Australian National University

Crawford School of Public Policy

TTPI

Tax and Transfer Policy Institute

Legal aid eligibility and court outcomes: a design-based double-machine-learning approach

TTPI - Working Paper 2/2024
February 2024

Fabio I. Martinenghi
Research Associate
UNSW

Abstract

The UN General Assembly recognises the right to equal treatment before the law as a universal human right. A minimum requirement to uphold this right is to guarantee access to a lawyer independently of one's income, and this is what legal aid does. While legal aid is necessary to reduce income-based inequality before the law, the extent to which this is achieved depends on the quality of its services. In a context where all defendants have access to a lawyer, I study the impact of refusing legal aid on court outcomes. This is a conservative estimate of the performance gap between privately and publicly-funded legal services. I do this by taking a double machine learning approach to a new administrative dataset linking data from legal aid application forms to their court outcomes in New South Wales, Australia. Access to the legal aid application forms, which include all the factors used in the decision to provide aid, allows me to learn the unknown treatment assignment function and identify treatment effects via the random forests algorithm. I find that aid applicants who fail the means test and hire a private lawyer are 10 p.p. less likely to be incarcerated than if they passed it and relied on legal aid. With an average positive incarceration length of almost 4 years, this gap is consequential. Finally, I argue that this approach, combining high quality administrative data with double machine learning, may allow policy-makers to track the performance of public programs that are hard to evaluate using standard econometric approaches.

JEL Codes: I30, K14, H44

Keywords: Indigent, defense crime, criminal justice

* *I am grateful to Don Weatherburn for facilitating contact with Legal Aid NSW, and to Legal Aid NSW for giving me access to their dataset and for their outstanding support. The views expressed in this paper are my own and not necessarily those of the Legal Aid NSW. I thank Richard Holden for his financial support, Fangzhou Yu for his helpful comments and all the participants at seminars and conferences at Bocconi University, the University of Melbourne, the University of Technology Sydney and the Australian National University for their feedback. Author contact* f.martinenghi@unsw.edu.au

Tax and Transfer Policy Institute

Crawford School of Public Policy

College of Asia and the Pacific

+61 2 6125 9318

tax.policy@anu.edu.au

The Australian National University

Canberra ACT 0200 Australia

www.anu.edu.au

The Tax and Transfer Policy Institute (TTPI) is an independent policy institute that was established in 2013 with seed funding from the federal government. It is supported by the Crawford School of Public Policy of the Australian National University.

TTPI contributes to public policy by improving understanding, building the evidence base, and promoting the study, discussion and debate of the economic and social impacts of the tax and transfer system.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

Legal aid eligibility and court outcomes: a design-based double-machine-learning approach

Fabio I. Martinenghi

Abstract

The UN General Assembly recognises the right to equal treatment before the law as a universal human right. A minimum requirement to uphold this right is to guarantee access to a lawyer independently of one's income, and this is what legal aid does. While legal aid is necessary to reduce income-based inequality before the law, the extent to which this is achieved depends on the quality of its services. In a context where all defendants have access to a lawyer, I study the impact of refusing legal aid on court outcomes. This is a conservative estimate of the performance gap between privately- and publicly-funded legal services. I do this by taking a double machine learning approach to a new administrative dataset linking data from legal aid application forms to their court outcomes in New South Wales, Australia. Access to the legal aid application forms, which include all the factors used in the decision to provide aid, allows me to learn the unknown treatment assignment function and identify treatment effects via the random forests algorithm. I find that aid applicants who fail the means test and hire a private lawyer are 10 p.p. less likely to be incarcerated than if they passed it and relied on legal aid. With an average positive incarceration length of almost 4 years, this gap is consequential. Finally, I argue that this approach, combining high-quality administrative data with double machine learning, may allow policy-makers to track the performance of public programs that are hard to evaluate using standard econometric approaches.
Keywords: Indigent Defense, Crime, Criminal Justice.

JEL: I30, K14, H44.

1 Introduction

The principle of equality before the law states that everyone's rights should be equally protected by the law without discrimination. It is an important and ancient democratic principle, dating at least back to the fifth century B.C., when it was mentioned by Thucydides (Thucydides, 1874) in praise of Athenian institutions. More recently, the United Nations General Assembly recognised it as universal human right (United Nations, 1948). One way in which governments implement this principle is by offering "legal advice, assistance and representation [...] at no cost for those without sufficient means"[1] (United Nations, 2012). This is also called legal aid.

While legal aid is necessary to achieve some equality before the law, the extent to which this equality is realised depends on the quality of the services that constitute the legal aid, i.e. the performance of publicly-funded compared to privately-funded legal representation. Intuitively, the more accessing privately-funded lawyers improves a defendant's outcomes in court compared to a publicly-funded lawyer, the less the justice system has achieved equality before the law.

In this paper, I study the effect of refusing legal aid on the applicant's case outcomes. This is a first in the literature. I use administrative data on legal aid applicants charged with serious crimes in New South Wales (NSW), Australia. For such cases, legal aid can only be refused—after failing a means test—if the applicant is found able to pay for a private lawyer. Such features of the data imply (i) that most applicants who are refused aid hire private lawyers (circa 90%), and (ii) that the privately-funded private lawyers in my sample are likely to charge lower fees than the average private lawyer, as they are hired by people whose income is low enough to potentially qualify for aid. Hence, under the assumption that higher-performing private lawyers charge higher legal fees, my estimates can be interpreted as providing a lower bound to the penalty associated with relying on publicly-funded legal representation.

Because applicants are not assigned aid at random, but rather via a means test, estimated treatment effects that rely on mean comparisons between successful and unsuccessful applicants are marred by selection bias and fail identification. While the relevant means tests are precisely defined in NSW, they are complemented by principle-based discretionary rules, meaning that the treatment assignment function is largely unknown. To address this issue, I take a double machine learning approach (Chernozhukov et al., 2018a). Specifically, I use the random forests algorithm (Breiman, 2001) on administrative data that include all the inputs of the treatment assignment function—i.e. all the 70+ factors derived from the legal aid application form and considered when deciding whether to grant aid. This allows me

[1] This is the definition of the right to counsel, which the government enforces via legal aid.

to flexibly learn the treatment assignment function and use the resulting propensity score to recover treatment effects. Indeed, double machine learning (DML) can work as an inverse propensity score estimator, a class of estimators which is unbiased and efficient under unconfoundedness (Hirano et al., 2003).

I find that aid refusal decreases the probability of a defendant going to jail by 9 p.p, compared to being represented by a public lawyer. This is a sizeable effect, especially considering that the mean incarceration length is 46.5 months—almost 4 years. Aid refusal does not significantly reduce incarceration time, but it decreases the chances of pleading guilty to the highest charge by 4.7 p.p. and increases the chances of being fined by 5.4 p.p. I focus on average treatment effects on the treated (ATT), rather than average treatment effects (ATE), as they are more stable in settings such as the present, where the treatment group is small relative to the control group. In a separate analysis, I further investigate the performance gap between the private and public legal services by dropping from my sample applicants who self-represented. I find that failing to receive aid and hiring of a private lawyer decreases the defendant's probability of incarceration by 10.2 p.p., decreases the chances of pleading guilty to the highest charge by 5.1 p.p. and increases those of being fined by 7.1 p.p.

I rely on two administrative datasets. First, the Legal Aid NSW dataset covers all applications for aid made between 2012 and 2021 and related to serious criminal cases. Second, each application is linked to its court data, stored in the Re-Offending Database, which is curated by the NSW Bureau of Crime Statistics and Research (BOCSAR). This linkage provides me with a wealth of information on the applicant (e.g. country of origin, number of dependants, year of birth), the application (e.g. time of submission, which offices processed it) and the case (e.g. trial length, sentence).

The contribution of this paper is two-fold, as it both adds to the indigent legal defence literature and the causal machine learning literature. On the indigent defence front (see Anderson and Heaton, 2012; Iyengar, 2007; Roach, 2014; Shem-Tov, 2022; Agan et al., 2021; Schwall, 2017; Lee, 2021), it is the first study to provide evidence on the effect of legal aid on court outcomes. The existing empirical research around indigent defence does not address this point, as it mostly leverages institutional designs and reforms in the public defence context to study how lawyers respond to different incentive schemes. One strand of the literature studies how the court outcomes of indigent defendants change when legal aid comes in the form of publicly-funded private lawyers (called "panel lawyers" in Australia, and "assigned counsels" in the US) compared to public lawyers ("in-house lawyers" in Australia, "public defenders" in the US; see Anderson and Heaton, 2012; Iyengar, 2007; Roach, 2014; Shem-Tov, 2022). This literature consistently shows that panel lawyers tend to perform worse than public lawyers. Because panel lawyers are simply private lawyers who have taken a public-funded case, these results could suggest that counterfactual private lawyers perform worse than public lawyers—although this depends on how panel lawyers are remunerated.

Relatedly, Porreca and McCannon (2023) study the introduction of the right to counsel from a historical perspective, finding difference-in-differences evidence that it led to an increase in conviction rates in 19th century London. Their study, while related to mine, speaks more to the issue of the professionalisation of criminal defense than that of legal aid provision —as they note, a publicly-funded legal aid society was only established in 1949.

Another strand of the literature further investigates how fee structure affects legal performance. They do this by focusing on reforms that change the compensation structure of a single type of lawyer, panel lawyers (Schwall, 2017; Lee, 2021). Schwall (2017) and Lee (2021) both use a difference-in-differences design and find that, compared to hourly rates, flat fees drastically reduce the number of hours worked by defence lawyers while either having no detectable impact on court outcomes (Schwall, 2017) or a large negative impact on the same (Lee, 2021). In-house lawyers work under a fixed salary, as opposed to the private and panel lawyers' hourly rates. Hence, this strand of the literature would suggest that in-house lawyers perform worse than private lawyers. It is, therefore, ex-ante unclear whether in-house lawyers (or public defenders) perform better or worse than private lawyers—an open empirical question that this paper sets to investigate.

Furthermore, this study contributes to the causal machine learning literature by providing a first example of ML-based credible identification. It also illustrates how to go beyond a black-box approach to DML by making use of its estimated propensity score. The literature applying ML methods for treatment effects estimation is still in its infancy. In RCT settings, Bertrand et al. (2021) and Davis and Heller (2020) leverage causal random forests (Wager and Athey, 2018) to recover individual treatment effects. Chernozhukov et al. (2018b) develop and apply an ML method to study treatment effect heterogeneity in an RCT context. Similarly, Knaus et al. (2022) develop and apply a procedure that combines ML and program evaluation techniques to heterogeneity in RCTs. In observational settings, only two published papers use ML for estimating treatment effects. Deryugina et al. (2019), first, use random forests and LASSO to predict the life-years lost due to air pollution. Then, they apply (Chernozhukov et al., 2018b) to study treatment effect heterogeneity. Knaus (2022) applies DML to longitudinal survey data to estimate the (dose-response) effects of musical practice on students' cognitive skills.

The present study adds to this literature for two reasons. First, it shows how the identification strategy relying on conditional independence (or *unconfoundedness*), generally considered weak in the economic literature, can be credible in the right context. This is the case in observational program evaluation settings where the variables used for assigning treatment are known and observed, but functional form of the assignment mechanism is not known. Indeed, while some policies make use of sharp rules for assigning treatment, many rely of fuzzier rules that leave space for discretion, or may be based on sharp rules in theory but not in practice. In such cases, relying on more standard methods such as regression discontinuity designs not only restricts the target parameter to a local average treatment effect, but it further focuses on the *complier* fraction of such population, leading to potential sample-size challenges. Instead, using DML with flexible algorithms such as random forests allows to credibly estimate the treatment assignment function while also targeting (global) ATEs and ATTs, serving as either a complement or a substitute of standard methods.

Moreover, a concern about the transparency of machine learning methods seems to be a barrier to their widespread adoption in economics research. While more familiar methods, such as difference in differences or regression discontinuity design, are associated with a wide array of tests to assess their robustness and the strength of their identifying assumptions, analogous tests are less present and known in the ML literature, thus leaving a black-box impression. To address this issue, I show that the DML Interactive Regression Model in Chernozhukov et al. (2018a) estimates a propensity score and this can be used, for instance, to assess the credibility of the common support assumption and the sensitivity of the results to including observations with low overlap. This is in a similar spirit to Knaus et al. (2022), which adapts a covariate balance test to DML.

From a policy implication perspective, the approach taken in this paper can be applied to monitor the performance gap between public and private legal services, and extended to other policy evaluation exercises. Moreover, the very ability to measure this gap opens new policy questions, allowing citizens to form an opinion over whether such a gap should exist, and potentially advocate for policies to reduce it. Because publicly-funded lawyers are typically paid less than private sector lawyers (see Forell et al., 2010, for NSW evidence), a simple policy to reduce the gap could be to increase their salary/rates. Indeed, the most recent available data show that public and private lawyers earned, respectively, A$86,700 and A$94,800 on average in 2008/09 (Forell et al., 2010), while public and private barristers earned, respectively, A$106,184 (Forell et al., 2010) and A$361,850 (ABS, 2008)[2] on average.

The remainder of the paper is organised as follows. Section 2 presents the policy, its eligibility rules, the organisation providing aid and the rules for its provision and some institutional context. Section 5 presents the empirical approach taken to estimate the treatment effects. Section 6 discusses the results and Section 8 concludes.

2 Context and Background

2.1 Legal Aid

In Australia, Dietrich v. The Queen (1992) establishes the right to a fair trail for federal cases, which include those involving serious crimes ("indictable crimes" in Australia, "felonies" in the US). In New South Wales, this has been interpreted as implying the right to counsel—the right to "legal advice, assistance and representation [...] at no cost for those without sufficient means" (United Nations, 2012)—so that defendants accused of serious crimes are offered the chance to apply for legal aid.

Legal Aid NSW is a government organisation providing legal assistance to disadvantaged people in New South Wales (NSW), Australia. It was founded in 1979 under the name of Legal Services Commission, and it now counts more than 1,600 staff and 25 offices, two satellite offices and 243 regular outreach locations spread across NSW (Legal Aid NSW, 2017).

Legal Aid provides such legal services either directly, via in-house lawyers and barristers, or by funding private lawyers and barristers at an agreed lower-than-market rate. Private lawyers working on Legal Aid matters are called "panel lawyers" ("assigned counsels" in the US), a useful term that avoids confusing them with private lawyers working on privately-funded matters. Legal Aid covers (with restrictions) matters in Family, Civil and Criminal Law —both at a state and at a federal (or "Commonwealth") level. In this paper, I focus on the Criminal Law Division of Legal Aid NSW, as it is the one covering Criminal matters.

[2] Estimate calculated from ABS data dividing the income generated by private barristers in the 2007-08 financial year (A$1.4b) by their number in the same period (3,869 private barristers). Public barristers are to be intended as senior public lawyers acting as barristers.

2.2 Eligibility

Legal aid determines whether an applicant is eligible for aid based on four criteria: the Jurisdiction test, the Means test, the Merit test and the Availability of Funds test (Legal Aid NSW, 2023). Matters related to indictable crimes automatically pass the Jurisdiction and the Merit tests —the former with some rare exceptions detailed in Appendix A. Historically, the availability of funds test has never been used. It follows that the Means test is the only test relevant to our context. Appendix A includes a brief review of the other Legal Aid tests for completeness.

