0.081 (0.0064)	-0.093 (0.0064)	-0.006 (0.0061)	-0.006 (0.0061)
Model III: Immigrant	-0.080 (0.0064)	-0.092 (0.0063)	-0.005 (0.0061)	-0.005 (0.0061)
Mean dep. variable	0.29	0.29	0.29	0.29
Court-by-year characteristics	No	Yes	No	Yes
Individual controls	No	No	Yes	Yes
No. of Immigrants	4,900	4,900	4,900	4,900
No. of Chileans	580,406	580,406	580,406	580,406

Notes: This table presents the results of the OVB test proposed in Section 4 using the semiparametric correction of Newey (2009) that uses series approximations to compute control function corrections. We implement the semiparametric correction following Low and Pistaferri (2015) where the first step uses Gallant and Nychka (1987) estimator to approximate the unknown density by third degree Hermite polynomial expansions and the second step controls for non-linear transformations of the density prediction. As in Low and Pistaferri (2015), we consider three models. Let \hat{f} denote the predicted density. The control function used in Model I is \hat{f} and its square, in Model II is $\Phi(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{f})$ and its square—where Φ is the normal cumulative distribution function and $(\hat{\alpha}_0, \hat{\alpha}_1)$ are the estimated coefficients of a Probit model of *Release* on a constant and \hat{f} —, and in Model III is $\lambda(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{f})$ and its square—where $\lambda(x) = \phi(x)/\Phi(x)$ is the inverse Mills ratio and ϕ the normal density. We report the point estimate for the immigrant indicator (i.e., the coefficient $\hat{\beta}_D^o$) of equation (8) and its standard error. Standard errors are computed using bootstrap with 500 repetitions to account for the fact that the density is estimated in the first stage. Both sets of controls (*individual controls* and *court-by-year controls*) are always included in the selection equation, but the columns vary in their inclusion in the outcome equation. Judge and attorney controls are defined as in Table 2 and are excluded from the outcome equation. *Individual controls* are defined as in Table 2. To avoid saturating the nonlinear first-stage with court-by-year fixed effects they are replaced in the regressions by court-by-year time varying covariates—namely, the average number of judges, the average pretrial release rate, and the number of prosecutions (within a court in a given year).

Table A.II: Crime Type Distribution

	Chilean		Immigrant	
	%	N	%	N
Homicide	0.01	6,384	0.01	69
Sexual offense	0.02	12,275	0.03	188
Theft or robbery	0.26	180,086	0.17	1,104
Other property crime	0.18	126,335	0.16	1,038
Drug offense	0.12	85,936	0.21	1,306
White-collar or tax crime	0.02	11,946	0.02	142
Crime against public trust	0.06	44,031	0.07	464
Crime against people's freedom and privacy	0.29	198,834	0.29	1,816
Other crimes	0.04	27,179	0.04	235

Notes: This table presents the crime type distribution, by nationality, for the estimation sample. Shares are calculated to sum 100% within Chileans and immigrants.

B Data appendix

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

B.1 Sources

We merge two 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 by the PDO alone or those handled in coordination with a private attorney (as opposed to cases handled only by a private attorney), which covers more than 95% of the universe of criminal cases in Chile. The unit of analysis is a criminal prosecution and contains defendants characteristics (ID, name, gender, nationality, and place of residence, among other characteristics) and case characteristics (case ID, court, public attorney assigned, initial and end dates, different categories for the 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 information on arraignment judges and their assigned cases for arraignment 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 other than their names and IDs. This data was shared by the Department of Studies at the Chilean Supreme Court (<https://www.pjud.cl/corte-suprema>).

B.2 Estimation sample

The initial sample contains 3,571,230 cases and covers all the cases recorded by the PDO that had an arraignment hearing 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 we also drop observations where the length of pretrial detention is greater than the length of the case. After these adjustments, the sample size reduces to 3,559,019 (i.e., the number of cases reduces by 12,211).

