Estimating the Impact of Public Defence on Court Outcomes:
credible identification using double machine learning

Fabio I. Martinenghi

University of New South Wales

Legal Aid: important & understudied

- Def: Govt provision of legal services at no/little cost to defender. Can be in-house or panel (private)
- It enforces the right to counsel, which in the Western tradition dates back at least to the Napoleonic Code 1808 (Zacharis, 2008 via McCannon & Porreca, 2023)
- In Australia, Dietrich v. The Queen [1992] establishes the right to counsel for federal (Commonwealth) cases

Legal Aid: important & understudied

- It is essential from an equity perspective because indigent defendants, absent aid, either:

  1. self-represent

     ⇒ aid promotes equality in front of the law

  2. hire a private lawyer (if they manage to fund it)

     ⇒ aid turns likely private debt into public expenditure

     - I focus on this latter case, as it covers 97% of the defendants in my dataset

⇒ Notwithstanding its great importance, we lack statistical evidence on how legal aid affects defendant outcomes

Empirical contribution

I study the question:

> What is the effect of refusing aid on court outcomes? Does it cause harm?

This is ultimately a question about the **performance** of public lawyers against their *counterfactual* private lawyers

This is the first paper addressing this question

Applied Metrics contribution

- I show how DML is a credible “research design” of its own, given the right setup

- The new DML tools (Chernozhukov et al, 2018) have been around for a while but have not won applied folks over

- This is because they are often framed as simply enabling better “control” on observables than linear regression

Administrative Data

- Legal Aid NSW - Crime Division: applications for legal aid between 2012-2021, in the context of serious crimes
- NSW Bureau of Crime Statistics and Research's ROD: court outcomes of applicants
- Indigenous defendants omitted due to restrictive privacy laws
- linked either via a police ID code or via name + DOB + court dates
- 33k cases and over 200k charges
- Info on defendants and their applications, cases and charges

Results preview

I find that being refused aid and hence hiring a private lawyer:

- Pr(incarceration): -11.3 p.p. $\rightarrow$ extensive margin
- Incarceration spell: + 5 months $\rightarrow$ intensive margin
- Pr(guilty plea): -5.3 p.p.
- Pr(fine): -4.8 p.p.
- Share of guilty charges: -1.7 p.p.
- ATT $<$ ATE, suggesting that TEs are stronger for legal aid recipients.

DML

The DML identification recipe

Credible ML-driven identification =

Data on all inputs of the treatment assignment function

+

Flexible algo (random forest) to learn the treatment assignment function

Let's unpack it

- treatment assignment function $\rightarrow$ propensity score
- Then DLM works as an inverse PS weighting estimator
- Hirano, Imbens & Ridder (2003)'s ECMA: IPSW is unbiased and efficient under **unconfoundedness**

Key identifying assumption: unconfoundedness

- Conditional independence assumption or, originally, *strongly ignorable treatment assignment* (Rosenbaum, 1983)
- characterises assignments mechanisms that
  1. do not depend on potential outcomes (e.g. no selection into treatment)
  2. are probabilistic—no unit is assigned to treatment or control with probability one (overlap)
  3. are individualistic (no network/spillover effects)

DML learns the treatment assignment function

- Two birds with one stone:

  1. credibly estimate treatment assignment function—all factors determining treatment are observed

     - a mix of sharp policies, principles and discretion

     $\Rightarrow$ likely highly non-linear

     $\Rightarrow$ use random forests

  2. Doubly robust! Just need to get either treatment or outcome model right

Fully heterogeneous treatment effects

- Interactive regression models (IRM): X interacted with D, hence allowing fully heterogeneous TEs

$$\begin{aligned} Y_i &= g_0(D_i, X_i) + U_i, \quad \mathbb{E}(U_i \mid X_i, D_i) = 0 \\ D_i &= m_0(X_i) + V_i, \quad \mathbb{E}(V_i \mid X_i) = 0 \end{aligned} \tag{1}$$

- Analysis at the case level, i
- Target parameters: ATT, ATE
- For OLS zealots, there is an alternative model which uses FWL and where treatment is additively separable (homog TEs)

Results

- I start from a naive specification, pooling refused + terminated applications in the treatment group and in-house + panel cases in the control group
- I then explore the robustness of these results
- and finally report the results of the preferred model

