

CRIME AND JUSTICE BULLETIN

NUMBER 269 | NOVEMBER 2025

Examining gender bias in the policing of intimate partner violence

Adam Teperski and Ilya Klauzner**AIM**

To examine whether police officers exhibit gender bias in their decision to charge persons of interest (POIs) in intimate partner violence (IPV) assault incidents involving heterosexual partners. This is relevant to the concern that police may misidentify female victims of domestic violence as offenders.

METHOD

We begin by constructing a dataset of 52,423 police events of IPV assault involving heterosexual partners (including 38,413 male and 14,010 female POIs reported from July 2010 to February 2023). Across these events, 46% of female POIs and 64% of male POIs were proceeded against by police. Court outcomes were then used to test the strength of police cases against POIs in our sample to determine whether police bring weaker charges to court against women relative to men. We assume that police officers have an implicit threshold for the strength of evidence needed to charge someone with an IPV assault offence, aiming to prove the charge beyond reasonable doubt. If the evidence meets this conviction threshold, they will proceed with the charge; if not, they won't. These conviction thresholds are estimated separately for men and women by employing an instrumental variables strategy. This involves exploiting the quasi-random allocation of police officers to IPV events, and the variation in the willingness of individual police officers to initiate charges. We approximate marginal POIs as POIs who would have been charged by an officer with a higher propensity to charge but not charged by an officer with a lower propensity to charge. If officers apply a lower conviction threshold for marginal female POIs relative to marginal male POIs (i.e., they are more inclined to charge women and require less evidence) this suggests police charge decisions are biased against women.

RESULTS

Conviction rates (and therefore the quality of evidence) were similar for marginal male and female POIs charged with IPV assault. Marginal female defendants had a slightly higher conviction rate (81.7%) than marginal males defendants (67.2%), although this difference was not statistically significant. These findings suggest that police officers apply a consistent threshold of evidence when deciding to charge male and female POIs in IPV incidents.

CONCLUSION

This study finds no evidence of gender bias in police charging decisions for IPV assault, specifically in the form of a lower evidentiary threshold applied to female POIs. However, these findings are limited to officers with substantial experience in responding to domestic violence incidents. Additionally, while this analysis focuses on police decision-making, it does not rule out the possibility of bias elsewhere in the criminal justice system, including in court processes. Misidentification remains a complex and difficult-to-detect issue that may still occur in individual cases.

KEYWORDS

Intimate partner violence

Domestic and family violence

Policing

Offenders

Victims

Mischarging

Misidentification

INTRODUCTION

Intimate Partner Violence (IPV) is a significant public health and criminal justice issue impacting many Australians. Recent surveys suggest that 23% of women and 7% of men have experienced IPV (defined as physical and/or sexual violence from an intimate partner)¹ since the age of 15 (Australian Institute of Health and Welfare [AIHW], 2024). The flow-on effects of these experiences can be far-reaching, including worsened mental health, substance abuse, suicide, housing instability and unemployment (Adams et al., 2024; Australia's National Research Organisation for Women's Safety [ANROWS], 2019; Morgan & Boxall, 2022; Webster, 2016). These harmful outcomes are known to be particularly pronounced for women, and to adversely impact the wellbeing and development of children who witness or are exposed to violence (Webster, 2016). The economic costs arising from gender-based violence are significant, estimated to be around \$26 billion a year in Australia (KPMG, 2016). Despite significant government investment to address the high prevalence and costs of IPV, rates of self-reported Domestic Violence (DV) victimisation have remained largely unchanged across Australia and within New South Wales (NSW) over the last decade (AIHW, 2024; Freeman, 2023).

While the self-reported incidence of DV has remained stable over time in NSW (Freeman, 2023), there has been a long-term upward trend in the number of domestic violence (DV) incidents recorded by NSW police. Police recorded crime data shows an increase from December 2014 to December 2024 of 3.1% per annum for DV assaults (from 29,113 to 38,272), 7.9% per annum for DV intimidation, stalking and harassment behaviours (from 13,000 to 24,125), and 7.9% per annum for breaches of Apprehended Domestic Violence Orders (ADVO) (from 13,082 to 24,866) (NSW Bureau of Crime Statistics and Research [BOCSAR], 2024).

This rise in recorded incidents could reflect greater public willingness to call police or a broadening definition of what constitutes DV. But a proportion is also likely the result of a shift towards stronger police enforcement in response to DV incidents. This is evidenced by the rate at which the NSW Police Force initiate legal proceedings against offenders involved in reported DV incidents, which grew by 6.6 percentage points (p.p.) between 2007 and 2016 (from 58.2% to 64.8%; Holmes & Fitzgerald, 2017) and has remained close to 65% since 2017 (BOCSAR, 2024). In this same period, several other initiatives were also introduced to strengthen the NSW Police Force response to DV. These included doubling the number of NSW police domestic violence specialists over the five-year period to December 2022 (Audit Office of NSW, 2022) and boosting the number of ADVO compliance checks conducted by NSW police by 85% between 2019 to 2024 (BOCSAR, 2024).

There are conflicting opinions as to whether the increased policing of DV has benefited victims. Some legal experts and advocates welcome stronger police enforcement (NSW Law Reform Commission, 2010; Wangmann, 2014). However, Nancarrow et al., (2020) argue that incident-based policing may lead to unintended negative consequences for the victim. A frequently raised concern is that police may incorrectly "misidentify" IPV victims as offenders, particularly in cases where there are conflicting or mutual claims of violence (Nancarrow et al., 2020). This can occur when police fail to recognise the victim's actions as occurring in self-defence, or in complex disputes where it may be easier to charge both parties and let the courts determine the criminality of individuals involved. While mischarging is not inherently gender-specific, it is thought to affect women more often than men for a few reasons. The first is that women commonly use self-defensive and retaliatory violence to protect themselves and others from abuse (Boxall et al., 2020; Mackay et al., 2018), and in these cases, women may be more likely to use weapons (such as household items) to counter their physical strength disadvantage (Nancarrow et al., 2020). The second is that women do not always behave in a way that fits with the stereotype of the 'ideal victim,' who is submissive and powerless (Nancarrow et al., 2020). The third is that women may be less willing to cooperate with police because they are fearful of reprisal by the predominant aggressor or fear losing custody of their children (Nancarrow et al., 2020). Mischarging of women may be one reason why the

¹ Intimate relationships include marriages, couples who live together, and dating relationships. In NSW, IPV offences are recorded as a subset of domestic violence offences.

proportion of female offenders has risen from 19% of all offenders in 2014 to 22% of offenders in 2023 (Donnelly, 2024).

Police mischarging has a significant impact on victims and may be becoming more prevalent. Impacts of mischarging include worsening employment prospects, poorer mental health, reduced access to children and use of government services, the potential for continued abuse during court proceedings and the negative consequences of having a criminal record (see Larance et al., 2019; Reeves, 2019; Reeves, 2021; Shuard, 2021). Police may also apply for an ADVO, which if breached could result in further criminal proceedings. Additionally, the anticipation of mischarging could act as a barrier to victims reporting DV behaviours (Wolf et al., 2003). Mischarging may have become more prevalent over the last decade in NSW, as evidenced by the six-fold increase in the proportion of cross-Intimate Partner Violence (IPV) assault events (where both heterosexual partners are charged) over the last ten years (from 0.4% of all IPV events in July 2010 to 2.5% in February 2023; Wang & Klauzner, 2025). In light of this, understanding the police decision-making process and the extent of mischarging in IPV matters may inform policymakers about appropriate responses to the rise in cross-IPV events.

Mischarging of victims as offenders in intimate partner violence matters

There is limited research either internationally or within Australia that investigates the extent to which IPV mischarging/misidentification occurs. The only Australian prevalence estimate available comes from an internal analysis by Victoria Police in 2020 which found that female victims were misidentified as perpetrators in 12 per cent of domestic and family violence (DFV) cases (Victoria Police, 2021). Most Australian literature on misidentification has utilised small scale surveys and focus groups to investigate factors that influence misidentification, particularly for female victims. The most extensive qualitative research was a study conducted by Nancarrow et al. (2020) investigating the circumstances in which female victim-survivors are misidentified in DFV civil law matters in Queensland. They conducted focus groups and interviews with 39 police officers, 39 specialist DFV service providers, 17 women with lived experience of DFV and five magistrates. The authors identified several factors that may drive misidentification of females as perpetrators in criminal incidents and cross-applications for protection orders (i.e., where each party to a protection order is the subject of an order against them). This included women's use of resistive violence combined with a focus in the investigation on injury (i.e., taking the injured party to be the victim), misinterpretation of victim distress as evidence of perpetration or guilt, and the use of protection orders as a means of retaliation. Further, Nancarrow et al. found that police practices tended to consider only the actions associated with the event at hand, rather than the entire history of abuse that has occurred in the relationship including any evidence of coercive control. Many police officers interviewed also reported adopting a risk-averse approach to decision-making, preferring to leave it to the courts to determine the primary aggressor. This was heightened in situations where officers were inexperienced or had limited time/resources to investigate the history of the relationship.

Nancarrow et al.'s (2020) results are supported by other qualitative studies of mischarging/misidentification conducted in Australia. For instance, Boxall et al. (2020) analysed 153 police narratives of DV involving a female POI. Half of the episodes involved self-defensive or retaliatory violence, defined as incidents where the female POI either: 1) perpetrated DV in immediate response to abusive or threatening behaviour from her male partner; or 2) had a documented history of being victimised by her male partner in prior DV episodes. This may suggest that the narrow legal definition of self-defence contributes to misidentification of female POIs as offenders. Separately Iliadis et al. (2024) surveyed 109 female and seven male² DFV victim-survivors to understand their perceptions of and experience with body-worn cameras (BWCs) in the policing of DFV. This work aimed to explore concerns that the increasing use of BWCs in routine policing in Australia may encourage police to take a narrow incident-based focus in DFV incidents, thereby increasing the risk of mischarging. While their research found that victim-survivors were generally supportive of the use of the new technology, nearly one-third raised concerns that BWCs could result in more victim-survivors being identified as the primary aggressor. Some

² Although the study surveyed participants across all Australian states, around 55% of participants resided in Queensland.

participants suggested that this is because a victim-survivor's trauma responses during the incident might be misinterpreted by officers, who have preconceived notions about how a credible victim should behave. Consequently, these victim-survivors felt they must present themselves in a certain way to ensure state agents acknowledge their abuse experiences, or risk being mischarged. Others were concerned that the primary aggressor could manipulate the situation by reporting false narratives, such as self-harming to appear as the victim or suggesting they have been attacked. The benefits of such manipulation to the offender may be greater when BWCs are in use.

Other Australian evidence has focused on the experiences of people with victim protection orders. In particular, Wangmann (2010) examined the use of cross-ADVO applications in NSW, drawing on interviews with legal professionals and women involved in cross-applications ($n = 27$ and $n = 10$ respectively), along with documentary analysis of 234 court files and 73 court observations. Wangmann found that cross-applications can be used by men as a tactic to bring about mutual withdrawal, noting that many cross-applications are lodged by men after an ADVO was initiated by the police to protect a woman following a DV incident. Separately, Reeves (2023) conducted semi-structured interviews with 18 legal practitioners who had experience representing women who had been misidentified as primary aggressors on family violence intervention orders (FVIOs) in Victoria. This study found that the FVIO hearings often operate under the assumption that police will bring legitimate applications to the court and respondents will consent to the order. Respondents were rarely able to present their own experience of victimisation to the magistrate unless they contested the order at a later date. However, respondents could be placed on interim FVIOs while awaiting their contest hearing(s), and the delay between the initial hearing and contest hearing could be longer than six months. By the time the contest hearing occurred, the respondent may have been subject to a large proportion of the final order duration. As a result, contesting was often seen as an unnecessary extension of the court process which could result in further victimisation.

A separate U.S literature investigates the impact of mandatory arrest policies for IPV. Mandatory arrest policies aim to standardise the police response to IPV by reducing police discretion in arrest decisions. However, across jurisdictions that have implemented mandatory arrest policies, the average probability of arrest was only 50%, leaving considerable room for police to apply discretion even after the policies were in effect (see Eitle, 2005). While mandatory arrest policies aim to increase arrest rates overall, there is some concern that they may, unintentionally, also result in increased mischarging of female victim-survivors through more instances of dual arrest (where both the victim and perpetrator are arrested), and 'retaliatory arrests' (where victims are arrested due to a false or exaggerated complaint filed by the perpetrator) (see Frye et al., 2007; Rajah et al., 2006).