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 available to the judges is likely to be different.
- We drop cases involving juvenile defendants (254,243 observations). We do this because the juvenile criminal justice system works differently, and the mandated selection rule and the preventive measures differ between systems (see [Cortés et al., 2019](#) for details).
- We drop cases where the defendant hires a private attorney as their 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 that are longer than two years in duration (55,495 observations).
- For defendants that are accused of more than one crime in a given case and the records provide multiple observations, we consider the most severe crime (see below for the severity definition). In this step we drop 200,412 observations. To be clear, we do not drop defendants, only cases. We do this to have at most one case/defendant pair per day of arraignment hearing.
- We drop cases where the detention judge ID is missing (66,975 observations).
- We drop the types of crime with a likelihood of pretrial detention that is less than 5% (942,677 observations). We do this because we want to study the decisions of judges 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 (2,848 observations). We also resolve to only consider cases in which the assigned public attorney has defended at least 10 cases previously. It was not necessary to drop any data because of this restriction.

After all these adjustments the sample size is 699,368, which is consistent with the figure in [Table 1](#).

B.3 Variables

Many of the variables used in our estimations are directly contained in the administrative records. In what follows we describe how we construct the other variables.

- **Severity:** we proxy crime severity by computing the share of cases within the type of crime in which the defendants are detained pretrial.
- **Criminal record:** we can track all the 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 identification ID that were initiated before the current case. For individuals with no previous prosecutions, these variables are set to zero. To build these variables, we can track cases from 2005 onwards.
- **Pretrial misconduct:** pretrial misconduct is an indicator variable that takes value one if the defendant does not return to a scheduled hearing or is engaged in pretrial recidivism, or both. Nonappearance in court is recorded in the administrative data. Pretrial recidivism is built by looking at the arrests associated to the particular defendant's ID with an initial date that is between the initial and end dates of the current prosecution.
- **Attorney quality and judge leniency:** as in [Dobbie et al. \(2018\)](#), we use the residualized (against court-by-time fixed effects) leave-out mean release rate.
- **Court-by-year of prosecution fixed effects:** we consider the initial date to set the fixed effects.

Crime categories We classify crimes following the PDO classification and group them in the following nine categories.

- **Homicides:** considers all homicides, including specific categories such as parricide and femicide, among other specific types.
- **Sexual offenses:** examples include sexual abuse, pedophilia, and rape, among other sex crimes.
- **Thefts and robberies:** includes robbery, burglary, theft, and larceny.
- **Other property crimes:** examples include receiving or possession of stolen goods, arson, and criminal damages.

- Drug offenses: includes illegal consumption, drug trafficking, and drug production.
- White-collar and tax crimes: examples include economic fraud and the falsification of money, checks, or credit cards.
- Crimes against public trust: examples include falsification of public, official, and commercial documents, forgery of private documents, falsification of certificates, and identity theft.
- Crimes against the freedom and privacy of people: considers threats against citizens, but also includes threats to police officers and trespassing.
- Other crimes: examples include gun possession and intellectual property theft.

C Kitawaga-Oaxaca-Blinder decompositions

Let $\bar{R}_g = \mathbb{E}[\text{Release}_i | I_i = g]$, with $g \in \{0, 1\}$. With KOB decompositions, group differences, $\Delta_R = \bar{R}_1 - \bar{R}_0$, can be explained using a vector of observed covariates, X_i . To do this, we first run an OLS projection for each group, $\text{Release}_i = X_i' \beta_g^R + \epsilon_{ig}^R$, where X_i includes a constant and the individual controls of the benchmark equation, and ϵ_{ig}^R is the OLS projection error. By construction, OLS fits group means, so $\bar{R}_g = \bar{X}_g' \beta_g^R$, with $\bar{X}_g = \mathbb{E}[X_i | I_i = g]$. Then

$$\Delta_R = \bar{X}_1' \beta_1^R - \bar{X}_0' \beta_0^R = (\bar{X}_1 - \bar{X}_0)' \beta_1^R + \bar{X}_0' (\beta_1^R - \beta_0^R). \quad (\text{C.I})$$

The first term accounts for differences in release rates given differences in observables. The second term accounts for differences in release rates between defendants with the same observables (i.e., for differences in the estimated coefficients). When X_i includes all the relevant characteristics that matter for the release decision, the second term can be interpreted as discrimination. If there are unobserved variables that correlate with I_i and matter for misconduct potential, however, the OLS coefficients will capture their effect and the latter term will mistakenly be interpreted as discrimination because of omitted variable bias (OVB).