The Means test focuses on the applicant's income and assets. It seeks to determine whether they are eligible for aid and how much they would need to contribute towards their legal expenses if their application were successful. It consists of three sub-tests—income, assets and ability to pay legal costs tests. To pass the income test, an applicant must earn less than $400 in *net assessable income*[3]. This is calculated as gross income minus some benefits and deductions, most of which are a function of the number of the applicant's dependants (see Legal Aid NSW, 2022, for details).

An implication of this test is that if an applicant is receiving the maximum amount of welfare transfers, they pass the test. The asset test sets the eligibility threshold at 100$ in net assets, after taking an applicant's (gross assessable) assets and deducting, subject to caps, a range of items—such as home and business equity, motor vehicle equity and furniture value. Then, assets and income are considered together to assess whether the applicant can afford to hire private legal representatives and to secure a loan to pay for them. The Means test is not applied in certain types of matters (see Legal Aid NSW, 2022).

Finally, every application passes the Availability of Funds test as long as Legal Aid has the funding to provide its services. Historically, this has always been the case.

2.3 Serious crimes

I focus on serious crimes, which are called *indictable offences* in Australia (and *felonies* in the US), and which hence exclude the less serious *summary offences* (*misdemeanors* in the US). In NSW, criminal offences can be catalogued as either Summary, Table 1, Table 2 and Strictly Indictable. Table 1 and 2 offences can be deemed as either summary—and hence

[3] In Appendix C, I attempt to use a fuzzy regression discontinuity design on this threshold, but am unable to produce credible treatment estimates due to several issues.

processed by a magistrate in the Local Court—or indictable—hence dealt with by a judge in the District or Supreme Court. Instead, Summary Offences are always summary and Strictly Indictable are always indictable.

Within indictable matters, legal aid is available for committals, bail applications, mentions and adjournments, sentence matters, variations or revocations of a Community Correction Order or Conditional Release Order and criminal trials (Legal Aid NSW, 2023).

2.4 Daily operations

A defendant can apply for legal aid by filling in a form and either posting, emailing, or hand-delivering it to a Legal Aid office. Alternatively, they can have an in-house Legal Aid lawyer or a private lawyer submit it for them—the former free of charge, the latter according to their rate.

However, the location where the application is lodged does not determine who will handle the case. A case is assigned based on the court jurisdiction under which it fell. Each Legal Aid office is responsible for at least one court. Not all courts are covered by Legal Aid offices though, and Legal Aid has to rely entirely on panel lawyers when a case falls into a non-covered court. Cases are also assigned to panel lawyers whenever the relevant Legal Aid office has a conflict of interest or if it is at capacity. Finally, whenever a case is assigned to a public (in-house) lawyer, the lawyer is allocated on an as-good-as-random basis.

2.5 The role of Discretion

There is an important exception to the application of the eligibility tests. As stated in Legal Aid Commission Act (1979) ss 30, Legal Aid reserves the right to provide aid even when the applicant fails the tests outlined above. In daily operations, this is called *discretion*. In practice, if an applicant fails the Means test, the lawyer who was assigned the case can ask for discretion if the case meets a range of loosely defined conditions related to the financial situation of the applicant. This request is then reviewed, and mostly granted, by another member of staff.

3 Data

The data used in this analysis covers the universe of New South Wales criminal cases where the defendant applied for legal aid between 1 January 2012 and 31 December 2021 and is not recorded as Aboriginal nor Torres Strait Islander[4]. I obtain this by merging two administrative datasets.

The first dataset was provided by Legal Aid NSW and includes information on applicants to legal aid and their applications (e.g. whether it was successful) and the lawyer assigned by Legal Aid—the latter if aid is granted. It includes rich information on the applicants, such as income, assets, number of dependants and welfare receipts, as these factors are used to determine the outcome of the application.

The second dataset is the Re-Offending Database (ROD), developed and maintained by the New South Wales Bureau of Crime Statistics and Research (BOCSAR). From this database, I extract charge-level data on indictable matters. These include information court outcomes such as fine amounts, guilty pleas incarcerations. They also include some demographic characteristics of the defendant, such as gender and age, and details on the court that processed the case.

BOCSAR was in charge of linking Legal Aid applications to their court proceedings information across the two datasets. Most of the cases are linked using a case-specific identifier code. Whenever this information is lacking, they are linked to the date in which the application for aid was submitted and the name and date of birth of the applicant/defendant. Such identifying information was not included in the final dataset I received for privacy purposes.

In my analysis, I focus on seven case-level outcomes: "reduced charges", which indicates if the number of finalised charges is less than that of initial charges; "guilty share", recording the share of charges for which the defendant is found guilty; "reduced seriousness", equal if a defendant was not found guilty of their most serious initial charge. Whether a defendant was incarcerated ("incarcerated") and for how many months ("incarceration") are also used as outcomes. Finally, I look at whether the defendant pleaded guilty to their highest charge, whether they were fined.

[4] Access to Aboriginal and Torres Strait Islander data has additional ethics requirements that I could not fulfil within my timeline.

4 Descriptive analysis

Table 1 compares the mean values of key covariates across treatment groups. For some covariates, I show their unconditional mean (by treatment group), the share of non-zero values[5] and the mean of their positive values. I do this for covariates that have a large share of zero values, which makes their unconditional mean very imprecise.

We can infer from Table 1 that a mean comparison between the outcomes of treatment and control groups would likely be strongly affected by selection bias. Indeed, applicants whose application for legal aid was refused or terminated have on average higher incomes, more valuable assets, spend more on housing and are less likely to have a disability. These are all factors that increase their probability of facing a rejection or termination, and these factors are also likely to be affecting their outcomes—for instance via how much they are able to spend on legal fees. On the other hand, applicants whose application for legal aid was refused or terminated have more dependants on average (i.e. household members relying on their income), which decreases their probability of a rejection. Finally, the two groups are similar across age and share of females.

Inspecting the (unadjusted) mean outcomes for the whole sample and by treatment group can also be informative. Table 2 reports the mean value for each outcome plus an extra variable, positive incarceration time, which will be useful to interpret the study findings. Columns 2 and 3 report the mean values for the control and treatment groups, respectively, while columns 4 and 5 report the mean differences and their p-value, respectively. Column 5 reports the mean values for the whole sample, unconditional on treatment status. It should be noticed that the size of the treatment group (cases where aid was refused or terminated) is 13.5% of that of the control group (cases where aid was granted), implying that the unconditional means underrepresent the treatment group.

Inspecting Table 2, we can see that the share of charges for which defendants are found guilty does not differ across treatment groups. The same applies to the probability for a defendant to have their charges reduced. Most charges are finalised[6] (66%) and it is relatively rare for the most serious charge to be withdrawn or dismissed (11%). Lawyers manage to reduce the charges of the defendants 57% of the time and legal aid lawyers achieve that 2 p.p. more frequently than private lawyers and self-represented defendants.

[5]These variables do not have negative values—they are bounded at zero.
[6]Finalising a charge means that the defendant is found guilty of the charge.

Table 1. Covariate balance table for legal aid applicants

	Granted		Refused/Terminated		Difference	
Variable	Mean	Obs.	Mean	Obs.	Value	p-value
Weekly gross income (A$)	134.40	29,207	280.73	3,946	-146.33	0.0000
	(262.19)		(665.18)		(10.70)	
Any gross income	0.34	29,207	0.44	3,946	-0.09	0.0000
	(0.47)		(0.50)		(0.01)	
Positive gross income (A$)	391.10	10,037	644.43	1,719	-253.32	0.0000
	(315.66)		(884.02)		(21.55)	
Net assets (A$)	435.66	29,207	12,666.64	3,946	-12,230.98	0.0000
	(14,873.93)		(136,274.12)		(2,171.12)	
Any net assets (A$)	0.01	29,207	0.08	3,946	-0.07	0.0000
	(0.11)		(0.27)		(0.00)	
Positive net assets (A$)	35,345.45	360	155,225.33	322	-119,879.87	0.0000
	(129,463.72)		(453,905.55)		(26,199.30)	
Weekly Housing (A$)	44.75	29,207	77.60	3,946	-32.85	0.0000
	(91.96)		(126.23)		(2.08)	
Any weekly Housing (A$)	0.28	29,207	0.38	3,946	-0.10	0.0000
	(0.45)		(0.49)		(0.01)	
Positive weekly Housing (A$)	162.03	8,067	203.74	1,503	-41.71	0.0000
	(107.77)		(127.02)		(3.49)	
Motor vehicle (A$)	587.54	29,207	1,921.19	3,946	-1,333.66	0.0000
	(3,636.04)		(9,287.10)		(149.37)	
Any motor vehicle (A$)	0.10	29,207	0.18	3,946	-0.09	0.0000
	(0.30)		(0.39)		(0.01)	
Positive motor vehicle (A$)	6,095.98	2,815	10,588.03	716	-4,492.06	0.0000
	(10,179.63)		(19,595.68)		(757.04)	
No. of Dependants	0.15	29,207	0.23	3,946	-0.08	0.0000
	(0.63)		(0.73)		(0.01)	
Female	0.12	29,207	0.13	3,946	-0.01	0.2246
	(0.33)		(0.34)		(0.01)	
Age	34.77	29,207	35.22	3,946	-0.45	0.0414
	(12.32)		(13.19)		(0.22)	
Disability	0.15	29,207	0.12	3,946	0.03	0.0000
	(0.35)		(0.32)		(0.01)	

Notes: this table reports the mean covariate values by treatment group and their difference, for a selection of covariates. Column 1 reports the mean covariate values for legal aid recipients. Column 2 reports the number of observations for this group. Column 3 reports the mean covariate values for legal aid applicants to whom aid was either refused, or granted but then terminated. Column 4 reports the number of observations for this group. Column 5 reports the difference between control and treatment groups—Columns 1 and 3, respectively. The p-value of this difference is reported in column 6. For "Gross income", "Net asstes", "Weekly housing", and "Motor vehicle", I report the unconditional mean, the share of positive values and the mean of positive values. As these variables are zero-inflated, their unconditional means are imprecise. Notice also that these variables are bounded at zero. Standard errors are reported between parentheses.

The rates of incarceration sit at 57%, and is 58% for legal aid cases and 43% for private and self-represented. Legal aid defendants are expected to spend 27 months in jail, 6 p.p. more frequently than their counterpart. However, conditional on going to jail, legal aid defendants spend 42 months in custody —3 months less than privately- and self-represented defendants. This could suggest that private lawyers are better at keeping out of jail those defendants whose expected time in jail is relatively low. Indeed, the unconditional incarceration length (which includes zero values) is higher for legal aid defendants (27 months) versus non-legal aid (21 months). The share of guilty pleas to the highest charge is high in both groups —65% and 63% for legal aid and non-legal aid defendants, respectively —and about 3 p.p. higher for legal aid defendants. Finally, the share of fined defendants is 6% and, unsurprisingly, legal aid defendants are less likely to be fined than their counterparts.

5 Empirical Strategy

In this section, I present the empirical strategy used to identify the effect of legal aid on court outcomes. I use a double machine learning estimator on administrative data to learn the unknown treatment assignment function and estimate the average treatment effect on the treated. In Section 5.1, I introduce the key assumptions under which DML estimators identify the average treatment effect and the average treatment effect of the treated. I follow Imbens and Rubin (2015)'s exposition of the unconfoundedness assumption (Rosenbaum and Rubin, 1983), on which the Interactive Regression Model—which I use—relies for identification purposes.

Then, I present the Interactive Regression Model in Section 5.2. This is one of the models introduced in Chernozhukov et al. (2018a). It relies on propensity scores and allows the treatment variable to interact with all the other covariates, and hence does not assume homogeneous treatment effects. While this approach allows me to identify both the average treatment effect (ATE) and the average treatment effect on the treated (ATT), I focus on the ATT in my analysis. Indeed, the ATT can be estimated more reliably in cases such as mine, where the treatment group is much smaller than the control group. Intuitively, in such cases propensity score estimators can more easily find propensity score matches for treatment group units in the wider control group.

Before delving into the technical part, stating the key idea concisely can be helpful. Legal Aid decides on aid allocation based on the information contained in a form. All this

Table 2. Balance-type table for outcomes

	$D_i = 0$	$D_i = 1$	Diff.	p-value	All
Reduced charges	0.57	0.55	0.02	0.0124	0.57
	(0.49)	(0.50)	(0.01)		(0.49)
Guilty share	0.66	0.66	0.00	0.8740	0.66
	(0.36)	(0.37)	(0.01)		(0.36)
Reduced seriousness	0.11	0.10	0.01	0.1907	0.11
	(0.31)	(0.30)	(0.01)		(0.31)
Incarceration (months)	27.13	21.16	5.97	0.0000	26.42
	(44.78)	(44.22)	(0.75)		(44.75)
Incarceration > 0 (months)	46.42	49.21	-2.79	0.0495	46.67
	(50.35)	(56.28)	(1.42)		(50.92)
Incarcerated	0.58	0.43	0.15	0.0000	0.57
	(0.49)	(0.50)	(0.01)		(0.50)
Guilty plea to highest charge	0.65	0.63	0.03	0.0015	0.65
	(0.48)	(0.48)	(0.01)		(0.48)
Fined	0.05	0.11	-0.06	0.0000	0.06
	(0.23)	(0.32)	(0.01)		(0.24)

Notes: this table reports the mean outcome values by treatment group, their difference, and the means for the whole sample. Column 1 reports the mean outcome values for legal aid recipients. Column 2 reports the mean covariate values for legal aid applicants to whom aid was either refused, or granted but then terminated. Column 3 reports the difference between control and treatment groups—Columns 1 and 2, respectively. The p-value of this difference is reported in column 4. Column 5 reports the means for the whole sample, unconditional on treatment status. Notice that legal aid recipients are over-represented compared to their counterpart, as the size of the treatment group is 13% of that of the treatment group, of legal aid recipients. Standard errors are reported between parentheses.

information is included in my dataset[7]. The allocation of aid follows sharp rules only partly, as loose principle-based rules play an important role via discretion. My claim is that the flexibility of double machine learning with random forests allows me to credibly recover the treatment assignment function and hence identify both ATE and ATT while addressing the two main threats to identification. First, such flexibility effectively tackles potential bias due to functional-form misspecification. Second, the complete access to the form data used for assigning treatment greatly reduces omitted variable bias concerns.

5.1 Unconfoundedness

Unconfoundedness is the key assumption required to identify ATE and ATT using DML. Referred to as *strongly ignorable treatment assignment* in the original Rosenbaum and Rubin (1983) article, unconfoundedness characterises assignments mechanisms that (i) do not depend on potential outcomes, (ii) are probabilistic—no unit is assigned to treatment or control with probability one—and (iii) are individualistic. I break down this definition in the paragraphs below. Notice that after I introduce an assumption, I will hold it as true when introducing the following one. Moreover, I indicate matrices and column vectors in bold as in Imbens and Rubin (2015).

An *assignment mechanism* is a row-exchangeable function $\Pr(\mathbf{D} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$ which returns values in the interval $[0, 1]$ and satisfies

$$\sum_{\mathbf{D} \in \{0,1\}^N} \Pr(\mathbf{D} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = 1,$$

for all $\mathbf{X}, \mathbf{Y}(0)$, and $\mathbf{Y}(1)$.

