Legal Representation in Immigration Courts: The Impact of Randomly Assigned Observers

Lenni B. Benson, Decio Coviello, Nicola Persico, Petra E. Todd*

July 27, 2025

Abstract

Legal representation is essential for noncitizens seeking relief in U.S. immigration courts, yet the federal government does not provide counsel. This paper presents findings from a randomized controlled trial in which law students were assigned to observe immigration judges presiding over removal hearings. We investigate whether and to what extent the presence of an observer affects the likelihood that a respondent seeks and secures legal representation. The results show that judicial observation significantly increases the likelihood that judges adjourn initial hearings to give respondents time to secure counsel, and subsequently increases the probability that respondents are represented by an attorney in later hearings. Importantly, we document substantial heterogeneity across judges in their baseline propensity to grant continuances and show that this variation shapes how judges respond to being observed.

Keywords: immigration courts, access to justice, legal help, randomized experiment, observers.
JEL Classification: K00, K37, K40, J15, D63.

*Lenni Benson (New York Law School). Decio Coviello (HEC Montreal). Nicola Persico (Northwestern University). Petra E. Todd (University of Pennsylvania). We thank Austin Kocher and participants at NYLS Asylum Conference and at Queen's University for helpful comments. We thank Diego Ciccia, Roberto Vacante, and Matteo Vasca for outstanding assistance. We are also indebted to Professor Amelia Frank-Vitale who supervised a group of court observers at Barnard College. We also could not have completed these courtroom observations without the cooperation of the NY City Immigration Court administrators and the hundreds of volunteer observers. This paper is undertaken, in part, thanks to funding from the Canada Research Chairs program. Experiment registered: AEARCTR-0011751. E-mails: lenni.benson@nyls.edu; decio.coviello@hec.ca; nicola@nicolapersico.com; petratodd@gmail.com.

1 Introduction

Every day, thousands of individuals appear before one of the 72 U.S. immigration courts—often accompanied by spouses and young children.[1] The vast majority do not speak English, have little understanding of immigration law, and arrive at their first hearing without legal representation. Yet legal counsel is crucial: according to our data, individuals with legal representation are 3 to 8 times more likely to be granted permission to remain in the United States than those without.[2] One major reason for this disparity is that many noncitizens are unaware of the one-year deadline for filing an asylum claim, which often passes even before their first court appearance. Attorneys are familiar with exceptions to this deadline and can help preserve their clients' eligibility.

The federal government does not provide free legal representation, so those who cannot afford an attorney, or are not aware of the benefits of having one, must navigate the system without legal assistance.[3] Additionally, during periods of intensified immigration law enforcement, there can be shortages of affordable, qualified attorneys, making it difficult to find legal counsel. In our data, for example, only 8% of respondents are represented in their first hearing and only 30%, on average, over all hearings. To ensure fairness, it is essential that judges grant sufficient time for respondents to secure legal counsel.

This paper reports results on a new intervention in immigration court proceedings that has the potential to increase access to legal representation. In particular, we carried out a randomized experiment that allocated non-expert observers (lawyers and law school students) to judges' hearings in the three New York City immigration courts. The judges were told that the observer was present to learn about the immigration hearing process. Depending on the treatment arm, some judges were also informed that the court administration was aware that the courtroom was being observed on that day. The observers were mainly volunteers from New York Law School and Barnard College; they were given brief training about the vocabulary and procedure used in immigration courts. They were instructed to introduce themselves to the judge and to sit through the hearing, doing nothing except possibly taking notes.

The experiment took place over ten months, from July 2023 to May 2024. A weekly randomized protocol assigned the observers to judges' master calendar hearings, as described in detail below. The

[1] See U.S. Government Accountability Office Report (GAO).

[2] People who have no counsel rarely apply for any statutory relief. Immigration Courts have no general equitable power. Their authority is limited to the forms of relief specified in the immigration statutes. The government attorneys can grant forms of prosecutorial discretion but not the judge. All of these permissions to remain require some affirmative steps by the respondent. A judge acting alone cannot protect a respondent from removal. These statistics are computed using the EOIR public data set from 2000 to April 2025. The data reflects a range of differential success from a low of 300% to a high of 800% looking at data from 2018 to 2022. We present this data to emphasize that it takes time to adjudicate applications before the court. The process is complex. Table A.1 reports statistics for the years between 2017 and 2022.

[3] See INA § 292, 8 U.S.C. § 1362.

hearing schedule was obtained from the calendar circulated to the New York area attorneys by the Immigration and Customs Enforcement prosecutors who are part of the Office of the Principal Legal Advisor (OPLA). We chose to limit our observation (experimental treatment) to first hearings, i.e., to the first time the respondent appears in court in a given case, as opposed to later case hearings. This choice was made both because these hearings are often very consequential and, also, to limit concerns about non-random selection of the cases being observed (i.e., experimentally treated).

We merge the data that we collected on which hearings were observed (treated) with anonymized proceedings data from U.S. Immigration Courts (EOIR) to investigate whether, and to what extent, the observer's presence affected the likelihood that a respondent secures legal representation. The EOIR database contains information on the reason why each hearing is adjourned. We find that an observer's presence in the first hearing of a case without an attorney leads to a 19% increase in the likelihood that a judge will adjourn to allow the respondent time to hire an attorney relative to non-observed hearings. The observer's presence also increases the likelihood that the respondent *actually secures an attorney* later in the proceeding. In particular, it increases the likelihood of having an attorney by 13.7% within our post-treatment observation period. As previously noted, our experiment included two randomly assigned treatment arms, one in which the observers merely introduced themselves to the judge and another in which the observer also mentioned that the court administration was informed of the observation. Only the second treatment arm had a statistically significant effect, suggesting that the judges' response is stronger in the presence of hierarchical motives/considerations.

We also investigate whether our treatment had heterogeneous effects for different types of judges and respondents. Our results show that judge identity is an important determinant of initial case outcomes, as has also been found in related literature (see section two). When we examine heterogeneity in how judges respond to the presence of an observer (i.e. treatment), we find the strongest effects for judges who, prior to our experiment, showed a higher baseline propensity to grant continuances to seek representation. With regard to respondent characteristics, the most common countries of origin for the individuals in our dataset are: Ecuador (19%), Venezuela (11%), China (9%) and Columbia (6%). There are significant observer effects for respondents of all nationalities except China. Interestingly, Chinese respondents arrive at the first master calendar hearing with much higher rates of legal representation, so they may be better informed than others about the importance of having legal counsel and/or better able to afford counsel.

The final disposition of most (56%) of our cases was not known at the time of this writing, so we cannot assess how our intervention affected the final disposition of cases. Nonetheless, we believe it is

important to share these findings now rather than delay dissemination until all cases reach resolution, which can take years. The intervention we implemented is highly scalable, and our results have immediate implications for enhancing fairness and transparency in immigration court proceedings. To our knowledge, this is the first experimental study to examine the effect of courtroom observation on the likelihood that a respondent secures legal representation. Although it is known that having legal counsel generally leads to better outcomes for respondents, a key policy challenge has been identifying cost-effective strategies to increase access to counsel. Our findings offer timely insights into one such approach.

The paper develops as follows. Section two discusses related literature, section three provides further information on immigration court proceedings and on the EOIR database that we use, and section four describes our experiment in greater detail and provides balance test evidence in support of the randomization. The key empirical results are presented in section five and section six concludes.

2 Related literature

Although we are not aware of any other studies that analyze the effects of court observers on judge's behavior, there is a related non-experimental literature that estimates the causal effect of having representation on immigration court decisions. Eagly and Shafer (2015) empirically examines how having a lawyer affects outcomes of immigration court decisions. Analyzing data on over 1.2 million deportation cases (from 2007-2012), they find that only 37% of all immigrants and 14% of detained immigrants secured representation.[4] Immigrants with representation were much more likely to seek relief and to obtain relief from removal. Barriers to legal representation were more severe in rural areas and small cities, where almost one-third of cases were adjudicated. Another study by Ryo and Peacock (2021) also explores how legal representation affects the outcomes of immigration cases, analyzing data on 1.9 million removal cases adjudicated between 1998 and 2020. They find that the effect of representation is greater when there is a female judge and/or a more experienced judge. They also observe that the representation effect is larger in times of increasing judge caseloads. Using nonexperimental data, Ryo and Humphrey (2022) analyze the demand and supply-side factors that affect whether an immigrant respondent obtains legal representation. Controlling for the availability of practicing immigration lawyers in close proximity to a respondent's place of residence, other significant predictors of representation include their geographic location, primary language, and the size of co-national social networks. Ryo and Humphrey (2022) argue that addressing the problem of low rates of legal representation requires consideration of linguistic and

[4]Only 2% obtained pro bono representation.

social isolation factors. A challenge in these observational studies is the endogeneity of the decision to acquire legal representation with respect to anticipated immigration court outcomes.

