CRIME AND JUSTICE
Bulletin

NSW Bureau of Crime
Statistics and Research

Contemporary Issues in Crime and Justice *Number 206*

This bulletin has been independently peer reviewed.

September 2017

Evaluation of the 2015 Domestic Violence Evidence-in-Chief (DVEC) reforms

Steve Yeong & Suzanne Poynton

Aim: *To investigate the impact of the 2015 Domestic Violence Evidence-in-Chief (DVEC) reforms on court outcomes for Domestic Violence (DV) assaults in New South Wales (NSW).*

Method: *Court outcome data from the NSW Bureau of Crime Statistics and Research's Reoffending Database (ROD) and DVEC data from the NSW Police Force's Computerised Operational Policing System (COPS) were extracted for the period January 2014 to June 2016. These data were used to assess the impact of DVEC using two different identification strategies. First, an Instrumental Variables approach was employed in order to estimate the causal effect of a DVEC statement on the probabilities that an individual defendant: i) has at least one proven offence (a conviction) or ii) pleads guilty to at least one offence at their court appearance. We also use this approach in order to estimate the impact of a DVEC statement on the time taken to finalise a DV assault matter conditional on a guilty plea being entered. Second, a dynamic panel data model with fixed effects was employed in order to investigate the relationship between each Local Area Command's (LAC's) monthly conviction and guilty plea rate for DV assaults and its uptake of DVEC.*

Results: *We find no evidence to indicate that DVEC has had a significant impact on the probability of a guilty plea or the time to finalisation for matters resulting in a guilty plea. We do, however, find limited evidence to indicate that the presence of a DVEC statement may increase the probability of a conviction by about two percentage points. With respect to the monthly LAC level analysis, we find no evidence to indicate that a LAC's DVEC uptake rate has any significant effect on its rate of convictions or guilty pleas. Both of these results are extremely robust to a variety of specifications.*

Conclusions: *Once factors like seasonality and systematic differences between LACs are taken into account, there is limited evidence of a significant impact of the NSW DVEC reforms on court outcomes for DV assaults.*

Keywords: *DVEC, domestic violence, pre-recorded evidence, generalised method of moments, forward orthogonal deviation, instrumental variables, impact evaluation, two-stage least squares*

INTRODUCTION

The *NSW Crimes (Domestic and Personal Violence) Act 2007* defines a Domestic Violence (DV) offence as '*personal violence committed by a person against another person with whom the person who commits the offence has or has had a domestic relationship.*' The Act defines a domestic relationship as occurring with another person if the person:

- Is or has been married to that person/is or has been the de facto partner of that person;
- Has or has had an intimate personal relationship with that person;
- Is living or has lived in the same household/residential facility as that person;
- Has or has had a relation involving his/her dependence on the ongoing (un)paid care of that person; and/or
- Is or has been the relative of that person.[1]

Less than 50 per cent of DV incidents are reported to police (see: Australian Bureau of Statistics [ABS], 2006; Birdsey & Snowball, 2013; Grech & Burgess, 2011). Based on interviews with recent DV assault victims who did not report the DV to police, Birdsey and Snowball (2013) note the most commonly cited reasons for under reporting include; fear of revenge, further violence, embarrassment/shame, reporting taking too much time and reporting being too much of a bother.

Hoyle (1998) argues that under-reporting occurs as a result of victim's cost-benefit analysis of providing evidence.[2] Specifically,

Suggested citation: Yeong, S. & Poynton, S. (2017). *Evaluation of the 2015 Domestic Violence Evidence-in-Chief (DVEC) reforms* (Crime and Justice Bulletin No. 206). Sydney: NSW Bureau of Crime Statistics and Research.

in interviews with victims of DV, Hoyle (1998) found that in many cases DV victims determined the emotional cost involved in providing evidence outweighed the probability of a beneficial outcome resulting from the criminal justice process. Elliott (1989) and Ford and Regoli (1993) suggest that unwillingness on the part of DV victims to give evidence in court serves to (at least partially) explain historically low prosecution and arrest rates for DV offences identified in the literature (see Dutton, 1995; Fagan, 1989; Ford & Regoli, 1993; Schmidt & Steury, 1989). Pre-recording victim evidence provides a way for victims/complainants to provide evidence while reducing the stress and trauma associated with a court appearance.

NSW DOMESTIC VIOLENCE EVIDENCE-IN-CHIEF (DVEC) REFORMS

In 2015 NSW became the first Australian jurisdiction to allow pre-recorded evidence to be admissible for DV matters. The Domestic Violence Evidence-in-Chief (DVEC) reforms came into effect for all DV charges commencing on or after 1 June 2015. DVEC involved changing the way in which DV complainants are able to provide evidence. Essentially DVEC removes the 'hearsay rule of evidence' as it applies to DV complainants in criminal proceedings. The implication is that statements made out of court can be used as evidence to the existence of a fact. DVEC enables a recorded video and/or audio statement of a DV complainant to be admissible as evidence-in-chief in proceedings for DV offences.[3] The policy provides similar safeguards to those already available to children and the cognitively impaired in similar situations.

DVEC was intended to:

- Reduce the trauma for complainants associated with recounting events in front of the offender.
- Reduce the difficulty in recalling events when giving evidence (for cross examination) as they would be able to see their previously recorded statement.
- Illustrate the demeanour and experience of the complainant proximate to the time of the event.
- Reduce the capacity of the defendant to intimidate the complainant to change or recant their evidence.
- Increase both the conviction rate and the number of guilty pleas.