Our intuition is that if differences in unobservables matter for the release decision in a nondiscriminatory fashion, then they should be relevant to explain differences in pretrial misconduct rates when released. Formally, let $\overline{PM}_g = \mathbb{E}[PM_i | I_i = g, \text{Release}_i = 1]$, where PM_i is an indicator that takes value one if defendant i engages in pretrial misconduct. Using the same logic as before, we can estimate $PM_i = X_i' \beta_g^P + \epsilon_{ig}^P$ in the sample of released defendants and write

$$\Delta_{PM} = \bar{X}_1' \beta_1^P - \bar{X}_0' \beta_0^P = (\bar{X}_1 - \bar{X}_0)' \beta_1^P + \bar{X}_0' (\beta_1^P - \beta_0^P). \quad (\text{C.II})$$

Then, one way of testing if unobservables are important for interpreting our results is checking whether differences in observables are capable of explaining differences in pretrial misconduct rates. If the second component is large in the release equation but small in the outcome equation, then we can conjecture that unobservables are only playing a small role in explaining release rate disparities in a statistical sense, and we can therefore confidently interpret the benchmark estimations as evidence of discrimination.

Table C.I shows the results. For each dependent variable, we present two versions of the KOB decompositions. Columns labeled as *raw* present the standard KOB decomposition using the individual controls of the benchmark regressions as the vector of observables. Columns labeled as

Table C.I: Kitawaga-Oaxaca-Blinder Decomposition

	Release		Pretrial Misconduct	
	Raw (1)	Residualized (2)	Raw (3)	Residualized (4)
Total Difference	0.067 (0.011)	0.030 (0.011)	0.050 (0.009)	0.076 (0.007)
Explained: Due to difference in characteristics	-0.036 (0.003)	-0.056 (0.003)	0.065 (0.003)	0.076 (0.004)
Unexplained: Due to differences in coefficients	0.103 (0.011)	0.085 (0.010)	-0.014 (0.008)	-0.001 (0.004)
No. of Immigrants	6,362	6,362	4,900	4,900
No. of Chileans	693,006	693,006	580,400	580,400

Notes: This table presents the Kitawaga-Oaxaca-Blinder decomposition for release and pretrial misconduct estimation, considering raw and residualized covariates (residualized against court-by-year fixed effects).

residualized include dependent variables and the vector of individual controls residualized against court-by-year fixed effects. The release equation (columns 1 and 2) shows that there is a large share of the variation in release rates that cannot be explained by observed characteristics. In absolute value, the unexplained component of the average release gap is between two and three times larger than the share of variation explained by observables. Moreover, and consistent with the analysis so far, both components shift unconditional release disparities in opposite directions. In the absence of relevant unobserved variables, this reinforces the hypothesis of discrimination suggested by the benchmark regressions; however, it could also reflect the presence of OVB.

Columns 3 and 4 replicate the analysis using observed pretrial misconduct rates (among released defendants) as dependent variable. Note that in this case the complete average difference in pretrial misconduct rates can be explained by the same observed characteristics. This suggests that there are no relevant unobservables that explain misconduct potential, which reinforces the idea that the main benchmark regressions are not affected by OVB. Moreover, because observables do a good job of explaining actual misconduct rates, it reinforces the idea that our specific vector of observables makes the analysis closer in spirit to the disparate impact perspective, in the sense that this particular set of variables seem to do a reasonably good job of explaining observed misconduct rates. Or, at least, relevant omitted variables do not seem to be correlated with immigration status.

D Outcome test

This appendix describes the outcome test, specifically the observational implementation proposed by [Grau and Vergara \(2021\)](#), and provides suggestive evidence that the proposed identification argument is valid in our setting.

Outcome test The outcome test identifies a combination of biased beliefs and taste-based discrimination using observed outcomes of marginally selected individuals. Formal proofs are provided in [Arnold et al. \(2018\)](#), [Grau and Vergara \(2021\)](#), and [Hull \(2021\)](#). In what follows, the intuition for the outcome test is presented.

Recall in [Section 3](#) that the release decision can be conceptualized as follows:

$$Release_i = 1 \{p(I_i, Z_i) \leq t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)\}. \quad (\text{D.I})$$

The outcome test examines whether the effective thresholds, $t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)$, are systematically different between immigrant defendants and Chilean defendants. Put formally, it tests whether

$$\mathbb{E} [t_{j(i)}(1, Z_i) - b_{j(i)}(1, Z_i)] - \mathbb{E} [t_{j(i)}(0, Z_i) - b_{j(i)}(0, Z_i)] \quad (\text{D.II})$$

is different from zero, where the expectation is taken across Z_i and $j(i)$. If the difference is not zero, then defendants with equal “true” pretrial misconduct probabilities will be detained at different rates. Notably, if a group is, on average, more prone to be engaged in pretrial misconduct, this does not affect the results of the outcome test. Differences in pretrial misconduct potential will affect the “LHS” of [\(D.I\)](#); the OT estimates differences in the “RHS”. If a group is systematically more risky, then it will cross the threshold more often, which is different from having a different threshold. That is why the outcome test identifies a notion of discrimination that abstracts from accurate sources of statistical discrimination. If judges are engaged in accurate statistical discrimination, then the outcome test should not be rejected ([Hull, 2021](#)).