Unrelated to court observers and, also, to immigration courts, some studies have analyzed the effects of randomly provided legal counsel. For example, Grenier and Pattanayak (2011) report the results of a randomized evaluation of a legal assistance program provided by a law school, which was targeted at unemployment claimants, assisting them in appealing their case to state administrative judges. They present both intent-to-treat (ITT) effects of the offer of legal assistance as well as treatment-on-the-treated (TT) effects for those that took up the offer. The study finds that the offer of assistance delays the duration of the proceeding but that neither the offer nor receipt of assistance had a statistically significant effect on whether the claimant prevailed in their appeal.

Somewhat related, there are several papers highlighting the heterogeneity of dispositions across immigration judges. For example, a provocatively-titled study ("Refugee Roulette," Ramji-Nogales et al. (2007)) analyzes a database that includes 133,000 decisions rendered by 884 immigration judges. They document substantial heterogeneity in judge's decisions, even for judges within the same office who receive a similar caseload, depending on which judge adjudicates the case and on the immigrant's nationality. They find a tendency for male judges and more experienced judges to grant asylum less often. Lastly, they show that proceeding durations are longer when the respondent has legal representation but that the effect of having representation on the final outcome is ambiguous. Hausman (2016) also observes wide disparities across judges and that, even within the same immigration court, some judges are up to three times more likely than their colleagues to order a respondent deported. Analyzing a database of almost five million immigration cases, Hausman (2016) examines whether the appeals process promotes uniformity across inferior court judges. He concludes that it generally does not, because harsher judges often have immigrants deported early in the process, before they are able to obtain legal representation or to file for asylum (or any other kind of relief). Also, immigrants without lawyers are less likely to appeal the decision.

Finally, Chen and Eagel (2017) show that judge identity is a key determinant of immigration court outcomes. They use machine learning methods (random forests) applied to data from 492,903 asylum hearings rendered by 441 judges over the years 1981-2013 to study what variables best predict court decisions. The results show that granting asylum is mainly driven by trends and judicial characteristics with about one third driven by case information, news events and court information. They document

substantial autocorrelation in judges' rulings.[5]

Another aspect of the immigration crisis in recent years has been an increase in the fraction of immigration removals taking place outside the purview of the immigration court system.[6] As discussed in Koh (2016), the merits of many cases are never discussed in court. Such cases would include expedited removals at the border, removals based on prior removal orders, removal orders for non-lawful residents with aggravated felony convictions, and stipulated removal orders following waivers of the right to a court hearing. Orders are typically signed by immigration judges, with a large fraction issued *in absentia*. Koh (2016) also presents evidence that the judges removal decisions vary depending on when the judge was appointed and depending on the President in office (Presidents Bush, Jr., Obama or Trump).[7]

Again, to our knowledge, this is the first paper to experimentally assess the impact of court observers on the probability that the respondent secures an attorney in immigration proceedings.

3 Institutional background, data sources, and descriptive statistics

We next provide an overview of the ecosystem around immigration courts in the US. We also describe our data source with special attention to New York City immigration courts, where our experiment took place.

3.1 Background on immigration court proceedings and importance of representation

Immigration court proceedings are administrative hearings; as such, they are not governed by the Federal Rules of Civil Procedure, nor are they subject to standardized Rules of Evidence. Instead, the Executive Office for Immigration Review (EOIR), a division of the Department of Justice (DOJ), regularly revises the immigration court practice manual that includes all the guidelines for immigration judges.[8] Immigration judges are government attorneys employed by the Department of Justice's Executive Office for Immigration Review (EOIR). They are on probation for a minimum of the first three years of service and do not have life tenure. There are roughly 700 judges spread across 75 immigration courts throughout

[5]There are a few studies that examine how immigration court decisions respond to current events. For example, Brodeur and Wright (2019) find that the Sept 11, 2001 terrorist attacks led to a 3.2 percentage point decrease in the likelihood that applicants from Muslim-majority countries are granted asylum. Peacock and Ryo (2022) show that Chinese respondents as well as East and Southeast (E/SE) Asian respondents experienced a significantly higher removal rate during the early pandemic period.

[6]See Benson (2017) about the growth of using removal procedures outside the immigration court.

[7]Kim and Semet (2019) also find that the president in office appears to exert an influence over the outcomes of immigration procedures. Analyzing 780,000 custody decisions by immigration judges, they find that nonimmigrants were less likely to obtain bond during the Trump administration than during the two prior administrations.

[8]See See Immigration Court Practice Manual

the US. The Department of Homeland Security (DHS) is represented in court by an attorney of the Office of the Principal Legal Advisor (OPLA).

An immigration case officially begins when a branch of DHS files a Notice to Appear (NTA) (See Figure A.1). Once an NTA is issued, a new case is opened whose first proceeding is generated by the NTA, and a spot is automatically reserved on a judge's *master calendar hearing*. A master calendar hearing is a hearing that is intended to lay out the timing for the particular proceeding and to allow the parties to review the charges or pleadings. If an individual is pursuing relief from removal, the applicant must have time to present the evidence and legal arguments; the hearings that follow the master calendar hearing are called "Merits" or "Individual" hearings. Our observers were randomly assigned to attend master calendar hearings.

Cases are assigned to each immigration judge's *master calendar* on a random rotational basis, using an automatic calendaring process.[9] Occasionally, a case may have more than one proceeding attached to it, but the first one to be heard is the one related to the NTA: our empirical analysis analysis focuses on that first "master calendar" hearing and the proceeding generated from it.

During the master calendar hearing, immigration judges evaluate the evidence presented by the respondent and decide what we call a proceeding *outcome*. The outcome could be a determination that the respondent has the right to stay in the U.S. (relief granted), or that s/he should be removed. The most common outcome, however, is that the judge adjourns the proceeding to a reset master calendar hearing or to an individual calendar hearing. EOIR requires the judges to code the reason for adjournment. Occasionally, judges administratively close proceedings, temporarily suspending removal without terminating the case. This effectively removes the case from the active docket of a judge with no additional hearings.[10] If a respondent fails to appear for a scheduled hearing, judges may issue an in absentia order of removal. The relative frequency of different outcomes will be described later and shown in Table 1.

Immigration cases sometimes take a long time to decide, almost three years on average, but with a large variance.[11] It is generally thought that respondents benefit from the delaying the decision because, in the interim, they are allowed to stay in the country. Starting in 2021, a subset of judges in New York City were required to use a new dedicated dockets process.[12] Dedicated dockets are specialized dockets

[9] See Chapter 3, Section 1 of the Uniform Docketing System Manual

[10] In 2018, the Trump administration restricted the use of administrative closure, and in 2021, the Biden administration restored the authority of immigration judges to use this type of closure.

[11] In New York we compute that, conditional on being decided during our sample period, a case took 3.6 years to decide and the variance of the durations is 4 years.

[12] The EOIR administrators initiated dedicated dockets in ten cities: Denver, Detroit, El Paso, Los Angeles, Miami, Newark, New York City, San Diego, San Francisco, and Seattle. Source: Dedicated Docket

designed to expedite the proceedings of selected recently arrived families, and are designed so that cases in these dockets can be decided within 300 days of the initial hearing.[13]

Respondents in immigration court may be represented by an attorney, but only at their own expense. Government-provided legal representation is extremely limited and generally reserved for individuals with mental impairments or for children held in government detention. In New York, some funding is available for nonprofit attorneys to represent individuals in removal proceedings, primarily through support from the City of New York and the New York State Office for New Americans. Some individuals are also able to secure private or pro bono legal counsel. As previously noted, legal representation has a substantial impact on the outcome of immigration cases (see Ramji-Nogales et al. (2007), Miller et al. (2015), Eagly and Shafer (2016), Musalo et al. (2024) and Ryo and Peacock (2021)). Our data reveals a stark disparity: a respondent who has secured counsel is 300% to 800% more likely to remain in the United States compared to one who has not.

One key reason for this effect is that most respondents are unaware of critical procedural deadlines—most notably, the requirement to file for asylum within one year of entering the U.S. [14] By the time of their first hearing, many have already missed this deadline. Legal representation is especially important in such cases, as attorneys are familiar with jurisdictional nuances and the exceptions to the one-year rule.[15] More broadly, immigration court proceedings are legally complex, and most individuals are not equipped to navigate the process or meet its procedural demands without professional assistance.