Recognising the harms associated with mischarging and the risk that mandatory DfV arrest policies may lead to more female victim-survivors being mischarged, a number of US states introduced 'primary aggressor' policies. These policies encourage police to consider the context of the violence, including any history of prior IPV or evidence of self-defence, when determining whether legal action should be taken: the policy objective being to reduce the incidence of dual arrests involving females and, potentially, reduce female arrests overall. Hirschel et al. (2021) evaluated the impact of these policies using a dataset of 1,300,000 criminal incidents reported to police from 36 US states during 2000-2009. They found that primary aggressor legislation had no impact on dual arrests. Instead, police officers in jurisdictions adopting this approach became more reluctant to make any arrest in IPV incidents after the policy came into effect, even in circumstances where an arrest may have been warranted. Beeble et al. (2022) employed a similar methodology using a dataset of 311,558 incidents involving IPV across 36 US states in 2015 and found no evidence to suggest that primary aggressor policies had any impact on arrest outcomes (including dual arrests). Beeble et al. (2022) did however confirm the results of earlier studies in finding that jurisdictions with mandatory DfV arrest policies had significantly higher odds of both male-only arrests and female-only arrests than other jurisdictions. Additionally, the authors found that mutual violence in IPV incidents involving male and female partners increased the odds that both partners were arrested. However, female partners experienced an increase in the odds that they were arrested by nearly five times as much as the increase in the odds that a male partner was arrested. This suggests that women are disproportionately impacted in complex IPV cases where police find it difficult to determine the primary aggressor.

The current study

The literature outlined above identifies police decision making as one of the key factors contributing to the mischarging/misidentification of female victims as perpetrators (Nancarrow et al., 2020; Reeves, 2019). In NSW, clear guidelines exist for police officers responding to a DV report or attending a DV incident. The *Code of practice for the NSW Police Force response to Domestic and Family Violence* (NSW Police, 2018a) directs police to undertake thorough investigations to determine the primary aggressor at the incident, including consideration of witness statements, the behaviours of involved parties, whether there are injuries and any prior history of DV and ADVOs of involved parties. Police are also discouraged from charging both parties involved in a DV incident, barring some exceptions (such as if a victim commits a property damage offence after being assaulted) (NSW Police, 2018a). However, the extent to which these policies and procedures are effectively applied in practice has not been evaluated, and the research outlined above suggests that this policy position may not always align with the experiences reported by victim-survivors.³

To investigate the role of gender in charge decision-making, this study leverages the quasi-random allocation of responding police officers to criminal events to understand whether police officers are more likely to charge women following an IPV assault compared to men with the same probability of conviction.⁴ To investigate whether women are being mischarged for DV offences, we use court outcomes to test the strength of police cases. This allows us to infer whether police bring weaker charges to court against women relative to men. This approach assumes that police apply a threshold when deciding whether to legally proceed against a person for a DV offence. That is, a level of evidence that they believe is sufficient to be able to successfully prove the charge beyond reasonable doubt. Above this conviction threshold, a police officer will charge the person, and below this conviction threshold they won't charge the person. If police decision-makers are equally likely to charge women and men, we would expect that they apply the same conviction threshold regardless of the gender of the POI. However, if police officers apply a lower threshold for conviction in cases where a female is accused of committing DV (i.e. they require less evidence to charge) then this indicates that officers are mischarging women. Note that this method also assumes that police and court decision-makers are not biased in the same direction (e.g. against women), as this would limit the extent to which police mischarging is later corrected by courts.

METHOD

One way to estimate the extent of mischarging would be to estimate differences in police legal action rates for male and female POIs following an IPV assault incident, controlling for factors that may influence the arrest decision (such as the prior offending history of POIs, and the quality of evidence available to police). However, a concern with this approach is that it could underestimate the extent to which police are more inclined to charge women with similar characteristics as men, as it cannot control for factors that are not collected in administrative datasets and are thus unobservable to researchers (e.g., factors relating to the quality of evidence available to police such as behaviours of the involved parties after the incident and the extent of injuries of parties involved), or differences in underlying rates of DV perpetration by gender. For example, evidence suggests that women are less likely to perpetrate DV than men (see for example, the 2021 Australian Bureau of Statistics [ABS] Personal Safety Survey (ABS, 2023) which indicates that 7.5 million adults have experienced family, domestic and sexual violence (FDSV) from a male perpetrator and just 2.2 million adults have experienced FDSV from a female perpetrator; AIHW,

³ See also, Jillard and Mansour (2014) who found that two-thirds of women with ADVOs made against them stated that they were the primary victim of violence in their relationship with the person in need of protection (PINOP). Similarly, during interviews with 17 women with lived experience of DV, Nancarrow et al., (2020) found that women are often being misidentified as perpetrators in DV cases, despite being the primary victims. This could suggest that NSW police may not be receiving sufficient training on the dynamics of DFV.

⁴ An alternative way to assess gender bias in police charging decisions is to use a measure of police leniency to estimate how being charged affects reoffending outcomes, separately for men and women. However, internal analysis and consultation with NSW Police indicate that once someone is charged in a DV incident, they become known to police, and this may increase the likelihood that any future offending will be detected. In this case, the police leniency instrumental variable violates the exclusion restriction, as police leniency directly influences both charge decisions and reoffending outcomes.

2024). As a result, a simple regression approach may find that women are less likely to be charged in a DV incident than men with similar observable characteristics (e.g., prior offending histories and number of concurrent offences). Consequently, this approach would likely suggest that police are less likely to charge women with similar characteristics as men, which may not hold if all unobservable characteristics could be controlled for.

We overcome this concern by implementing an outcomes test (Arnold et al. 2018). An outcomes test is a type of analysis where bias in decision-making is inferred not from differences in the rates at which groups receive a particular decision but disparities in the outcomes flowing from those decisions. It can be used wherever a decision (charging a person with an offence) is made in order to achieve a particular objective (conviction). This strategy has been used to examine racial bias in bail decisions in the US (Arnold et al. 2022) and in Australia (Klauzner, 2023). We implement the outcomes test in our study by exploiting the quasi-random allocation of police officers with different charge propensities to IPV events. This allows us to estimate the difference in conviction rates among men and women with the same likelihood of being charged. If women are convicted at lower rates than men with the same likelihood of being charged, this suggests that women are more likely to be mischarged as the primary aggressor in IPV incidents. That is, it would suggest that police officers are, on average, applying a lower threshold (in terms of factors that could influence conviction) when deciding to charge women with IPV assault compared to men. This approach is not biased by the influence of unobserved confounders related to the decision to charge made by police, as we estimate conviction thresholds for POIs who are effectively randomly assigned into being charged or not depending on the charge propensity of the officer assigned to the IPV event. As a result, we would expect these POIs to have similar observable and unobservable characteristics.

Note that since the conviction decision can be determined by judicial officers (in cases that are not resolved by guilty plea), this outcomes test approach assumes that magistrates/judges accurately evaluate whether a person is guilty of an offence and do not themselves exhibit bias in decisions involving IPV offences. While there is a scarcity of Australian studies into gender bias in conviction decisions, US research investigating this topic finds little evidence of gender bias in DV or sexual assault cases (see Belknap et al., 2000; Felson & Pare, 2007). Other US studies find that men are consistently treated more severely at every stage of the court process even when legal and extra-legal variables are controlled for (see Henning & Feder, 2005; Starr, 2015). Australian research reports that police believe courts to be more accurate in determining whether a POI involved in a DV incident was a primary aggressor (see Nancarrow, 2019; Reeves, 2019). However, Nancarrow (2019) highlights that due to high caseloads and time constraints, magistrates often rely heavily on the evidence provided by the police when making conviction decisions. This suggests that police misidentification may not always be later corrected by magistrates. For this reason, we also undertake further exploratory analyses to investigate whether there may be evidence of gender bias in conviction decisions. We do this by examining whether, conditional on being proceeded against, raw conviction disparities between men and women can be explained by gender differences in: 1) case characteristics; 2) prior offending; and (3) other demographic variables. This additional analysis is presented in Appendix A.⁵ We find that conditional on being proceeded against, 75% of men and 71% of women were convicted, reflecting a 4 p.p. difference. Half of this disparity (2 p.p.) was attributed to men having more extensive prior offending histories and having more concurrent offences than women. The remaining 2 p.p. of the conviction disparity was not explained by differences in offender characteristics, which could reflect court bias against men or omitted variable bias. Although this analysis is associative, rather than causal, it suggests there is limited evidence of gender bias in our sample.

⁵ We take this approach instead of directly controlling for judge leniency or court location in our IV regression, as these variables are only recorded for POIs who are proceeded against. Including these variables as a control in the second stage of our IV regression drops POIs who are not proceeded against, which prevents us from estimating our parameter(s) of interest.

Data

Data in the study were derived from two sources.

- 1. NSW Police Force's Computerised Operational Policing System (COPS) data:** An extract of all DV police events from COPS that occurred between 1 July 2010 and 28 February 2023, including information on the POI, victim, and type of incident(s) attached to the event. This extract also details the sociodemographic characteristics of POIs and victims, including their age, gender, Aboriginality, relative socioeconomic disadvantage, and remoteness of area associated with their residential address at the time of the event.
- 2. BOCSAR's Re-Offending Database (ROD):** An extract from ROD which provides details on all youth justice conferences, police cautions, and court appearances finalised in NSW from January 1994. The extract of ROD used in this study provides information regarding dates of offending, offence type and court outcomes.

To track court outcomes for individuals following a police recorded IPV incident, we linked all COPS DV events reported between 1 July 2010 and 28 February 2023 to ROD records using a combination of probabilistic matching and direct matching on source system linkage keys. We then restricted the sample to events which met the following criteria: 1) there was at least one DV assault incident within the event;⁶ 2) there was one unique man and one unique woman involved in the event (recorded by police as a victim or POI);⁷ 3) police had flagged the relationship between the two parties as being intimate; 4) the event did not occur in prison; 5) there was a single POI per event, except for cases of cross-DV (where each person involved in the event appeared as both a victim and an offender). This left us with a sample of 147,553 IPV DV assault events between heterosexual partners, involving 113,241 unique POIs. Finally, we removed observations where the charging officer had attended too few IPV assaults for us to construct a measure of police propensity to charge (detailed further below in the method section). Specifically, each officer needed to have attended at least five IPV events with male POIs and five IPV events with female POIs, and there needed to be at least two officers for each Police Area Command (PAC) by month-year fixed effect.⁸ These criteria dropped 88% of observations or 13,778 police officers (from 15,731 to 1,953 officers). As the remaining 12% of police officers responded to multiple IPV events, the number of IPV events in our analysis sample fell by only 64% (from 147,553 to 52,423) after applying these criteria. We compare the characteristics of POIs excluded from the sample via these criteria to those included in the sample in Appendix B. We find that aside from a small difference in charge rates (59 p.p. vs 67 p.p., respectively), offender and offence characteristics are quite similar across the groups. This does however mean that officers included in our analysis sample are the most experienced in responding to DV assault incidents, relative to other officers. Our final analysis sample included 52,423 events involving 46,013 unique POIs.

Variables

Outcome variable

The outcome variable considered in this study is a binary variable measuring whether or not an offender who was proceeded against for IPV assault was convicted in court by verdict or guilty plea, and zero otherwise.

⁶ An incident is DV-flagged if the NSW police decide that DV was a factor in the incident. This involves determining whether parties involved in an incident had a domestic relationship, as per the *Crimes (Domestic and Personal Violence) Act 2007*. Assault is defined as the direct and immediate infliction of force, injury or violence upon a person or people, or the direct and immediate threat of force, injury or violence where there is an apprehension that the threat could be enacted (see ABS, 2011).

⁷ In this study, we use police records of gender recorded at each police incident. Gender was recoded as male in 2% of incidents when recorded by police as unknown.

⁸ Without the requirement that officers attend at least five IPV events with male POIs and five IPV events with female POIs, our police leniency instrument is weak. This is because we need to observe several decisions for the same officer before that officer's gender-specific leniency stabilises around an average value. Without the requirement of two officers for each PAC by month-year fixed effect, we are unable to construct our residualised measure of police leniency (see the method section for more details).

Explanatory variables

The primary explanatory variable in our study is whether a recorded IPV assault resulted in a police officer commencing legal proceedings against a person of interest. In addition, we consider the following control variables that may influence a police officer's decision to charge a POI for an IPV assault (and which we can measure).

1. Sociodemographic characteristics

- Age: the POI's age at the incident (coded as under 18, 18-24, 25-34, 35-44, and 45+).
- Age difference: the difference between the age of the POI and victim at the time of the incident (coded as 11+ years younger, 6-10 years younger, 2-5 years younger, 1 year younger/older, 2-5 years older, 6-10 years older, and 11+ years older).
- Aboriginality: whether a POI self-identified as an Aboriginal or Torres Strait Islander person at the time of the incident (coded as Aboriginal, non-Aboriginal, and unknown).
- Socioeconomic disadvantage: the ABS' Socio-Economic Indexes for Areas' (SEIFA) Index of Relative Disadvantage (ABS, 2016a) associated with the postcode of a POI's residence at the time of the incident, segmented into quartiles (coded as Q1, Q2, Q3, Q4, and unknown).
- Remoteness of residence: the ABS remoteness of area associated with the postcode of an offender's residence at the time of the incident (coded as Major cities, Inner regional, Outer regional, Remote or Very remote, and missing), based on the Accessibility and Remoteness Index of Australia (ARIA; ABS, 2016b).