The set $\mathbb{D} = \{0, 1\}^N$ is the set of all vectors of length N in which each element is equal to either 0 or 1, where N is the given (finite) population size. $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ are also N-vectors, while $\mathbf{X}$ is a $N \times K$ matrix of covariates. Here, $\mathbf{X}$ is made of N *row* vectors X_i, each X_i including all the K covariates of unit i. Row-exchangeable means that it does not matter how one orders the elements within a vector or matrix. Notice that the bold notation here indicates vectors (or matrices in the $\mathbf{X}$ case), so that $\Pr(\mathbf{D} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1))$ is the probability that a given treatment assignment vector $\mathbf{D}$ occurs, where each $\mathbf{D}$ specifies the treatment status of each unit in N. Such assignment vectors, 2^N in total, are assumed

[7] The only piece of information that is in the form but not in the dataset is the text in which the applicant describes what happened.

to sum up to one.

While in this general definition treatment assignment is allowed to depend on the potential outcomes, this is not true under unconfoundedness. Unconfoundedness requires that the assignment mechanism is *independent of the potential outcomes*:

$$\Pr(\mathbf{D} \mid \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1)) = \Pr\left(\mathbf{D} \mid \mathbf{X}, \mathbf{Y}'(0), \mathbf{Y}'(1)\right)$$

for all $\mathbf{D}, \mathbf{X}, \mathbf{Y}(0), \mathbf{Y}(1), \mathbf{Y}'(0)$, and $\mathbf{Y}'(1)$, which allows dropping the potential outcomes from the assignment mechanism notation, now written as $\Pr(\mathbf{D} \mid \mathbf{X})$. This is the first of three assumptions embedded into the unconfoundedness concept.

Now, we can define the probability p_i that unit i is assigned to treatment as

$$p_i(\mathbf{X}) = \sum_{\mathbf{W}: D_i = 1} \Pr(\mathbf{D} \mid \mathbf{X}) \tag{1}$$

meaning that one can find the assignment probability of a unit by calculating the share of assignment vectors $\mathbf{D}$ where it is assigned to treatment. Notice that in the present application, the $\mathbf{X}$ matrix (exclusively) includes the information contained in the Legal Aid application form.

Then, taking the average probability of being assigned to treatment for units with covariates $X_i = x$, one can find the *propensity score*. Formally,

$$e(x) = \frac{1}{N(x)} \sum_{i=1}^{N} p_i(\mathbf{X})$$

where $N(x) = \sum_{i=1}^{N} \{\mathbf{1}_{X_i = x}\}$ counts the number of units that have $X_i = x$. If $N(x) = 0$ for some x, the propensity score is defined to be zero at that x. We can now go back to presenting the second and third elements of the unconfoundedness assumption.

The second assumption is that the assignment mechanism is *individualistic*, i.e. it is such that, for some function $q(\cdot) \in [0, 1]$,

$$p_i(\mathbf{X}) = q\left(X_i\right), \text{ for all } i = 1, \ldots, N,$$

and

$$\Pr(\mathbf{D} \mid \mathbf{X}) = c \cdot \prod_{i=1}^{N} q\left(X_i\right)^{D_i}\left(1 - q\left(X_i\right)\right)^{1-D_i}$$

for $(\mathbf{D}, \mathbf{X}) \in \mathbb{A}$, for some set $\mathbb{A}$, and zero elsewhere, where constant c is such that the probabilities sum up to one. Hence, individualistic assignment allows to simplify the propensity score to:

$$e(x) = \frac{1}{N(x)} \sum_{i:X_i=x} q\left(X_i\right)$$

or, in short,

$$e(x) = q\left(x\right)$$

implying (a) that the assignment probability of each unit i is not allowed to depend on the covariate values of other units—but only on its own—and (b) that $q(.)$ now only depends on X_i, making it equal to propensity score $e(x)$. Finally, notice that the independence from potential outcomes combined with the probabilistic assignment assumption implies that the propensity score can be interpreted as the unit-level assignment probability, rather than the average probability of assignment for units with $X_i = x$.

Third, an assignment mechanism is said to be *probabilistic* if each unit is assigned to treatment with probability

$$0 < p_i(X_i) < 1, \text{ for each possible } X_i \text{ for all } i = 1, \ldots, N$$

, so that no unit is treated or not treated with certainty. This is also called the *common support* assumption.

In sum, unconfoundedness imposes that the treatment is assigned to each i independently of the characteristics and potential outcomes of the other units, as well as the potential outcomes of i itself. Violations of the former imply, among other things, network/spillover effects, while violations of the latter imply selection into treatment.

5.2 Interactive Regression Model

Chernozhukov et al. (2018a) applies the general DML theory to four different models, two of which use instrumental variables. Either of the remaining two models can be used in the present application: the Partially Linear Regression model (PLR) and the Interactive Regression Model (IRM). I focus on the IRM as (i) it subsumes PLR as a special case, and

(ii) it implies estimating a propensity score, which is useful to unpack and interpret the results. The only advantage of choosing PLR over IRM is efficiency and this is not a prime concern in the present application.

Given a binary treatment variable, the Interactive Regression Model estimates fully heterogeneous average treatment effects. Here, "fully heterogeneous" means that each of the two response curves —the potential outcomes as a function of $\mathbf{X}$—is allowed to be a different nonparametric function. Vectors $\mathbf{Y}$ and $\mathbf{D}$ are modelled such that

$$\mathbf{Y} = g_0(\mathbf{D}, \mathbf{X}) + \mathbf{U}, \quad \mathbb{E}(\mathbf{U} \mid \mathbf{X}, \mathbf{D}) = \mathbf{0}, \tag{2}$$

$$\mathbf{D} = m_0(\mathbf{X}) + \mathbf{V}, \qquad \mathbb{E}(\mathbf{V} \mid \mathbf{X}) = \mathbf{0}. \tag{3}$$

This model allows me to estimate not only the ATE, here written as:

$$\tau_{ATE} = \mathbb{E}\left[g_0(\mathbf{1}, \mathbf{X}) - g_0(\mathbf{0}, \mathbf{X})\right]$$

but also the ATT:

$$\tau_{ATT} = \mathbb{E}\left[g_0(\mathbf{1}, \mathbf{X}) - g_0(\mathbf{0}, \mathbf{X}) \mid \mathbf{D} = \mathbf{1}\right].$$

Variables $\mathbf{X}$ are allowed to affect both the outcome $\mathbf{Y}$ and the treatment $\mathbf{D}$, each via a different function—$m_0(\mathbf{X})$ and $g_0(\mathbf{X})$, respectively. In particular, $m_0(\mathbf{X})$ is the propensity score and as such, its estimate can be inspected and used to test assumptions such as the common support assumption. This can help dissipate the black-box-type concerns typical of ML applications.

Moreover, both $m_0(\mathbf{X})$ and $g_0(\mathbf{X})$ are unknown and likely to be highly non-linear in this application, where the decision to assign a unit to treatment is made via a mix of sharp rules, loose principles and discretion. Hence, the benefits of estimating them flexibly using ML methods (Athey et al., 2019, here random forests,) are potentially large.

Another useful property of DML is double robustness, which states that the estimator is consistent for τ_{ATE} and τ_{ATT} as long as either the outcome model (2) or the treatment model (3) are correctly specified. In the present application, the emphasis is put on the correct specification of the treatment model, as all the variables used to assign treatment are observed and included (see Section 5.4).

In my preferred specification, I add court-level fixed effects to Equations 2 and 3 using a within transformation. This is to model how cases are assigned to Legal Aid offices and thus control for potential variation in the performance of Legal Aid offices—given as-good-as-random case allocation within office.

5.3 Random Forests

This section briefly introduces random forests and regression trees, on which random forests are based. They are introduced as DML requires choosing an ML method and I use random forests in the present application. For an econometrics textbook introduction, see Hansen (2022) (which I follow closely), while for a more general-audience introduction, see Krzywinski and Altman (2017).

A regression tree is a nonparametric regression that relies on many step functions to approximate an unknown conditional expectation function. More formally, regression trees aim at estimating $m(x) = \mathbb{E}[y_i \mid \mathbf{X_k} = x]$ for each scalar $y_i \in \mathbf{Y}$ and vector $\mathbf{X_k} \in \mathbf{X}$, the $N \times K$ matrix of covariates.

The first step is to "grow the tree". This involves sequentially partitioning, or splitting, the sample into increasingly small subsamples, or *branches*. The sample is split sequentially and at least once per each covariate $\mathbf{X_k}$. At each iteration, the algorithm selects a covariate and estimates a threshold for that covariate, which will determine how the sample is split. The algorithm's goal is to generate subsamples that are maximally homogeneous in $\mathbf{Y}$ and heterogeneous across each other. In an example where $\mathbf{Y}$ is a categorical variable with three values, each subsample is generated so as to include as many observations as possible of a given category while including as few observations as possible of the other categories.

More formally, the splits are determined by estimating the following model using non-linear least squares:

$$\begin{aligned} \mathbf{Y} &= \mu_1 \mathbb{1}\{\mathbf{X}_k \leq \gamma\} + \mu_2 \mathbb{1}\{\mathbf{X}_k > \gamma\} + \mathbf{e} \\ \mathbb{E}[\mathbf{e} \mid \mathbf{X}] &= 0 \end{aligned}$$

where covariate index k and parameter γ are free parameters—the latter setting the threshold of the covariate at which the split occurs.

This initial procedure leads to a large tree with a small estimation bias and high overfitting. Overfitting is then addressed by the second step, the *pruning* step, where the number of subsamples is reduced based on an information criterion.

Regression trees are not very good at approximating smooth functions and are sensitive to small changes in the data. These issues are addressed by random forests, which have the further advantage of performing better out-of-sample. Random forests, introduced by (Breiman, 2001), work by model averaging, or *ensembling*, many different regression trees. Such trees are constructed by using random subsamples of the data and of the covariates—a form of bootstrapping. The use of random subsets of covariates is warranted because it removes the correlation between the bootstrapped regression trees, thus decreasing the variance of the bootstrapped conditional mean, $m(x)$. Notice that both regression trees and random forests, like most ML methods, aim at predicting rather than estimating. The insight of DML is that by using such prediction algorithms once on the outcome model and once on the treatment model we can consistently estimate treatment effects. An important role is also played by (i) a procedure called *cross-fitting* and (ii) the use of Neyman-orthogonal moments, the combination of which removes the regularisation bias and overfitting that otherwise make ML estimators unhelpful in estimation settings.

5.4 Identification

Observational studies in Economics rarely require careful thinking about the treatment assignment mechanism. The present study is different in this regard, and its focus on treatment assignment should not be intended as a limitation. As Angrist and Pischke (2009) point out: "the propensity score rightly focuses researcher attention on models for treatment assignment [...]. This view seems especially compelling when treatment assignment is the product of human institutions or government regulations, while the process determining outcomes is more anonymous".

Moreover, while comparing randomised control trials and observational studies, Rosenbaum and Rubin (1983) note: "[...] in a randomized trial, x [in $e(x)$, the propensity score] is known to include all covariates that are both used to assign treatments and possibly related to the [potential outcomes]. This condition is usually not known to hold in a nonrandomized experiment [i.e. observational studies]". They also point out is that in RCTs the propensity is a known function while in observational studies it needs to be estimated.

However, thanks to rich administrative data and a DML approach, I am able to credibly fill the gap between RCTs and observational studies. On the one hand, my dataset includes all the variables used to determine assignment to treatment, which likely affects the outcomes too. They codify all the information an applicant has to write down in the Legal Aid application form. On the other hand, the Random Forests algorithm allows me to flexibly estimate the propensity score without imposing strong functional form assumptions. The result is likely to be a good approximation of the true propensity score.

In other words, studies that rely on propensity score estimation to achieve identification are vulnerable to omitted variable bias (OVB). This can be either due to the confounding effect of (i) variables that affect both outcome and treatment but are not included in the regression(s), and (ii) functional-form misspecification of $m_0(\mathbf{X})$ and $g_0(\mathbf{X})$. Point (i) is addressed by the rich dataset, as explained above, while point (ii) is addressed via the use of random forests—which allows estimating functions far more complex than the standard linear models.

6 Results

I present the results starting with the estimates on the effect of refusing (or terminating) aid on court outcome. First, I take a naive approach that makes use of the whole sample at the expense of some potential bias (Table 3). Then, I test the robustness of these results and refine them, leading to the preferred specification (Table 7). Because defendants in serious crime cases are in practice protected by the right to counsel in Australia, 90% of those who are denied of terminated aid have a private lawyer. This should be kept in mind when interpreting the results. Finally, I re-estimate the model after dropping those cases where the defendant is self-represented, allowing me to compare the performance of public representation against the counterfactual private representation[8].

6.1 Naive estimates

Table 3 reports the ATT estimates for the naive specification. ATE estimates and charge-level results are reported in Appendix D. Table 3's specification includes, in the control group, cases for which legal aid was granted and where the defendant was represented either by an in-house or panel lawyer—i.e. either by a Legal Aid lawyer, or by a publicly-funded private lawyer. In the treatment group, it includes cases for which aid was either denied or terminated, and where the defendant was either represented by a private lawyer or self-represented. It is naive in the sense that panel lawyers introduce an element of selection, as

[8] *Counterfactual private representation* is to be intended as the kind of private representation a defendant would have gotten if they were denied aid

Table 3. DML: ATT of Not Receiving Aid on Court Outcomes, allowing for heterogenity

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.034	0.008	-0.051	-0.018	0.0000	34668
Guilty share	0.016	0.006	0.004	0.028	0.0092	34600
Reduced seriousness	-0.012	0.005	-0.022	-0.002	0.0243	30172
Incarcerated - extensive	-0.115	0.008	-0.130	-0.100	0.0000	34665
Incarceration (mth.)	-4.169	0.677	-5.495	-2.842	0.0000	34668
Incarceration - intensive (mth.)	2.917	1.368	0.235	5.599	0.0330	18844
Guilty plea to highest charge	-0.037	0.008	-0.052	-0.021	0.0000	34668
Fined	0.084	0.006	0.073	0.095	0.0000	34655
Fine (AUD)	215.299	248.443	-271.640	702.238	0.3862	2689

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to the whole sample. The chosen ML method is Random Forests.

they can choose which cases to run, and terminated cases introduce a source of endogeneity in the treatment group, as these applications might have been terminated because of the applicant's misconduct.

Starting with the two most important outcomes, I find that denying or terminating aid decreases the chances of incarceration by 11.5 p.p. and time in jail by 4.2 months. As I show below, the sign of these estimates is robust across multiple specifications. Other robust findings are that denying or terminating aid reduces the chances of pleading guilty by 3.6 p.p. and increase the probability of being fined by 8.3 p.p. These estimates make sense given most treated defendants are privately represented and provide preliminary evidence that a penalty for relying on legal aid exists.

Table 3 also shows that denying or terminating aid: reduces the chances of some charges being dropped of withdraw; leads to an increase in the share of charges for which the defendant is found guilty; and reduces the defendant's chances to have their most serious charge dropped or withdrawn. However, I show below that these findings are not robust.