The national rate of legal representation for all proceedings, whether still pending or closed, has dropped dramatically from 61% in 2017 to 18% in 2025. In the New York courts during this period, the rates of representation declined from a high of 84% to 29%. The diminishing representation rate may be partly reflect that respondents do not have sufficient time to secure legal counsel.[17]

[13] As previously noted, these specialized dockets have been criticized for potentially rushing cases and limiting the fairness of the immigration process. After the completion of this study, the EOIR announced yet another specialized docket for "recent arrivals: see New Dedicated Docket

[14] U.S. law requires that applicants for asylum file within one year of entry of the United States in almost all situations. People frequently do not have a first hearing before the immigration court for six month or longer. Some individuals will try to meet the one year deadline by finding a way to prepare and file the basic asylum application before that deadline.

[15] Attorneys serve other important functions, perhaps most obvious in the data, is that represented people return to court to appear at later hearings. The court has the power to issue in absentia orders.[16] Attorneys explain the negative consequences and encourage people to continue with their applications before the court.

[17] As our study only examined the responses of observed judges in New York, it may be that the higher rates of representation in New York also influenced judges to be patient as it is much more likely to find counsel in New York that in areas with fewer resources.

3.2 Data sources, sample selection, and descriptive statistics

Our data comes from the EOIR case database, which contains anonymized proceeding data from U.S. immigration courts. These data are released monthly as part of the Electronic Reading Room and pursuant to Freedom of Information Act requests.[18] This dataset contains information on all proceedings to US immigration courts between January 2000 and April 2025. The data include detailed information on each proceeding, such as the nationality of the respondent, the nature of the charges, and judge's decisions. The EOIR data are organized at the proceeding-hearing level and include 34.6M observations/hearings for the entire court system.

To construct the *baseline sample* for our hearing-level analysis, we start from the EOIR case database and restrict attention to hearings over which a New York City immigration court has jurisdiction.[19, 20] Essentially all of these hearings (97.6%) pertain to removal cases; we drop all other hearings to ensure homogeneity. We then drop hearings with adjournment reasons "Data Entry Error"[21]. We then focus on first hearings and drop those with adjournment reason "Completion prior to hearing".[22] We further restrict attention to first hearings held during our sample period (July 10, 2023 to May 10, 2024) of which we can match the identity of the judges to our experimental data.[23] At this point, the dataset includes 96 judges and 129,664 hearings. We restrict attention to the first hearing of these cases (as opposed to subsequent hearings) to ensure that the hearings we analyze are homogenous and, also, to avoid hearing and proceeding selection issues (this first hearing is necessarily a master calendar hearing). We call this our baseline sample.

Table 1 presents descriptive statistics for hearing-level variables in the baseline sample. Approximately 14% of hearings were assigned an observer, and 6% were effectively observed (compliance with the experimental protocol will be discussed below). Attorneys were present in 8% of hearings, and around 6% of

[18]The data is available here: EOIR case database

[19]Specifically, we keep the aprroximately 3.6M observations/hearings with *base_city_code* equal NYC, NYB, NYV.

[20]There are three immigration courts in New York City. These courts are housed within multipurpose federal buildings and the court system is one tenant in the building. Entry to these buildings requires people to pass through a metal detector and their possessions and bodies are screened by security personnel. The largest court is labeled in the data NYC and is located on the 12^{th} and 14^{th} floors of 26 Federal Plaza. This court is commonly referred to as 26 Federal Plaza. This court covers 73% of the hearings in our data and it currently staffs 36 immigration judges. NYB, commonly called 290 Broadway, is the second largest court (14% of the hearings) and is located on multiple floors of the federal building at 290 Broadway. As of October of 2024 the court has a staff of 27 judges. The court labelled NYV, commonly called Varick, is the third court (12% of the hearings) and is located at the 5^{th} floor of 201 Varick street. It currently staffs 15 judges. Source: Immigration Courts. Over the years these courts represent 3.4M hearings and 299 judges in the EOIR dataset.

[21]This step drops 98,630 hearings.

[22]Thus dropping a further 56,727 hearings.

[23]For the 10,000 cases with Initial Appearance Docket (IAD) judge we replaced the IAD code with the code of the judge in the next hearing. We drop 1,402 cases assigned to visiting judges (i.e. ij_code="V11") as they are not part of the randomization protocol. We finally drop 500 observations, of which we cannot match the identity of the judges, and 46 observations where the NTA date post-dated the first master calendar hearing

hearings were dedicated dockets. The data indicates that asylum applications were filed in 0.2% of the master calendar hearings. The most frequent nationalities among respondents were Ecuadorian (19%), Venezuela (11%), Chinese (8.74%), and Colombian (6%). In total, 205 nationalities were represented in the data.

Table 1: Descriptive Statistics

	N	Mean	SD
Observer(Assigned)	129,664	0.1367	0.3435
Observer	129,664	0.0587	0.2351
Attorney in hearing	129,664	0.0815	0.2736
Dedicated docket	129,664	0.0558	0.2296
Asylum application	129,664	0.0023	0.0482
Respondent from Ecuador	129,664	0.1902	0.3924
Respondent from Venezuela	129,664	0.1117	0.3150
Respondent from China	129,664	0.0874	0.2825
Respondent from Colombia	129,664	0.0616	0.2404
Last hearing	129,664	0.1033	0.3044
Absentia	129,664	0.1551	0.3620
Administrative Closure	129,664	0.0042	0.0645

Notes: Baseline sample: all master calendar hearings between 10^{th} July 2023 - 10^{th} May 2024.

4 The court observation experiment

Prior to our experiment, Professor Lenni Benson, working with law students from New York Law School and a group from NYU Law School, began attending immigration court hearings to gather qualitative information on how unrepresented people were being treated as they navigated the unprecedented lengthy lines to enter the federal buildings. They also sought information on how judges were handling cases where Customs and Border Patrol had listed random New York city area nonprofit addresses as the respondent's address rather than a home address. Students selected the judge and date and the hearing time and filled out reports of their observations. Information gathered during this process was used by Professor Benson and New York community advocates to alert the community to systemic access problems as well as problems with government documents being served and to note how the judges were providing advisals about the right to counsel. Based on this preliminary project, Professor Benson and the co-authors designed the current project and, in July of 2023, began to randomly assign observers to court hearings and gather the quantitative data used in our study.

4.1 The experimental design

We recruited 83 observers from the New York Law School and Barnard College. Students volunteered to be part of the experiment and received an initial online training to understand how immigration hearings typically work and the general project goals.[24] During the ten-month experimental period spanning from July 10^{th}, 2023, to May 10^{th}, 2024, participating students received a weekly email containing a link to a Qualtrics survey. This survey asked them to indicate their available days for the following week. This information was then used to randomly assign students to judges during that weeks' master calendar hearings. The assignment process occurred weekly, matching available students with the judges' calendars provided by the Office of the Principal Legal Advisor (OPLA). Students were randomly allocated to all three New York immigration courts. Because each master calendar hearing day typically involves 30 proceedings, observers were able to observe multiple cases per judge-day of assignment. Throughout the sample period, students were assigned to a total of 462 hearings, for a total of 18,117 individual proceedings.

Observers received detailed instructions for the observation via email, including information about the assigned judge, the court's address, and the specific message they were to deliver. The experiment involved two randomly assigned treatment arms, each with a distinct message:

- Message 1: *"Thank you for having me in your court, your honor. My name is (...), I am participating in the immigration court observation project coordinated by Professor Benson at New York Law School. I appreciate this opportunity to learn about the immigration courts."*.
- Message 2: *"Thank you for having me in your court, your honor. My name is (...). Professor Benson at New York Law School asked me to thank you for this opportunity to learn. We have also let the Court administration know I am observing."*

Although the court was informed of the list of observers for the week, the specific message-observer-judge combination remained unknown before the hearings to the judges and the head of the court. After each master calendar hearing, observers were asked to complete a post-observation survey on Qualtrics. The surveys were not always completed: refer to Section 4.3. If completed, the surveys were uploaded to our server.[25]

[24] Slides are available at this link: Observer training
[25] The questionnaire is available at this link: Observer survey

4.2 Check that observers were randomly assigned

The observers were randomly assigned to master calendar hearings in each week. Our baseline sample (see Section 3.2) restricts attention to a subset of these hearings, i.e., to those cases for which the master calendar hearing was also the case's first hearing.[26]

We can check that the assignment of observers is random within our baseline sample by examining whether any judge/court/respondent characteristics predict the treatment status, that is, whether a hearing is observed. We consider as potential predictors the fraction of judge's hearings held in the main court (NYC), the gender of the judge, and whether the judge was confirmed (more than 3 years of service) and the gender of the respondent. At the hearing level, the predictor variables include whether an attorney is present, whether the hearing was conducted in a foreign language, the nationality of the defendant. Under random assignment, all of these variables should not predict treatment status. The randomization was performed weekly, so we examine whether these covariates predict treatment status week-by-week.