2. Prior criminal offending

- Prior court appearances: the number of proven court appearances for the POI in the five years prior to the incident start date (coded as 0, 1-2, 3-5, and 6+).
- Prior prison sentences: the number of finalised court appearances for the POI in the five years prior to the incident start date where a full-time prison sentence or juvenile control order was imposed (coded as 0, 1-2, 3-5, and 6+).
- Prior breaches of violence orders: the number of proven breaches of a violence order⁹ for the POI in the five years prior to the incident start date (coded as 0, 1-2, 3-5, and 6+).
- Prior DV offences: the total number of proven DV-flagged offences for the POI in the five years prior to the incident start date (coded as 0, 1-2, 3-5, and 6+).
- Prior DV assault offences: the number of proven DV-flagged assault offences for the POI which are in the five years prior to the incident start date (coded as 0, 1-2, 3-5, and 6+).
- Prior victimisation: the number of COPS incidents in the five years prior to the index incident start date where the POI was recorded as the victim and legal action was initiated in relation to the event (coded as 0, 1-2, 3-5, and 6+).
- Prior DV victimisation: the number of proven proceeded DV-flagged incidents in the five years prior to the index incident start date where the POI was recorded as the victim and legal action was initiated in relation to the event (coded as 0, 1-2, 3-5, and 6+).

3. Time and location of the DV event

- Policing region: a collection of indicator variables representing the PAC or Police Division (PD) that responded to the index incident.
- Time of event: a vector of indicator variables representing the month and year in which the event was reported to police.

⁹ Defined as offences with an Australia and New Zealand Standard Offence Classification (ANZSOC) of 1531 (ABS, 2011).

We used the time and location of the DV event to construct PAC/PD fixed effects by month-year. This allows us to control for PAC/PD-level factors that differ between PAC/PDs and time (e.g. fixed budgeting allocations, PAC culture or commander preferences).

Empirical approach

Following the methodology of Arnold et al. (2018) and Klauzner (2023), we measure gender differences in charging decisions by comparing conviction rates for men and women at the margins of being charged for an IPV assault. This approach involves estimating the implicit conviction thresholds police apply when deciding whether to legally proceed against a POI in relation to an IPV incident. After evaluating the evidence and circumstances of an event, police will charge an offender if they believe the offence can be successfully proven beyond reasonable doubt. This implies that police have a threshold of conviction where their decision to charge shifts from not charging to charging the POI. It also implies the existence of marginal POIs, who are legally proceeded against by officers who have a higher propensity to charge but not charged by officers with a lower propensity to charge. By identifying conviction thresholds for marginal defendants by gender, we can infer the average implicit threshold used by individual officers for each gender. If police use the same threshold when deciding whether to charge men and women, this suggests a lack of evidence for gender bias in police decision-making. However, if the threshold is lower for women, this suggests that charging decisions are negatively biased against women POIs, as police officers proceed against women in IPV cases where the evidence available is not sufficient to secure a conviction.

Constructing a police propensity to charge instrumental variable

To estimate these implicit conviction thresholds, we utilise an instrumental variables (IV) strategy. This involves exploiting variation in police officers' propensity to charge a male or female POI and the quasi-random allocation of officers to IPV incidents to estimate the impact of being legally proceeded against on the probability of conviction. The propensity to charge index is created for each police officer in our sample and separately for men and women using a similar strategy to Arnold et al. (2018) and Klauzner (2023). First, on the subset of male POIs we regress a binary variable equal to one if a POI is proceeded against on PAC by month by year fixed effects. Using the residuals of this regression, we calculate a leave-one-out mean of officer propensity to charge by summing the residuals for each officer in a month, excluding decisions related to the defendant for whom the index is calculated, and dividing by the number of decisions made by that officer. The leave-one-out approach addresses potential endogeneity between the propensity to charge measure and the treatment variable.¹⁰ Further, the residualised measure of propensity to charge removes variation in police charge decisions influenced by unobserved factors related to the PAC, seasonality, and calendar time including changes in command leadership.¹¹ We then repeat this approach for the subset of female POIs. Intuitively, this measure of propensity to charge represents the deviation between a given officer's likelihood to proceed against offenders of a given gender for IPV assault and the average proceeding rate in the same PAC at the same time for that same gender, in all events unrelated to the index individual's event.

Instrumental variable method

After constructing our gender-specific police propensity to charge measure, we implement an IV regression using two-stage least squares (2SLS), calculated separately for subsamples of male and female POIs. For a given subsample, both stages are estimated together using the STATA package `ivreghdfe` (Correia, 2018). The first stage of the analysis estimates the probability of being proceeded against based on the propensity to charge of the officer assigned to that event:

$$P_{ic} = \alpha_0 + \alpha_1 * L_{ict} + \beta X_i + v_{ic} \quad (1)$$

¹⁰ A simple measure of propensity to charge would be the proportion of all decisions made by a police officer where a POI was charged. In this case the value of the IV would be the same for all IPV events which are responded to by the same police decision-maker. However, at the individual level, this measure of a police officer's leniency would include the charge decision for each individual's own IPV event. This creates endogeneity between the IV and the treatment variable, as at the individual level, the IV influences charging, but charging also influences the IV. This endogeneity is eliminated by using a leave-one-out approach to construct the IV.

¹¹ For further discussion of residualisation of instrumental variables in a criminal justice context, see Rahman (2019).

Where P_{ic} is a binary variable equal to one if police officer c proceeds against POI i , and L_{ict} is our measure of police officer propensity to charge for officer c when attending an IPV assault incident involving POI i . Additionally, X_i are control variables (including sociodemographic characteristics, prior offending variables, and PAC by month dummies) and u_{ic} is the error term. The first stage model is estimated separately over the subsample of males and females. The predicted probabilities in the first stage ($\hat{P}_{ic}$) are then used to estimate the impact of being charged on the probability of conviction:¹²

$$C_{ic} = \gamma_0 + \gamma_1 * P_{ic} + \delta X_i + u_{ic} \quad (2)$$

Where C_{ic} is a binary variable equal to one if POI i is convicted in court and zero if they are charged and not convicted or they are not charged, and u_{ic} is the error term. The parameter of interest is γ_1 , which is estimated separately for males and females (denoted γ_1^m and γ_1^f) respectively. These parameters measure the impact of being charged on the probabilities of conviction for male and female defendants on the margins of being legally proceeded against (i.e. defendants who would be charged by a police officer who has a higher propensity to charge but not charged by a police officer with a lower propensity to charge). These parameters can also be interpreted as gender-specific conviction thresholds police apply when deciding whether to legally proceed against a POI in relation to an IPV incident. We test for the presence of police gender-bias by comparing the difference between γ_1^m and γ_1^f . Specifically, if $\gamma_1^f < \gamma_1^m$, this would indicate that women on the margins of being charged have a lower probability of conviction than men on the margins of being charged. This would be evidence in favour of police bias against women. Conversely, if $\gamma_1^f > \gamma_1^m$, this would indicate that men on the margins of being charged have a lower probability of conviction than women on the margins of being charged. This would be evidence in favour of police bias against men. Finally, if $\gamma_1^f = \gamma_1^m$, we would conclude that there is sufficient evidence to suggest that there are no differences in police decisions to charge women and men on the margins of being charged for IPV assault. These interpretations of the coefficients rest on several assumptions, including that police charge offenders in order to secure a conviction outcome, and magistrates exhibit no gender bias when determining guilt.

Instrumental variable assumptions

Our IV strategy also relies on several statistical assumptions. Firstly, the randomisation assumption requires that the assignment of officers to IPV assault incidents is as good as random. This assumption is satisfied in this instance since officers are assigned to DV incidents based on their location and time of the incident, and our propensity to charge measure accounts for the time and location of the incident. Assignment of police officers to IPV assault incidents should therefore be as good as random. We test this assumption in Appendix C1 by regressing our propensity to charge index on the control variables and fixed effects. This shows that our propensity to charge index is not correlated with observable POI characteristics, after controlling for the time and location of the incident.

Our instrument must also satisfy the exclusion restriction, which assumes that police officers do not affect conviction rates except through their decision to charge. While police officers do not make the ultimate decision to convict (as this rests with the defendant themselves, if they plead guilty, or the magistrate/judge, if the defendant doesn't plead guilty) police officers could affect conviction through other actions they take as part of the investigation (such as collecting better evidence, communicating with involved parties, and convincing victims to cooperate, or through providing testimony to the magistrate). The exclusion restriction would be violated if officers who have higher charging propensities are also better at building cases for evidence-led prosecutions. We test this by re-running our analysis using alternative measures of proceeding: 1) ever proceeded after the IPV incident; 2) proceeded against within 90 days of the IPV incident, and 3) proceeded against within 180 days of the IPV assault incident and present these results in Appendix C3. We argue that these proceeding windows are proxies for the strength of evidence against the POI. In particular, shorter windows imply stronger cases, as victims are better able to recall specific details of the incident, which can strengthen the evidence presented in court. We show that our results are consistent using these alternative measures of proceeding. This indicates that the exclusion restriction is likely to hold.¹³

¹² We cluster standard errors at the PAC level to account for the fact that observations are correlated between PACs.

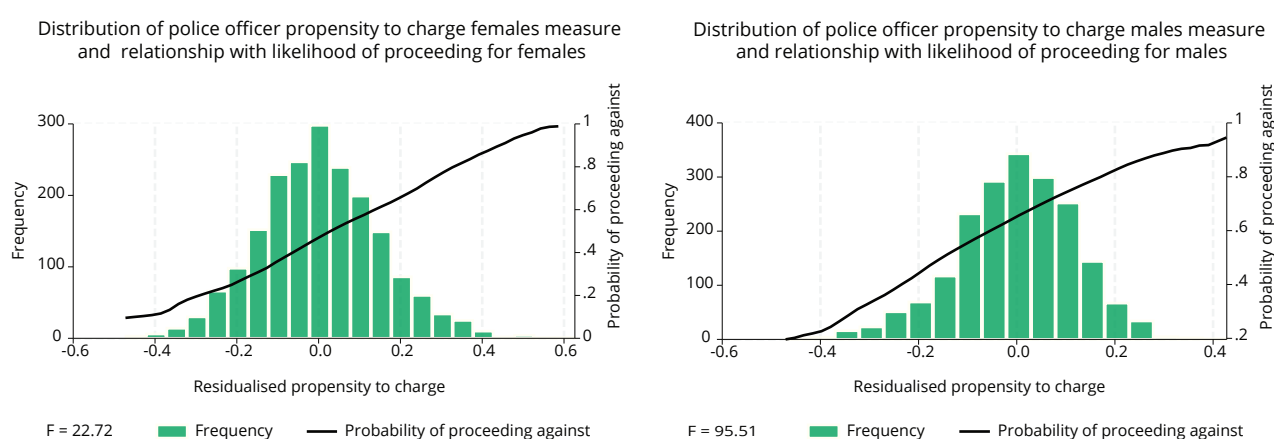
¹³ While this test provides some evidence that the exclusion restriction is satisfied, this assumption is not formally testable, as it implies the instrument is uncorrelated with the error term in Equation 2, which is unobservable.

Next, our instrument must satisfy the monotonicity assumption, which states that for a given incident, any POI charged by an officer with a lower propensity to charge would have also been charged by an officer with a higher propensity to charge, and any POI not charged by an officer with a higher propensity to charge would have also not been charged by an officer with a lower propensity to charge. While this assumption cannot be directly tested, monotonicity implies that the impact of police leniency on the probability that a POI is charged should be non-negative for any subsample. We test this by estimating the impact of our gender-specific propensity to charge measure on the probability that different subsamples of male and female POIs will be proceeded against in Appendix Table C2 (including subsamples over indicators for the sociodemographic characteristics and prior criminal offending variables listed in the Variables section). This shows that the impact of police propensity to charge on the probability of proceeding is positive and significant for almost all subgroups, suggesting the monotonicity assumption is satisfied.

Finally, the relevance assumption states that the propensity to charge of the police officer strongly influences whether a defendant will be charged or not. We test this assumption by examining the first stage regression defined in Equation 1, which models the impact of our propensity to charge measure on the probability a POI is proceeded against. We present estimates of the coefficient on L_{ict} from first stage regressions for men and women in Appendix Table C2. The coefficient on our propensity to charge measure is highly significant for both subsamples, with Olea-Pflueger average effective F-statistics of 92.57 and 18.96 for the male and female groups respectively (see Olea & Pflueger, 2013).¹⁴ This shows that the propensity to charge of the police officer is strongly associated with whether a POI will be charged or not.