Eight points with respect to the implementation and operation of the policy are of note. First, police officers are required to receive special training before they are able to take DVEC recordings. The rate at which DVEC training was provided to officers varied between each Local Area Command (LAC) in part due to decisions made by LAC Commanders regarding the allocation of their officers to DVEC training. Information on when and which officers received training are unavailable. Second, police must have consent from the complainant in order to take a DVEC statement. Third, although the decision about whether or not to play the recording in court is made in consultation with the complainant, the Police Prosecutor can in principle proceed to play the recording against the wishes of the complainant. Fourth, the recording must be made as soon as practical after the time of the incident. This can be at the scene or at the police station, depending on the situation. Fifth, defendants must be given the opportunity to view/listen to the recording at least once prior to any court appearance. Sixth, evidence of how an accused person behaves in front of police while watching/listening to the recording is inadmissible with two exceptions. First, if the viewing took place during the investigation and second, if the accused commits an offence during or immediately after viewing the recording. Seventh, the accused is under no circumstances allowed to obtain a copy of the recording. Finally, the recording may be used in place of a written statement for both summary and committal proceedings in addition to indictable offences that are related to the DV incident.

Implementation of the policy across the various NSW LACs makes identifying the causal effect of the policy difficult. Specifically, because information on when and which officers received the training required to take DVEC recordings is unavailable, we are unable to differentiate between treated and control LACs in a conventional Difference-in-Differences framework. Therefore, when a DVEC statement is *not taken*, we cannot tell whether it was because the victim refused to consent to the recording, or because the responding officer didn't have the training. This means we cannot simply compare outcomes for charges/appearances with and without DVEC statements as an unobserved third factor may bias our estimate of the policy's impact. For example, a cooperative victim is likely to give a DVEC statement when prompted by police. Similarly, a cooperative victim could plausibly provide better evidence and thus have a higher likelihood of a conviction, with or without DVEC. In this case, a simple comparison would overestimate the impact of DVEC on the probability of a conviction as we cannot observe and control for witness cooperation. For this reason we employ the two identification strategies (outlined later) that aim to alleviate this problem of endogenous selection. Before we describe these strategies, we briefly review relevant research.

LITERATURE REVIEW

In an Australian context pre-recorded evidence has been reserved for matters involving sexual assault victims, children and complainants who are cognitively impaired. In 1992, Western Australia (WA) became the first Australian jurisdiction to allow pre-recorded evidence. Since then similar provisions have emerged for vulnerable victims in all Australian jurisdictions. For example, the Northern Territory (NT) and South Australia (SA) allow pre-recorded evidence for victims of sexual assault deemed to be at a disadvantage if required to give evidence in

court. The Australian Capital Territory (ACT), WA, Queensland (QLD) and Victoria (VIC) allow victims of sexual assault, those who are cognitively impaired and/or those who are under the age of 18 to provide the entirety of their evidence-in-chief at a pre-trial hearing.[4] During such hearings the accused and his/her legal practitioner are present in one room, while the complainant provides their evidence from a separate room. The accused is able to see the complainant via Closed Circuit Television but the complainant is unable so see or hear the accused. Both legal practitioners are able to question the complainant and the entire process is recorded for viewing at the subsequent court appearance for the jury.[5] In NSW, following the proclamation of the NSW Evidence (Children) Act 1997 on 1 August 1999, children and the cognitively impaired who have been subject to sexual and/or severe physical abuse have been able to provide part or the entirety of their evidence-in-chief through a pre-recorded video and/or audio statement.[6]

Much of the research surrounding pre-recorded evidence has focussed on the experiences of the relevant stakeholders in the criminal justice process.

McConachy (2002) evaluated the introduction of electronic recording of evidence-in-chief for children in NSW. She interviewed legal practitioners, children and their parents/carers as well as officers from the NSW Police Force and Department of Community Services. Her interviews indicate a predominately positive stakeholder response. In particular, legal practitioners reported that pre-recorded evidence increased the level of accuracy and detail present in the evidence. They described two channels for this effect. First, it encouraged a natural free flowing conversation (i.e. narrative style of interviewing) that resulted in a more detailed testimony. This is because the experience is less daunting for children relative to public testimony in front of the offender. Second, because court appearances often occur months after the offence, pre-recorded evidence reduces the likelihood of forgetting details or an alterations to the story.

Cashmore and Trimboli (2005) conducted a comprehensive evaluation of a pilot of a specialist court designed to deal specifically with child sexual assault cases. The pilot involved a battery of measures designed to improve the NSW justice system's response to child sexual assault cases. Part of the evaluation included results from interviews with stakeholders in relation to pre-recorded evidence. Cashmore and Trimboli (2005) report an overwhelmingly positive response from stakeholders. In addition to the benefits outlined earlier, defence lawyers cited another benefit of pre-recorded evidence; namely an increase in the time available to prepare for an appearance. Parents and carers also had positive responses to the facility for pre-recorded evidence, reporting that it lowered the stress and trauma associated with testimony for children.

With respect to victims, carers, parents and advisors, the overwhelming consensus is positive. All of these stakeholders report strong reductions in the stress, trauma, anxiety and apprehension about providing evidence (for examples see Burton, Evans, & Sanders, 2006; Cashmore & Trimboli, 2005; Davies & Hanna, 2013; Hamlyn, Phelps, Turtle, & Sattar, 2004; McConachy, 2002).

Westera et al. (2011) investigated how pre-recorded evidence and interview style impacted police assessment of complainant credibility in a survey of 136 police officers in New Zealand. Westera et al. (2011) found that both the narrative style of interview (generally associated with pre-recorded evidence) and pre-recorded evidence itself, to be positively associated with police's perception of complainant credibility and thus propensity to proceed with charges.