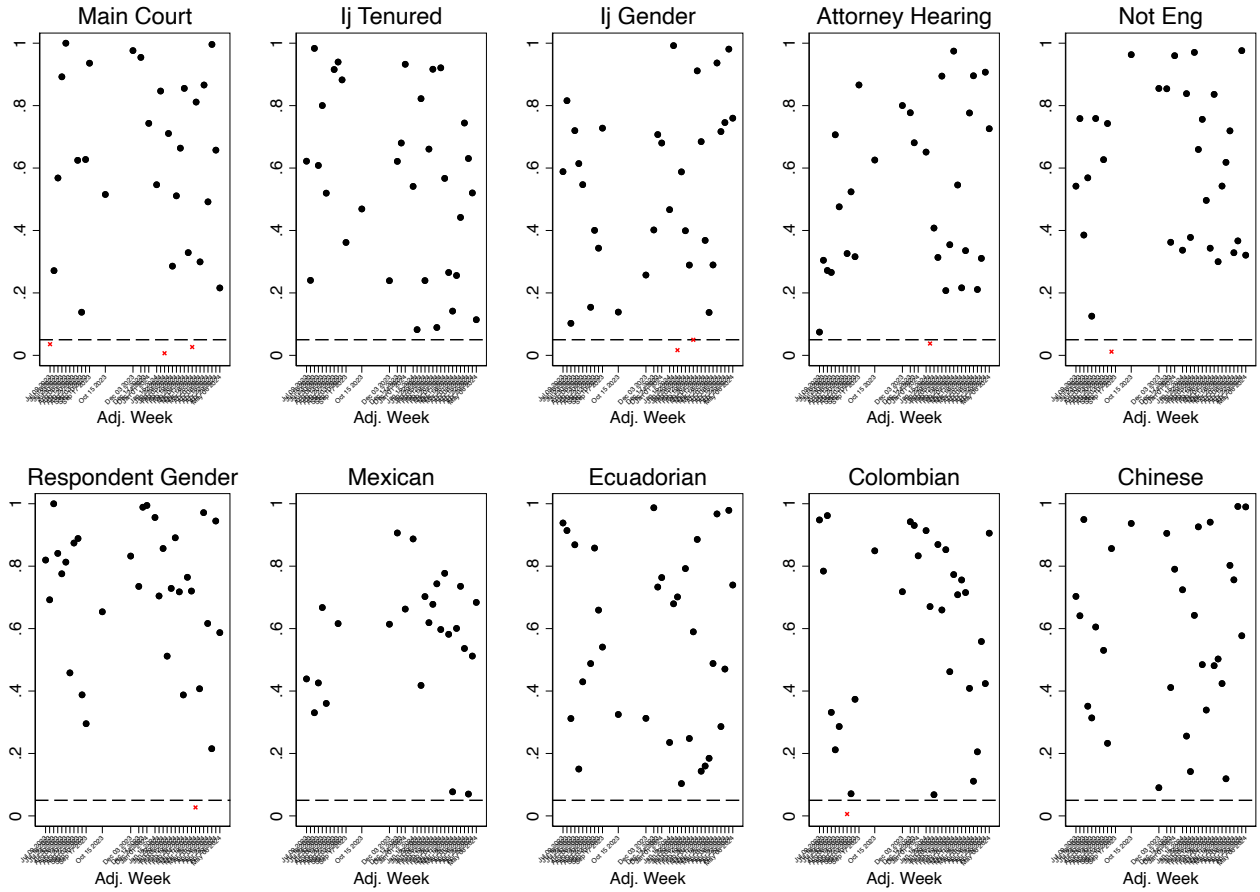
Figure 1 presents the p-values from logistic regressions where the dependent variable is whether an observer was assigned in a given week and the independent variables are the following (entered one at a time in the logistic model):

- Whether the hearing was at the main court (NYC)
- Indicator variables for tenured judges
- Indicator variables for both the judge and the respondent's gender
- Indicator variable for attorney present at the hearing
- Indicator variable for hearings not conducted in English
- Indicator variables for four nationalities with the greatest number of immigrants: Mexican, Ecuadorian, Colombian, and Chinese.

In the figure, the dashed line indicates a 5% significance level. Over the 35 weeks of random assignment, two weeks (Nov 20 and Dec 11) had only one observer, and one week (Jan 8) lacked schedule reports. Red crosses highlight 9 instances where the null hypothesis of random assignment was rejected, suggesting an

[26] We also drop cases from judges who received no cases in the 12 months preceding our experiment. The final sample is of 118,302 hearings. This step ensures the comparability of our final estimation sample across different specifications.

Figure 1: Balance test.



Notes: Dots are the p-values from logistic regressions, where the dependent variable is whether an observer was randomly assigned in a given week and the independent variables include respondent, case, judge, attorney characteristics. Dashed line indicates a 5% significance level. Red crosses highlight instances where the null hypothesis of random assignment was rejected. Sample: all hearings between July 10^{th} 2023, and May 10^{st} 2024.

approximate 3% rejection rate. The results are consistent with random assignment of the observer status within each week.[27]

4.3 Implementation accuracy

Although our experiment assigned students to observe court proceedings, the students were not always able to attend the assigned proceedings – mostly because of judge schedule changes, cancellation of hearings, or difficulties accessing the building. We refer to all these obstacles collectively as "implementation accuracy."

[27]We also performed balancing tests using a weekly t-test on the same covariates, which gave a similar rejection rate.

We check the extent of implementation accuracy by estimating the following OLS regression model:

$$Observer_{pjt} = \alpha + \beta Observer(Assigned)_{pjt} + \mu_j + \gamma_t + \varepsilon_{pjt} \quad (1)$$

where, $Observer_{pjt}$ is an indicator variables that takes value one if a post-observation survey was uploaded for proceeding p of judge j in week t. $Observer(Assigned)_{pjt}$ is an indicator variables that takes value one if an observer was randomly assigned to proceeding p of judge j in week t. The model includes week of hearing fixed effects (γ_t), because our randomization was done weekly. In addition, our specification includes judge fixed effects (μ_j) and week of notice-to-appear (NTA) fixed effects, where the latter are intended to control for any macroeconomic events affecting immigration composition or immigration court outcomes. With a randomized experiment, it is not strictly necessarily to control for additional covariates, as they are exogenous with respect to the random assignment. However, their inclusion often leads to greater precision in the estimated treatment effects, as discussed in Hahn (1998). Standard errors are clustered at judge-week of the hearing level.

Due to the obstacles previously described and as seen in Table 2, compliance with our random assignment protocol was imperfect. Although the first-stage F-statistic is strong, randomly assigning an observer to a master calendar hearing increases the probability of that hearing being observed by only 34 percentage points.[28] Therefore, when analyzing the impact of an observer on court outcomes we will present both OLS-ITT (Intention-to-Treat) and IV-TT (Treatment on Treated) estimates. The IV-TT estimates use the random assignment observer status as an instrument for the actual observer status. That is, the randomization can be viewed as making a hearing randomly eligible to have an observer. Note that an IV estimator with a discrete instrument usually identifies a LATE (local average treatment effect) estimate, which is the treatment effect for the subgroup of compliers-those who receive an observer only because of our random assignment. In our context, though, there are no hearings that would receive an observer in the absence of our random assignment (i.e., no "always-takers"), so the group of compliers corresponds to the treated group.(see the definitions of these subgroups in Imbens and Angrist (1994))

As previously mentioned, the noncompliance in our experiment is attributable to a few different factors, including last minute changes in the judge's schedule and cancellations of some originally scheduled hearings. Some noncompliance may also be due to the fact that we relied on volunteer (unpaid) law

[28] In all specifications, the first-stage F-statistic of the strength of the assignment variable as an instrument (equivalent to the square of the t-test since we have one instrument) is well above 10. The table also shows that the implementation accuracy is not related to the treatment type, because the coefficients $Observer_{pjt}(Assigned) : Message1$ and $Observer_{pjt}(Assigned) : Message2$ are the same.

school students as observers, who may not have been highly motivated. We view some noncompliance as unavoidable and not necessarily a drawback of our experiment. For example, scaling up the intervention using lowly compensated observers would likely result in less than 100% compliance with the experimental protocol. Additionally, last minute changes in the judges schedule are probably inevitable.

Table 2: First stage estimates.