Figure 1 plots the distribution of our police officer propensity to charge measure and the relationship between police officer propensity to charge and the probability a POI is charged, separately for males and females. This figure shows that, for both genders, a higher measure of police propensity to charge is correlated with a higher likelihood of proceeding, confirming the strong first stage regressions reported in Table C2. Further, the distribution of police officer propensity to charge approximately shows that most officers have a propensity to charge close to zero (meaning they are just as likely to charge as the average officer in the same PAC at the same time), while a small proportion of officers have a very high/ low propensity to charge.

Figure 1. Distribution of police officer propensity to charge and relationship with probability of proceeding, by gender



¹⁴ These F-statistics were estimated via STATA's weakivtest command (see Pflueger & Wang, 2015). The simulation study of Andrews et al. (2019) suggests that as a rule of thumb, an instrument is not weak if the Olea-Pflueger average effective F-statistic is greater than 10.

RESULTS

Descriptive statistics

We begin by examining the characteristics of female and male POIs identified in IPV assault events, and the difference between these groups is shown in Table 1. All differences are significant at the 5% significance level. We find that men were 2.7 times more likely to be identified as the POI in an IPV assault relative to women ($n = 38,413$ vs. $n = 14,010$). Female POIs were also less likely to be proceeded against by 18.4 percentage points (p.p.). When comparing sociodemographic characteristics of POIs in the entire sample, female POIs were younger than both male POIs and male victims during IPV assaults. For instance, female POIs were 1.3 p.p. more likely to be under 18 years old, 4.3 p.p. more likely to be in the 18-24 age range and were much more likely to be more than 11 years younger, 6-10 years younger or 2-5 years younger than the male victim of an IPV assault (by 9.9 p.p., 12.0 p.p., and 15.6 p.p. respectively). Male and female POIs were quite similar with respect to Aboriginality, socioeconomic disadvantage, and remoteness of residence. Across the entire sample, around 10-11% of POIs were identified as Aboriginal by police, close to 65% were from socioeconomically disadvantaged areas (in SEIFA quartile Q1 or Q2), and roughly 65% resided in major cities. However, there were large differences in the offending histories of both groups in the whole sample. In particular, female POIs were 16.9 p.p., 9.2 p.p., 5.8 p.p., 9.0 p.p. and 5.3 p.p. less likely to have one or more prior proven offences, sentenced prison episodes, proven ADVO breaches, proven DV offences or proven DV assault offences, respectively, within the last five years, relative to male POIs. Relative to male POIs, female POIs were also 18.5 p.p. more likely to have experienced prior victimisation and 25.9 p.p. more likely to have experienced prior DV victimisation within the last five years.

The fact that female POIs are less likely to be proceeded against does not necessarily indicate that police apply different standards of evidence when deciding whether to charge men and women for IPV assault. This is because decisions to legally proceed against a person are likely to be influenced by observable POI characteristics (such as prior offending history), as well as unobservable POI and case characteristics (such as offence seriousness, or the behaviours of the involved parties during the incident), and the distribution of these characteristics may differ across genders. In the following section, we present estimates from our IV regression which exploits the allocation of police officers to IPV incidents. We first estimate the proportion of men and women on the margins of being charged and then compare their probability of conviction. This removes any selection bias as the assignment of police officers to IPV incidents is random, conditional on the time and location of incident.

Table 1. Differences in the criminal justice and sociodemographic characteristics of POIs involved in IPV incidents

		Female (n=14,010)	Male (n=38,413)	Difference (F-M)	
Panel A. POI police outcomes (whole sample)					
Proceeded against by police (%)		45.91	64.29	-18.38	***
Panel B. POI offence characteristics (whole sample)					
Concurrent offences	1	16.54	22.36	-5.83	***
	2+	6.00	15.61	-9.61	***
	None	77.47	62.03	15.44	***
Panel C. POI sociodemographic characteristics (whole sample)					
Age categories (%)	Under 18	2.77	1.45	1.32	***
	18-24	19.44	15.17	4.27	***
	25-34	32.93	32.57	0.36	***
	35-44	27.44	29.54	-2.1	***
	45+	17.42	21.27	-3.85	***
Aboriginality (%)	Aboriginal	10.08	10.69	-0.62	*

Table 1. Differences in the criminal justice and sociodemographic characteristics of POIs involved in IPV incidents (..... continued)

		Female (n=14,010)	Male (n=38,413)	Difference (F-M)	
	Non-Aboriginal	73.60	73.14	0.46	*
	Unknown	16.32	16.17	0.16	*
Age difference between POI and victim categories (%)	11+ younger	12.62	2.75	9.87	***
	6-10 younger	17.43	5.46	11.97	***
	2-5 younger	27.73	12.11	15.62	***
	1 year younger-1 older	23.82	25.04	-1.23	**
	2-5 older	11.24	28.59	-17.36	***
	6-10 older	4.61	15.89	-11.28	***
	11+ older	2.56	10.15	-7.59	***
SEIFA quartile (%)	Q1 (most disadvantaged)	30.43	33.32	-2.89	***
	Q2	33.68	32.19	1.50	**
	Q3	21.98	21.19	0.79	**
	Q4 (least disadvantaged)	10.51	8.81	1.70	***
	Unknown	3.39	4.49	-1.10	***
Remoteness (%)	Inner regional	24.21	23.40	0.81	***
	Major city	65.43	65.49	-0.05	***
	Outer regional and remote	7.58	7.33	0.25	***
	Unknown	2.78	3.78	-1.01	***
Panel D. POI offending and victimisation within prior five years (whole sample)					
Proven offences in prior 5 years (%)	1-2	18.83	28.12	-9.29	***
	3-5	6.40	12.02	-5.63	***
	6+	1.87	3.85	-1.98	***
	None	72.91	56.01	16.89	***
Sentenced prison episodes in prior 5 years (%)	1	2.02	7.54	-5.52	***
	2+	0.82	4.47	-3.65	***
	None	97.16	87.99	9.17	***
Proven ADVO breach offences in prior 5 years (%)	1	3.30	7.60	-4.31	***
	2+	1.00	2.54	-1.54	***
	None	95.70	89.86	5.84	***
Proven DV offences in prior 5 years (%)	1	8.07	14.00	-5.92	***
	2+	2.31	5.34	-3.02	***
	None	89.61	80.67	8.95	***
Proven DV assault offences in prior 5 years (%)	1	6.45	10.87	-4.41	***
	2+	0.76	1.60	-0.85	***
	None	92.79	87.53	5.26	***
Proven victimisation in prior 5 years (%)	1	22.08	21.32	0.76	***
	2+	31.81	14.04	17.77	***
	None	46.11	64.64	-18.53	***
Proven DV victimisation in prior 5 years (%)	1	18.79	8.98	9.81	***
	2+	18.94	2.87	16.07	***
	None	62.27	88.15	-25.88	***

Note. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Instrumental variable estimates

Table 2 presents the results from our IV regression.¹⁵ The first two columns show our estimates of γ_1^f and γ_1^m , which represent the impact of charging a marginal POI for IPV assault involving heterosexual partners on their probability of a conviction, for females and males respectively. The table also shows the difference between these estimates, as well as the p-value of a statistical test to determine whether the estimates are statistically different. This involved a two-sample comparison of coefficients, using a bootstrapped Wald-type test that accounts for potential dependence across samples.¹⁶ If γ_1^f is lower than γ_1^m , it indicates that the probability of conviction is lower for marginal females than marginal males. This suggests bias in decision-making. That is, police officers may be applying a lower conviction threshold when considering whether to legally proceed against a female POI for an IPV assault. The first row in Table 2 presents our estimates while adjusting for our fixed effects (e.g. PAC and time fixed effects), while the second row adjusts for these fixed effects as well as the other controls described in the methods section (e.g. offender characteristic, prior offending). From Table 2 we can see that after controlling for all observable characteristics and fixed effects, marginal female POIs who are charged with IPV assault have a probability of conviction of 81.7% and marginal male POIs charged with IPV assault have a probability of conviction of 67.2%. This 14.5 p.p. difference between the two groups in the probability of conviction is not statistically significant at the 5% level ($p = 0.40$). This result suggests that police apply the same conviction threshold for males and females when deciding whether to take legal action in IPV assault matters. In other words, women and men with the same probability of being charged are convicted at the same rate.

Note that the interpretation of γ_1^f and γ_1^m as conviction thresholds police apply when deciding whether to charge female and male POIs involved in IPV assault rests on the assumption that magistrates exhibit no gender bias when determining guilt. In Appendix A, we investigate whether our IV results may be influenced by bias in conviction decisions. We find that conditional on being proceeded against for DV assault, 75% of men are convicted and 71% of women are convicted, reflecting a 4 p.p. difference. However, after controlling for prior offending, concurrent offending, demographics, judge fixed effects and time fixed effects, females were 2-3 p.p. less likely to be convicted. This could reflect a slight court bias in favour of women or omitted variable bias. This may suggest that our IV regression is overestimating γ_1^m and underestimating γ_1^f .

In Appendix D we also estimate the relationship between being proceeded and being convicted via Ordinary Least Squares (OLS,) separately by gender. This approach captures the impact of being charged with IPV assault on conviction for average rather than marginal defendants. We find that relative to our IV estimates, OLS underestimates the impact of charging on conviction for women and overestimates the impact of charging on conviction for men. This is likely due to inframarginality concerns with OLS. Specifically, females may be more likely to be low-risk POIs (who are more likely to never be charged), and males may be more likely to be high-risk POIs (who are more likely to always be charged). In this case, OLS estimates may reflect unobserved differences in case characteristics (e.g. evidence available to police), rather than the true effect of charging. The IV approach helps address inframarginality concerns by estimating the impact of charging on conviction for marginal POIs who are effectively randomly assigned into being charged or not depending on the charge propensity of the officer assigned to the IPV incident. As we would expect marginal POIs to have similar observable and unobservable characteristics, the IV approach allows us to compute more credible estimates of the impact of charging on conviction that are not influenced by unobserved differences in case characteristics.

¹⁵ Standard errors are clustered at the PAC level to account for the fact that observations are independent between PACs but correlated across a given PAC.

¹⁶ As men and women in the sample were drawn from the same PAC clusters, their estimation errors may be correlated. To account for this, the difference in coefficients γ_1^f and γ_1^m and its p-value were calculated using a nonparametric bootstrap procedure. This involved repeatedly re-sampling PAC clusters and re-estimating the model separately for each gender in 1,000 replicates, preserving the clustered data structure and any cross-sample correlation in the errors. It was not possible to implement the standard `suest` command as it is not compatible with `reghdfe`.

Table 2. 2SLS estimates of the impact of being proceeded against on the probability of being found guilty, by gender

		p-value of					R-squared	R-squared	Observations	Observations
		Female	Male	Difference	difference	Controls	Female	Male	Female	Male
Outcome										
Model 1	Probability of conviction	.821***	.709***	0.112	0.43	No	0.55	0.52	11,692	37,231
Model 2	Probability of conviction	.817***	.672***	0.145	0.40	Yes	0.55	0.52	11,692	37,231

Note. Standard errors in parentheses. Stars indicate statistical significance at a variety of conventional thresholds of statistical significance: * – 10%, ** – 5%, *** – 1%.

DISCUSSION

This study examined whether gender bias can be detected in experienced police officers' decisions to proceed against POIs in IPV assault matters involving heterosexual partners. Gender bias is exhibited if police apply different standards of evidence when choosing whether to charge female POIs compared to male POIs. To do this, we estimated the impact of being charged on the probability of conviction separately for male and female defendants on the margins of being legally proceeded against. These estimates can be interpreted as the difference in the implicit risk thresholds that experienced police officers apply when deciding whether to proceed against male and female POIs for IPV assault. Specifically, we find that marginal female defendants had a slightly higher conviction rate (81.7%) than marginal male defendants (67.2%), although this difference was not statistically significant.. Because the conviction rate of marginal women and men is effectively the same, we find no evidence of gender bias in police decisions to take legal action in relation to an IPV assault.