The response of legal practitioners has been mixed. On one hand, practitioners from both sides report advantages consistent with those outlined earlier by Cashmore and Trimboli (2005); and McConachy (2002) (also see Burton et al., 2006; Hamlyn et al., 2004). However with respect to pre-recorded evidence-in-chief, practitioners report the main disadvantage as being that live direct examination provides a warm up for cross-examination. Hence, victims may experience increased difficulties when recalling events for cross-examination (see Burton et al., 2006). With respect to the entirety of a complainant's testimony being provided via pre-recording, practitioners cited the main disadvantages as being: loss of immediacy and drama at trial, emotionally distancing the jury from the victim, concerns about the ability of jurors to maintain concentration during long recordings and defence counsel concern about their inability to see the jury's response prior to cross-examination (see Burton et al., 2006; Davies & Hanna, 2013; Hamlyn et al., 2004).

There is very little evidence bearing on the question of whether pre-recorded evidence increases guilty plea or conviction rates. The only Australian study of relevance is one carried out by Birdsey and Smith (2012). They evaluated the Domestic Violence Intervention Court Model (DVICM) in two pilot sites. The DVICM commenced on 12 September 2005 in Campbelltown Local Court (which draws cases from Campbelltown and Macquarie Fields LACs), and on the 10 October 2005 in Wagga Wagga Local Court (which draws cases from the Wagga Wagga LAC). The objective of DVICM was to improve the criminal justice system response to DV. One aspect of DVICM of interest to our study was the provision of digital and video cameras to police to help them better capture evidence of property damage and physical injury. Birdsey and Smith (2012) investigated a large number of court and police outcomes by comparing the pilot sites to the rest of NSW in a Difference-in-Differences style setup. They found only limited evidence to indicate that the DVICM increased the propensity for police to proceed with

charges after recording/detecting the offence and no evidence that it increased overall conviction rates or the willingness of defendants to plead guilty.

THE PRESENT STUDY

The aim of this report is to investigate the impact of DVEC on convictions and guilty pleas. As outlined earlier, the possibility of omitted variable bias prevents us from a simple comparison of court outcomes with/without or before/after the introduction of DVEC. Victims who are more credible witnesses, for example, may be more willing to avail themselves of the opportunity to give pre-recorded evidence. In order to address this problem we use two different identification strategies to answer two sets of related questions.

We begin by exploiting individual court appearance level information available in the data to investigate how appearances with and without DVEC statements compare in terms of court outcomes.

The first set of questions is:

1. Did the presence of a DVEC statement increase the probability that a defendant has at least one offence proven against them (which we refer to as a conviction) at their court appearance?
2. Did the presence of a DVEC statement increase the probability that a defendant pleads guilty to at least one offence at their court appearance?
3. Conditional on the outcome of the appearance being a guilty plea, did the presence of a DVEC statement reduce the time between the initial charge by police and the subsequent finalisation? We refer to this difference as 'plea delay'.

The second set of questions exploit variation in the take up of DVEC across Local Area Commands (LACs) in order to estimate how changes in the uptake rate may have impacted monthly conviction and guilty plea rates.

The second set of questions is:

4. Did an increase in a LAC's monthly DVEC uptake rate generate an increase in that LAC's monthly conviction rate?
5. Did an increase in a LAC's monthly DVEC uptake rate generate an increase in that LAC's monthly rate of guilty pleas?

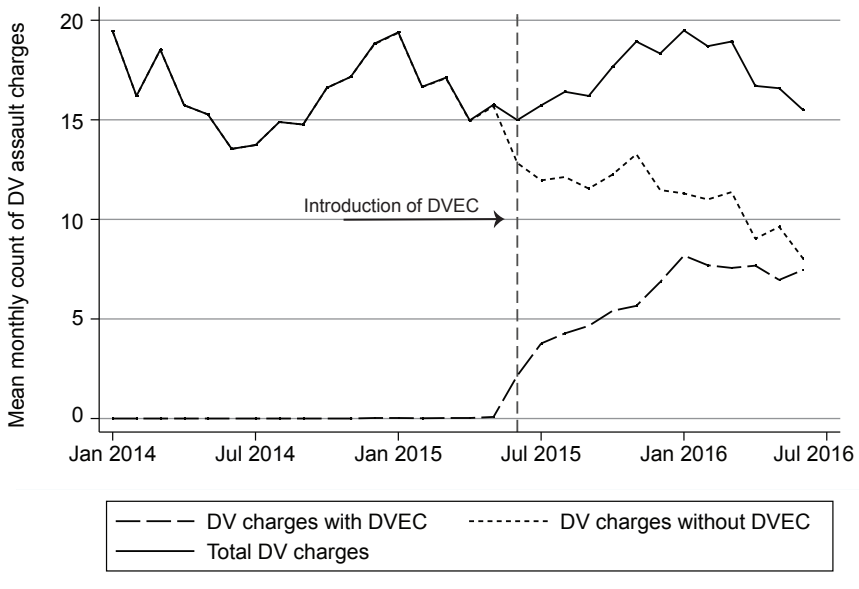
METHOD

DATA SOURCES

Data from the NSW Police Force's Computerised Operational Policing System (COPS) are used in this analysis. COPS data contains information on every DV assault incident recorded by NSW police. For each LAC, the data identifies when an incident was recorded, if police chose to proceed with charges, when those charges were laid, and whether or not a DVEC statement was recorded on the charge. These data allow us to generate a balanced monthly panel consisting of information on the number of DV assault charges and DVEC statements associated with these charges for 76 LACs over the period January 2014 to January 2017.[7, 8]

Figure 1 plots three monthly series generated from these COPS data: (1) the mean count of DV assault charges (the dark solid line) (2) the mean count of DV assault charges without a DVEC recording (the short dashed line) and (3) the mean count of DV assault charges with a DVEC recording (the long dashed line). The introduction of DVEC is shown by the dashed vertical line commencing in June 2015. Two points are of note with respect to Figure 1. First, the dark solid line illustrates the high degree of seasonality present in the number of DV assault charges. DV assault charges appear to rise and fall in the summer and winter months, respectively. We must take this seasonality into account when estimating the causal effect of DVEC. Second, after the introduction of the policy we can see a sharp rise in the number of DV assault charges with a DVEC statement (and a corresponding decline in the number without a DVEC statement). This indicates a strong overall uptake of the policy.