Dep.Var.	Observer (1)	Observer (2)	Observer (3)	Observer (4)
Observer(Assigned)	0.343***		0.337***	
	(0.025)		(0.025)	
Observer(Assigned): Message 1		0.327***		0.329***
		(0.035)		(0.036)
Observer(Assigned): Message 2		0.357***		0.345***
		(0.033)		(0.033)
Observations	118,302	118,302	108,094	108,094
Judge FE	Yes	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes	Yes

Notes: Coefficient (standard error in parentheses) of the effect on having an Observers on Observer (Assigned) the day of the master calendar hearing to a judge. Columns 3 and 4 report estimates in the estimating sample sub-sample of cases with no attorneys in the first hearing, which represents our main estimation sample in Sections 5.1 and 5.2. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings between July 10^{th} 2023, and May 10^{st} 2024.

4.4 No anticipation effects

To check against the possibility (unlikely, in our view) that judges had advance knowledge of which hearings were going to be observed, we check whether judges were more likely to decide certain proceedings in advance of the observed hearing, as opposed to during it. If, hypothetically, the judge had advance knowledge of which hearings were going to be observed and wanted to decide a particular proceeding without being observed, s/he could make a decision in advance. For this analysis, we add to the baseline sample described in Section 3.2 the adjournment reason "Completion prior to hearing", and we estimate eq. 2 considering this outcome. Table A.3 shows that, while some proceedings (3%) are decided "ahead of schedule," the likelihood is not higher for observed hearings suggesting, as expected, that judges did not know in advance which of their hearings would be observed. We provide details of our estimated model in next section.

5 Empirical results

In this section, we first examine how the presence of an observer at a case's first hearing affects that hearing's outcome in a key dimension: whether the judge adjourns the hearing to give the respondent time to secure an attorney (Section 5.1). Then, in Section 5.2, we show that the observer's presence improves the likelihood that a respondent actually secures an attorney in the case's future hearings.

5.1 Hearing-level impact: adjournment to "seek attorney representation"

Because we are focusing on the respondent's ability to secure representation having started without one, we drop from our sample proceedings in which the respondent already has an attorney at the start. This brings the sample size down to 108,094. We refer to this sample as to our "estimation sample."

We estimate the following linear probability model:

$$y_{pjt} = \alpha + \beta \text{Observer}_{pjt} + \mu_j + \gamma_t + \varepsilon_{pjt} \tag{2}$$

where: y_{pjt} is an indicator for whether the first master calendar hearing in proceeding p for judge j in week t is adjourned to see attorney representation; Observer_{pjt} is an indicator variable that takes value one if the first hearing of proceeding p was observed (and not merely if an observer was scheduled to attend); μj and γ_t are judge- and week-of-the-hearing fixed effects. By including judge fixed effects, we control for the known variability across judges (see Figure 2) and rely in estimation on within-judge variation in observation status. As previously noted, we also include week of master calendar hearing and week of the NTA fixed effects, because our randomization was done weekly and to control for broader time effects. Standard errors are clustered at the judge-week level.[29]

Table 3 shows the estimates when $y_{pjt} = 1$ denotes that the reason for adjournment is "respondent to seek representation," else $y_{pjt} = 0$. This outcome means that, at the end of the hearing, the judge gives the respondent time to hire an attorney and schedules a second hearing some months later. Column 1 presents the OLS estimates of β from eq: (2). Without accounting for observer compliance issues, the effect of having an observer in the courtroom increases the likelihood that a hearing outcome is "adjourned for respondent to seek representation" by 14%, compared to non-treated. Column 2 is the reduced form (ITT): the effect of randomly assigning an observer on outcome y_i. The effect, 6%, is somewhat smaller than in col. 1 due to partial compliance (some observers, for the reasons discussed

[29] Recall that our procedure assigns observers to judges randomly weekly.

in Section 4.3, fail to show up at the assigned courtroom). The best estimate, in our view, is column 3. This estimate uses the observer's experimental assignment $\text{Observer}_{pjt}(Assigned)$ to a hearing to instrument for Observer_{pjt}, i.e., for whether the hearing was actually observed (TT). This estimate (0.061pp/.320=19%) is somewhat higher than the OLS estimate due, perhaps, to the fact that some observers ended up observing judges other than the one they were assigned to observe, a shift that may be correlated with judge characteristics:[30] the TT estimate corrects for this endogeneity. In any case, all the estimates are statistically and economically significant and, roughly, in the same ballpark quantitatively. The estimates are robust to controlling for the length of the time interval between receiving the notice to appear and the first hearing.

Table 3: Impact of having an observer on seeking representation.

Dep.Var.	SeekRep	SeekRep	SeekRep
Model	OLS	OLS-ITT	IV-TT
	(1)	(2)	(3)
Observer	0.046***		0.061**
	(0.013)		(0.031)
Observer(Assigned)		0.020**	
		(0.010)	
Observations	108,094	108,094	108,094
Judge FE	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes
Mean Y	0.320	0.320	0.320
F-first			183.9

Notes: Coefficient (standard error in parentheses) of the effect of having an Observers or an Observer (Assigned) the day of the master calendar hearing to a judge on adjournment is "respondent to seek representation". Regressions include Judge, Week of the hearing and Week of the NTA FEs. F-first is the first-stage F-statistic of the significance of the instrument. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings with no attorneys in the first hearing between July 10th 2023, and May 10st 2024.

Mechanism The bulk of the effects on "seeking representation" seem to arise in our treatment 2 (refer back to page 11), i.e., when the judge is told that the court administration has been made aware that the judge is being observed on that day. Absent this communication (i.e., in treatment 1), the effects are similar in magnitude but are not statistically significant in our preferred specification (IV-TT): see

[30] Certain judges may be more inclined to deny access and, also, be differentially responsive to the treatment. In some cases, the judge was ill and court personnel moved the observer to a new courtroom.

Table 4. This observation suggests that part of the judge's response to the observer is due to the judge's perception of being monitored.

Table 4: Impact of different messages on seeking representation.

Dep.Var.	SeekRep	SeekRep	SeekRep
Method	OLS	OLS-ITT	IV-TT
	(1)	(2)	(3)
ObserverXMess.1	0.036**		0.053
	(0.018)		(0.040)
ObserverXMess.2	0.064***		0.077**
	(0.018)		(0.036)
Observer(Assigned)XMess.1		0.018	
		(0.014)	
Observer(Assigned)XMess.2		0.028**	
		(0.013)	
Observations	108,094	108,094	108,094
Judge FE	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes
Mean Y	0.320	0.320	0.320
Diff_Eff	0.0283	0.00971	0.0236
pval	0.237	0.578	0.628
F-first			105.6

Notes: Coefficient (standard error in parentheses) of the effect of having an Observers or an Observer (Assigned) the day of the master calendar hearing to a judge on adjournment is "respondent to seek representation". Messages 1 and 2 are described in Section 4.1. Diff_Eff (pval) is the difference of the estimated coefficients (the p-value of the test for the difference). F-first is the first-stage F-statistic of the significance of the instrument. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings with no attorneys in the first hearing between July 10^{th} 2023, and May 10^{st} 2024.

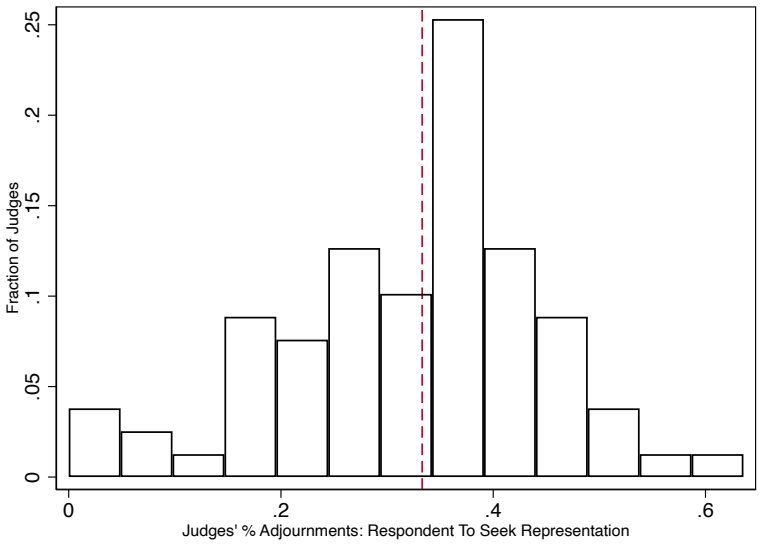
Heterogeneous effects by judge Next, we explore whether our treatment has heterogeneous effects across judges. We classify judges according to their 12-month-pre-treatment probability of adjourning unrepresented first hearings for reason of "seeking representation." We interpret this probability as the judge's individual propensity to facilitate representation.