To explore the sensitivity of these estimates, I start by plotting their associated propensity score in Figure 1. Indeed, Chernozhukov et al. (2018a)'s Interactive Regression Model involves the estimation of a propensity score—here via random forests. We can visually inspect it—I break it down by treatment group—to check that the common support assumption is reasonable and to explore whether there are areas where the two propensity score densities do not overlap. As Figure 1 shows, no propensity score is equal to one or zero, providing evidence in favour of the common support assumption. However, the high density around [0.05, 0.075] interval warrants caution, as the estimator assigns high weight to these observations. Moreover, while there is a good degree of overlap between the propensity scores of the treated and control groups, the treated group's density has a long right tail where it has little to no overlap with the corresponding density for the control group. While this is simply reflecting the imbalance between treatment groups, it is important to investigate how sensitive the results are to the exclusion of this tail.

6.1.1 Sensitivity analysis

In this section, I run the DML analysis in different subsamples to investigate whether my estimates are generated by particular features of the data. I find that both the signs and the magnitudes of my estimated treatment effects are remarkably stable for some outcomes (*reduced charges*, *incarceration*, *incarcerated*, *guilty plea to the highest charge* and *fined*), but not others (*guilty share*, *reduced seriousness*).

One concern around the naive estimates is that they might be driven by the right tail of the propensity score for the treatment group. To explore this, I first restrict my sample to applicants whose income is 600 AUD per week or less (Table 4), as the treatment propensity is a function of income (see Figure 2). The stability of the estimates should reassure that they are not driven by individuals with high incomes, who are hence likely to be denied aid (i.e. high treatment propensity) and hire high-performance lawyers.

That same concern is also addressed by directly dropping cases that are part of that tail, here defined as those with a treatment propensity higher than 0.34. The results are again stable, providing evidence that they are not a product of cases with low overlap (see Table 5). This approach does not produce valid standard error though, due to selection issues. A more formal and standard approach is also taken, where the propensity score is trimmed outside the $[0.1, 0.9]$ interval. The estimates remain similar, as shown in Table 6.

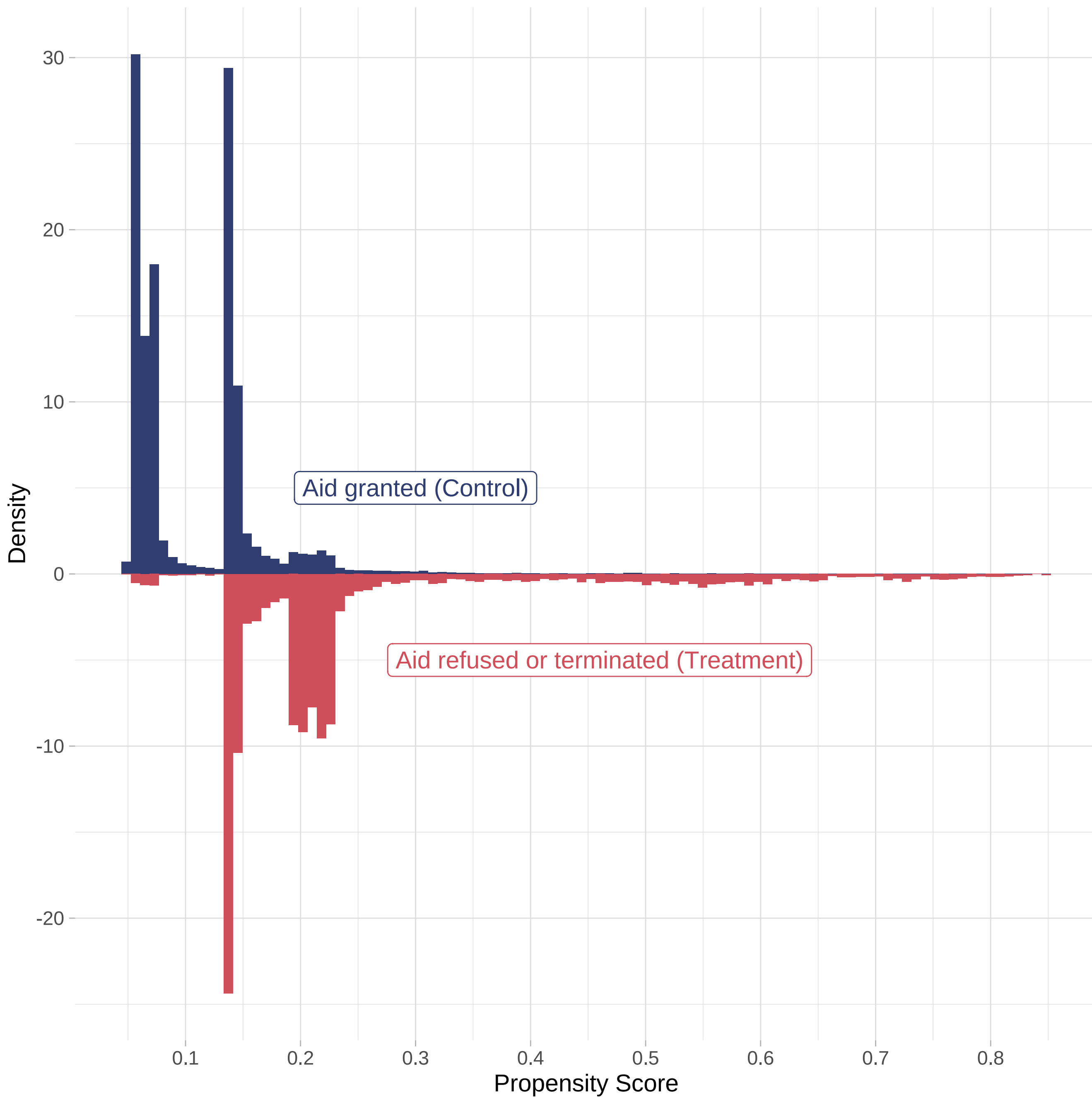


FIGURE 1. PROPENSITY SCORE

Notes. This figure plots the density of propensity score of the naive DML specification by treatment group. The propensity score is calculated using the random forests algorithm.

Table 4. DML: ATT of Not Receiving Aid on Court Outcomes when income 600$ p.w. or less

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.04	0.01	-0.05	-0.02	0.00	34045
Guilty share	0.02	0.01	0.00	0.03	0.02	33977
Reduced seriousness	-0.01	0.01	-0.02	0.00	0.03	29685
Incarcerated	-0.12	0.01	-0.13	-0.10	0.00	34042
Incarceration (months)	-4.08	0.74	-5.52	-2.64	0.00	34045
Incarceration - intensive (mth.)	3.00	1.44	0.18	5.83	0.04	18694
Guilty plea to highest charge	-0.04	0.01	-0.06	-0.03	0.00	34045
Fined	0.08	0.01	0.07	0.09	0.00	34032
Fine (AUD)	296.76	289.53	-270.72	864.23	0.31	2597

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the defendant earns a net assessable income of 600$ per week or less. The chosen ML method is Random Forests.

Table 5. DML: ATT of Not Receiving Aid on Court Outcomes when propensity score is 0.34 or less

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.039	0.009	-0.057	-0.022	0.0000	33631
Guilty share	0.014	0.006	0.002	0.027	0.0283	33563
Reduced seriousness	-0.010	0.006	-0.021	0.001	0.0638	29375
Incarceration (months)	-3.920	0.739	-5.369	-2.471	0.0000	33631
Incarcerated	-0.121	0.008	-0.137	-0.104	0.0000	33628
Guilty plea to highest charge	-0.050	0.009	-0.067	-0.033	0.0000	33631
Fined	0.085	0.006	0.072	0.097	0.0000	33618

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the propensity score is lower or equal to 0.34. This is the interval of the propensity score with stronger overlap. The chosen ML method is Random Forests.

Table 6. DML: ATT of Not Receiving Aid on Court Outcomes when propensity score is between 0.1 and 0.9

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.031	0.009	-0.048	-0.013	0.0005	18384
Guilty share	0.012	0.007	-0.001	0.025	0.0656	18336
Reduced seriousness	-0.010	0.006	-0.021	0.001	0.0759	16099
Incarceration (months)	-2.553	0.709	-3.943	-1.164	0.0003	18384
Incarcerated	-0.100	0.008	-0.115	-0.084	0.0000	18382
Guilty plea to highest charge	-0.037	0.008	-0.053	-0.020	0.0000	18384
Fined	0.076	0.006	0.064	0.087	0.0000	18375

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the propensity score is between 0.1 and 0.9. This is reduce the disproportionate impact of high propensity score weights. The chosen ML method is Random Forests.

6.2 Preferred estimates

The preferred estimates are reported in Table 7 and refine the above results by addressing two key issues. First, terminated cases could be biasing the estimates. For instance, public lawyers could be dropping misbehaving clients, who would then have to find a private lawyer or self-represent, biasing the outcomes of the treated group downwards. Second, while case assignment is as-good-as-random among in-house lawyers within the same office, panel lawyers can self-select into particular cases based on their expertise, hence introducing unmodelled heterogeneity.

I address these issues by re-running the DML analysis after dropping cases (i) where the application for aid was initially granted, but then terminated, and (ii) where defendants who were granted aid were represented by panel lawyers. I also introduce court-level fixed effects via a within transformation to model the random assignment of cases to legal aid lawyers at the office level.

The effects on "reduced charges" and "incarceration" reduce in magnitude to -1.2 p.p. and -0.8 months, respectively, and become statistically insignificant. The treatment effects on "guilty share" and "reduced seriousness" remain precisely null—i.e. statistically insignificant and with small standard errors. The other coefficients remain statistically significant and consistent in sign and magnitude with the naive estimates. Denying legal aid to a defendant

Table 7. DML: ATT of Not Receiving Aid on Court Outcomes in a sample without terminated grants nor panel lawyers

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.012	0.013	-0.039	0.014	0.3676	16854
Guilty share	0.002	0.008	-0.014	0.017	0.8424	16807
Reduced seriousness	-0.007	0.005	-0.017	0.004	0.2134	14875
Incarceration (months)	-0.524	0.918	-2.323	1.275	0.5681	16854
Incarcerated	-0.088	0.013	-0.114	-0.063	0.0000	16852
Guilty plea to highest charge	-0.044	0.011	-0.065	-0.022	0.0001	16854
Fined	0.054	0.008	0.038	0.070	0.0000	16846
Fine (AUD)	-39.896	63.610	-164.569	84.777	0.5305	1464

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to all cases where aid was not terminated and where the defendant was not represented by a panel lawyer. The chosen ML method is Random Forests. Finally, court-level fixed effects are included via a within transformation.

Table 8. DML: ATT of Not Receiving Aid on Court Outcomes in a sample without terminated grants nor panel lawyers

	Coef.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.023	0.014	-0.050	0.004	0.0964	16339
Guilty share	0.005	0.008	-0.010	0.020	0.5088	16337
Reduced seriousness	-0.011	0.005	-0.021	-0.001	0.0278	14416
Incarceration (months)	-0.868	0.968	-2.765	1.029	0.3700	16339
Incarcerated	-0.105	0.017	-0.140	-0.071	0.0000	16338
Guilty plea to highest charge	-0.048	0.009	-0.066	-0.031	0.0000	16339
Fined	0.071	0.010	0.050	0.091	0.0000	16338
Fine (AUD)	-88.413	65.332	-216.462	39.636	0.1760	1272

Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to all cases where aid was not terminated and where the defendant was not represented by a panel lawyer, nor self-represented. The chosen ML method is Random Forests. Finally, court-level fixed effects are included via a within transformation.

Table 9. DML: ATE of Not Receiving Aid on Court Outcomes in a sample without terminated grants nor panel lawyers

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.020	0.012	-0.043	0.003	0.0887	16339
Guilty share	-0.004	0.007	-0.017	0.009	0.5617	16337
Reduced seriousness	-0.009	0.004	-0.017	-0.002	0.0141	14416
Incarceration (months)	-1.950	0.815	-3.548	-0.352	0.0168	16339
Incarcerated	-0.133	0.011	-0.156	-0.111	0.0000	16338
Guilty plea to highest charge	-0.067	0.008	-0.082	-0.051	0.0000	16339
Fined	0.075	0.010	0.056	0.095	0.0000	16338
Fine (AUD)	4.931	73.787	-139.688	149.550	0.9467	1272

Notes: The table presents Double Machine Learning ATE estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to all cases where aid was not terminated and where the defendant was not represented by a panel lawyer nor self-represented. The chosen ML method is Random Forests. Finally, court-level fixed effects are included via a within transformation.

reduces their chances of being incarcerated by -9 p.p. It decreases their probability of pleading guilty to the highest charge 4.7 p.p. and increases their chances of being fined it by 5.4 p.p.

6.3 Private vs. Public

Once I drop self-represented defendants (circa 500 cases), I can focus on the effect of hiring a private lawyer, where being represented by a public lawyer is the baseline. Here, it is useful to note that the kind of private lawyer hired by someone who had a positive chance to qualify for aid would not be an expensive one (see *Weekly gross income* in Table 1), but likely one cheaper than average.

The treatment effect estimates for the effect of hiring a private lawyer are consistent with and slightly larger than those for the effect of refusing aid (Section 6.2). The effect on "reduced seriousness" becomes significant but is small, with private lawyers making their client more likely to be found guilty of their most serious charge by 1.2 p.p. The defendant is also more likely to be fined by 7.1 p.p. Nonetheless, privately represented defendants are 10.2 p.p. less likely to be incarcerated and 5.1 p.p. less likely to plead guilty of their most serious charge.

Comparing ATTs and ATEs in Tables 8 and 9, respectively, it should be noted that the average treatment effects are larger than the average treatment effects on the treated, implying that the untreated population—i.e. those who were granted aid—might reap larger benefits from private lawyers than those who can afford them.

7 Discussion

Australian law protects the right to counsel, meaning that those who are denied aid can typically afford a private lawyer—this is precisely why they are denied aid. On the one hand, I find that such private lawyers provide their clients with the benefit of lowering their chances of incarceration by about 10 p.p. The high-stake nature of these trials for serious crimes, where the mean (positive) incarceration length is 46.5 months—almost 4 years —makes this 10 p.p. effect particularly consequential. On the other hand, the effect on incarceration length is small and not significant. This suggests that, for serious charges, private lawyers are better at avoiding jail for their clients, rather than lowering their jail time.

Looking at the impacts on other outcomes, the negative effect (-5 p.p.) of private representation on the probability of pleading guilty could be explained by the pressure on public lawyers to move on to the next case and/or by the incentive of private lawyers to extend the duration of the case. Moreover, it is not surprising that private representation cause an increase in fines, as the judge would see it as a sign that that defendant can pay for them. Finally, while statistically significant, the importance of the findings that private lawyers are less likely to lead to a drop/withdrawal of charges, and increase in the probability to be found guilty the most serious charge, should not be overstated. These are small coefficients and are secondary compared to outcomes related to incarceration.

7.1 Channels

There are two main mechanisms that could be driving the 10 p.p. incarceration gap: self-selection and workload. According to the self-selection channel, the greater financial rewards of privately-funded work could be attracting the more experienced lawyers and barristers. According to the workload channel, the heavy workload of Legal Aid (in-house) legal practitioners could be forcing them to reduce the time they spend on each case, thus impacting their clients' outcomes in court. Both channels are supported by the qualitative and quantitative evidence gathered in a recent report commissioned by National Legal Aid (Millane et al., 2023). This highlights both the increasing gap in remuneration between privately- and publicly-funded cases and the heavy workload that legal aid staff has been facing over the years.