Figure 1. Uptake of DVEC across NSW



Data extracted from BOCSAR's Reoffending Database (ROD) are also used in this analysis. ROD data contains individual charge level information; including information on the defendant's age, sex, Indigenous status, area of residence, measures of area-level socioeconomic disadvantage, whether or not the defendant was held on remand, had legal representation, the defendant's prior criminal history and the number of concurrent charges at the defendant's court appearance. Importantly, the ROD data also contain the outcome of each charge (e.g. a proven outcome; a guilty plea) as well as an identifier that allows us to link each charge with the police data indicating whether or not a DVEC statement was taken. When we aggregate these data we obtain a balanced monthly panel consisting of information on court outcomes for 76 LACs over the period January 2014 to June 2016.

COURT APPEARANCE LEVEL ANALYSIS

To investigate the impact of DVEC on court outcomes at the individual appearance level we assume the structural relation presented in Equation (1) below.[9]

$$\Pr(y_{icjt}=1) = \beta * D_{icjt} + \mathbf{X}_{icjt} + \alpha_c + \delta_j + \lambda_t + u_{icjt} \qquad (1)$$

Where y_{icjt} is the outcome of appearance i, in court c, where the defendant was charged by officers from LAC j, at time t. D_{icjt} is a binary variable taking value one for appearances with a DVEC statement, zero otherwise. $\mathbf{X}_{icjt}$ is a vector of controls that includes the defendant's age, gender, number of prior proven court appearances, Socio-Economic Indexes for Areas (SEIFA) percentile, Indigenous status, as well as whether or not the defendant had legal representation and was held on remand at finalisation. We also include court (α_c), LAC (δ_j) and time (λ_t) Fixed Effects (which are described in further detail in the proceeding subsection). In this part of the analysis we are interested in estimating β; which can be interpreted as the causal effect a DVEC statement has on the outcome of interest.

We use three outcome measures. The first takes value one if the appearance resulted in at least one proven offence (a conviction), zero otherwise. The second takes value one if the defendant pleaded guilty to at least one charge at the appearance, zero otherwise. For these two outcomes β is to be interpreted as the average change in the probability that the defendant is convicted of, or pleads guilty to, at least one charge at their court appearance. The third is the number of days between the date on which police charged the defendant and the date they entered into a guilty plea. We refer to this outcome as plea delay. With respect to plea delay, β is to be interpreted as the average change in the number of days between charge and finalisation as a result of a DVEC statement.

The problem with simply estimating Equation (1) is that our estimate of β may suffer from omitted variable bias. For example, the willingness of victims to cooperate with police is a factor that (at least partially) determines the outcome of a court appearance but is also plausibly correlated with whether or not a DVEC statement is taken. Left unaccounted for, omission of this factor would generate an upward bias of our estimates of the effect of a DVEC statement on the outcomes we are investigating. To get around this problem we employ an Instrumental Variables (IV) approach.

The IV approach allows us to obtain a consistent estimate of the Local Average Treatment Effect, which in our case is roughly equal to the Average Treatment Effect on the Treated.[10] The fact that DVEC training was in part a result of decisions made by LAC Commanders provides us with a potentially valid instrument because the decision to provide the training is arguably unrelated to the effect of a DVEC statement on the outcomes of interest. We discuss this issue in considerable depth shortly.

To construct the instrument, we divide the monthly count of DV assault charges with a DVEC recording in each LAC by the total number of DV assault charges for that month. The result is a measure of the DVEC uptake rate in each LAC in each month. We use this in a Two-Stage Least Squares (2SLS) IV regression. We estimate both stages using Ordinary Least Squares (OLS)[11] and report heteroskedasticity robust standard errors clustered at the court level.[12] In the first stage we estimate the Linear Probability Model (LPM) in Equation (2) below.

$$\Pr(D_{icjt}=1) = \pi_1 \mathbf{Z}_{jt} + \mathbf{X}_{icjt} + \alpha_c + \delta_j + \lambda_t + v_{icjt} \qquad (2)$$

Where $\Pr(D_{icjt}=1)$ is the probability that appearance i, in court c, for a DV offence occurring under the jurisdiction of LAC j, at time t will involve a DVEC recording. $\mathbf{Z}_{jt}$ is our instrument, and all other variables have the same definitions as in Equation (1). We follow Angrist and Pischke (2008) and recode the DVEC uptake rate into a binary variable equal to one if the LAC's DVEC uptake rate is above the NSW average for that month, zero otherwise.[13, 14] We then proceed by using the predicted probabilities of a DVEC statement from the first stage, $\widetilde{D}_{icjt}$, to estimate the causal effect of a DVEC statement on our three outcome measures. Equation (3) shows the second stage regression:

$$\Pr(y_{icjt}=1) = \beta^{IV} * \widetilde{D}_{icjt} + \mathbf{X}_{icjt} + \alpha_c + \delta_j + \lambda_t + \epsilon_{icjt} \qquad (3)$$

β^{IV} is the coefficient of interest and our estimate of DVEC's impact on the probability that an appearance results in at least one proven outcome (a conviction) or guilty plea. In the plea delay regressions, β^{IV} is the causal effect of a DVEC statement on the number of days between charge and finalisation conditional on the outcome of the appearance being a guilty plea.