As seen in Figure 2, judges differ in their propensities to grant continuances for purposes of securing legal counsel. The median judge has a pre-treatment probability equal to 0.31, and the standard deviation is relatively large: 0.13. This large standard deviation implies that being randomly assigned to one judge or another causes a predictable advantage or disadvantage for the respondent in his/her ability to secure

representation. This disparity has been termed "representation roulette."[31]

As shown earlier, for the average judge, the presence of an observer increases the baseline probability of adjournment for seeking representation by about 20 percent (col. 3 of Table 3). We next estimate the heterogeneous effect of this treatment, where judges are grouped into two categories by whether their propensities lie above or below the median. In col. 3 of Table 5, we see that observer effect arises from judges above the median distribution (40 percent of the control mean); the treatment has no impact on judges with below-median pre-experimental propensity to adjourn for "seeking representation." This suggests that the the observer's presence increases the chances for the respondent to secure represenation but also increases, to some degree, the variance implied by random assignment to judges. We will return to this point in our concluding remarks.

Figure 2: Representation roulette



Notes: This figure shows the distribution of judges distinguished by their pre-experiment propensity to adjourn for "Respondent to seek representation". The dashed line represents the median distribution. For graphical presentation, extreme values of the distribution were excluded.

Next, we analyze the relationship between the estimated judge fixed effects and observable characteristics related to the judge that are available in our data. In particular, we regress the estimated judge fixed effects on the following variables: indicator for female, age, indicator for hearing at the main court, years since judge was appointed, indicator for > 2 years since appointment (indicating whether past probation status), number of cases judge handled prior to our experiment, and an indicator for whether the chief immigration officer at that time was appointed by a Republican President. As seen in Table

[31] See the earlier discussion on page 5.

Table 5: Impact of having an observer on seeking representation.

Dep.Var.	SeekRep	SeekRep	SeekRep
Method	OLS	OLS-ITT	IV-TT
	(1)	(2)	(3)
Observer	0.016		-0.029
	(0.020)		(0.051)
ObserverXJudgeAboveMed	0.045*		0.132**
	(0.025)		(0.060)
Observer(Assigned)		-0.011	
		(0.017)	
Observer(Assigned)XJudgeAboveMed		0.047**	
		(0.021)	
Observations	108,094	108,094	108,094
Judge FE	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes
Mean Y	0.320	0.320	0.320
Sum_Eff	0.0607	0.0354	0.103
pval	0.001	0.004	0.005
F-first			63.26

Notes: Coefficient (standard error in parentheses) of the effect of having an Observers or an Observer (Assigned) the day of the master calendar hearing to a judge on adjournment is "respondent to seek representation". JudgeAboveMed are judges with pre-experiment propensity to adjourn "Respondent to seek representation" above the median. Sum_Eff (pval) is the sum of the estimated coefficients (the p-value of the test for the sum). F-first is the first-stage F-statistic of the significance of the instrument. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings with no attorneys in the first hearing between July 10^{th} 2023, and May 10^{st} 2024.

B.3, none of the observables are significant predictors. Thus, there is substantial heterogeneity across judges in their propensities to adjourn for reasons of seeking representation, but this heterogeneity is not well explained by observable characteristics.[32]

Heterogeneous effects by case type and respondent nationality Finally, in Appendix B, we also explore the heterogeneous effects of our treatment by different proceeding types, including "dedicated

[32]Lim et al. (2016) analyze criminal sentencing decisions in Texas state district courts using data on approximately half a million criminal cases from 2004 to 2013. They also find substantial heterogeneity in sentencing behavior across judges but that the heterogeneity cannot be explained by judges' ethnicity, gender, or political orientation, after conditioning on geographic factors. Cohen and Yang (2019) analyze data on sentencing decisions by criminal district court judges, who are appointed by the president, and find that judges appointed by Republican presidents are harsher with Black and male defendants compared to Democratic-appointed judges. These patterns, they find, cannot be explained by other characteristics of the judges. Blasingame et al. (2024) leverage immigration data, similar to our own, and the Trump administration's (first) quota policy to investigate how immigration judges, appointed by either Republican or Democratic administrations, responded to this policy. They find that the quota policy led judges to issue more removal orders, particularly among judges who were previously less inclined to rule against respondents. Vojta and Peacock (2025) document that immigration judges who have completed their probationary period tend to be less harsh and less productive compared to judges who are still in probation status.

dockets" (proceedings of recently arrived families that are slated for a relatively fast decision, refer to Section 3.1), asylum applications, and most-common respondent nationalities. The observer effects do not differ by case type (dedicated dockets or asylum applications). However, we do find evidence of heterogeneous observer effects for the most common respondent nationalities (Ecuador, Venezuela, China, Colombia). Specifically, we find statistically significant observer effects for the Latin American countries but not for China. Interestingly, respondents from China are much more likely to have an attorney at the time of the first hearing than those from the other countries (see the rates shown in Table B.1 in appendix B).

5.2 Proceeding-level impacts: securing an attorney

In Section 5.1, we showed that experimentally assigning an observer to a first hearing substantially increases the probability that the judge adjourns the case specifically to give the respondent more time to "seek representation." Here, we examine whether our intervention caused an *actual increase* in attorney representation, as seen 12 months after the end of our experimental observation period (April, 2025). To this end, we re-estimate eq. (2) with the dependent variable y_{pjt} now being equal to 1 if the respondent *eventually* secures an attorney, i.e., if the respondent initially appeared without an attorney at the master calendar hearing and then later is recorded in the EOIR database as having attorney representation any time after.

For this proceeding-level analysis, we follow the 108,094 proceedings in the original estimation sample until they are adjudicated or, if not adjudicated, up to twelve months after the conclusion of the experiment (until April, 2025). At that time, 45% have been adjudicated. Thus, at the proceeding level, there is some right censoring in that respondents in unresolved cases may still eventually obtain counsel, which could downward bias in our estimates. Also, we note that the judge hearing a case is not necessarily the same judge that heard that case's first master calendar hearing. To avoid potential concerns about endogeneity of the judge fixed effects, in our estimation, the judge fixed effects pertain to the judge who heard the case at its first master calendar hearing (i.e., the time when our randomization took place).

Table 6 reports the estimates of β from eq. (2) and variants of it. Column 1 shows that the presence of an observer increases the likelihood of the respondent eventually securing an attorney by 7% (0.02/0.298=7%), compared to hearings without an observer. Column 2 displays the reduced form, or Intention-to-Treat (ITT) estimate, which assesses the effect of randomly assigning an observer: the estimated ITT effect is 4.7%, somewhat smaller than the OLS estimate in Column 1, a difference attributable

Table 6: Impact of having and observer on having an attorney

Dep.Var	Attorney	Attorney	Attorney	Attorney	Attorney	Attorney
Method	OLS	ITT	IV-TT	OLS	ITT	IV-TT
	(1)	(2)	(3)	(4)	(5)	(6)
Observer	0.020**		0.041**			
	(0.010)		(0.020)			
Observer(Assigned)		0.014**				
		(0.007)				
ObserverXMess.1				0.005		0.022
				(0.015)		(0.027)
ObserverXMess.2				0.034**		0.042*
				(0.015)		(0.024)
Observer(Assigned)XMess.1					0.008	
					(0.009)	
Observer(Assigned)XMess.2					0.015*	
					(0.009)	
Observations	108,094	108,094	108,094	108,094	108,094	108,094
Judge FE	Yes	Yes	Yes	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean Y	0.298	0.298	0.298	0.298	0.298	0.298
Diff_Eff				0.0290	0.00781	0.0200
pval				0.147	0.505	0.539
F-first			183.9			105.6

Notes: Coefficient (standard error in parentheses) of the effect of assigned observers or effective observers on whether the proceeding has a respondent's attorney within one year after the end of our randomization period (Attorney). Messages 1 and 2 are described in Section 4.1. Diff_Eff (pval) is the difference of the estimated coefficients (the p-value of the test for the difference). F-first is the first-stage F-statistic of the significance of the instrument. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings with no attorneys in the first hearing between July 10th 2023, and May 10st 2024.

to partial compliance with the experimental assignment. Column 3 reports what we view as the best estimate, derived using the experimental assignment of an observer to a hearing as an instrument for actual observer presence. This TT estimate is 0.041, which represents a 13.7% increase (0.041/0.298=13.7%) in the likelihood that a respondent eventually secures an attorney. The estimates are robust to controlling for the length of the time interval between receiving the notice to appear and the first hearing.