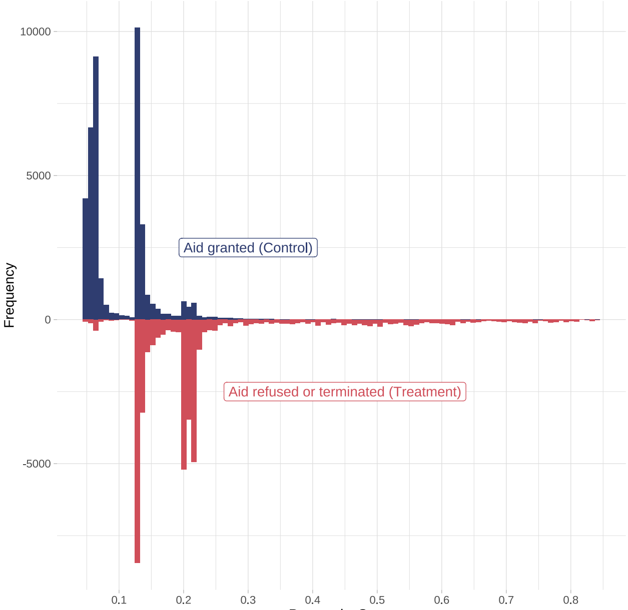

10000
5000
0
-5000
Frequency
Aid granted (Control)
Aid refused or terminated (Treatment)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

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

	Coeff.	s.e.	l-CI	r-CI	p-value	Obs.
Reduced charges	-0.027	0.006	-0.039	-0.014	0.0000	33153
Guilty share	0.010	0.005	0.001	0.019	0.0382	33153
Reduced seriousness	-0.010	0.004	-0.018	-0.002	0.0113	28849
Incarceration (months)	2.601	1.004	0.632	4.569	0.0096	18770
Incarcerated	-0.120	0.006	-0.131	-0.108	0.0000	33151
Guilty plea to highest charge	-0.020	0.006	-0.032	-0.008	0.0010	33153
Fined	0.059	0.004	0.052	0.066	0.0000	33149
Fine (AUD)	523.272	300.748	-66.185	1112.728	0.0819	2016

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: 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	-0.03	0.01	-0.04	-0.01	0.00	32587
Guilty share	0.01	0.00	0.00	0.02	0.05	32587
Reduced seriousness	-0.01	0.00	-0.02	0.00	0.02	28409
Incarceration (months)	2.71	1.02	0.71	4.70	0.01	18620
Incarcerated	-0.12	0.01	-0.13	-0.11	0.00	32585
Guilty plea to highest charge	-0.02	0.01	-0.03	-0.01	0.00	32587
Fined	0.06	0.00	0.05	0.07	0.00	32583
Fine (AUD)	584.41	316.67	-36.25	1205.07	0.06	1953

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: 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
Reduced charges	-0.026	0.007	-0.039	-0.013	0.0001
Guilty share	0.008	0.005	-0.001	0.018	0.0901
Reduced seriousness	-0.010	0.004	-0.018	-0.002	0.0181
Incarceration (months)	2.784	1.053	0.720	4.848	0.0082
Incarcerated	-0.120	0.006	-0.132	-0.108	0.0000
Guilty plea to highest charge	-0.022	0.006	-0.034	-0.010	0.0005
Fined	0.059	0.004	0.052	0.067	0.0000
Fine (AUD)	697.413	384.567	-56.324	1451.150	0.0698

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: 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
Reduced charges	-0.007	0.009	-0.024	0.010	0.4231
Guilty share	-0.007	0.006	-0.019	0.005	0.2477
Reduced seriousness	-0.006	0.006	-0.017	0.005	0.3020
Incarceration (months)	6.180	1.208	3.814	8.547	0.0000
Incarcerated	-0.120	0.008	-0.136	-0.104	0.0000
Guilty plea to highest charge	-0.041	0.008	-0.057	-0.025	0.0000
Fined	0.062	0.005	0.051	0.072	0.0000
Fine (AUD)	366.016	321.952	-264.999	997.031	0.2556

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.