While the dataset used in the present study does not allow me to disentangle the two channels[9], they are both likely to be driving the gap and are both rooted in a funding shortage. Indeed, funding has declined by 3 per cent per capita in the last decade, and the gap between the market rate and the rate paid by Legal Aid to private lawyers has only widened (Millane et al., 2023). This goes against the recommendations of a federal Productivity Commission, which identified this issue almost a decade ago (Productivity Commission, 2014)

The pay-gap issue extends to (in-house) public lawyers and barristers. Indeed, data from the 2008/09 financial year show that public and private lawyers earned, respectively, A$86,700 and A$94,800 on average (Forell et al., 2010), while public and private barristers earned, respectively, A$106,184 (Forell et al., 2010) and A$361,850 (ABS, 2008)[10] on average.

As mentioned, the workload that needs to be managed by public lawyers is also likely to be playing a role. Anecdotal evidence from Millane et al. (2023) and Jacobsen (2011) supports the view that such heavy workload represents an issue. This could explain why Legal Aid clients are 5 p.p. more likely to plead guilty, as guilty pleas are an efficient solution in cases where the probability of an incarceration is high. Moreover, no applicant was denied aid due to a lack of funding in the present study sample. This is consistent with the idea that public lawyers are responding to the scarcity of funding by prioritising access to legal aid over the quality of the service. Ultimately, only an increase in funding would allow for an increase in quality whitout compromising the current level of access.

8 Conclusion

Legal aid to indigent defendants is omnipresent in developed countries. It is commonly justified on social justice grounds, but this should not preclude a quantitative investiga-

[9]In particular, the lack of data on Aboriginal and Torres Strait Islander cases and on the number of legal practitioners working for Legal Aid prevents me from analysing how outcomes correlate with workload, hence disentangling this channel from the self-selection channel.

[10]Estimate calculated from ABS data dividing the income generated by private barristers in the 2007-08 financial year (A$1.4b) by their number in the same period (3,869 private barristers). Public barristers are to be intended as senior public lawyers acting as barristers.

tion. Thus far, it has remained largely unstudied. Measuring its impact is important from an accountability point of view —i.e. how public defence compares to private defence, or even self-representation—but also to understand its broader role within the criminal justice system.

This paper makes the first moves in filling this gap, studying the effect of refusing legal aid on court outcomes. Thanks to rich administrative data from New South Wales, Australia, a double machine learning approach can credibly learn the unknown treatment assignment function and identify average treatment effects and average treatment effects on the treated.

I find that, since most defendants who are denied aid hire a private lawyer, denying aid reduces a defendant's probability of incarceration by about 9 p.p., which goes up to 10 p.p. after excluding the self-represented. These gaps are particularly consequential given that defendants in the study population who end up in jail are sentenced to almost 4 years on average. These estimates provide evidence that a performance gap exists between private and public legal representation—at least when representing legal aid applicants.

Beyond the magnitude and sign of these estimates, such a DML approach to impact evaluation gives policy-makers a new tool to study the effects of public programs, either substituting traditional econometric methods when their assumptions are not credible, or complementing them. This promotes accountability and can open new policy questions. In this case: how large should the performance gap between public and private representation, if any?

References

ABS (2008). 8667.0 - Legal Services, Australia, 2007-08. Technical report. Accessed 2 November 2023. [Cited on pages 6 and 29.]

ABS (2016). Australian standard geographical classification (asgc). *Statistical Geography*, 1. [Cited on page 40.]

Agan, A., Freedman, M., and Owens, E. (2021). Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense. *The Review of Economics and Statistics*, 103(2):294–309. [Cited on page 3.]

Anderson, J. M. and Heaton, P. (2012). How much difference does the lawyer make: The effect of defense counsel on murder case outcomes. *Yale Law Journal*, 122(1):154–217. [Cited on pages 3 and 4.]

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press. [Cited on page 19.]

Angrist, J. D. and Rokkanen, M. (2015). Wanna get away? regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association*, 110(512):1331–1344. [Cited on pages 42 and 46.]

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148 – 1178. [Cited on page 17.]

Barreca, A. I., Guldi, M., Lindo, J. M., and Waddell, G. R. (2011). Saving Babies? Revisiting the effect of very low birth weight classification*. *The Quarterly Journal of Economics*, 126(4):2117–2123. [Cited on pages 39 and 40.]

Bertrand, M., Crépon, B., Marguerie, A., and Premand, P. (2021). Do workfare programs live up to their promises? experimental evidence from cote d'ivoire. Working Paper 28664, National Bureau of Economic Research. [Cited on page 4.]

Blandhol, C., Bonney, J., Mogstad, M., and Torgovitsky, A. (2022). When is tsls actually late? Working Paper 29709, National Bureau of Economic Research. [Cited on page 39.]

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32. [Cited on pages 2 and 19.]

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2019). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210. [Cited on page 37.]

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2022). *rdrobust: Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs*. R package version 2.1.1. [Cited on page 37.]

Canay, I. A. and Kamat, V. (2017). Approximate Permutation Tests and Induced Order Statistics in the Regression Discontinuity Design. *The Review of Economic Studies*, 85(3):1577–1608. [Cited on page 40.]

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018a). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68. [Cited on pages 2, 5, 12, 16, and 21.]

Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I. (2018b). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. Working Paper 24678, National Bureau of Economic Research. [Cited on page 4.]

Davis, J. M. and Heller, S. B. (2020). Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs. *The Review of Economics and Statistics*, 102(4):664–677. [Cited on page 4.]

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12):4178–4219. [Cited on page 4.]

Dietrich v. The Queen (1992). *F.C. No. 92/044*. `http://www8.austlii.edu.au/cgi-bin/viewdoc/au/cases/cth/HCA/1992/57.html`. [Cited on page 6.]

Forell, S., Cain, M., and Gray, A. (2010). Recruitment and retention of lawyers in regional, rural and remote new south wales. Working paper, Law and Justice Foundation of NSW, Sydney. [Cited on pages 5, 6, and 29.]

Hahn, J., Todd, P., and der Klaauw, W. V. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209. [Cited on page 37.]

Hansen, B. (2022). *Econometrics*. Princeton University Press. [Cited on page 18.]

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189. [Cited on page 3.]

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475. [Cited on page 37.]

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press. [Cited on pages 12 and 14.]

Ishihara, T. and Sawada, M. (2022). Manipulation-robust regression discontinuity designs. *arXiv preprint arXiv:2009.07551.* [Cited on page 38.]

Iyengar, R. (2007). An analysis of the performance of federal indigent defense counsel. Working Paper 13187, National Bureau of Economic Research. [Cited on pages 3 and 4.]

Jacobsen, G. (2011). Legal aid lawyers buckle under work stress. *Sydney Morning Herald.* `https://www.smh.com.au/national/nsw/legal-aid-lawyers-buckle-under-work-stress-20110905-1ju9v.html`. [Online; accessed 27-11-2023]. [Cited on page 29.]

Kline, P. (2011). Oaxaca-blinder as a reweighting estimator. *American Economic Review*, 101(3):532–37. [Cited on pages 42 and 46.]

Knaus, M. C. (2022). Double machine learning-based programme evaluation under unconfoundedness. *The Econometrics Journal*, 25(3):602–627. [Cited on page 4.]

Knaus, M. C., Lechner, M., and Strittmatter, A. (2022). Heterogeneous employment effects of job search programs. *Journal of Human Resources*, 57(2):597–636. [Cited on pages 4 and 5.]

Krzywinski, M. and Altman, N. (2017). Classification and regression trees. *Nature Methods*, 14(8):757–758. [Cited on page 18.]

Lee, A. J. (2021). Flat fee compensation, lawyer incentives, and case outcomes in indigent criminal defense. [Cited on pages 3 and 4.]

Legal Aid Commission Act (1979). Legal aid commission act 1979 no 78 (nsw). `https://legislation.nsw.gov.au/view/whole/html/inforce/current/act-1979-078`. [Cited on page 8.]

Legal Aid NSW (2017). Who we are. `https://www.legalaid.nsw.gov.au/about-us/who-we-are`. [Online; accessed 28-03-2023]. [Cited on page 6.]

Legal Aid NSW (2022). Policies in brief. `https://www.legalaid.nsw.gov.au/__data/assets/pdf_file/0008/7757/Policies-in-Brief.pdf`. [Brochure; accessed 28-03-2023]. [Cited on pages 7 and 35.]

Legal Aid NSW (2023). Policy online. `http://www.legalaid.nsw.gov.au/for-lawyers/policyonline`. [Online; accessed 28-03-2023]. [Cited on pages 7 and 8.]

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714. The regression discontinuity design: Theory and applications. [Cited on page 38.]

Millane, E., Jackson, A., and Blane, N. (2023). Justice on the brink: Stronger legal aid for a better legal system. https://www.nationallegalaid.org/wp-admin/admin-ajax.php?cdm-download-file-id=MTA3MHwyMDIzLTExLTIxIDE3OjE2OjIwfG5vdmVtYmVyLTIwMjMtanVzdGljZS1vbi10aGUtYnJpbmstZmluYW =Link. Report prepared by Impact Economics and Policy for National Legal Aid. [Online; accessed 27-11-2023]. [Cited on page 29.]

Porreca, Z. and McCannon, B. C. (2023). The right to counsel: Criminal prosecution in 19th century london. *Available at SSRN 4505535*. [Cited on page 4.]

Productivity Commission (2014). Access to justice arrangements: Productivity commission inquiry report, volume 2. https://www.pc.gov.au/inquiries/completed/access-justice/report/access-justice-volume2.pdf. [No. 72, 5 September 2014; Online; accessed 27-11-2023]. [Cited on page 29.]

Roach, M. A. (2014). Indigent Defense Counsel, Attorney Quality, and Defendant Outcomes. *American Law and Economics Review*, 16(2):577–619. [Cited on pages 3 and 4.]

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55. [Cited on pages 12, 14, and 19.]

Schwall, B. (2017). Getting what you paid for: how payment methods affect attorneys' behaviors. Working paper, Clemson University working paper. [Cited on pages 3 and 4.]

Shem-Tov, Y. (2022). Make or Buy? The Provision of Indigent Defense Services in the United States. *The Review of Economics and Statistics*, 104(4):819–827. [Cited on pages 3 and 4.]

The Parliament of Australia (2002). Criminal proceeds confiscation act 2002 - sect. 104. http://classic.austlii.edu.au/au/legis/qld/consol_act/cpca2002285/s104.html. [Cited on page 35.]

Thucydides (1874). *History of the Peloponnesian War*. Retrieved via Project Gutemberg. Translated by Richard Crawley (1874). [Cited on page 2.]

United Nations (1948). *Universal Declaration of Human Rights*. [Cited on page 2.]

United Nations (2012). United nations principles and guidelines on access to legal aid in criminal justice systems. *General Assembly resolution*. https://www.unodc.org/documents/justice-and-prison-reform/UN_principles_and_guidlines_on_access_to_legal_aid.pdf. [Cited on pages 2 and 6.]

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242. [Cited on page 4.]

A Other eligibility tests

The Jurisdiction test checks whether the applicant's matter is in a jurisdiction or area of law covered by Legal Aid. Focusing on Criminal Law, aid is available for most matters: Local Court, District Court, Supreme Court, Court of Criminal Appeal and High Court, Children's and Prisoner's matters. Aid is not available for cases where the applicant has been caught in the act of committing violent crimes, for Local Court hearings where there are no prospects of success, for cases related to monies and/or property used or intended to be used to commit a crime ("tainted", see The Parliament of Australia, 2002) to start trials in the Local Court, except for Apprehended Domestic Violence Orders trials.

The Merit test assesses what are the net benefits of providing or denying legal aid to the applicant, and whether the applicant has a reasonable chance to win the case. In practice, this test is used to deny aid for matters that are not serious and/or have no chance of success. Serious criminal matters automatically pass the Merit test. For other cases exempt from the Merit test, refer to (Legal Aid NSW, 2022, see).

B Formal and Informal Income

Legal Aid services can be accessed only under a 400$ per week income cutoff, which raises the question of whether and how income can be misreported. It is useful to distinguish between two types of income: formal and informal.

Informal income ("cash-in-hand") is not recorded in any official record, and even if it were, it could be easily misreported. Moreover, because informal income is not legal, it does not enter the Legal Aid application and the applicant has no incentive in disclosing a potential income of this type.

Instead, formal income is easy to track and verify, most notably through bank account statements. Hence, it is hard for the applicant to hide it. For this reason, I use formal income as the running variable in my FDR analysis, rather than total income. More precisely, I use "net assessable income", which is formal income net of a range of items such as the number of dependants and the welfare benefits received.

The choice of formal income as the running variable, instead of total income (formal plus informal), poses no threat to the FDR identification strategy. What matters is that there is *a* running variable and a discontinuity (plus all the assumptions detailed in Section 5). The main impact of choosing a running variable over another is on the definition of the population that the treatment effect estimate will target. Hence, here the focus is on the effect of having aid denied or terminated on a population with (net assessable) formal incomes close to 400 AUD per week.

C Fuzzy Regression Discontinuity

In this section, I present a fuzzy regression discontinuity analysis of the local average treatment effect of being denied aid on charge-level outcomes.

C.1 Empirical stretegy

As mentioned above, some applicants who do not pass the means test, earning $400 or more per week in formal income[11], but receive legal aid anyway by *discretion*. This implies that the threshold rule is probabilistic rather than deterministic, and hence that a fuzzy RD approach is warranted. I use a fuzzy regression discontinuity (FRD) design to estimate the effect of being represented by a private rather than public lawyer on the marginal applicant for the population at the cutoff, c. In this FRD, the running variable is formal income and the cutoff value is 400$ p.w.

[11]For a discussion on formal income as a running variable, rather than informal or total income, see Appendix B

Formally, the target causal parameter is defined as

$$\tau_{FRD} \equiv \lim_{e \to 0^+} \mathbb{E}\left[Y_i(1) - Y_i(0) \mid X_i = c, D_i(T_i = 1) + D_i(T_i = 0) = 1\right] \tag{4}$$

where $T_i = \mathbb{1}\left(X_i \geq c\right)$ is a dummy indicating the theoretical treatment assignment based on the cutoff rule, $D_i = \{0, 1\}$ is a dummy indicating actual treatment, $Y_i(D_i)$ are potential outcomes, X_i is the running variable and $D_i(T_i = 1) + D_i(T_i = 0) = 1$ is the conditioning event defining the *complier* group.

As shown in Hahn et al. (2001), the following Wald-type estimator identifies the causal parameter τ_{FDR},

$$\tau_{FRD} = \frac{E[Y_i \mid X_i = c^+] - E[Y_i \mid X_i = c^-]}{E[D_i \mid X_i = c^+] - E[D_i \mid X_i = c^-]} \tag{5}$$

where $E[Y_i \mid X_i = c^{\pm}] \equiv lim_{X \to c^{\pm}} E[Y_i \mid X_i = x]$. Parameter τ_{FRD} is a local average treatment effect (LATE; Imbens and Angrist, 1994), and this follows from the fact that Equation 5 is an instrumental variable estimator, where D_i instruments T_i, the numerator is the *reduced form* and the denominator is the *first stage*.