The insight produced by the outcome test is that although $t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)$ is not observed, for defendants that were released on a borderline decision $p(I_i, Z_i) = t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)$, and therefore the average misconduct rates of marginally released defendants identify $t_{j(i)}(I_i, Z_i) - b_{j(i)}(I_i, Z_i)$. Then, the outcome test is reduced to a difference in means that tests whether misconduct rates are different between marginally released immigrant defendants

and marginally released Chilean defendants.

The main identification challenge, then, is to identify marginal individuals. [Arnold et al. \(2018\)](#) provide a quasi-experimental approach that relies on the quasi-random assignment of judges. That approach is not applicable in our setting given the small share of immigrant defendants (the required instrument is underpowered). [Grau and Vergara \(2021\)](#) propose an observational approach that does not require instruments, which is the one implemented in this paper.

P-BOT The prediction-based outcome test (P-BOT) proposed by [Grau and Vergara \(2021\)](#) uses the propensity score to identify marginal individuals. More specifically, [Grau and Vergara \(2021\)](#) provide sufficient conditions under which released defendants with lower propensity scores are more likely to be marginal given their observables. The implementation of the outcome test then proceeds as follows. First, we estimate the propensity score and compute the predicted values. Second, we rank released defendants according to their predicted release probabilities and define as marginal the released defendants at the bottom of the distribution. Third, we implement differences in means for pretrial misconduct rates between immigrant defendants and Chilean defendants who were marginally released.

Identification requires three assumptions and here we present a high-level discussion of these sufficient conditions (for technical details see [Grau and Vergara \(2021\)](#)). First, we need a common support assumption on the distribution of latent risk that allows us to claim that “the more marginals” are effectively marginals. Second, the result relies on a separability assumption between observables and unobservables (by the econometrician) in the release equation. This implies that the effect of observables on the likelihood of being released is not affected by unobservables. Because the release decision is based on pretrial misconduct probabilities, this also implies a similar pattern in the outcome equation. Third, the result allows for unrestricted correlation between observables and unobservables but puts restrictions on the patterns of heteroskedasticity.

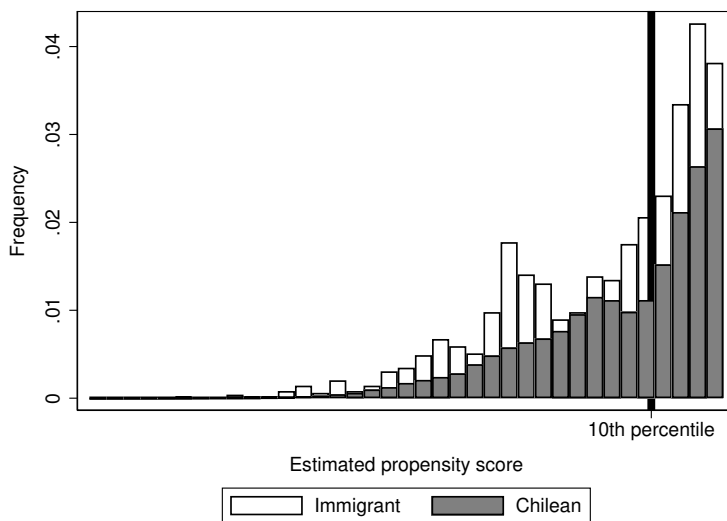
In what follows we present suggestive evidence that these assumptions are met in this setting implementing the tests as set out in [Grau and Vergara \(2021\)](#).

P-BOT implementation For implementing the test, we estimate the propensity score using the same set of observables included in the main benchmark estimations: the immigrant indicator, individual controls, judge and attorney controls, and court-by-time fixed effects. Then, we rank released defendants according to the predicted values and define the bottom 10% of the distribution as marginals. Standard errors are bootstrapped considering that the sample selection rule is based

on an estimated value.