Valid identification of β^{IV} hinges on two assumptions we must make about our instrument:[15]

1. Relevance assumption: Our instrument must be related to the probability that a charge has a DVEC statement.

2. Exclusion restriction: Our instrument must be otherwise (conditionally) unrelated to any one of our three outcome measures.

The relevance assumption is readily testable and likely to hold in our case. By construction, a LAC's monthly DVEC uptake rate is directly related to the conditional probability of a DVEC statement being taken by officers from that LAC.

We cannot explicitly test the exclusion restriction but argue our instrument meets this assumption conditional on the Fixed Effects. For example, one may be tempted to argue that the exclusion restriction is violated because high DVEC uptake LACs systematically differ from low DVEC uptake LACs in their conviction rates because of officer experience. That is, officers from LACs with high rates of DV would not only have more experience with DV and thus be better at collecting evidence, but also have a stronger incentive to opt into DVEC training. However, provided that this experience is approximately constant over the time span of our sample (June 2015 - June 2016), our instrument should be conditionally exogenous to this factor.

Another concern regarding the exclusion restriction, and our analysis more generally, is that DVEC statements may induce police to prosecute offenders that they wouldn't have otherwise have proceeded against in the absence of a DVEC statement. They may, for example, overestimate the effectiveness of DVEC statements on the likelihood of a conviction. This would have the effect of increasing the number of court appearances involving DV assaults but not necessarily the number of proven outcomes, thus actually reducing the conviction rate. We address both of these problems in our robustness checks finding no meaningful change to the main results. Nevertheless, to guard against the possibility that our exclusion restriction is violated in some other way, we estimate the impact of DVEC on court outcomes using a second strategy, which we now turn to.

AGGREGATE LAC LEVEL ANALYSIS

In the aggregate LAC level analysis we are interested in investigating the relationship between each LAC's uptake of the policy and its rate of convictions and guilty pleas. A model that estimates this relationship should account for the following four confounding factors (1) the inherent seasonality in DV assaults, conviction and guilty plea rates (see Figures 1, 2 and 3, respectively) (2) any systematic differences between conviction and guilty plea rates between the various LACs (3) any dynamic elements in the policy's impact on conviction and guilty plea rates (e.g. inertia in crime rates, autocorrelation in conviction and guilty plea rates, or lagged changes in guilty plea/conviction rates) and (4) unobserved differences between cases where DVEC statements are employed and cases where they are not.

In order to account for the first three of these confounding factors we employ the Dynamic Panel Data Model given by Equation (4) below:

$$y_{jt} = \alpha y_{jt-1} + \beta * Z_{jt} + \lambda_t + \delta_j + \epsilon_{jt} \quad (4)$$

Where y_{jt} is the monthly conviction or guilty plea rate for DV assault charges, where charges were made by officers from LAC j, at time t.[16] To be clear, we are modelling two separate outcomes (conviction and guilty plea rates) using the same structural relation. The conviction rate is constructed as the monthly count of DV charges resulting in at least one proven offence, divided by the total monthly count of DV charges for that LAC. The guilty plea rate is constructed as the monthly count of DV charges resulting in at least one guilty plea, divided by the total monthly count of DV charges for that LAC. We include a Lagged Dependant Variable (LDV) in order to account for inertia and autocorrelation in our outcomes. Z_{jt} is the DVEC uptake rate at LAC j, in month t. Recall from the previous subsection that we define the DVEC uptake rate as the monthly count of DV assault charges with a DVEC recording, divided by the total number of DV assault charges. We also include LAC (δ_j) and month (λ_t) Fixed Effects (FEs) in order to account for systematic differences between LACs and different months of the year, respectively. The month FEs account for the inherent seasonality in conviction and guilty plea rates common across LACs, while the LAC FEs account for time invariant differences between LACs that may affect conviction and plea rates. Finally, ϵ_{jt} is an independently and identically distributed (i.i.d) error term. β is the coefficient of interest and can be interpreted as the average percentage point change in the monthly conviction or guilty plea rate resulting from a one percentage point increase in the monthly DVEC uptake rate.

In order to estimate Equation (4) we employ the difference Generalised Method of Moments (GMM) approach for estimating a Dynamic Model with FEs. We employ difference instead of system GMM because the assumption underlying system GMM is likely to be violated in our case. System GMM assumes that the (first differenced) instruments are exogenous to the LAC FE. This is unlikely to hold in our case because changes to the DVEC uptake rate are plausibly correlated with the LAC FE. For example, a LAC with a pro-active commander is likely to experience larger changes in its DVEC uptake rate relative to a LAC that has a less pro-active commander.

Instead of using the First Difference (FD) transformation on Equation (4), we remove the LAC FE using the Forward Orthogonal Deviation (FOD) transformation.[17, 18]

Employing the FOD transformation on Equation 4 we arrive at Equation 5 below.[19]

$$y^*_{jt} = \alpha y^*_{jt-1} + \beta^{GMM} * Z^*_{jt} + \lambda^*_t + \epsilon^*_{jt} \quad (5)$$

Lagged endogenous variables are used as instruments for the explanatory variables in Equation (5). The idea is to instrument for explanatory variables using the second lag of

those variables. We apply the Backward Orthogonal Deviations (BOD) transformation on the instruments. The resulting instruments are both relevant and exogenous, thus alleviating unobserved heterogeneity. The appropriate number of lags for the instruments is determined using the Arellano and Bond (1991) test of autocorrelation in the FD residuals.[20] That is, an insignificant partial autocorrelation coefficient at lag two on the FD residual indicates that, starting from lag two, we can use endogenous variables as instruments. We employ two-step GMM estimators with standard errors that are robust to arbitrary forms of heteroskedasticity and have been given the Windmeijer (2005) finite sample correction.