Our results are not consistent with Nancarrow et al.'s (2020) qualitative findings suggesting that women are more likely to be misidentified in DV matters. This may be because Nancarrow et al. (2020) were primarily concerned with misidentification in applications for civil law DV protection orders, which operate under a 'balance of probabilities' standard of proof, whereas our analysis is limited to criminal legal proceedings, which require proof beyond reasonable doubt. Because the issuance of protection orders requires less evidence, misidentification may be more likely to occur in civil proceedings. The divergence in findings also stems from the fact that our analysis is more limited in scope. In this study we test whether police respond differently to male and female POIs when initiating criminal legal proceedings in IPV matters. Our analysis does not answer the much broader question of whether police are accurately distinguishing the person perpetrating the violence from the person in need of protection. One concern often raised in the context of IPV is that the criminal law has traditionally taken a narrow definition of DV, which focuses on discrete incidents and visible forms of violence, rather than the entire history of abuse (both reported and unreported) suffered by the victim (Wangmann, 2012). By restricting the kinds of IPV behaviours that police are able to respond to, police may fail to recognise patterns of coercive control (experienced by up to 28% of DV victims; MacDonald et al., 2024) and self-defensive and retaliatory violence (used in up to half of DV incidents; see Boxall et al., 2020; Mackay et al., 2018; Wangmann, 2012). This is supported by Wang and Klauzner's (2025) analysis of 194 police narratives of IPV assault events involving heterosexual partners who were both legally proceeded against. The authors found that although the vast majority (67%) of partners shared a previous history of DV¹⁷, this appeared to play an ancillary role in police charging decisions. In addition, in cases where parties differed in their accounts of the circumstances surrounding the IPV incident, police often considered the presence of physical injury to determine the primary aggressor rather than the context in which the violence was committed. This

17 This is measured by considering 1) if one or both parties were defendants on an ADVO that protected the other party in the relationship; 2) if there were any previous DV events recorded where parties were involved as either the POI or victim (including verbal arguments); and 3) whether any unreported DV history was recorded in the police narrative.

focus on injury may contribute to misidentification of female victim-survivors as perpetrators in cases where women have used self-defensive violence (Nancarrow et al., 2020), or where the primary aggressor has self-harmed to appear to be the victim (Iliadis et al., 2024). Other research has similarly found that when responding to DV cases, the severity of violence strongly influences investigative effort and police charge decisions (Dowling et al., 2018), and that the absence of tangible evidence of physical violence can discourage police from continuing investigations (NSW Ombudsman, 2006). Given the criminal law's narrow definition of DV, and the degree to which police investigations and proceeding decisions focus on evidence of physical violence, our result that police apply similar conviction thresholds when deciding whether to proceed against male or female POIs is perhaps unsurprising.

Considering our results, the six-fold increase in the proportion of DV assault events where both heterosexual partners are charged over the last decade in NSW (from 0.4% of all IPV events in July 2010 to 2.5% in February 2023; Wang & Klauzner, 2025) may be influenced by factors beyond police bias in charge decisions. Instead, this may be an unintended consequence of greater police enforcement of DV and higher DV reporting rates, as evidenced by the 2.8% annual increase in the number of DV related incidents and the 4.4 p.p. increase in the 30-day legal action rate for DV assault incidents over the last decade (BOCSAR, 2024; Ramsey et al., 2022). If it is the case that police often focus on physical injury, increased rates of cross-IPV could reflect police proceeding against more victim-survivors who have used self-defensive or resistant violence during the index event (see Nancarrow et al., 2020), or alternatively, an increased use of violence by victim-survivors in response to more severe abuse (see Bair-Merritt et al., 2010).

Our study has several limitations. Firstly, the methodology used does not account for any other forms of bias that may exist in the criminal justice system (such as in NSW courts). For instance, Epstein and Goodman (2018) remark that judges often discount women's stories of abuse due to misunderstanding trauma responses and courtroom demeanour. Similarly, Nancarrow et al. (2020) suggest that courts may struggle to identify perpetrators in DV cases where women do not fit the stereotypical image of a victim. If police and court decision-makers are biased in the same direction (e.g. more inclined to charge women and more inclined to convict women), this would limit the extent to which police mischarging is later corrected by courts. Consequently, our outcomes test would overestimate the probability of conviction for women on the margin of being arrested and underestimate the degree of police bias. The degree of police bias would be similarly underestimated if police and courts are both biased against men. Secondly, we are only able to estimate implicit charge thresholds for more experienced officers, since our IV approach excludes police officers with a small number of decisions from our sample. This may lead to an underestimate of police bias in charge decisions, as past research indicates that more experienced officers are less likely to dually arrest victims and perpetrators (see Morris, 2009). This underestimate may be significant, as we excluded observations from 88% of police officers in order to implement our chosen method. Although it is worth noting that the officers that were included in the analysis attended one-third of all IPV assaults recorded in the period covered by the study. Thirdly, our findings may be influenced by omitted payoff bias¹⁸ if police officers charge POIs involved in DV cases for reasons other than to prove a charge beyond reasonable doubt. For example, police may consider arresting and charging POIs to diffuse a heated situation or to prevent the spike in DV revictimisation that occurs soon after a DV incident is reported to police (see Amaral et al., 2023). If police are more likely to charge and arrest male (or female) POIs with weaker cases in order to prevent revictimisation, then our outcomes test would underestimate the impact of charging on conviction for marginal male (or female) POIs. In NSW, police guidelines recommend that officers give consideration to arresting DV offenders if there is an immediate threat to the victim's safety (see NSW Police, 2018b).¹⁹ However, to charge and subsequently arrest an offender, police must first be confident that an offence has been committed and that the charge can be proved beyond reasonable doubt.

¹⁸ For more details, see Canay et al., (2024), Gelbach (2021) and Hull (2021).

¹⁹ In our sample, around 85% of POIs charged for DV assault are subsequently arrested.

Another limitation is that our empirical strategy cannot detect accurate statistical discrimination, where individuals are treated differently based on the average characteristics of the group they belong to (Arnold et al., 2018). In the context of our study, this could involve a police officer being more likely to proceed against men for DV assault because they (correctly) believe that men, in general, are more likely to be found guilty by plea or verdict. Alternatively, statistical discrimination could arise if police are more likely to charge women since they (correctly) anticipate that women are more likely to plead guilty. If police bias is driven entirely by statistical discrimination, an outcomes test would find no evidence of bias. However, the outcomes test that we utilise can detect taste-based discrimination (where police officers are more or less likely to proceed against a POI simply because of their gender) and inaccurate statistical discrimination, (where charge decisions are influenced by incorrect gender stereotypes about the probability of guilt).²⁰

While we find no evidence that experienced police officers are more inclined to charge women with IPV assault compared with similar men (i.e. gender bias), this does not mean that misidentification does not occur. Adopting an incident-based approach to DV may fail to accurately distinguish between the person most commonly perpetrating the violence and the person more often in need of protection (see Wangmann, 2010; Nancarrow, 2019; Nancarrow et al, 2020). Partially in light of such concerns, in July 2024 the *Crimes Legislation Amendment (Coercive Control) Act 2022* (NSW) commenced in NSW which criminalised coercive controlling behaviours occurring as a pattern of behaviour. This coincided with the introduction of a new legislative definition of domestic abuse in the *Crimes (Domestic and Personal Violence) Act 2007* (NSW) which emphasises that DV can take the form of both a single act and a course of conduct. Our research predates these legislative changes.²¹ It is possible that the introduction of the coercive control offence, the new understanding of domestic abuse, along with the significant police, legal sector, judicial, DFV sector training and community awareness campaigns accompanying these changes, may have altered the justice response to DFV. Further research should be undertaken to evaluate the extent to which these policies are effective in helping to identify the person in need of protection and ensuring victim safety.

ACKNOWLEDGEMENTS

The authors would like to extend thanks to Suzanne Poynton, Jackie Fitzgerald, Stewart Boiteux, Jonathan Gu and Nicole Adams from BOCSAR and three anonymous peer reviewers for helpful comments. We would also like to thank Bonnie Ross for proofreading and Mark Ramsay and Derek Goh for helping prepare the data used in this bulletin. Florence Sin undertook desktop publishing of this report.

²⁰ See Klauzner (2023) for more information on types of discrimination that can be detected via an outcomes test.

²¹ These changes took effect on 1 July 2024.

REFERENCES

- Adams, A., Huttunen, K., Nix, E., & Zhang, N. (2024). The dynamics of abusive relationships. *The Quarterly Journal of Economics*, 139(4), 2135-2180. <https://doi.org/10.1093/qje/qjae022>.
- Amaral, S., Dahl, G. B., Endl-Geyer, V., Hener, T., & Rainer, H. (2023). *Deterrence or backlash? Arrests and the dynamics of domestic violence* (No. w30855). Retrieved from the National Bureau of Economic Research website at: <https://doi.org/10.3386/w30855>.
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11(1), 727-753. <http://dx.doi.org/10.1146/annurev-economics-080218-025643>.
- Arnold, D., Dobbie, W., & Hull, P. (2022). Measuring racial discrimination in bail decisions. *American Economic Review*, 112(9), 2992-3038. <https://doi.org/10.1257/aer.20201653>.
- Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics*, 133(4), 1885-1932. <https://doi.org/10.1093/qje/qjy012>.
- Audit Office of NSW. (2022). *Performance audit: Police responses to domestic and family violence*. Retrieved from the Audit Office of NSW website at: <https://www.audit.nsw.gov.au/our-work/reports/police-responses-to-domestic-and-family-violence>
- Australia's National Research Organisation for Women's Safety. (2019). *Domestic and family violence, housing insecurity and homelessness*. Retrieved from the Australia's National Research Organisation for Women's Safety website at: https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2019/03/08034450/DV-Housing-Homelessness-Synthesis-2.Ed_.pdf.
- Australian Bureau of Statistics. (2011). *Australian and New Zealand Standard Offence Classification (ANZSOC)* (cat. no. 1216.0). Retrieved from the Australian Bureau of Statistics website at: <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-offence-classification-anzsoc/latestrelease>.
- Australian Bureau of Statistics. (2016a). *Census of population and housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2016* (cat. no. 2033.0.55.001). Retrieved from the Australian Bureau of Statistics website at: <https://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001>.
- Australian Bureau of Statistics. (2016b). *Australian Standard Geographical Classification (ASGC)* (cat. no. 1216.0). Retrieved from the Australian Bureau of Statistics website at: <https://www.abs.gov.au/ausstats/abs@.nsf/%20mf/1270.0.55.005>.
- Australian Bureau of Statistics. (2023). *Personal Safety, Australia*. Retrieved from the Australian Bureau of Statistics website at: <https://www.abs.gov.au/statistics/people/crime-and-justice/personal-safety-australia/latest-release>.
- Australian Institute of Health and Welfare (2024). *Family, domestic and sexual violence*. Retrieved from the Australian Institute of Health and Welfare website at: <https://www.aihw.gov.au/family-domestic-and-sexual-violence/resources/fdsv-summary>.
- Bair-Merritt, M. H., Shea Crowne, S., Thompson, D. A., Sibinga, E., Trent, M., & Campbell, J. (2010). Why do women use intimate partner violence? A systematic review of women's motivations. *Trauma, Violence, & Abuse*, 11(4), 178-189. <https://doi.org/10.1177/1524838010379003>.
- Beeble, M. L., Adams, A. E., & Lee, K. (2022). Situational, community, and state policy-level factors associated with arrest in incidents of intimate partner violence. *Crime and Justice Review*, 48(3), 377-402. <https://journals.sagepub.com/Doi/Full/10.1177/07340168221115407>.

Belknap, J., Graham, D., Hartman, J., Lippen, V., Allen, G., & Sutherland, J. (2000). *Factors related to domestic violence court dispositions in a large urban area: The role of victim/witness reluctance and other variables*. Cincinnati, OH, USA: University of Cincinnati. Retrieved from the US Department of Justice Office of Justice Programs website at: <https://www.ojp.gov/pdffiles1/nij/grants/184232.pdf>.

Boxall, H., Dowling, C., & Morgan, A. (2020). *Female perpetrated domestic violence: Prevalence of self-defensive and retaliatory violence* (Trends & Issues in Crime and Criminal Justice No. 584). Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti04176>

Canay, I. A., Mogstad, M., & Mountjoy, J. (2024). On the use of outcome tests for detecting bias in decision making. *Review of Economic Studies*, 91(4), 2135-2167. <https://doi.org/10.1093/restud/rdad082>.

Clarke, D., Romano, J. P., & Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in Stata. *The Stata Journal*, 20(4), 812-843. <https://doi.org/10.1177/1536867X20976314>.

Correia, S. (2018). Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix. Retrieved from: <https://scoreiria.com/research/singletons.pdf>.

Donnelly., (2024). *Trends in female offending in New South Wales: 2014 to 2023* (Bureau Brief No. 173). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/bb/bb151-200/BB173-Report-female-offending-trends.pdf>.

Dowling, C., Morgan, A., Boyd, C., & Voce, I. (2018). *Policing domestic violence: A review of the evidence* (Research Report No. 13). Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/rr198147>.

Eitle, D. (2005). The influence of mandatory arrest policies, police organizational characteristics, and situational variables on the probability of arrest in domestic violence cases. *Crime & Delinquency*, 51(4), 573-597. <https://doi.org/10.1177/0011128705277784>.