Common support Figure D.I shows the (estimated) propensity score distributions for released defendants, separating immigrant defendants and Chilean defendants. The figure suggests that the continuity and full support assumptions are met in our setting.

Figure D.I: Immigrant Released Defendant and Chilean Released Defendant Propensity Score Histograms (Zoom up to 20th Percentile)



Notes: This plot shows the propensity score histograms for immigrant released defendants and Chilean released defendants. The vertical line represents the 10th percentile of the distribution. For presentation purposes, we only show each histogram up to the 20th percentile. However, the histogram is calculated considering the entire population of released defendants.

Monotonicity One way to assess this assumption is to check whether the coefficients of a regression of $Release_i$ on covariates are stable (in terms of sign) when considering subsamples with (probably) different unobservables. Likewise, monotonicity also implies that the coefficients of a regression of pretrial misconduct on covariates are stable (in terms of sign) when considering subsamples with (probably) different unobservables.

Tables D.I and D.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. Thereafter, rows are paired by mutually exclusive sample categories that are (probably) characterized by different unobservables. For example, row 2 shows the results for the immigrant subsample, and row 3 shows the results for the Chilean subsample. Then, rows 4 and 5 split the sample by gender, and so on. The results strongly support

the monotonicity assumption. In all but eight cases (i.e., 90% 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.

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

<i>Estimation sample</i>	Previous case	Previous pretrial misconduct	Previous conviction	Severity previous case	Severity current case
All	-0.004 (0.003)	-0.026 (0.001)	-0.013 (0.003)	-0.136 (0.004)	-1.009 (0.003)
Immigrant	0.044 (0.033)	-0.027 (0.017)	0.061 (0.032)	-0.129 (0.058)	-1.082 (0.034)
Chilean	-0.006 (0.003)	-0.026 (0.001)	-0.014 (0.003)	-0.136 (0.004)	-1.007 (0.003)
Male	-0.004 (0.003)	-0.027 (0.001)	-0.012 (0.003)	-0.119 (0.004)	-1.014 (0.003)
Female	0.011 (0.007)	-0.010 (0.003)	-0.030 (0.007)	-0.341 (0.013)	-0.958 (0.009)
Low income	-0.006 (0.004)	-0.021 (0.002)	-0.015 (0.004)	-0.131 (0.006)	-1.019 (0.005)
High income	-0.003 (0.003)	-0.029 (0.001)	-0.012 (0.003)	-0.140 (0.005)	-1.002 (0.004)
Low judge leniency	-0.002 (0.004)	-0.028 (0.002)	-0.016 (0.004)	-0.151 (0.005)	-1.050 (0.004)
High judge leniency	-0.005 (0.004)	-0.023 (0.002)	-0.011 (0.003)	-0.121 (0.005)	-0.968 (0.004)
Low attorney quality	-0.004 (0.004)	-0.025 (0.002)	-0.016 (0.004)	-0.144 (0.005)	-1.073 (0.004)
High attorney quality	-0.005 (0.004)	-0.026 (0.002)	-0.010 (0.003)	-0.127 (0.005)	-0.940 (0.004)
Small court (No. of cases)	0.002 (0.004)	-0.022 (0.002)	-0.017 (0.004)	-0.153 (0.005)	-1.084 (0.004)
Big court (No. of cases)	-0.009 (0.004)	-0.028 (0.002)	-0.011 (0.003)	-0.126 (0.005)	-0.943 (0.004)
Small court (No. of judges)	0.003 (0.004)	-0.022 (0.002)	-0.017 (0.004)	-0.149 (0.005)	-1.077 (0.004)
Big court (No. of judges)	-0.009 (0.004)	-0.028 (0.002)	-0.012 (0.003)	-0.127 (0.005)	-0.946 (0.004)
Low severity court	-0.005 (0.003)	-0.019 (0.001)	-0.010 (0.003)	-0.111 (0.005)	-0.879 (0.004)
High severity court	-0.002 (0.004)	-0.028 (0.002)	-0.019 (0.004)	-0.158 (0.005)	-1.132 (0.004)

Notes: 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 parentheses are standard errors.