Estimator (5) is implemented by estimating both its numerator and denominator using local linear regression. This is equivalent to implementing a sharp regression discontinuity (SRD) design to study the effect of crossing the cutoff $X_i = c$ on Y_i (reduced form) and on D_i (first stage). The same estimation procedure applies to the first stage, but here I focus on the reduced form for illustrative purposes.

I follow the approach by Calonico et al. (2019, 2022) and choose an MSE-optimal bandwidth, h,[12] that allows for robust bias-corrected inference and point estimation. I run the same linear regression of Y_i on $X_i - c$ in the interval $X_i \in [c - h, c)$

$$\left(\widehat{\alpha}_\ell, \widehat{\beta}_\ell\right) = \underset{\alpha,\beta}{\arg\min} \sum_{i: X_i \in [c-h, c)} \left(Y_i - \alpha - \beta\left(X_i - c\right)\right)^2 K(X_i - c) \tag{6}$$

and in the interval $X_i \in [c, c + h]$:

$$\left(\widehat{\alpha}_r, \widehat{\beta}_r\right) = \underset{\alpha,\beta}{\arg\min} \sum_{i: X_i \in [c, c+h]} \left(Y_i - \alpha - \beta\left(X_i - c\right)\right)^2 K(X_i - c). \tag{7}$$

where $K(.)$ is a triangular kernel function. This results in the following sharp RD estimate

[12]More correctly, I allow the optimal bandwidth to vary on the two sides of the cutoff. The fiction of a single bandwidth is for ease of exposition.

for the effect of crossing the threshold on outcome Y_i:

$$\begin{aligned}\widehat{\tau}_{SRD} &\equiv E[Y_i \mid X_i = c^-] - E[Y_i \mid X_i = c^+] \\ &= \widehat{\alpha}_\ell - \widehat{\alpha}_r.\end{aligned} \tag{8}$$

C.1.1 Identification

The identification of local average treatment effects in fuzzy RDDs requires three main assumptions: continuity, monotonicity and exogeneity.

The continuity assumption states that the conditional expectation function (CEF) of the potential outcomes is continuous at the cutoff. Formally, it states that $\mathbb{E}\left[Y_i(D_i) \mid X_i = x\right]$ is continuous in X at c. This assumption can be invalidated by manipulations of the running variable (with the exception of noisy manipulation, as shown in Ishihara and Sawada, 2022). In the present paper's context, manipulation would occur if either the lawyer processing the application or the applicant were able to systematically misreport the true formal income (running variable), for instance underreporting it to give aid to those not passing the means test. The analyst would not be able to observe the true scores and determine which were manipulated. Manipulation is unlikely to be an issue here, as formal income is hard to doctor. While some applicants do not pass the means test and still receive aid, I can identify them. It is not due to misreporting of their earning, it is due to discretion—resulting in a fuzzy design.

I test the continuity assumption primarily relying on the McCrary (2008) test of discontinuity in the density function of the running variable around the cutoff. Under the auxiliary assumption of one-sided manipulation, this is the only valid test of the continuity assumption (Ishihara and Sawada, 2022). The one-sided manipulation assumption states that all noise-free (or *precise*) manipulation must be guided by either an incentive to treat or untreat but not both simultaneously—as can happen when manipulators differ in their motivations. While I also use balance tests and run sharp RDDs using pre-treatment covariates as outcomes, I do this while noting that balance tests are informative but may not be valid without passing the density test (see Ishihara and Sawada, 2022).

Adapted to the present context, where I define treatment as not getting aid, the monotonicity assumption formally states that there exists ε such that $D_i(c+e) \geq D_i(c-e)$ for all $0 < e < \varepsilon$. This means that, for any defendant, earning marginally more than 400$ p.w. income cutoff does not make them less likely to hire a private lawyer.

The exogeneity assumption, typical of IV estimation, is local in RDDs. It states that $(\tau_i, D_i(X_i))$ is jointly independent of X_i near c. This implies that, in the neighbourhood of c, not being eligible for legal aid (the instrument) affects court outcomes only via its effect on the probability of hiring a private lawyer (the treatment)—the exclusion restriction assumption. It is hard to imagine how a marginal difference in the weekly income of an applicant would affect anything other than the probability of receiving aid, or, equivalently, of hiring a private lawyer. This assumption also implies no manipulation, i.e. that a defendant's formal income (the running variable) does not depend on any potential knowledge of the treatment effects.

Treatment assignment is unconditionally locally random and hence the estimated treatment effects are not affected by the identification issues outlined in Blandhol et al. (2022). These affect identification in two-stage-least square regressions where the identification of the ATE relies on a conditionally random assignment to treatment.

C.1.2 Treatment and treatment group

For consistency with the RDD literature, I define treatment as being on the right side of the cutoff. In this FRD application, it implies that (i) the "theoretical" treatment assignment T_i is "having a net assessable income of more than 400 AUD p.w.", and (ii) actual treatment D_i is "being represented by a private lawyer". In my sample, defendants who hire a private lawyer were refused aid. This means that they applied for aid but it was either not granted or it was initially granted but later terminated before the end of the proceedings. I also consider an alternative treatment group, where I drop those applicants whose aid was eventually terminated and only keep those who were refused aid. This renders the treatment group more homogeneous.

C.1.3 Donut extension

As motivated in Section 6, I take a donut approach to the FRD, pioneered in Barreca et al. (2011). It implies stretching the regression discontinuity assumptions marginally further away from the cutoff. Intuitively, treated units are assumed to be assigned to treatment at random *after dropping those closest to the cutoff.* This approach is usually taken in contexts in which sorting into treatment cannot be excluded for units in a neighbourhood of the cutoff, but can be excluded for those just outside of it.

C.2 FRD Results

I start the regression discontinuity analysis with a test for the continuity of the density of the running variable, net assessable income. As visualised in Figure 3a, there is an abnormal concentration of defendants earning 400 AUD per week in net assessable income. However, Figure 4 shows that this is part of a general pattern where assessable income is abnormally frequent at each 50$ mark, as highlighted by the red vertical lines. Hence, such patterns seem to be exogenous to the provision of aid and due to rounding, rather than a manipulation of the running variable. Nonetheless, as a measure of prudence, I drop the observations within a 2 AUD radius from the 400 AUD cutoff, and run the analysis taking a donut hole approach, pioneered in Barreca et al. (2011). This restores the continuity of the running variable's density, as shown in Figure 3b. The p-value of the binomial tests, reported on Figures 3a and 3b and rounded to the third decimal, provide formal evidence of this visual intuition.

I estimate the effect of treatment on case characteristics as a balance test (Table 10). I include several defendant and case characteristics. "Year of birth" records the defendant's year of birth, "Female" is a binary variable equal to one if a defendant's sex is "female", and zero otherwise. "L.O.T.E." indicates whether a defendant's primary language is not English. The "Remoteness index" variable reports the Accessibility/Remoteness Index of Australia (ARIA+; see ABS, 2016), which is a continuous index ranging from 0 to 15, from high accessibility to high remoteness, respectively. Variables that qualify as defendant/case/charge characteristics that were available but either (i) are virtually constant across the whole sample (e.g. if a case is related to domestic violence) or that (ii) are used to compute the net assessable income, are not included in the balance test.

I find that, overall, charges are similar across the cutoff. "Female" is an exception, —there are more females eligible for aid than not eligible. Re-running the analysis including this covariate does not affect my main treatment effect estimates (not shown). Moreover, although graphical inspection of discontinuities in individual covariates is useful, recent papers have cautioned against sole reliance on tests of covariate-by-covariate equality of means in RD settings, as some could by random chance appear significant. Canay and Kamat (2017) noted that testing several individual hypothesis provides useful insights but can lead to "spurious rejections" due to multiple hypothesis testing.

Table 10. Balance Test

	Mean Below	Mean Above	Diff. in Means	Robust p-value	Bandwidth	N Above	N Below
Year of birth	1978.77	1975.43	-3.35	0.40	22.72	979	721
Female	0.05	0.20	0.15	0.03	34.06	1638	1208
L.O.T.E.	-0.01	0.01	0.01	0.73	26.60	1248	888
Remoteness index	1.44	1.28	-0.16	0.80	27.39	1156	747
Private firm submission	0.26	0.31	0.05	0.48	38.84	1885	1325
Submission year	2018.01	2017.22	-0.79	0.27	32.22	1524	1153

Notes: The table presents sharp RD estimates for case-level characteristics. For consistency with the rest of the analysis, I use charge-level data, and permforming the same balance tests using case-level data, clustering standard errors at the case level. Nonetheless, repeating this exercise with case-level data does not change the outcome of the hypothesis tests. The first column is the sharp RD effect for each variable. The second column reports p-values based on robust bias correction inference methods. The third reports the coverage-error-rate (CER) optimal bandwidths, while the last two show the effective sample sizes used to the left and to the right of the cutoff. Standard erros are clustered at the case level.

Table 11. First Stage: Impact of Overshooting Legal Aid's Eligibility Cut-off on Probability of Not Receiving It

$\hat{\tau}_{fs}$	95% Robust CI	Bandwidth	N Above	N Below
0.25** (0.08)	[0.1,0.42]	31.22	1513	1117

Notes: The table presents first stage estimates for all outcomes. The first column reports the effect on receiving legal aid on the marginal applicant at the cutoff. The second column reports 95% confidence intervals based on robust bias correction inference methods. The third reports the coverage-error-rate (CER) optimal bandwidths, while the last two show the effective sample sizes used to the left and to the right of the cutoff. The values in round are standard errors clustered at the case level.

C.2.1 LATE Estimates

The impact of overshooting the Legal Aid eligibility threshold on the share of Legal Aid applicants that do not get aid around the threshold is estimated at around +25 p.p. (see Figure 5 and Table 11) and the F-statistic is 61.9. This is evidence that the 400 AUD threshold indicator is a strong instrument for the refusal of aid (or its negation, legal aid provision).

The reduced form estimates, reported in Figure 12, provide evidence that overshooting the eligibility access for legal aid does not have a detectable impact on court outcomes, allowing us to rule out large effects. Figure 6 provides a visualisation of these results.

Predictably, statistically insignificant reduced form estimates lead to statistically insignif-

Table 12. Reduced Form: Impact of Overshooting the Legal Aid Eligibility on Court Outcomes

	$\hat{\tau}_{RF}$	95% Robust CI	Bandwidth	N Above	N Below
Guilty	0.05 (0.12)	[-0.19,0.28]	19.4	723	612
Guilty plea	0.02 (0.09)	[-0.15,0.21]	33.22	1372	1055
Incarcerated	0.02 (0.1)	[-0.17,0.21]	20.99	874	663

Notes: The table presents reduced form estimates for all outcomes. The first column reports the effect on receiving legal aid on the marginal applicant at the cutoff. The second column reports 95% confidence intervals based on robust bias correction inference methods. The third reports the coverage-error-rate (CER) optimal bandwidths, while the last two show the effective sample sizes used to the left and to the right of the cutoff. The values in round are robust t-statistics.

icant FRD estimates and large standard errors, as reported in Table 13.

C.2.2 LATE away from the cutoff

In this section I try to generalise the main estimates away from the cutoff while also increasing the sample size, following Angrist and Rokkanen (2015). I do this using the Oaxaca-Blinder linear reweighting estimator—as studied in Kline (2011). However, such generalisation is valid only if the conditional independence assumption holds, i.e. if net assessable income (the running variable) is orthogonal to a given court outcome, conditional on case characteristics. Angrist and Rokkanen (2015) study the testable implication of this assumption in two ways, which I follow.

The first, shown in Figure 7, is a plot of the residualised values of each outcome (after controlling for case characteristics) over the running variable. A flexible polynomial is then fit on these values. The test is passed if the fitted line is flat, except for a possible jump around the cutoff, which would be marking the presence of a treatment effect. Such a flat line is evidence of conditional orthogonality between the running variable and the outcome. Results are reported in Figure 7. Incarcerated is flat within a $50 bandwidth, while this is not the case (to the right of the cutoff) when expanding the bandwidth to the $100 and the $150 mark. Outcomes "guilty" and "guilty plea" remain fairly flat up to their $150

Table 13. Fuzzy RDD: Impact of Losing Legal Aid Eligibility on Court Outcomes

	$\hat{\tau}_{FDR}$	95% Robust CI	Bandwidth	N Above	N Below
Guilty	0.28 (0.71)	[-1.14,1.62]	19.78	723	612
Guilty plea	0.06 (0.31)	[-0.51,0.69]	33.53	1372	1055
Incarcerated	0.03 (0.45)	[-0.87,0.91]	21.73	951	691

Notes: The table presents fuzzy RD estimates for all outcomes. The first column reports the effect on receiving legal aid on the marginal applicant at the cutoff. The second column reports 95% confidence intervals based on robust bias correction inference methods. The third reports the coverage-error-rate (CER) optimal bandwidths, while the last two show the effective sample sizes used to the left and to the right of the cutoff. The values in round brackets are standard errors clustered at the case level.

bandwidths.

The second, shown in Table 14, is a formal test of the slope of each outcome on each side of the cutoff—the sample is split in two around the cutoff and the slopes are estimated. The test is passed if the slope coefficient is zero. Most coefficients are zero or close to zero, providing evidence that the conditional independence assumption largely holds.

Notwithstanding that these estimators target a different parameter from DML, it is interesting to notice the consistency in sign between these LATE estimates away from the cutoff (Table 15) and the preferred DML estimates (Table **??**). Indeed, the probability of being incarcerated is negative (although only significant at a 10% level) for compliers with net incomes between \$350 and \$450 and it is negative and significant for being guilty and for pleading guilty across the \$50, \$100 and \$150 bandwidths. While interesting, such estimates are rough extrapolations and should be considered more informative for their sign than for their point estimate.

Table 14. LATE Away From the Cutoff

	$D_i = 0$	$D_i = 1$
Guilty		
Win.= 50		

Net Income	-0.004***	-0.001
Std. Error	(0.001)	(0.001)
Obs.	2755	2075
Win.= 100		
Net Income	0.000	0.001**
Std. Error	(0.000)	(0.000)
Obs.	6451	3645
Win.= 150		
Net Income	-0.000	0.000
Std. Error	(0.000)	(0.000)
Obs.	13184	4356
Guilty plea		
Win.= 50		
Net Income	-0.003***	-0.002***
Std. Error	(0.001)	(0.001)
Obs.	2754	2075
Win.= 100		
Net Income	-0.000	-0.001***
Std. Error	(0.000)	(0.000)
Obs.	6390	3645
Win.= 150		
Net Income	-0.000***	-0.000
Std. Error	(0.000)	(0.000)
Obs.	13104	4356
Incarcerated		
Win.= 50		
Net Income	-0.001**	0.000
Std. Error	(0.001)	(0.000)
Obs.	2751	2075
Win.= 100		
Net Income	-0.000	0.001***
Std. Error	(0.000)	(0.000)
Obs.	6447	3645
Win.= 150		

Net Income	-0.000*	0.001***
Std. Error	(0.000)	(0.000)
Obs.	12760	4356

Notes: Table testing the conditional independence assumption away from the cutoff.