Epstein, D., & Goodman, L. A. (2018). Discounting women: Doubting domestic violence survivors' credibility and dismissing their experiences. *University of Pennsylvania Law Review*, 167(2), 399-461. Retrieved from the University of Pennsylvania website at: https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=9644&context=penn_law_review.

Felson, R. B., & Pare, P. P. (2007). Does the criminal justice system treat domestic violence and sexual assault offenders leniently? *Justice Quarterly*, 24(3), 435-459. <https://doi.org/10.1080/07418820701485601>.

Freeman, K. (2023) *Domestic and Family Violence trends in NSW, July 2010 to June 2022: update* (Bureau Brief No. 167). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/bb/bb151-200/bb167-report-dfv-trends-nsw.pdf>.

Frye, V., Haviland, M., & Rajah, V. (2007). Dual arrest and other unintended consequences of mandatory arrest in New York City: A brief report. *Journal of Family Violence*, 22, 397-405. <https://doi.org/10.1007/s10896-007-9094-y>.

Gelbach, J. B. (2021). *Testing economic models of discrimination in criminal justice* (working paper). Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3784953.

Henning, K., & Feder, L. (2005). Criminal prosecution of domestic violence offenses: An investigation of factors predictive of court outcomes. *Criminal Justice and Behavior*, 32(6), 612-642. <https://doi.org/10.1177/0093854805279945>.

Hirschel, D., McCormack, P. D., & Buzawa, E. (2021). A 10-year study of the impact of intimate partner violence primary aggressor laws on single and dual arrest. *Journal of Interpersonal Violence*, 36(3-4), 1356-1390. <https://doi.org/10.1177/0886260517739290>.

- Holmes, J., & Fitzgerald, J. (2017). *Trends in NSW Police clear up rates* (Bureau Brief No. 129). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/bb/bb101-150/bb129-report-trends-in-nsw-police-clear-up-rates.pdf>.
- Hull, P. (2021). *What marginal outcome tests can tell us about racially biased decision-making* (Working Paper No. 28503). Retrieved from the National Bureau of Economic Research website at: <https://www.nber.org/papers/w28503>.
- Iliadis, M., Harris, B., Vakhitova, Z., Woodlock, D., Flynn, A., & Tyson, D. (2024). *How police body-worn cameras can facilitate mischarging in domestic and family violence responses* (Trends & Issues in Crime and Criminal Justice No. 684). Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti77277>.
- Jillard, A., & Mansour, J. (2014). Women victims of violence defending intervention orders: The latest developments in practice and policy in NSW. *Alternative Law Journal*, 39(4), 235–240. <https://doi.org/10.1177/1037969X1403900407>.
- Klauzner, I. (2023). *Investigating bias towards Aboriginal people in police bail decisions* (Crime and Justice Bulletin No. 256). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/cjb/cjb251-300/cjb256-report-investigating-bias-in-police-bail-decisions-2023.pdf>.
- KPMG. (2016). The cost of violence against women and their children in Australia. *Canberra: Australian Government Department of Social Services*. Retrieved from the Australian Women's Health Alliance website at: https://australianwomenshealth.org/wp-content/uploads/2024/03/the_cost_of_violence_against_women_and_their_children_in_australia_-_summary_report_may_2016.pdf.
- Larance, L. Y., Goodmark, L., Miller, S. L., & Dasgupta, S. D. (2019). Understanding and addressing women's use of force in intimate relationships: A retrospective. *Violence Against Women*, 25(1), 56–80. <https://doi.org/10.1177/1077801218815776>.
- MacDonald, J. B., Willoughby, M., Gartoulla, P., Cotton, E., March, E., Alla, K., & Strawa, C. (2024). What the research evidence tells us about coercive control victimisation. Retrieved from the Australian Institute of Family Studies website at: https://aifs.gov.au/sites/default/files/2024-02/2311_CFCA_Coercive-control-victimisation.pdf.
- Mackay, J., Bowen, E., Walker, K., & O'Doherty, L. (2018). Risk factors for female perpetrators of intimate partner violence within criminal justice settings: A systematic review. *Aggression and Violent Behavior*, 41, 128–146. <https://doi.org/10.1016/j.avb.2018.06.004>.
- Morgan, A., & Boxall, H. (2022). *Economic insecurity and intimate partner violence in Australia during the COVID-19 pandemic* (Research report, 02/2022). Australia's National Research Organisation for Women's Safety. Retrieved from the ANROWS website at: <https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2022/01/28160540/Economic-insecurity-and-IPV-during-the-C19-RR2.pdf>.
- Morris, P. W. (2009). *Dual arrest in intimate partner violence incidents: The influence of police officer, incident, and organizational characteristics* [Doctoral dissertation, The Graduate Center, CUNY]. Available from Proquest Dissertations and Theses database (UMI No. 3378607).
- Nancarrow, H. (2019) *Unintended consequences of domestic violence law: Gendered aspirations and racialised realities*. Cham, Switzerland: Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-27500-6>.
- Nancarrow, H., Thomas, K., Ringland, V., & Modini, T. (2020). *Accurately identifying the "person most in need of protection" in domestic and family violence law* (Research report, 23/2020). Australia's National Research Organisation for Women's Safety. Retrieved from the ANROWS website at: <https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2019/10/25104930/Nancarrow-PMINOP-RR.3.pdf>.

NSW Bureau of Crime Statistics and Research. (2024). *NSW Trends in Domestic & Family Violence – Quarterly Report: Data to December 2024*. Retrieved from the NSW Bureau of Crime Statistics and Research website at: https://bocsar.nsw.gov.au/documents/topic-areas/domestic-violence/Domestic_Violence_Quarterly_Trends_Report.pdf.

NSW Law Reform Commission (2010). *Family Violence —A National Legal Response*. Retrieved from the Australian Law Reform Commission website at: https://www.alrc.gov.au/wp-content/uploads/2019/08/ALRC114_WholeReport.pdf.

NSW Ombudsman. (2006). *Domestic violence: Improving police practice*. Retrieved from the NSW Ombudsman website at: <https://cmsassets.ombo.nsw.gov.au/assets/Reports/Domestic-violence-improving-police-practice-Special-Report-to-Parliament-December-2006-.pdf.pdf>.

NSW Police. (2018a). *Code of practice for the NSW Police Force response to domestic and family violence*. Retrieved from the NSW Police website at: https://www.police.nsw.gov.au/_data/assets/pdf_file/0016/165202/Code_of_Practice_for_the_NSWPF_response_to_Domestic_and_Family_Violence.pdf.

NSW Police. (2018b). *Domestic and family violence policy*. Retrieved from the NSW Police website at: https://www.police.nsw.gov.au/_data/assets/pdf_file/0007/862063/Domestic_and_Family_Violence_Policy_2018.pdf.

Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, 61(1), 5-21. [https://doi.org/10.1016/0304-4076\(94\)90074-4](https://doi.org/10.1016/0304-4076(94)90074-4).

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358-369. <https://doi.org/10.1080/00401706.2013.806694>.

Pflueger, C. E., & Wang, S. (2015). A robust test for weak instruments in Stata. *The Stata Journal*, 15(1), 216-225. <https://doi.org/10.1177/1536867x1501500113>.

Rahman, S. (2019). *The marginal effect of bail decisions on imprisonment, failure to appear, and crime* (Crime and Justice Bulletin No. 224). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/cjb/cjb201-250/cjb224-report-bail-decisions-on-imprisonment-2019.pdf>.

Rajah, V., Frye, V., & Haviland, M. (2006). "Aren't I a Victim?": Notes on identity challenges relating to police action in a mandatory arrest jurisdiction. *Violence Against Women*, 12(10), 897-916. <https://doi.org/10.1177/1077801206292872>.

Ramsey, S., Kim, M.T., & Fitzgerald, J. (2022). *Trends in domestic violence-related stalking and intimidation offences in the criminal justice system: 2012 to 2021* (Bureau Brief No. 159). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/bb/bb151-200/bb159-report-dv-related-stalking.pdf>.

Reeves, E. (2019). Family violence, protection orders and systems abuse: views of legal practitioners. *Current Issues in Criminal Justice*, 32(1), 91-110. <https://doi.org/10.1080/10345329.2019.1665816>.

Reeves, E. (2021). 'I'm not at all protected and I think other women should know that, that they're not protected either': Victim-survivors' experiences of 'misidentification' in Victoria's family violence system. *International Journal for Crime, Justice and Social Democracy*, 10(2). <https://doi.org/10.5204/IJCJS.1992>.

Reeves, E. (2023). A culture of consent: Legal practitioners' experiences of representing women who have been misidentified as predominant aggressors on family violence intervention orders in Victoria, Australia. *Feminist Legal Studies*, 31(3), 369-390. <https://doi.org/10.1007/s10691-022-09506-5>.

Romano J. P., & Wolf M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100: 94-108. <https://doi.org/10.1198/016214504000000539>.

Shuard, J. (2021). Monitoring Victoria's family violence reforms: Accurate identification of the predominant aggressor. Retrieved from the Family Violence Reform Implementation Monitor website at: https://www.fvrim.vic.gov.au/sites/default/files/2021-12/FVRIM%20Predominant%20Aggressor%20December%202021_0.pdf

Starr, S. B. (2015). Estimating gender disparities in federal criminal cases. *American Law and Economics Review*, 17(1), 127-159. <https://www.jstor.org/stable/24735732>.

Teperski, A., & Rahman, S. (2023). *Why are Aboriginal adults less likely to receive cannabis cautions* (Crime and Justice Bulletin No. 258). Retrieved from the NSW Bureau of Crime Statistics and Research website at: <https://bocsar.nsw.gov.au/documents/publications/cjb/cjb251-300/cjb258-report-cannabis-cautioning-2023.pdf>

Victoria Police. (2021). *External agency briefing: Misidentification of the primary aggressor – A 2021 update* (not published).

Wang, S., & Klauzner, I. (2025). *The nature and extent of cross-intimate partner violence*. (Bureau Brief No. 178). Sydney: NSW Bureau of Crime Statistics and Research. Retrieved from <https://bocsar.nsw.gov.au/documents/publications/bb/bb151-200/BB178-Report-The-nature-and-extent-of-cross-intimate-partner-violence.pdf>

Wangmann, J. (2010). Gender and intimate partner violence: A case study from NSW. *University of New South Wales Law Journal*, 33(3), 945-969. Retrieved from the UNSW Law Journal website at: <https://unswlawjournal.unsw.edu.au/wp-content/uploads/2017/09/33-3-4.pdf>.

Wangmann, J. (2012). Incidents v context: *How does the NSW protection order system understand intimate partner violence?* *Sydney Law Review*, The, 34(4), 695-719. <https://search.informit.org/doi/abs/10.3316/INFORMIT.060410648833960>.

Wangmann, J. (2014). *Domestic violence orders need stronger enforcement*. Retrieved from The Conversation website at: <https://theconversation.com/domestic-violence-orders-need-stronger-enforcement-29910>

Weatherburn, D., Doyle, M., Weatherall, T., & Teperski, A. (2025). Does adjusting for legal factors make the racial gap in imprisonment risks disappear? *Journal of Criminology*, 26338076251329689. <https://doi.org/10.1177/26338076251329689>

Webster, K. (2016). *A preventable burden: Measuring and addressing the prevalence and health impacts of intimate partner violence in Australian women* (ANROWS Compass, 07/2016). Retrieved from the ANROWS website at: <https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2019/01/19025309/28-10-16-BOD-Compass.pdf>.

Wolf, M. E., Ly, U., Hobart, M. A., & Kernic, M. A. (2003). Barriers to seeking police help for intimate partner violence. *Journal of Family Violence*, 18, 121-129. <https://doi.org/10.1023/A:1022893231951>.