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

<i>Estimation sample</i>	Previous case	Previous pretrial misconduct	Previous conviction	Severity previous case	Severity current case
All	0.050 (0.004)	0.106 (0.002)	0.036 (0.003)	0.036 (0.005)	0.034 (0.005)
Immigrant	0.041 (0.040)	0.096 (0.020)	0.044 (0.039)	0.076 (0.073)	0.111 (0.045)
Chilean	0.050 (0.004)	0.106 (0.002)	0.036 (0.004)	0.036 (0.005)	0.033 (0.005)
Male	0.051 (0.004)	0.107 (0.002)	0.034 (0.004)	0.041 (0.006)	0.039 (0.005)
Female	0.049 (0.010)	0.096 (0.005)	0.046 (0.010)	-0.024 (0.020)	-0.012 (0.014)
Low income	0.045 (0.006)	0.097 (0.003)	0.038 (0.006)	0.036 (0.008)	0.075 (0.007)
High income	0.051 (0.005)	0.112 (0.002)	0.035 (0.004)	0.037 (0.007)	0.000 (0.006)
Low judge leniency	0.041 (0.005)	0.102 (0.002)	0.044 (0.005)	0.043 (0.008)	0.035 (0.007)
High judge leniency	0.059 (0.005)	0.110 (0.002)	0.027 (0.005)	0.029 (0.008)	0.034 (0.007)
Low attorney quality	0.047 (0.005)	0.110 (0.002)	0.042 (0.005)	0.048 (0.008)	0.033 (0.007)
High attorney quality	0.054 (0.005)	0.102 (0.002)	0.029 (0.005)	0.023 (0.008)	0.034 (0.007)
Small court (No. of cases)	0.039 (0.005)	0.102 (0.002)	0.035 (0.005)	0.043 (0.008)	0.094 (0.007)
Big court (No. of cases)	0.061 (0.005)	0.107 (0.002)	0.038 (0.005)	0.025 (0.007)	-0.024 (0.006)
Small court (No. of judges)	0.054 (0.005)	0.103 (0.002)	0.030 (0.005)	0.040 (0.008)	0.057 (0.007)
Big court (No. of judges)	0.047 (0.005)	0.107 (0.002)	0.042 (0.005)	0.030 (0.007)	0.013 (0.006)
Low severity court	0.045 (0.005)	0.101 (0.002)	0.039 (0.005)	0.042 (0.007)	0.053 (0.006)
High severity court	0.053 (0.005)	0.108 (0.002)	0.034 (0.005)	0.028 (0.008)	0.016 (0.007)

Notes: 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 parentheses are standard errors.

Ranking validity Assume that our set of observed variables, X_i , is a good approximation (up to some small well-behaved noise) of the (complete) information set of a judge. Under this 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 10% of the predicted probability distribution (among released defendants). Then, we omit one observable and (i) estimate the propensity score with the restricted set of observables 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.

We perform this exercise by using each of the 14 observables used in the estimation.¹ To compute the rank correlations, we discretize the nondiscrete regressors (using the median) to define $2^{(14-1)} = 8,192$ 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 D.III presents the results. We report both Spearman’s- ρ and Kendall’s- τ statistics for rank correlation. In all variables bar one (severity of current case) the correlations are very large. We interpret this as strong suggestive evidence of the validity of the identification argument.

¹The variables are 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, judge leniency squared, attorney quality, attorney quality squared, immigrant indicator, previous case indicator, previous pretrial misconduct indicator, and previous conviction indicator.

Table D.III: Rank Correlations

<i>Excluded predictor</i>	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	
	Spearman	Kendall	Spearman	Kendall
No of previous cases	0.972	0.948	-0.668	-0.549
Severity previous case	0.980	0.960	-0.672	-0.555
Severity current case	0.434	0.379	-0.264	-0.217
Average severity (year/court)	0.960	0.927	-0.671	-0.557
No of cases (year/court)	0.997	0.990	-0.666	-0.551
No of judges (year/court)	0.986	0.974	-0.675	-0.559
Judge leniency	0.988	0.976	-0.673	-0.556
Judge leniency square	1.000	0.999	-0.678	-0.560
Attorney quality	0.975	0.955	-0.676	-0.559
Attorney quality square	1.000	1.000	-0.675	-0.558
Immigrant	0.997	0.992	-0.806	-0.661
Previous case	0.998	0.996	-0.666	-0.551
Previous pretrial misconduct	0.989	0.981	-0.678	-0.562
Previous conviction	1.000	0.998	-0.670	-0.556

Notes: 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 Spearman's- ρ and 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., 13 predictors). The continuous predictors were transformed into binary variables using the median among released defendants as the threshold.