Table 15. LATE Away From the Cutoff

	50	100	150
First Stage	Guilty	Guilty	Guilty
Est.	0.058***	0.130***	0.154***
s.d.	(0.013)	(0.008)	(0.007)
Obs.−	2755	6451	13184
Obs.+	2075	3645	4356
Reduced Form	Guilty	Guilty	Guilty
Est.	-0.043***	-0.002	-0.018**
s.d.	(0.014)	(0.010)	(0.009)
Obs.−	2755	6451	13184
Obs.+	2075	3645	4356
FRD	Guilty	Guilty	Guilty
Est.	-0.730**	-0.018	-0.117**
s.d.	(0.344)	(0.079)	(0.059)
Obs.−	2755	6451	13184
Obs.+	2075	3645	4356
First Stage	Guilty plea	Guilty plea	Guilty plea
Est.	0.058***	0.128***	0.155***
s.d.	(0.012)	(0.008)	(0.007)
Obs.−	2754	6390	13104
Obs.+	2075	3645	4356
Reduced Form	Guilty plea	Guilty plea	Guilty plea
Est.	-0.018	-0.028***	-0.047***
s.d.	(0.015)	(0.010)	(0.008)
Obs.−	2754	6390	13104

Obs.+	2075	3645	4356
FRD	Guilty plea	Guilty plea	Guilty plea
Est.	-0.316	-0.216***	-0.307***
s.d.	(0.323)	(0.082)	(0.055)
Obs.−	2754	6390	13104
Obs.+	2075	3645	4356
First Stage	Incarcerated	Incarcerated	Incarcerated
Est.	0.058***	0.130***	0.151***
s.d.	(0.012)	(0.009)	(0.007)
Obs.−	2751	6447	12760
Obs.+	2075	3645	4356
Reduced Form	Incarcerated	Incarcerated	Incarcerated
Est.	-0.026**	0.004	0.007
s.d.	(0.011)	(0.007)	(0.006)
Obs.−	2751	6447	12760
Obs.+	2075	3645	4356
FRD	Incarcerated	Incarcerated	Incarcerated
Est.	-0.442*	0.030	0.045
s.d.	(0.230)	(0.057)	(0.043)
Obs.−	2751	6447	12760
Obs.+	2075	3645	4356

Notes: The table reports the first stage, reduced form LATE estimates away from the cutoff. The estimates are calculated using the Oaxaca-Blinder estimator as discussed in Kline (2011), adapted to a Fuzzy RDD framework following Angrist and Rokkanen (2015).

D Further estimates

In this section, I report Interactive Regression Model estimates not included in the main text.

Table 16. DML: Whole-sample Effect of Not Receiving Aid on Court Outcomes

	Coeff.	s.e.	l-CI	r-CI	p-value
Reduced charges	-0.016	0.009	-0.034	0.003	0.0909
Guilty share	0.003	0.007	-0.010	0.017	0.6183
Reduced seriousness	-0.006	0.006	-0.018	0.007	0.3739
Incarcerated - extensive	-0.111	0.009	-0.128	-0.093	0.0000
Incarceration (mth.)	-3.281	0.806	-4.860	-1.702	0.0000
Incarceration - intensive (mth.)	3.544	1.516	0.572	6.516	0.0194
Guilty plea to highest charge	-0.031	0.009	-0.049	-0.013	0.0006
Fined	0.061	0.006	0.050	0.073	0.0000
Fine (AUD)	522.195	455.704	-370.968	1415.358	0.2518

Notes: The table presents Double Machine Learning estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The PLM model is applied to the whole sample. The chosen ML method is Random Forests.

D.0.1 Partially Linear Model results

D.0.2 Other IRM results

Table 17. DML: Whole-sample Effect of Not Receiving Aid on Court Outcomes

	Coeff.	s.e.	l-CI	r-CI	p-value
Guilty	0.006	0.011	-0.016	0.027	0.6000
Convicted	-0.005	0.001	-0.007	-0.003	0.0000
Guilty plea	-0.039	0.011	-0.060	-0.018	0.0002
Incarcerated	-0.032	0.010	-0.052	-0.011	0.0026

Notes: The table presents Double Machine Learning estimates for the effect of not receiving aid (having aid denied or terminated) on charge-level court outcomes. The PLM model is applied to the whole sample. The chosen ML method is Random Forests. Standard errors are clustered at the case level using Chiang et al. (2021).

Table 18. DML: ATE of Not Receiving Aid on Court Outcomes, allowing for heterogenity

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.050	0.006	-0.062	-0.038	0.0000	34668
Guilty share	0.022	0.004	0.014	0.031	0.0000	34600
Reduced seriousness	-0.015	0.004	-0.022	-0.007	0.0001	30172
Incarcerated - extensive	-0.136	0.006	-0.147	-0.125	0.0000	34665
Incarceration (mth.)	-5.288	0.520	-6.307	-4.269	0.0000	34668
Incarceration - intensive (mth.)	2.150	0.993	0.204	4.097	0.0304	18844
Guilty plea to highest charge	-0.033	0.006	-0.044	-0.021	0.0000	34668
Fined	0.088	0.004	0.080	0.096	0.0000	34655
Fine (AUD)	316.080	202.236	-80.295	712.455	0.1181	2689

Notes: The table presents Double Machine Learning ATE estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to the whole sample. The chosen ML method is Random Forests.

Table 19. DML: Whole-sample Effect of Not Receiving Aid on Court Outcomes

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Guilty	0.005	0.008	-0.011	0.020	0.5665	202064
Convicted	-0.005	0.001	-0.007	-0.004	0.0000	202064
Guilty plea	-0.028	0.007	-0.043	-0.014	0.0001	202064
Incarcerated	-0.038	0.007	-0.052	-0.024	0.0000	201637

Notes: The table presents Double Machine Learning estimates for the effect of not receiving aid (having aid denied or terminated) on charge-level court outcomes. The irm model is applied to the whole sample. The chosen ML method is Random Forests. Standard errors are clustered at the case level using Chiang et al. (2021).

Table 20. DML: ATE of Not Receiving Aid on Court Outcomes in a sample without terminated grants nor panel lawyers

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Guilty	-0.013	0.013	-0.038	0.012	0.3253	98729
Guilty plea	-0.065	0.011	-0.088	-0.043	0.0000	98378
Incarcerated	-0.040	0.011	-0.061	-0.018	0.0003	98712

Notes: The table presents Double Machine Learning estimates for the effect of not receiving aid (having aid denied or terminated) on charge-level court outcomes. The IRM model is applied to all cases where aid was not terminated and where the defendant was not represented by a panel lawyer. The chosen ML method is Random Forests. Standard errors are clustered at the case level using Chiang et al. (2021).

Table 21. DML: ATE of Not Receiving Aid on Court Outcomes when income 600$ p.w. or less

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.05	0.01	-0.06	-0.04	0.00	34045
Guilty share	0.02	0.00	0.01	0.03	0.00	33977
Reduced seriousness	-0.01	0.00	-0.02	-0.01	0.00	29685
Incarcerated	-0.14	0.01	-0.15	-0.13	0.00	34042
Incarceration (months)	-5.35	0.53	-6.40	-4.31	0.00	34045
Incarceration - intensive (mth.)	2.21	1.01	0.23	4.18	0.03	18694
Guilty plea to highest charge	-0.04	0.01	-0.05	-0.02	0.00	34045
Fined	0.09	0.00	0.08	0.09	0.00	34032
Fine (AUD)	323.24	216.46	-101.02	747.49	0.14	2597

Notes: The table presents Double Machine Learning ATE estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the defendant earns a net assessable income of 600$ per week or less. The chosen ML method is Random Forests.

Table 22. DML: ATE of Not Receiving Aid on Court Outcomes when propensity score is 0.34 or less

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.051	0.006	-0.064	-0.039	0.0000	33631
Guilty share	0.022	0.005	0.013	0.031	0.0000	33563
Reduced seriousness	-0.014	0.004	-0.022	-0.007	0.0003	29375
Incarceration (months)	-5.313	0.547	-6.385	-4.241	0.0000	33631
Incarcerated	-0.136	0.006	-0.147	-0.124	0.0000	33628
Guilty plea to highest charge	-0.037	0.006	-0.049	-0.025	0.0000	33631
Fined	0.089	0.004	0.080	0.097	0.0000	33618
Fine (AUD)	428.341	264.886	-90.826	947.508	0.1059	2555

Notes: The table presents Double Machine Learning ATE estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the propensity score is lower or equal to 0.34. This is the interval of the propensity score with stronger overlap. The chosen ML method is Random Forests.

Table 23. DML: ATE of Not Receiving Aid on Court Outcomes when propensity score is between 0.1 and 0.9

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.030	0.008	-0.046	-0.014	0.0003	18384
Guilty share	0.005	0.006	-0.007	0.017	0.3956	18336
Reduced seriousness	-0.010	0.005	-0.020	0.000	0.0580	16099
Incarceration (months)	-3.190	0.622	-4.408	-1.972	0.0000	18384
Incarcerated	-0.129	0.008	-0.144	-0.114	0.0000	18382
Guilty plea to highest charge	-0.051	0.008	-0.067	-0.036	0.0000	18384
Fined	0.087	0.006	0.076	0.098	0.0000	18375
Fine (AUD)	257.823	218.640	-170.703	686.348	0.2383	1664

Notes: The table presents Double Machine Learning ATE estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to cases where the propensity score is between 0.1 and 0.9. This is reduce the disproportionate impact of high propensity score weights. The chosen ML method is Random Forests.

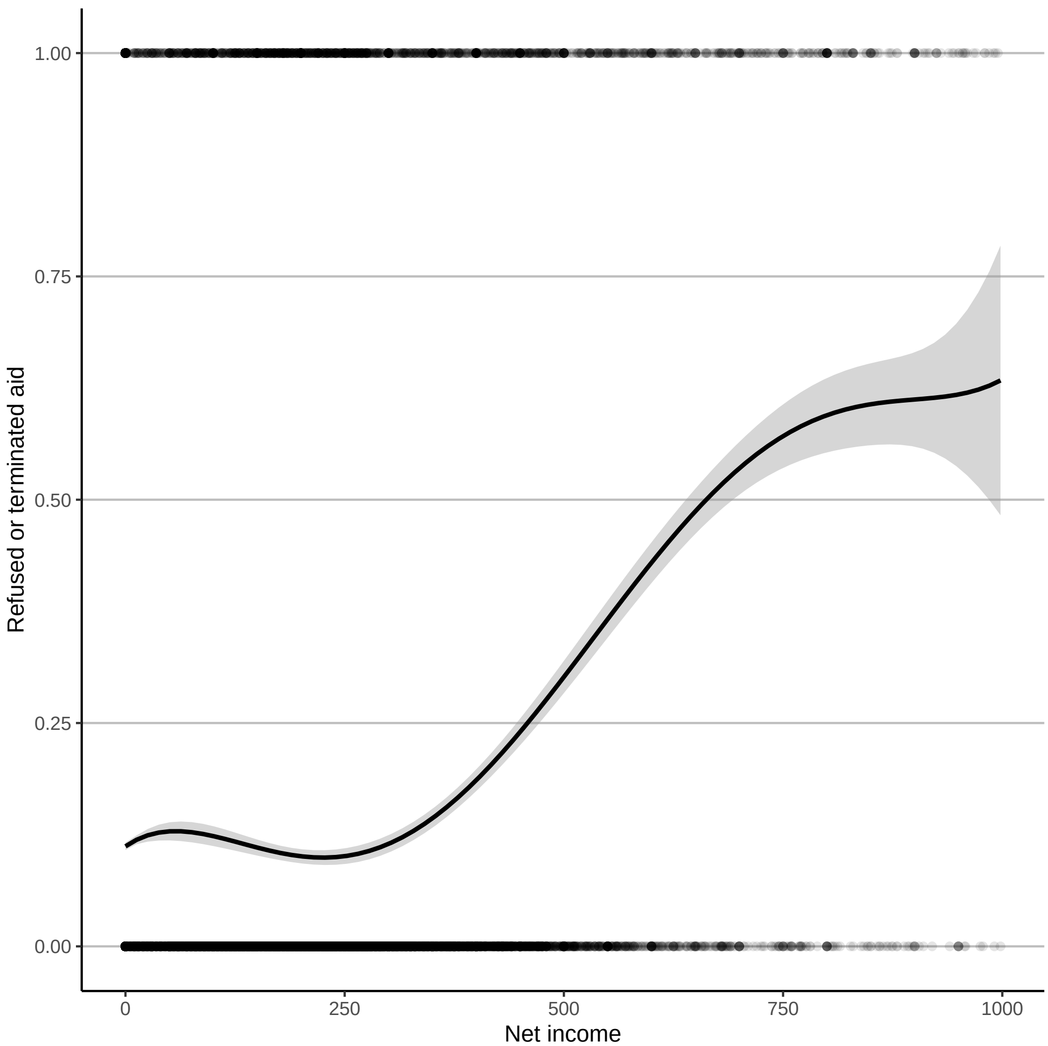
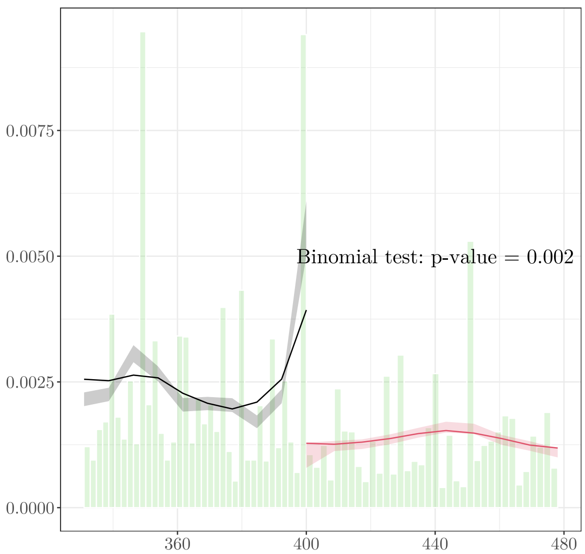
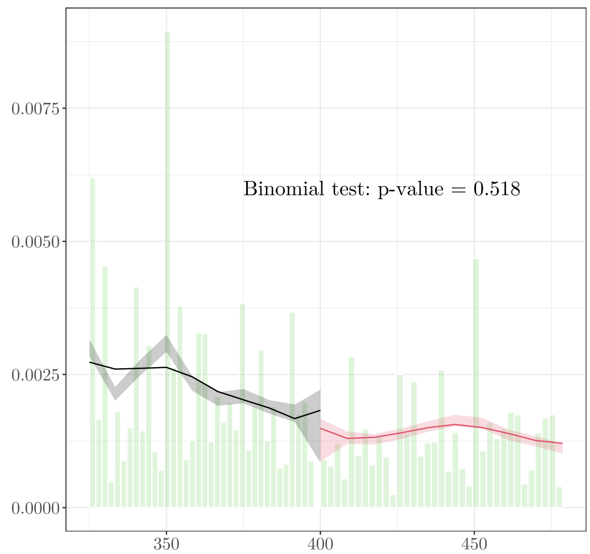


Figure 2. Treatment as a Function of Income

Notes. This figure shows a LOWESS curve fitting of the treatment, defined as refusing or terminating legal aid, over the income of the applicant net of welfare payments.