Columns 4-6 of Table 6 investigate heterogeneity in the estimated treatment effects depending on whether the judge is told that the court administration is aware of the observation (Message 1 versus Message 2, as described on page 11). The estimates indicate that the observer effect arises when the court administration is aware of the observation (Message 2). Estimates from treatment 1 are small and are not statistically significant in our preferred specification (IV-TT).

Thus, we find that our experimental intervention increases the likelihood that individuals gain at-

torney representation after the first hearing of an observed proceeding, particularly for the treatment arm (Message 2) that informed judges that the court administration was aware of the observation. The effect was largest among judges who, prior to our intervention, showed a greater inclination to allow respondents time to seek representation.

6 Conclusions

In U.S. immigration courts, individuals facing removal proceedings are not provided with government-appointed legal representation. Those unable to afford private counsel must navigate an exceptionally complex legal system without assistance. Given the technical nature of immigration law, legal representation is often critical for securing relief.

This paper evaluates a novel, low-cost, and scalable intervention aimed at improving access to counsel: courtroom observation. We report results from a randomized controlled trial in which volunteer observers—primarily law students—were assigned to attend immigration hearings. The study examines whether, and to what extent, the presence of these observers influenced judicial behavior and increased the likelihood that respondents subsequently secured legal representation.

We find that having an observer increased the likelihood that the judge adjourns the first hearing to allow the respondent time to hire an attorney. The impact occurs mainly for the treatment arm in which the judge was told that the court administration was informed about that day's observation, suggesting that part of the judge's behavioral response to being observed stems from hierarchical motives/considerations. We also investigate whether there are heterogeneous effects across judges. The data reveal substantial variance across judges in their propensity to grant continuances to secure legal counsel, consistent with the "immigration roulette" pattern documented in the literature. We find that observable judge characteristics do not explain the judges' differing propensities. In terms of their responses to our observation treatment, we find a greater response for judges who, prior to our experiment, were more inclined to adjourn "for reason of seeking representation." Lastly, we examine whether our experimental intervention led to an actual increase in attorney representation after the first hearing and find that it did. Initially unrepresented respondents were more likely to be represented in later hearings if an observer was allocated to their first hearing.

Given the substantial influence of legal representation on case outcomes, our findings suggest a straightforward strategy to enhance fairness in the immigration court system. Even a minimal inter-

vention—external observation—can produce meaningful effects. Our evidence indicates that judicial behavior responds to the pressure of being observed, particularly when that observation is known to court administration. This suggests that even modest monitoring mechanisms—such as those involving unpaid student volunteers—can improve procedural fairness and transparency. More broadly, our results underscore the critical role of accountability and public oversight in promoting integrity within immigration removal hearings.

This study focuses on a timely and pressing social issue: the challenges faced by unrepresented individuals navigating the complex US immigration court system. By shedding light on the impact of observation on court proceedings, the research contributes to the ongoing conversation about improving access to justice, reducing disparities, and ensuring fairness in immigration proceedings.

References

Benson, L. B. (2017). Immigration adjudication: The missing "rule of law". *Journal on Migration and Human Security*, 5(2):331–355.

Blasingame, E. N., Boyd, C. L., Carlos, R. F., and Ornstein, J. T. (2024). How the trump administration's quota policy transformed immigration judging. *American Political Science Review*, 118(4):1688–1703.

Brodeur, A. and Wright, T. (2019). Terrorism, immigration and asylum approval. *Journal of Economic Behavior & Organization*, 168:119–131.

Chen, D. L. and Eagel, J. (2017). Can machine learning help predict the outcome of asylum adjudications? In *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*, pages 237–240.

Cohen, A. and Yang, C. S. (2019). Judicial politics and sentencing decisions. *American Economic Journal: Economic Policy*, 11(1):160–91.

Eagly, I. and Shafer, S. (2016). Access to counsel in immigration court. *American Immigration Council*, 28.

Eagly, I. V. and Shafer, S. (2015). A national study of access to counsel in immigration court. *U. Pa. L. Rev.*, 164:1.

Grenier, J. and Pattanayak, C. W. (2011). Randomized evaluation in legal assistance: What difference does representation (offer and actual use) make. *Yale LJ*, 121:2118.

Hahn, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*, pages 315–331.

Hausman, D. (2016). The failure of immigration appeals. *University of Pennsylvania Law Review*, pages 1177–1238.

Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Kim, C. Y. and Semet, A. (2019). Presidential ideology and immigrant detention. *Duke LJ*, 69:1855.

Koh, J. L. (2016). Removal in the shadows of immigration court. *S. Cal. L. Rev.*, 90:181.

Lim, C. S., Silveira, B. S., and Snyder, J. M. (2016). Do judges' characteristics matter? ethnicity, gender, and partisanship in texas state trial courts. *American Law and Economics Review*, 18(2):302–357.

Miller, B., Keith, L. C., and Holmes, J. S. (2015). Leveling the odds: The effect of quality legal representation in cases of asymmetrical capability. *Law & Society Review*, 49(1):209–239.

Musalo, K., Law, A. O., Daher, A., Donato, K. M., and Meiners, C. (2024). With fear, favor, and flawed analysis: Decision-making in us immigration courts. *BCL Rev.*, 65:2743.

Peacock, I. and Ryo, E. (2022). A study of pandemic and stigma effects in removal proceedings. *Journal of Empirical Legal Studies*, 19(3):560–593.

Ramji-Nogales, J., Schoenholtz, A. I., and Schrag, P. G. (2007). Refugee roulette: Disparities in asylum adjudication. *Stan. L. Rev.*, 60:295.

Ryo, E. and Humphrey, R. (2022). Beyond legal deserts: Access to counsel for immigrants facing removal. *NCL Rev.*, 101:787.

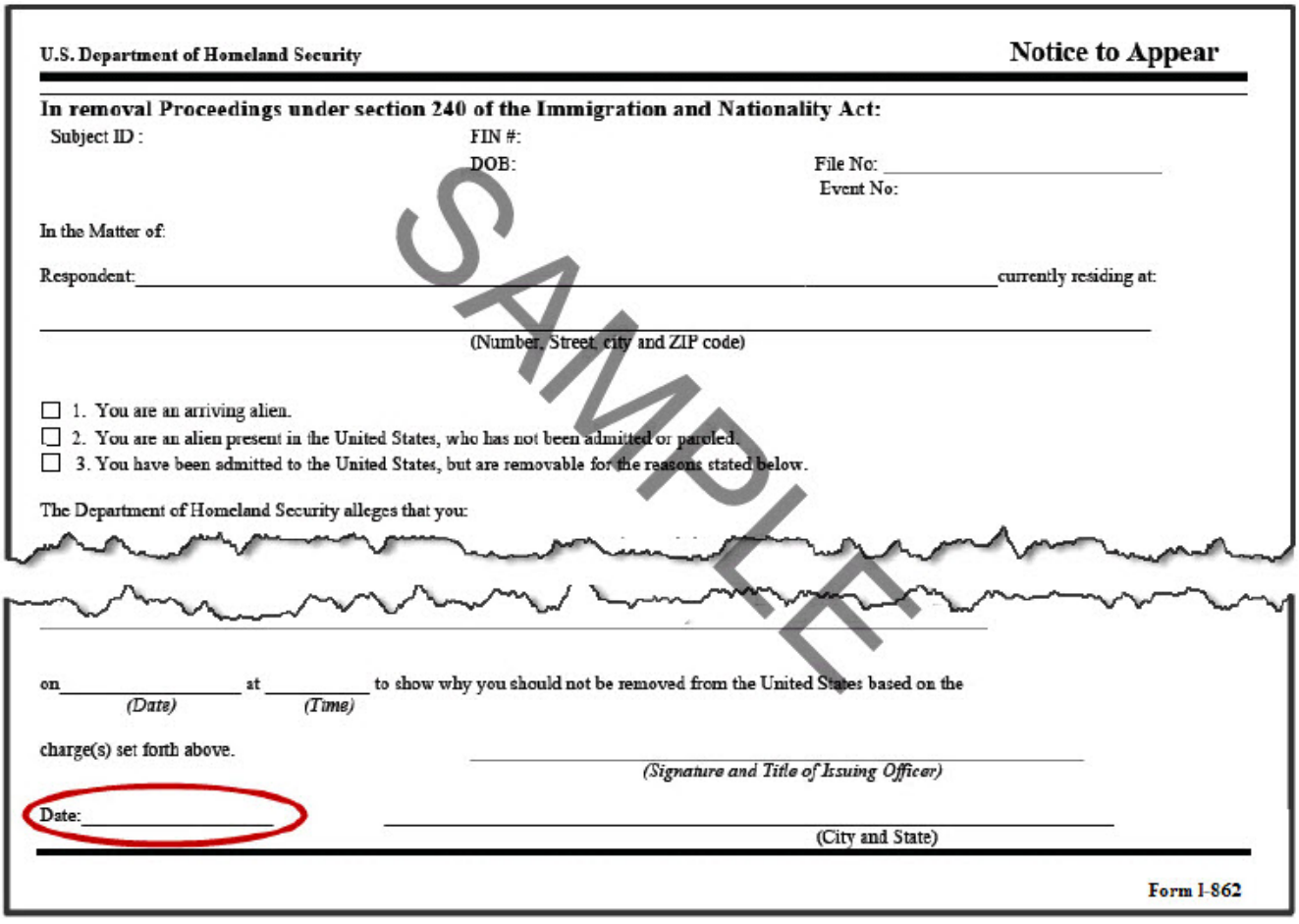
Ryo, E. and Peacock, I. (2021). Represented but unequal: The contingent effect of legal representation in removal proceedings. *Law & Society Review*, 55(4):634–656.