APPENDIX

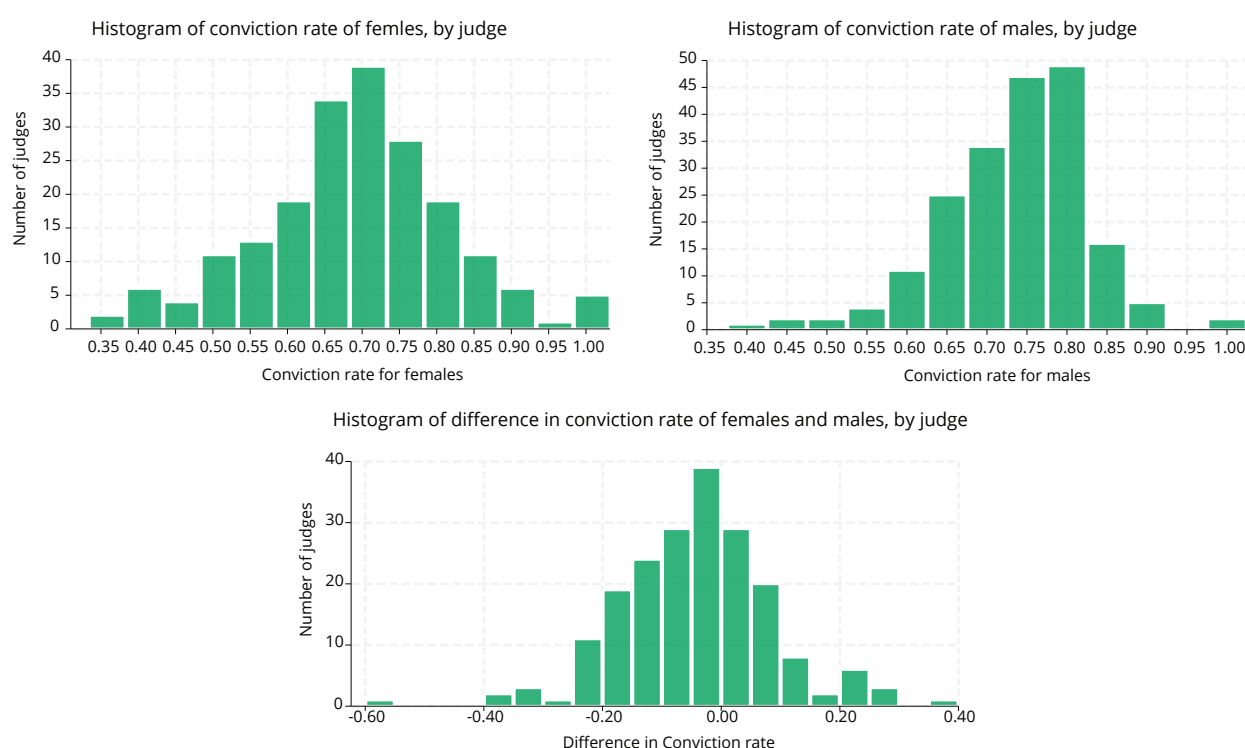
Appendix A – Investigating bias in conviction decisions, conditional on being proceeded against

The subsample of POIs in our sample who were proceeded against for DV assault includes 24,694 men and 6,432 women. Of these POIs, 18,499 (or 75% of) men and 4,572 (or 71% of) women were found guilty by guilty plea or verdict. In Figure A1, we plot the average conviction rate of female POIs by judge, the average conviction rate of male POIs by judge, and the difference in conviction rates of female and male POIs by judge. Note that for the purposes of creating these figures, of the 376 judges in our proceeded

subsample, we drop observations of 145 judges who have made less than five conviction decisions involving female POIs, and 33 judges who have made less than five conviction decisions involving male POIs. There is substantial variation in conviction rates of females around the median of 70%, and for males around the median of 74%. Further, judges are slightly more likely to convict men, as the one sample t-test that the difference in the conviction rate between females and males by judge is equal to zero is rejected at the 5% level ($p = 0.02$). However, this does not necessarily imply judges are biased against men, as it could reflect differences in case characteristics, prior offending histories, or concurrent offences of men versus women.

To investigate factors which may drive this difference in conviction rates, we use the Kitagawa-Oaxaca-Blinder (KOB) decomposition method. This splits disparities in conviction rates between men and women into an 'explained' and 'unexplained' component. The explained component estimates how much of the raw disparity in conviction rates between men and women can be attributed to differences in observable characteristics (including prior offending histories, concurrent offences, demographic factors, time and judge). The unexplained component measures how much of the raw conviction disparity can be attributed to omitted variable bias and/or discrimination.²² As a robustness test, we estimate a pooled decomposition which uses reference coefficients from a pooled male and female group (see Oaxaca & Ransom, 1994).²³ As shown in Table A1, this decomposition suggests that 2 p.p. of the disparity in conviction rates between men and women arises from men having more extensive prior offending histories and committing concurrent offences than women. The remaining 2 p.p. of the conviction disparity is unexplained, which could reflect court bias against men or omitted variable bias. This is consistent with US studies which find that men were consistently treated more severely at every stage of the court process even when legal and extra-legal variables were controlled for (see Henning & Feder, 2005; Starr, 2015).

Figure A1: Histograms of conviction rates of females, males and difference in conviction rates of females and males, by judge



²² For more details on the use of the Kitagawa-Oaxaca-Blinder decomposition in a justice setting, see Teperski and Rahman (2023) and Weatherburn et al. (2025).

²³ A separate regression of the impact of gender on conviction, while controlling for prior offending, demographics, judge fixed effects, concurrent offences and time fixed effects, also reaches a similar conclusion. Specifically, on average, women are 3 p.p. less likely to be convicted after controlling for all observable characteristics.

Table A1. Kitagawa-Oaxaca-Blinder results for proceeded subsample

	(1) Oaxaca	(2) Oaxaca Pooled
Overall		
Male conviction rate	0.75*** (0.01)	0.75*** (0.01)
Female conviction rate	0.71*** (0.01)	0.71*** (0.01)
Difference	0.04*** (0.00)	0.04*** (0.00)
Explained	0.02*** (0.00)	0.02*** (0.00)
Unexplained	0.02*** (0.00)	0.02*** (0.00)
Explained		
Prior offending	0.02*** (0.00)	0.02*** (0.00)
Demographics	-0.00*** (0.00)	0.00 (0.00)
Judge fixed effects	0.00 (0.00)	0.00 (0.00)
Concurrent offences	0.01*** (0.00)	0.01*** (0.00)
Month-year fixed effects	-0.00*** (0.00)	-0.00*** (0.00)
Unexplained		
Prior offending	0.26*** (0.00)	0.27*** (0.00)
Demographics	0.06*** (0.00)	0.06*** (0.00)
Judge fixed effects	1.78*** (0.00)	1.78*** (0.00)
Concurrent offences	0.20*** (0.00)	0.20*** (0.00)
Month-year fixed effects	-0.07*** (0.00)	-0.07*** (0.00)
Constant	-2.22 (.)	-2.22 (.)
<i>N</i>	30920	30920

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B – Comparison of events and officers in and out of sample

Table B1 compares the observable characteristics of POIs included in our analysis to those who were excluded. Of our cleaned sample of 147,553 observations, 95,130 observations were excluded as charging officers had attended too few DV assaults for us to construct our measure of police propensity to charge. The remaining 52,423 observations were included in our sample since charging officers had attended enough IPV assault events to construct our police propensity to charge measure. POIs included in our sample were less likely to be proceeded against, more likely to reside in socioeconomically disadvantaged and remote areas, and were slightly less likely to have experienced prior victimisation/prior DV victimisation.

Table B1. Differences in the criminal justice and sociodemographic characteristics of POIs involved in IPV incidents who were included in our sample vs those excluded from our sample

		Included in sample (n=52,423)	Excluded from sample (n=95,130)	Difference (T-C)	
Panel A. POI outcomes (whole sample)					
Proceeded against (%)		59.37	67.15	-7.77	***
Panel B. POI court outcomes (conditional on being proceeded against)					
Withdrawn (%)		17.85	19.29	-1.44	***
Guilty by verdict or plea (%)		74.14	72.11	-2.03	***
Guilty by plea (%)		49.31	48.86	0.45	***
Panel C. Offence characteristics (whole sample)					
Concurrent offences (%)	1	22.26	20.81	1.45	***
	2 or more	13.93	13.04	0.89	***
	None	63.81	66.15	-2.34	***
Panel D. POI sociodemographic characteristics (whole sample)					
Age categories (%)	Under 18	1.80	1.76	0.04	***
	18-24	16.31	15.46	0.86	***
	25-34	32.66	32.45	0.21	***
	35-44	28.98	28.98	0.00	***
	45+	20.24	21.35	-1.11	***
Aboriginality (%)	Aboriginal	10.53	12.09	-1.56	***
	Non-Aboriginal	73.26	73.81	-0.55	*
	Unknown	16.21	14.09	2.12	***
Age difference between POI and victim categories (%)	11+ younger	5.39	4.79	0.6	***
	6-10 younger	8.66	7.46	1.19	***
	2-5 younger	16.29	14.77	1.52	***
	1 year younger-1 older	24.72	24.58	0.13	***
	2-5 older	23.96	25.03	-1.07	***
	6-10 older	12.88	14.18	-1.31	***
	11+ older	8.12	9.18	-1.07	***
SEIFA quartile (%)	Q1 (most disadvantaged)	32.55	31.01	1.54	***
	Q2	32.59	26.21	6.38	***
	Q3	21.40	23.28	-1.87	***
	Q4 (least disadvantaged)	9.27	14.47	-5.21	***
	Unknown	4.19	5.03	-0.84	***
Remoteness (%)	Inner regional	23.62	20.78	2.84	***
	Major city	65.47	65.79	-0.32	***
	Outer regional and remote	7.40	9.56	-2.16	***
	Unknown	3.51	3.87	-0.36	***
Panel E. Offending within prior 5 years (whole sample)					
Proven offences in prior 5 years (%)	1-2	25.63	25.54	0.09	*
	3-5	10.52	10.67	-0.15	*
	6+	3.32	3.75	-0.43	***
	None	60.53	60.04	0.49	***
Sentenced prison episodes in prior 5 years (%)	1	6.07	6.46	-0.39	**
	2+	3.49	3.86	-0.37	***
	None	90.44	89.67	0.77	***
Proven ADVO breach offences in prior 5 years (%)	1	6.45	6.57	-0.11	***
	2+	2.13	2.36	-0.24	**
	None	91.42	91.07	0.35	*

Table B1. Differences in the criminal justice and sociodemographic characteristics of POIs involved in IPV incidents who were included in our sample vs those excluded from our sample (*continued*)

		Included in sample (n=52,423)	Excluded from sample (n=95,130)	Difference (T-C)	
Proven DV offences in prior 5 years (%)	1	12.41	12.53	-0.11	***
	2+	4.53	4.53	0.00	***
	None	83.06	82.94	0.12	***
Proven DV assault offences in prior 5 years (%)	1	9.69	9.85	-0.16	***
	2+	1.38	1.33	0.05	***
	None	88.94	88.82	0.12	***
Proven victimisation in prior 5 years (%)	1	21.52	20.65	0.87	***
	2+	18.79	16.43	2.36	***
	None	59.69	62.92	-3.23	***
Proven DV victimisation in prior 5 years (%)	1	11.61	10.37	1.24	***
	2+	7.16	5.31	1.85	***
	None	81.23	84.32	-3.08	***

Note. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Table B2 compares the experience of officers included in our sample to those who were excluded from our analysis. Experience is proxied by recording the earliest year an officer responded to an IPV event in our sample. For example, we observe that roughly 31.6% of officers included in our sample responded to their earliest IPV event in 2010, while around 19.0% of officers excluded from our sample responded to their earliest IPV event in the same year. Overall, officers included in our sample were more likely to respond to their first IPV event in an earlier year (closer to 2010), rather than a later year (closer to 2023). This indicates that on average, officers included in our analysis sample were more experienced.

Table B2. Differences in the year that officers responded to their first IPV event who were included in our sample vs those excluded from our sample

	Included in sample (n=1953)	Excluded from sample (n=13,820)	Difference (T-C)	
Year of first IPV event responded to by officer (%)				
2010	31.64	19.06	12.58	***
2011	18.33	15.82	2.51	**
2012	13.88	9.04	4.84	***
2013	9.01	7.07	1.94	**
2014	5.79	4.60	1.18	*
2015	6.04	4.51	1.53	**
2016	3.38	3.97	-0.59	**
2017	3.33	3.63	-0.30	**
2018	2.41	4.06	-1.65	***
2019	3.07	5.40	-2.33	***
2020	2.15	6.35	-4.20	***
2021	0.87	6.92	-6.05	***
2022	0.10	8.04	-7.94	***
2023	0.00	1.53	-1.53	***

Note. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Appendix C – IV assumption testing and diagnostics

This section tests the assumptions of our IV strategy (randomisation, relevance, monotonicity and exclusion).

Appendix C1 – Randomisation test

Table C1 shows results from a regression of our residualised measure of police propensity to charge on our controls. Since we assume the assignment of police officers to DV incidents is as good as random, most coefficients should be equal to zero, and the joint test of significance should not be rejected. The joint test of significance is rejected for both the male and female subsamples at the 5% significance level. However, this does not account for the risk of false positives when testing many coefficients simultaneously, or potential correlation between regressors. To account for this, we calculate Romano-Wolf stepdown adjusted p-values, which control the family-wise error rate (FWER), or the probability of making any type I error (see Romano & Wolf, 2005)²⁴. Nearly all coefficients are not statistically significant at the 5% significance level after Romano-Wolf adjustment. The exception is the coefficient of the no concurrent offences dummy in the male subsample, which is statistically significant after adjustment ($p = 0.00$). However, the estimated coefficient is very small (-0.01). This suggests that the assignment of police officers to DV incidents is as good as random after controlling for PAC and time.