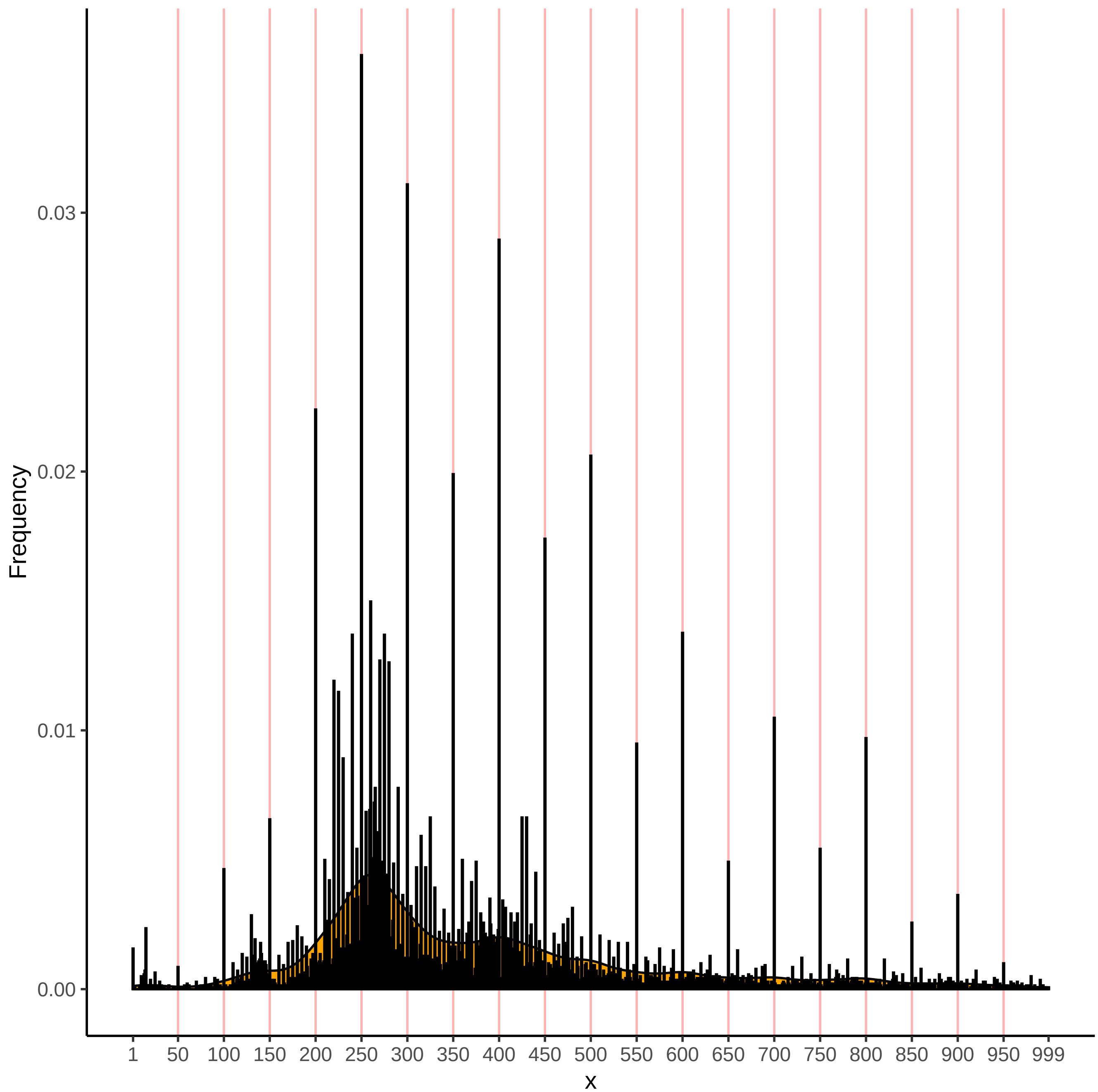


(a) Full-sample Density Test



(b) Donut-sample Density Test

Figure 3. Density Tests



h

FIGURE 4. HISTOGRAM AND DENSITY OF ASSESSABLE INCOME

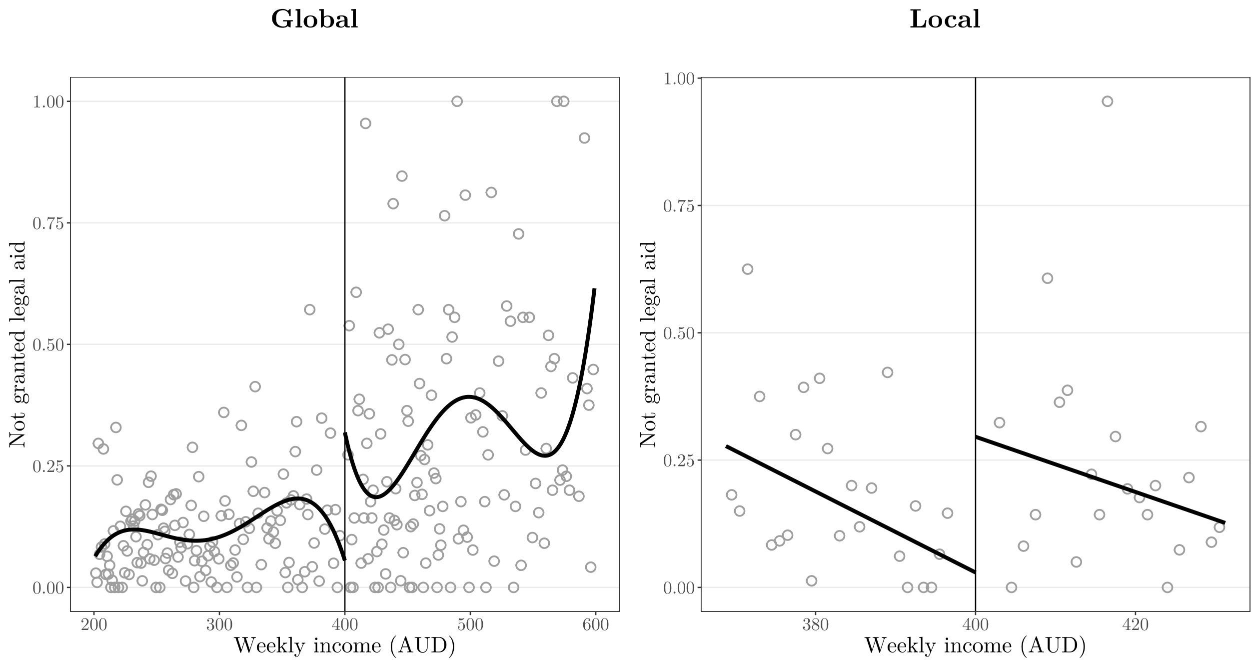


FIGURE 5. FIRST STAGE

Notes: The figure plots variance-mimicking binned values of an indicator variable for not receiving full aid (aid denied or terminated) over assessable net income. On the left, I include charges associated with applicants earning between 200-400 AUD and fit a polynomial of degree four for descriptive purposes. On the right, I restrict the sample to only include observations within the bandwidths used in the main analysis, and fit a polynomial of degree one (also as in the main analysis).

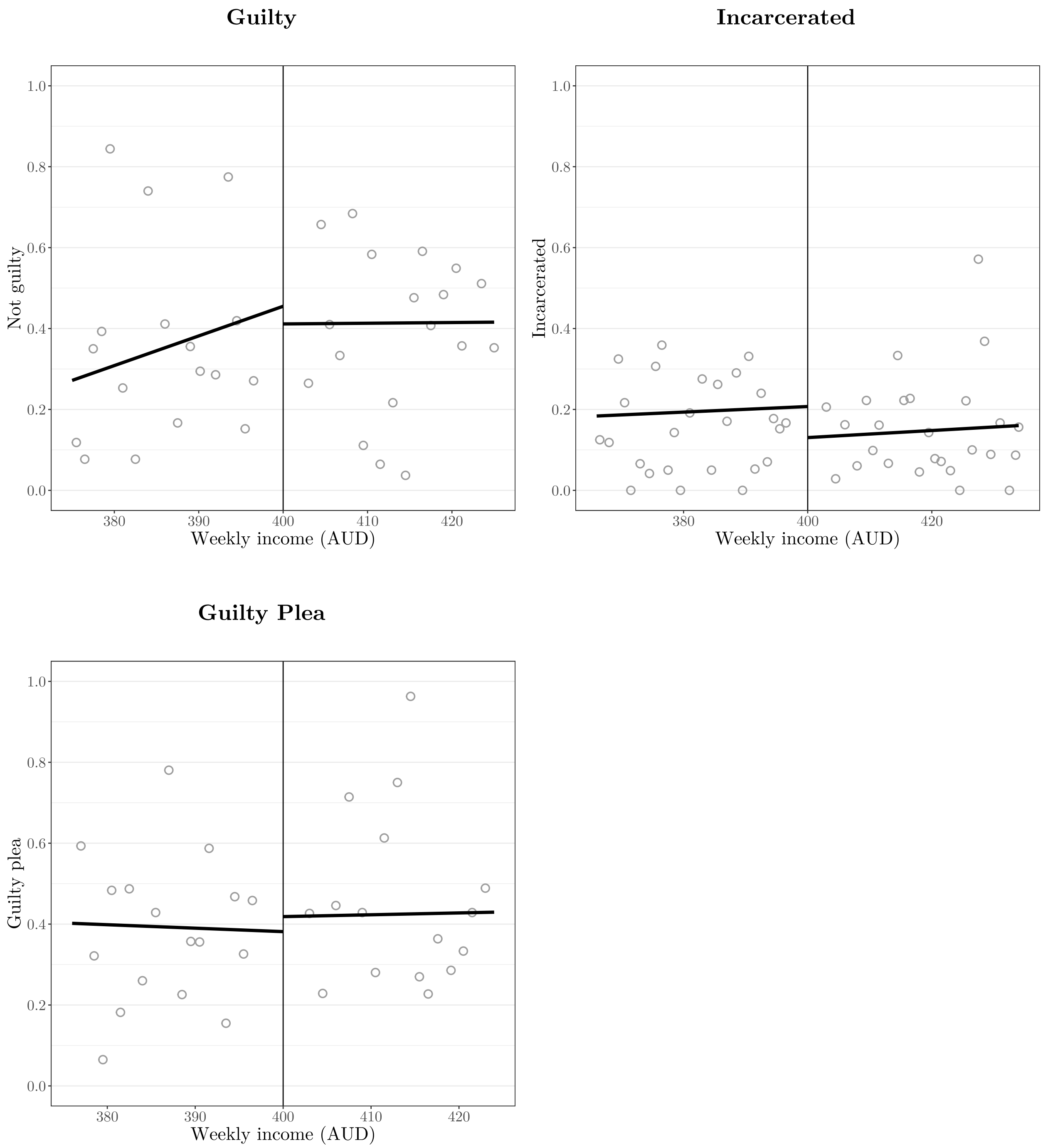


FIGURE 6. REDUCED FORM: EFFECT OF OVERSHOOTING LEGAL AID ELIGIBILITY ON COURT OUTCOMES

Notes: The figure plots variance-mimicking binned values of court outcomes over assessable net income, fitted with a polynomial of degree one. The bandwidths used are the same that are used to estimate the treatment effects in the FRD main specification.

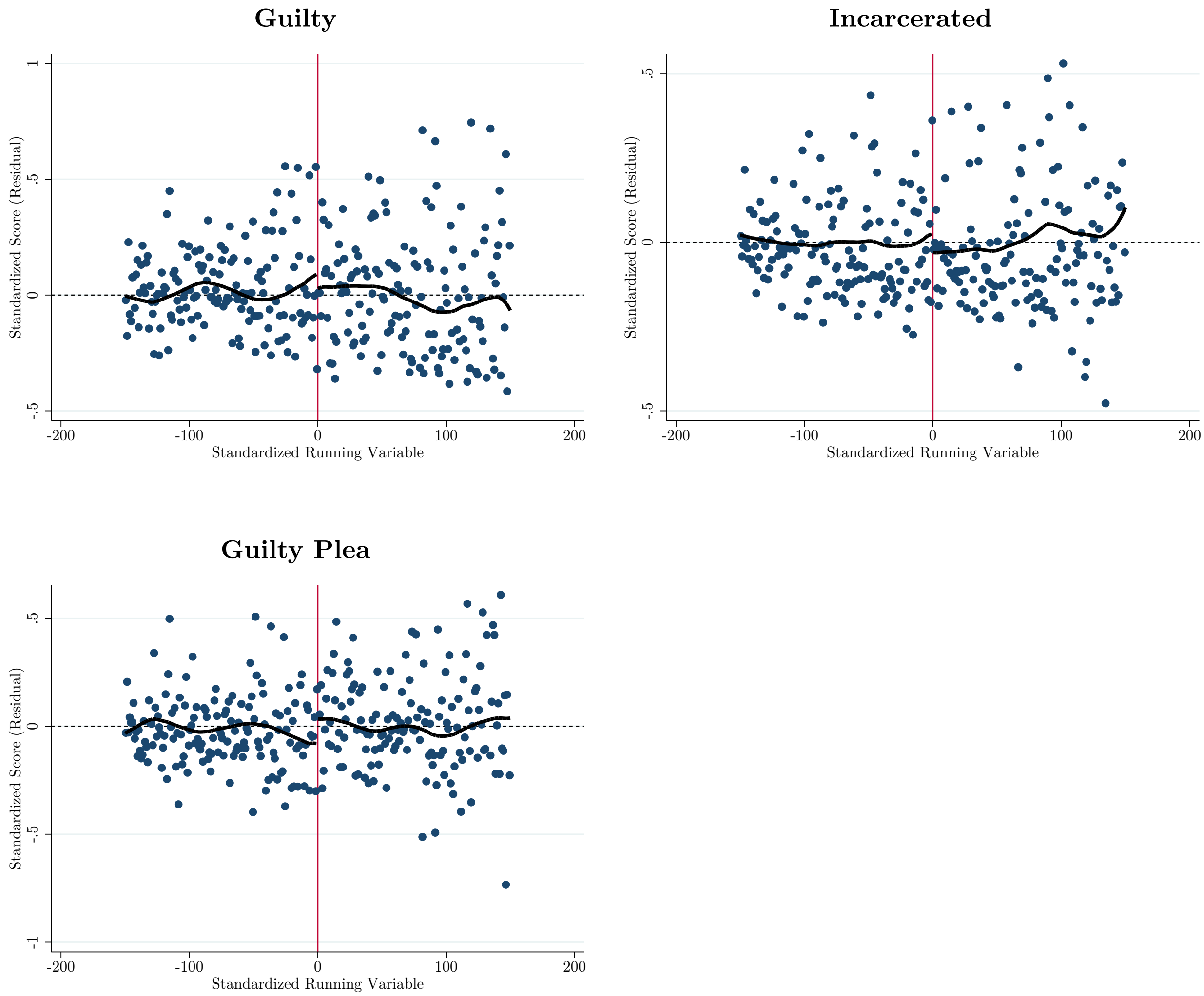


Figure 7. Reduced Form: Effect of Overshooting Legal Aid Eligibility on Court Outcomes

Notes: The figure plots variance-mimicking binned values of court outcomes over assessable net income, fitted with a polynomial of degree one. The bandwidths used are the same that are used to estimate the treatment effects in the FRD main specification.