As argued by Arellano and Bond (1991), dynamic models with FEs estimated through GMM are suitable for panels that have a relatively small T and large N, have dynamic elements, a linear relationship between the dependent and independent variables, and exhibit serial correlation and heteroskedasticity within but not across observations. The first two criteria are implicitly met by our data, while related literature and the inherent nature of the policy makes the case for the inclusion of dynamic elements. The final three criteria are testable and results from these tests (reported in Part B of the Appendix) support the utilization of our model.

RESULTS

DESCRIPTIVE STATISTICS

Before presenting the results from the 2SLS IV and dynamic panel data regression models we examine our key outcomes descriptively.

We begin by conducting a mean difference test comparing our three outcomes for court appearances, with and without DVEC statements. Recall from the previous section that our three outcomes are; the probability of at least one proven offence (a conviction) at the appearance, the probability of at least one guilty plea at the appearance, and finally the days between charge and finalisation conditional on the outcome being a guilty plea. From Table 1 we can see that the probability of a conviction for DV assault appearances with a DVEC statement is higher than for DV assault appearances without a DVEC statement; 78 per cent and 74 per cent, respectively. This four percentage point difference is statistically significant at the 1% level. The probability of a guilty plea for DV assault appearances with and without DVEC is about 59 per cent and 57 per cent,

Table 1. Comparison of average conviction, guilty plea and plea delay rates for cases with and without a DVEC statement

	Without DVEC	With DVEC	Difference
Conviction	0.740	0.782	0.042***
	(0.004)	(0.005)	(0.007)
Guilty plea	0.574	0.596	0.022***
	(0.005)	(0.006)	(0.008)
Plea delay (days)	79.404	79.545	0.141
	(1.055)	(1.210)	(1.639)
N	11,292	6,771	18,063

Note. DVEC = Domestic Violence Evidence-in-Chief, N for plea delay is 5,711 without DVEC and 3,609 with DVEC, standard errors in parentheses. $p<0.1$ *, $p<0.05$ **, $p<0.01$***

respectively. This two percentage point difference is statistically significant at the 1% level. The average time to finalisation for matters with at least one guilty plea is about 79 days, and this is roughly the same irrespective of whether or not a DVEC statement is taken.

Figures 2 and 3 plot the average monthly conviction and guilty plea rates for NSW, respectively. At first glance the policy appears to have had no substantial immediate or lagged impact on either the conviction or guilty plea rate. That is, if there was a large impact we would expect to see a spike in the post-policy levels at some point after the introduction of the policy on either Figure. We quantify this before/after comparison by conducting a mean difference test. Results from this test are presented in Table 2. Prior to the introduction of the policy, the average conviction rate was at about 75 per cent for DV assaults. After the introduction of DVEC the conviction rate is about 76 per

Figure 2. Before versus after: Mean conviction rate

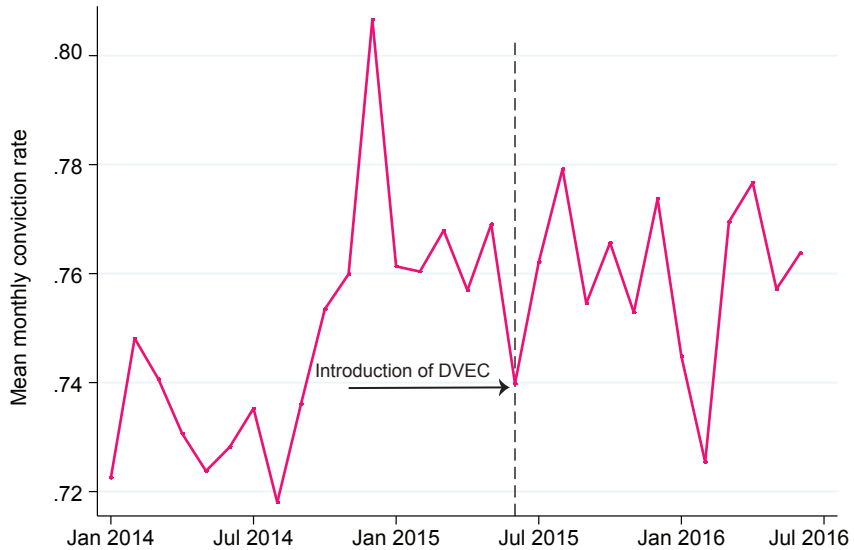


Figure 3. Before versus after: Mean guilty plea rate

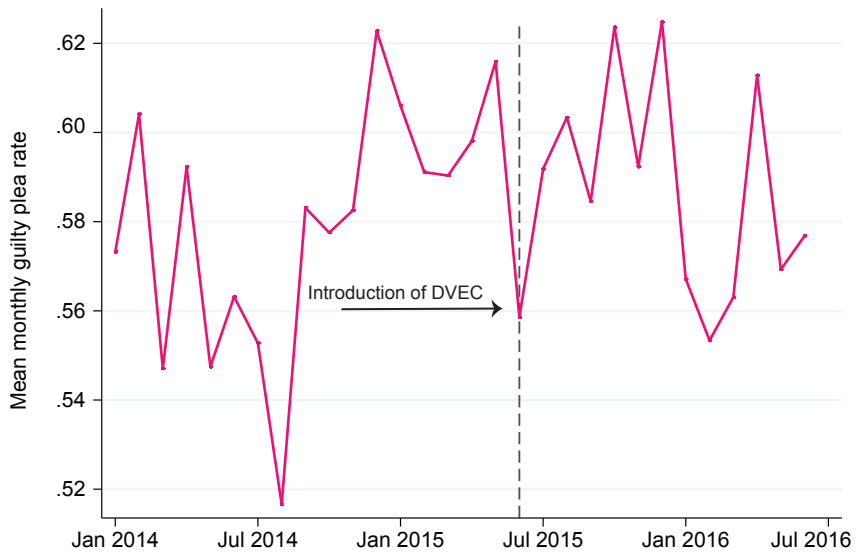


Table 2. Comparison of average conviction and guilty plea rates before and after the introduction of DVEC