Vojta, G. and Peacock, I. (2025). Economic incentives & judicial decision-making. *Available at SSRN 5291668*.

Online Appendix

A Appendix Figures and Tables

Figure A.1: Sample of Notice to Appear, Form I-862.



U.S. Department of Homeland Security

Notice to Appear

In removal Proceedings under section 240 of the Immigration and Nationality Act:

Subject ID :

FIN #:

DOB:

File No: _____

Event No:

In the Matter of:

Respondent: _____ currently residing at:

_____

(Number, Street, city and ZIP code)

☐ 1. You are an arriving alien.

☐ 2. You are an alien present in the United States, who has not been admitted or paroled.

☐ 3. You have been admitted to the United States, but are removable for the reasons stated below.

The Department of Homeland Security alleges that you:

_____

on _____ at _____ to show why you should not be removed from the United States based on the

(Date) *(Time)*

charge(s) set forth above.

_____

(Signature and Title of Issuing Officer)

Date: _____

_____

(City and State)

Form I-862

SAMPLE

Table A.1: Reliefs by Representation Status

Year	Relief for Represented	Relief for Un-represented
2019	0.17	0.05
2020	0.14	0.03
2021	0.18	0.02
2022	0.21	0.03

Notes: Statistics from the EOIR case database (cases filed from 2000 through April 2025). Relief represents the fraction of cases in which the immigration court's decision was "Relief Granted" conditional on having an application for asylum of the type: "ASYL", "ASYW", "WCAT", "245", "VD", "42A", "42B". Represented and Un-represented: are respondent with an attorney or not, respectively.

Table A.2: In Absentia by Representation Status

Absentia for Represented	Absentia Un-represented
0.02	0.17

Note: Statistics from the EOIR case database (cases filed from 2000 through April 2025). Absentia represents the fraction of cases in which the immigration court's decision was in absentia. Represented and Un-represented: are respondent with an attorney or not, respectively.

Table A.3: No anticipation/Placebo test: "IJ Completion prior to hearing".

Dep.Var.	Comp.Pre.	Comp.Pre.	Comp.Pre.
Method	OLS	OLS-ITT	IV-TT
	(1)	(2)	(3)
Observer	0.001		-0.001
	(0.003)		(0.007)
Observer (Assigned)		-0.001	
		(0.002)	
Observations	121,842	121,842	121,842
Judge FE	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes
Hearing FE	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes
Mean Y	0.0285	0.0290	0.0290
F-first			194.2

Notes: Coefficient (standard error in parentheses) of the effect of having an Observers or an Observer (Assigned) the day of the master calendar hearing to a judge on judge completion prior to the master calendar hearing. F-first is the first-stage F-statistic of the significance of the instrument. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings between July 10^{th} 2023, and May 10^{st} 2024

B Heterogeneous effects on different proceeding types.

Table B.1 shows the four respondent nationalities that are most common in our database - Ecuador (19%), Venezuela (11%), China (9%), and Colombia (6%). Other nationalities have a proportion less than 5%. The last column of the table shows the percentage of individuals that have an attorney at the first hearing, which is highest for Chinese respondents (22%). For the other three nationalities shown in the table, 2.5% or less have an attorney at the first hearing.

Table B.2 explores whether our observer treatment had heterogeneous effects for different types of proceedings and for respondents of different nationalities. Col. 1 shows that the treatment does not have a heterogeneous effect on "dedicated dockets," indicated by D.D., which are proceedings of recently arrived families that are slated for a relatively fast decision (refer to Section 3.1). Nor is there a heterogeneous effect for the very few proceedings which involve an asylum application (col 2.)[33]

Finally, we compute separate observer effects for the most common respondent nationalities. The observer effect is greatest for respondents from Colombia, Ecuador and Venezuela and is basically zero for Chinese respondents. Recall that respondents from the Latin American countries much lower representation rates at their first hearing in comparison to the Chinese respondents, (see Table B.1). This suggests that nationalities with low rates of attorney representation at the first hearing experience greater potential benefits from having an observer present.

Table B.1: Nationalities and Representation

Nationality	% Sample	% Attorney in hearing
Ecuador	19	2.5
Venezuela	11	1
China	9	22
Colombia	6	2.5

Notes: The table reports the distribution of the four main nationalities of the respondents in our sample. Other nationalities are presented in a proportion smaller than %5. % Attorney in hearing represents the proportion of master calendar hearings with an attorney present in the hearing.

[33]Only 301 of our observations feature an asylum application. This makes sense because most asylum applications are filed after the first hearing.

Table B.2: Heterogenous effects.

Dep.Var.	SeekRep	SeekRep	SeekRep
Method	IV-TT	IV-TT	IV-TT
	(1)	(2)	(3)
ObserverXD.D.Case	-0.109		
	(0.129)		
Observer	0.069**	0.061**	
	(0.031)	(0.031)	
D.D.Case	0.279***		
	(0.030)		
ObserverXAsylum Appl.		-0.809	
		(1.511)	
Asylum Appl.		0.256**	
		(0.104)	
ObserverXOther Nat.			0.055*
			(0.030)
ObserverXEcuador			0.140**
			(0.055)
ObserverXVen.			0.065
			(0.070)
ObserverXChina			0.007
			(0.035)
ObserverXColombia			0.080
			(0.075)
Ecuador Nat.			0.159***
			(0.007)
Ven. Nat.			0.144***
			(0.008)
China Nat.			-0.247***
			(0.006)
Colombia Nat.			0.154***
Observations	108,094	108,094	108,094
Judge FE	Yes	Yes	Yes
Week Ass. FE	Yes	Yes	Yes
Week NTA FE	Yes	Yes	Yes
Mean Y	0.320	0.320	0.320
Sum_Eff	-0.0396	-0.749	
pval	0.756	0.620	
F-first	86.42	91.93	25.29

Notes: Coefficient (standard error in parentheses) of the effect of having an Observers or an Observer (Assigned) the day of the master calendar hearing to a judge on adjournment is "respondent to seek representation". In col.3, Other Nat. is an indicator variable for respondents not from Ecuador, Venezuela, China or Colombia. Sum_Eff (pval) is the sum of the estimated coefficients (the p-value of the test for the sum). F-first is the first-stage F-statistic of the significance of the instrument. Regressions include Judge, Week of the hearing and Week of the NTA FEs. SEs are clustered at the judge-week level. Significance at 10% (*), 5% (**), and 1% (***). Sample: all hearings with no attorneys in the first hearing between July 10^{th} 2023, and May 10^{st} 2024.

Table B.3: Heterogenous effects.

	(1)
Dep.Var.	IJ FEs
Method	OLS
Female	0.032
	(0.037)
Age	-0.003
	(0.003)
Main Court	-0.034
	(0.042)
Years from appointment	0.001
	(0.003)
$>$ 2 Years from appointment	0.085
	(0.112)
N.Cases pre-exp	0.002
	(0.006)
Chief Immigration Judge Rep.	-0.012
	(0.036)
Observations	85

Notes: Coefficient (standard error in parentheses) of the regression of the judges FEs estimated during the experimental period on judges observable characteristics. "> 2 years from appointment" (indicating probation status), "N. Cases pre-exp" (number of cases prior to the experiment), and "Chief Immigration Judge Rep." (whether the chief immigration officer has been appointed by a Republican President). Significance at 10% (*), 5% (**), and 1% (***). Sample: all immigration judges in hearings with no attorneys in the first hearing between July 10^{th} 2023, and May 10^{st} 2024.