Preferred specification: FEs + only in-house + no terminated

- Add court-level fixed effects via within transformation
- Drop panel lawyers
- This is to model in-house lawyer randomisation at the court level
- Terminated cases dropped for treatment homogeneity

$$\begin{aligned} Y_i = g_0(D_i, X_i) + \alpha_{c(i)} + U_i, &\quad \mathbb{E}(U_i \mid X_i, D_i) = 0 \\ D_i = m_0(X_i) + V_i, &\quad \mathbb{E}(V_i \mid X_i) = 0 \end{aligned} \tag{2}$$

i: case; c: court

Table: 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	0.007	0.011	-0.014	0.029	0.4960	15937
Guilty share	-0.017	0.007	-0.031	-0.004	0.0122	15937
Reduced seriousness	-0.004	0.004	-0.012	0.004	0.3397	14067
Incarceration (months)	5.038	1.521	2.056	8.020	0.0009	8712
Incarcerated	-0.113	0.012	-0.136	-0.090	0.0000	15936
Guilty plea to highest charge	-0.053	0.008	-0.068	-0.037	0.0000	15937
Fined	0.048	0.008	0.031	0.064	0.0000	15936
Fine (AUD)	-8.351	64.018	-133.824	117.122	0.8962	1074

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. The chosen ML method is Random Forests. Finally, court-level fixed effects are included via a within transformation.

Table: DML: ATT 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	0.004	0.012	-0.020	0.028	0.7482	15937
Guilty share	-0.008	0.008	-0.025	0.008	0.3042	15937
Reduced seriousness	-0.003	0.006	-0.015	0.009	0.6213	14067
Incarceration (months)	5.266	1.422	2.479	8.053	0.0002	8712
Incarcerated	-0.083	0.017	-0.116	-0.050	0.0000	15936
Guilty plea to highest charge	-0.038	0.011	-0.059	-0.018	0.0003	15937
Fined	0.040	0.007	0.026	0.054	0.0000	15936
Fine (AUD)	-97.211	52.809	-200.714	6.293	0.0657	1074

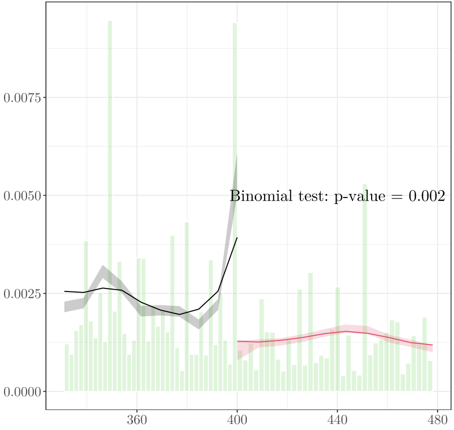
Notes: The table presents Double Machine Learning ATT estimates for the effect of not receiving aid (having aid denied or terminated) on case-level court outcomes. The IRM model is applied to all cases where aid was not terminated and where the defendant was not represented by a panel lawyer. The chosen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

RDD

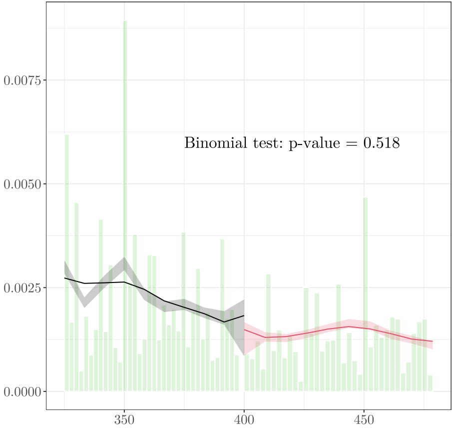
Empirical Strategies & Target Parameters: Fuzzy RDD

- Most important factor determining grant/refusal of aid is "net assessable income" (henceforth, "income")
  - income (incl. most welfare payments) minus income tax and deductions related to housing cost and dependants
  - cutoff: applicants earning **400$** p.w. or more are refused aid unless a *discretion* request is made and granted ⇒ **fuzzy RDD**
- Running variable: income
- Treatment: refusal or termination of legal aid
- Instrument: income $> 400\$$
- Target parameter is LATE in two ways: n-hood of 400 & compliers
- Compliers: applicants earning $\approx 400\$ p.w.$ who are granted aid if income ≤ 400 and refused otherwise