	Before DVEC	After DVEC	Difference
Conviction	0.749	0.760	0.011*
	(0.005)	(0.006)	(0.009)
Guilty plea	0.581	0.587	0.006
	(0.006)	(0.006)	(0.008)
N	1,272	972	2,244

Note. DVEC = Domestic Violence Evidence-in-Chief, standard errors in parenthesis.
p<0.1 *, p<0.05 **, p<0.01***

cent. This one percentage point increase is statistically significant at the 10% level. We find no significant difference for the guilty plea rate, which is roughly constant at 58 per cent. Note that the analyses in Tables 1 and 2 do not address the problem of omitted variable bias mentioned earlier or take into account possible seasonal effects and systematic differences between LACs.

COURT APPEARANCE LEVEL ANALYSIS

Here we present the main results from the court appearance analysis. That is, in this subsection we investigate how the presence of a DVEC statement at a defendant's court appearance affects the probability that the defendant receives at least one proven outcome (which we refer to as a conviction) or pleads guilty to at least one charge. We also investigate the possibility that DVEC shortens the time (measured in days) between police charging the offender and the offender entering a guilty plea. We refer to this outcome as plea delay. All standard errors presented in this section are clustered at the court level.[21]

Table 3 presents the coefficient(s) of interest for five different specifications that estimate the impact of DVEC on the probability of a conviction. Column 1 presents estimates from an OLS regression on Equation 1. The estimate indicates that the presence of a DVEC statement increases the probability of a conviction by about two percentage points. This effect is significant at the 1% level. The sign, significance and approximate size are retained in Column 2 using the Probit specification.[22] However we are yet to account for the possible omitted variable bias outlined earlier. That is, systematic differences between DVEC and non-DVEC appearances may generate a bias in our estimates.

Column 3 corrects for this potential bias using the Instrumental Variables (IV) procedure outlined in the method section. The coefficient on our DVEC indicator is now negative and insignificant. This suggests that an unobserved factor (such as the cooperativeness of the victim) may be causing an upward bias in the estimates presented in Columns 1 and 2.

Column 3 also presents three diagnostic checks on our IV model. First, we present the F-Statistic on the excluded instrument in the first stage of the 2SLS IV regression. The F-Statistic is 363.9, well above the conventional threshold of F=10 (see Staiger & Stock, 1994) and the critical values provided by Stock and Yogo (2005). However by clustering our standard errors at the court level, we are imposing the assumption of non-i.i.d errors in the model. We therefore also present the Kleibergen and Paap (2006) identification statistic. This statistic's value is 39.61 and is significant at the 1% level, thus indicating our instrument meets the relevance assumption. Finally, we also present the p-value for an 'endogeneity' or 'distance' test. That is we run two regressions, one where DVEC is treated as conditionally exogenous (an OLS regression of Equation 1) and the other where DVEC is treated as endogenous (a 2SLS IV regression), and then compute a heteroskedasticity cluster robust version of a difference of two Sargan-Hansen tests. The null hypothesis of this test is that DVEC is conditionally exogenous. The p-value associated with this test's C-Statistic is 0.259, indicating that we may be able to treat DVEC as conditionally exogenous.[23] If this is the case we should prefer our OLS estimate from Column 1 on efficiency grounds. However clearly the inference we can draw about the policy's impact on the probability of a conviction depends greatly on estimation method. We discuss this at length later in the bulletin.

Table 3. Court appearance level analysis: Main results for probability of a proven offence

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Probability of a proven offence	**OLS**	**Probit**	**Preferred IV**	**DV EXP**	**Propensity**
DVEC	0.020***	0.020***	-0.027	-0.026	-0.024
	(0.007)	(0.006)	(0.0392)	(0.039)	(0.039)
DV incidents				-0.001	
				(0.000)	
Propensity to charge					0.036
					(0.030)
Constant	0.578***	0.180			
	(0.135)	(0.649)			
Observations	17,115	17,032	14,582	14,582	14,582
Controls	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
LAC FE	YES	YES	YES	YES	YES
Court FE	YES	YES	YES	YES	YES
Number of clusters	137	123	137	137	137
First stage F-statistic			363.9	361.9	352.1
Identification statistic			39.61	39.70	39.74
C-statistic (p-value)			0.259	0.261	0.285

Note. OLS = Ordinary Least Squares, IV = Instrumental Variables, DV = Domestic Violence, EXP = Experience, FE = Fixed Effects, LAC = Local Area Command, clusters refer to courts, cluster robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Court appearance level analysis: Main results for probability of a guilty plea

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Probability of a guilty plea	**OLS**	**Probit**	**Preferred IV**	**DV EXP**	**Propensity**
Has DVEC	0.000	0.000	-0.046	-0.045	-0.040
	(0.007)	(0.020)	(0.044)	(0.044)	(0.044)
DV incidents				-0.001	
				(0.001)	
Propensity to charge					0.058*
					(0.035)
Constant	0.240**	-0.841**			
	(0.114)	(0.329)			
Observations	17,115	17,094	14,582	14,582	14,582
Controls	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
LAC FE	YES	YES	YES	YES	YES
Court FE	YES	YES	YES	YES	YES
Number of clusters	137	130	137	137	137
First stage F-statistic			363.9	361.9	352.1
Identification statistic			39.61	39.70	39.74
C-statistic (p-value)			0.240	0.246	0.290