Table C1. Regression of police instrumental variable on all observed characteristics

Variable	Male			Female		
	Residualised propensity to charge coefficient	Unadjusted p-value	Romano-Wolf FWER adjusted p-value	Residualised propensity to charge coefficient	Unadjusted p-value	Romano-Wolf FWER adjusted p-value
Concurrent offences						
1	-	-	-	-	-	-
2 or more	0.00	0.31	0.99	0.01	0.26	0.96
None	-0.01	0.00	0.00	0.00	0.40	1.00
Age categories						
18-24	-	-	-	-	-	0.98
25-34	0.00	0.99	1.00	0.01	0.30	0.28
35-44	0.00	1.00	1.00	0.01	0.03	0.28
45+	0.00	0.90	1.00	0.01	0.03	1.00
NA	0.00	0.84	1.00	0.04	0.60	1.00
Under 18	0.00	0.76	1.00	-0.01	0.50	0.98
Age difference between POI and victim						
1 year younger-1 older	-	-	-	-	-	-
11+ older	0.00	0.69	1.00	-0.01	0.32	0.99
11+ younger	0.00	0.65	1.00	0.00	0.74	1.00
2-5 older	0.00	0.99	1.00	0.00	0.92	1.00
2-5 younger	-0.00*	0.04	0.22	0.00	0.68	1.00
6-10 older	0.00	0.29	0.97	-0.01	0.57	1.00
6-10 younger	-0.01	0.02	0.09	0.00	0.86	1.00
NA	-0.01	0.45	1.00	-0.06	0.35	1.00
Aboriginality						
Aboriginal	-	-	-	-	-	-
Non-Aboriginal	0.01*	0.09	0.54	0.01	0.27	0.98
Unknown	-0.02	0.02	0.14	-0.01	0.29	0.98

²⁴ We estimate these p-values via the `rwolf2` command in STATA (see Clarke et al., 2020).

Table C1. Regression of police instrumental variable on all observed characteristics (*continued*)

Variable	Male			Female		
	Residualised propensity to charge coefficient	Unadjusted <i>p</i> -value	Romano-Wolf FWER adjusted <i>p</i> -value	Residualised propensity to charge coefficient	Unadjusted <i>p</i> -value	Romano-Wolf FWER adjusted <i>p</i> -value
Remoteness						
Inner regional	-	-	-	-	-	-
Major city	0.00	0.94	1.00	0.00	0.69	1.00
Outer regional and remote	0.01	0.83	1.00	0.04	0.00	0.15
Unknown	0.02	0.32	0.99	-0.04	0.02	0.28
SEIFA quartile						
Q1	-	-	-	-	-	-
Q2	-0.01	0.32	0.99	0.00	0.98	1.00
Q3	0.00	0.85	1.00	-0.01	0.09	0.64
Q4	-0.01	0.18	0.85	-0.01	0.13	0.80
Unknown	-0.02	0.27	0.96	0.00	0.00	0.15
Proven offences in prior 5 years						
1-2	-	-	-	-	-	-
3-5	0.00	0.81	1.00	0.01**	0.04	0.32
6+	0.01	0.10	0.58	0.02	0.31	0.98
None	0.00	0.05	0.30	0.00	0.45	1.00
Sentenced prison episodes in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.77	1.00	-0.03	0.19	0.92
None	0.00	0.07	0.45	-0.01	0.47	1.00
Proven ADVO breach offences in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.99	1.00	-0.02	0.44	1.00
None	0.00	0.21	0.89	-0.02	0.13	0.80
Proven DV offences in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.58	1.00	-0.02	0.17	0.89
None	0.00	0.49	1.00	0.01	0.35	1.00
Proven DV assault offences in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.56	1.00	0.02	0.40	1.00
None	0.00	0.87	1.00	0.00	0.88	1.00
Proven victimisation in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.97	1.00	-0.01	0.24	0.96
None	0.00	0.50	1.00	0.00	0.44	1.00
Proven DV victimisation in prior 5 years						
1	-	-	-	-	-	-
2+	0.00	0.83	1.00	0.01	0.44	1.00
None	0.00	0.90	1.00	-0.01	0.21	0.93
Constant	0.00			0.00		
F	8.57			14.64		
<i>p</i>-value of F	0			0		
R squared	0.19			0.32		
Observations	37,122			11,692		

Appendix C2 – Relevance and monotonicity test

Table C2 shows estimates of α_1 from our first stage regression defined in Equation 1, separately for males and females across our entire sample, and for various subsamples. The relevance assumption requires that police officer propensity to charge strongly influences whether a POI is proceeded against. As shown in the first row of Table C2, when estimating the first stage regression for the entire sample, our estimates of α_1 and the corresponding F-statistics are high for men and women ($\hat{\alpha}_1 = 0.43$ and $F = 267.38$ for men, while $\hat{\alpha}_1 = 0.50$ and $F = 113.70$ for women), which shows that the relevance assumption is satisfied. Next, an implication of the monotonicity assumption is that police officer propensity to charge should not be associated with a decrease in the probability a POI is proceeded against, across all subsamples. The rest of Table C2 shows estimates of our first stage regression across various subsamples for men and women. We find that coefficients are positive and significant for the majority of subgroups across both genders, which provides evidence that monotonicity is satisfied.

Table C2. Coefficient and significance of police officer propensity to charge in first stage regressions of proceeding for whole sample and subsamples

Group	Male				Female			
	Coefficient	F statistic	p-value	Observations	Coefficient	F statistic	p-value	Observations
Whole sample	0.31***	90.09	0.00	33,230	0.17***	25.37	0.00	11,041
Concurrent offences								
1	0.15**	4.91	0.03	5,182	0.34**	5.85	0.02	694
2 or more	0.01	0.09	0.77	2,830	-0.41	2.03	0.17	96
None	0.42***	82.89	0.00	20,132	0.16***	12.94	0.00	7,922
Age								
18-24	0.38***	24.99	0.00	2,775	0.06	0.35	0.56	949
25-34	0.22***	20.06	0.00	8,663	0.14*	3.59	0.06	2,229
35-44	0.40***	42.98	0.00	7,898	0.31***	21.33	0.00	1,695
45+	0.27***	10.49	0.00	5,297	0.22*	3.21	0.08	804
Under 18	1.19	1.67	0.21	61	-	-	-	22
Age difference between POI and victim								
1 year younger to 1 year older	0.18**	8.00	0.01	6,210	0.13	1.68	0.20	1,353
11+ years older	0.33**	5.15	0.03	1,602	-	-	-	24
11+ years younger	0.46	0.19	0.67	123	0.09	0.24	0.63	371
2-5 years older	0.27***	24.05	0.00	7,515	0.31*	4.00	0.05	371
2-5 years younger	0.64***	42.50	0.00	2,049	0.00	0.00	0.98	1,720
6-10 years older	0.41***	15.60	0.00	3,307	-0.08	0.02	0.90	77
6-10 years younger	-0.12	0.18	0.67	513	0.13	0.44	0.51	763
Aboriginality								
Aboriginal	0.21*	3.08	0.09	1,904	0.29	1.20	0.28	369
Non-Aboriginal	0.32***	71.09	0.00	24,161	0.15***	9.34	0.00	7,499
Unknown	0.19**	4.57	0.04	3,863	-0.03	0.02	0.89	1,002
Remoteness								
Inner regional	0.40***	37.68	0.00	7,019	0.19*	3.55	0.07	2,200
Major city	0.29***	58.95	0.00	21,681	0.18***	16.51	0.00	7,127
Outer regional and remote	0.13	3.09	0.10	1,853	0.03	0.03	0.86	505
Unknown	0.49	2.06	0.16	240	-	-	-	33

Table C2. Coefficient and significance of police officer propensity to charge in first stage regressions of proceeding for whole sample and subsamples (*continued*)

Group	Male				Female			
	Coefficient	F statistic	p-value	Observations	Coefficient	F statistic	p-value	Observations
SEIFA quartile								
Q1	0.33***	92.14	0.00	9,841	0.23***	6.24	0.02	2,623
Q2	0.20**	8.65	0.01	9,536	0.08	1.02	0.32	2,936
Q3	0.37***	22.78	0.00	5,505	0.21*	3.34	0.07	1,437
Q4	0.21**	4.98	0.03	1,826	0.07	0.14	0.71	516
Unknown	0.70***	11.75	0.00	396	0.28	0.29	0.60	66
Proven offences in prior 5 years								
1-2	0.36***	65.94	0.00	7,834	0.20*	3.01	0.09	941
3-5	0.04	0.12	0.74	1,210	-0.80*	4.16	0.05	84
None	0.31***	43.68	0.00	19,628	0.20***	21.00	0.00	7,682
Sentenced prison episodes in prior 5 years								
1	0.30	1.88	0.18	691	-	-	-	12
None	0.31***	84.98	0.00	31,060	0.17***	25.85	0.00	10,833
Proven ADVO breach offences in prior 5 years								
1	0.23	1.12	0.30	513	-	-	-	16
None	0.31***	89.89	0.00	31,401	0.17***	26.21	0.00	10,718
Proven DV offences in prior 5 years								
1	0.29***	10.47	0.00	2,607	0.16	0.19	0.66	189
None	0.31***	83.75	0.00	28,420	0.17***	21.38	0.00	9,929
Proven DV assault offences in prior 5 years								
1	0.34**	6.91	0.01	1,155	0.52	2.07	0.17	88
None	0.31***	84.97	0.00	30,252	0.18***	22.98	0.00	10,275
Proven victimisation in prior 5 years								
1	0.36***	37.36	0.00	4,811	0.15	1.49	0.23	1,119
2+	0.34***	10.59	0.00	2,332	0.19**	4.51	0.04	2,037
None	0.28***	46.93	0.00	21,206	0.11	4.08	0.05	3,946
Proven DV victimisation in prior 5 years								
1	0.26*	3.61	0.06	1,265	0.14	1.12	0.29	877
2+	0.78	1.41	0.25	148	0.21*	3.28	0.08	811
None	0.31***	91.43	0.00	29,282	0.18***	17.74	0.00	5,971

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates are omitted where subgroups were too small for an F-statistic to be calculated.

Appendix C3 – Exclusion restriction test

Table C3 compares the IV estimates presented in Table 2 to IV estimates which use alternative windows where offenders are proceeded against following the date of a DV event (within 90 or 180 days). The exclusion restriction may be violated if police officers affect whether an offender is found guilty beyond their charge decision (e.g., via investigative effort when responding to DV incidents). The estimates under the 'proceeded ever' column are the same as those presented in Table 2. These estimates are quite close to the estimates shown in the remaining columns which use alternative proceeding windows (within 90 or 180 days). As these proceeding windows are proxies for the strength of evidence against the POI, we argue that this provides evidence that the exclusion restriction is likely to hold.

Table C3. IV estimates of the impact of being proceeded against on the probability of being found guilty, by gender and proceeding window

	Proceeded ever	Proceeded within 90 days	Proceeded within 180 days
Female	.817***	.899***	.863***
Male	.672***	.654***	.663***
Controls	Yes	Yes	Yes
Difference	0.145	.245	.200
P-value of difference	0.40	0.43	0.992
R squared female model	0.53	0.50	0.52
R squared male model	0.52	0.49	0.50
Observations female	11,692	11,692	11,692
Observations male	37,231	37,231	37,231

Note. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Appendix D – OLS estimates of the impact of being proceeded on being charged

Table D1 estimates the impact of being proceeded against on being charged, separately for males and females. This regression uses the same controls as our instrumental variable regression described in our method section, and similarly clusters at the PAC level. This suggests that after controlling for observable characteristics, on average, proceeding against a woman results in a conviction 70.4% of the time, while proceeding against a man for DV assault results in a conviction 73.3% of the time. The difference in these rates is statistically significant at the 5% level. However, this does not necessarily imply that men are biased against in charge decisions. This is because this approach cannot control for unobservable factors (e.g., those relating to the quality of evidence available to police such as behaviours of the involved parties after the incident and the extent of injuries of parties involved).

Table D1. OLS estimates of the impact of being proceeded on being charged, by gender

Outcome			p-value of		Controls	R-squared		Observations	
	Female	Male	Difference	difference		Female	Male	Female	Male
Model 1 Probability of conviction	.710***	.755***	-.0443	0.000	No	0.71	0.61	11,692	37,231
Model 2 Probability of conviction	.704 ***	.733***	-.0298	0.002	Yes	0.71	0.62	11,692	37,231

Note. * $p<0.05$ ** $p<0.01$ *** $p<0.001$