Density Test



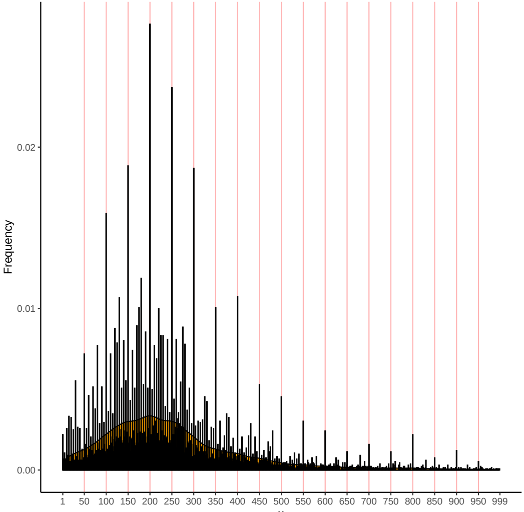
(a) Full-sample Density Test



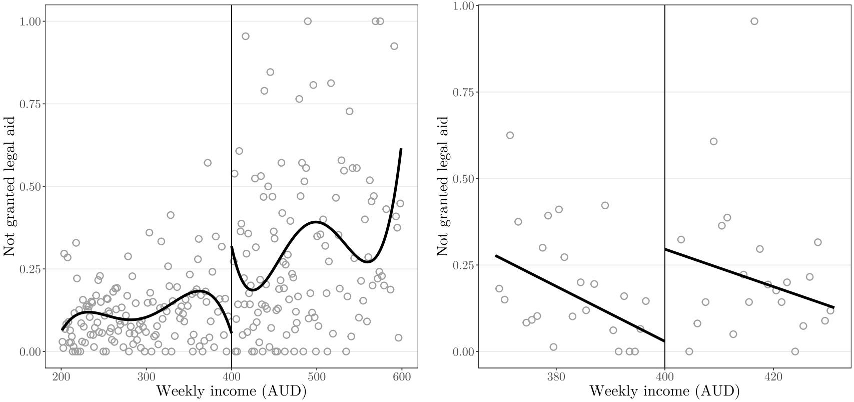
(b) Donut-sample Density Test

Bunching or Manipulation?

Bunching seems very much prevalent, but donut approach is more prudent

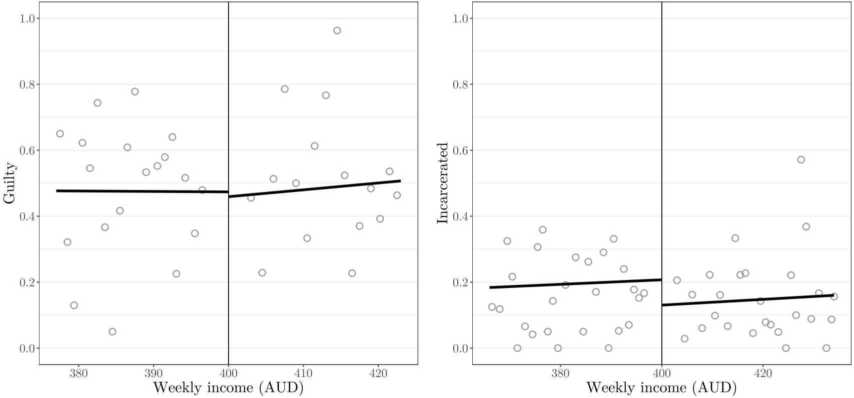


First Stage

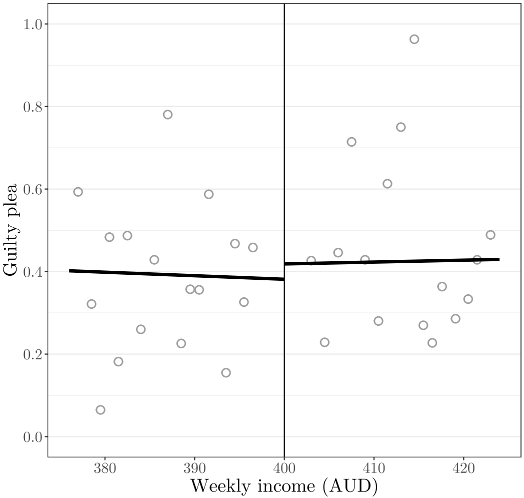


(a) Global (b) Local

Reduced Forms: I



Reduced Forms: II



Null reduced-form estimates

Table: 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.

… leading to null FRD estimates with huge CIs

Table: 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.

Conclusion

Refusal of aid based on current criteria does not harm case outcomes

- I study the **effect of refusing aid** on the court outcomes of serious crimes defendants
- Refused applicants hire a private lawyer, leading to insights on the difference in performance between private and public lawyers
- DML: Refusal of aid not harmful under current rules—private lawyers outperform in-house lawyers.
- RDD: Around the 400$ eligibility threshold, I find uninformative null estimated impacts on the marginal applicant (wide CIs)

DML can provide identification *and* policy relevance

- DML can be used profitably to study policy impacts using admin data
- This unlocks new policy-relevant questions.
- Here: how good should LA be? Is it meeting this target?
- Future work will study how the provision of aid affects recidivism and socio-economic outcomes

Treatment over Income - CEF