Note. OLS = Ordinary Least Squares, IV = Instrumental Variables, DV = Domestic Violence, EXP = Experience, FE = Fixed Effects, LAC = Local Area Command, clusters refer to courts, cluster robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Column 4 addresses the first concern regarding the exclusion restriction outlined earlier. It controls for the possibility that LACs with a high number of DV incidents may be both more likely to opt-into DVEC training and have higher conviction rates (e.g. because officers may be more experienced in dealing with DV cases). Specifically, in Column 4 we address the possibility that officer experience with DV may vary over the time span of our sample by adding a proxy for this experience. That is, in Column 4 we include each LAC's monthly count of recorded DV incidents. Inclusion of this proxy for officer experience results in no meaningful change to the estimate from Column 3. Moreover, the size and insignificance of the coefficient for this proxy indicates that the effect of officer experience in relation to DV is approximately constant over the time span of our sample and thus accounted for through our model's FEs.[24]

Column 5 addresses our second concern regarding the exclusion restriction. Specifically, that the presence of a DVEC statement induces officers to charge offenders that they wouldn't otherwise have charged without a DVEC. We attempt to control for this possibility by including in the model each LAC's monthly proportion of recorded DV incidents that result in a charge. That is, each LAC's monthly count of DV charges divided by its monthly count of recorded DV incidents. Inclusion of this control results in no meaningful change to the size, sign or significance of our DVEC indicator.

Table 4 presents our main results for the impact of DVEC on the probability of a guilty plea. Unlike Table 3, estimates of the policy's impact are consistent across all specifications. That is, DVEC has no significant impact on the probability of a guilty plea. In our IV regressions (Columns 3-5) the First Stage F and Identification statistics indicate that our instrument meets the relevance assumption. Results from the distance test indicate that we can treat DVEC as conditionally exogenous and should thus

Table 5. Court appearance level analysis: Main results for plea delay

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Plea delay	**OLS**	**Probit**	**Preferred IV**	**DV EXP**	**Propensity**
Has DVEC	-1.368	-0.074	0.048	0.391	0.804
	(1.821)	(0.123)	(9.706)	(9.733)	(9.876)
DV incidents				-0.136	
				(0.102)	
Propensity to charge					8.101
					(9.915)
Constant	125.5**	2.279**			
	(61.53)	(0.905)			
Observations	8,864	4,292	7,496	7,496	7,496
Controls	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
LAC FE	YES	YES	YES	YES	YES
Court FE	YES	YES	YES	YES	YES
Number of clusters	136	44	135	135	135
First stage F-statistic			214.6	213.0	210.4
Identification statistic			40.21	40.09	40.22
C-statistic (p-value)			0.802	0.774	0.746

Note. OLS = Ordinary Least Squares, IV = Instrumental Variables, DV = Domestic Violence, EXP = Experience, FE = Fixed Effects, LAC = Local Area Command, clusters refer to courts, cluster robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

prefer the estimate from Column 1. Overall Table 4 provides robust evidence to indicate DVEC has had no significant impact on the probability of a guilty plea.

In Table 5 we present estimates for the impact of DVEC on the time (in days) between charge and finalisation conditional on a guilty plea (plea delay). While the sign and size of the estimated coefficient appears to vary across specification, we find no significant impact from any model. Tests of instrument relevance and endogeneity are consistent with previous tables. That is, our instrument is strongly relevant but DVEC may be able to be treated as conditionally exogenous.

Recall that Table 3 presented mixed evidence with respect to the impact of DVEC on the probability of a conviction. Columns 1 and 2 indicated that DVEC has generated a significant two percentage point increase in the probability of a conviction. However, in Columns 3-5 the sign reverses and the effect becomes insignificant. Results from the endogeneity tests associated with these regressions indicate that we should prefer our OLS estimate on efficiency grounds. In Part A of the Appendix we present the results from three batteries of robustness checks.[25, 26, 27] For the conviction regressions, results from these checks are presented for both OLS and 2SLS IV estimates. We find the size, sign and (in)significance of our DVEC indicator to be retained within but not across estimation method. That is, when estimated through OLS, the effect is a robust two percentage point increase; however, when we switch to IV the effect becomes negative and insignificant in all specifications. For the guilty plea and plea delay regressions, only OLS estimates are presented as the effect is insignificant across all specifications and estimation methods.

In order to better understand the (in) significant impact of DVEC our outcome measures, we partition our sample into appearances where the most serious charge listed is; an Actual Bodily Harm (ABH), Grievous Bodily Harm (GBH), and finally, common assault DV charge. Details from this investigation are left for interested readers in Part A of the Appendix. Three points are worth noting with respect to this investigation. First, the insignificance of DVEC on all outcome measures is retained across ABH, GBH and common assaults within the IV regressions. Second, the insignificance of DVEC on both guilty pleas and plea delay is retained across all specifications and estimation methods. Finally, if DVEC is having an impact on the probability of a conviction, it is driven by an increase in convictions for common assaults. That is, the impact of DVEC is only significant for common assaults in the OLS regressions for convictions.

AGGREGATE LAC LEVEL ANALYSIS

The simple comparison presented earlier found a significant one percentage point increase in the mean NSW conviction rate after the introduction of the policy. However, this comparison fails to control for a variety of confounding factors such as seasonality and systematic differences between LACs that may influence conviction rates. We account for these issues using the dynamic panel data model outlined in the previous section.

Tables 6 and 7 present the main results from our estimates of the DVEC uptake rate on conviction and guilty plea rates, respectively. With respect to both Tables, Column 1 presents OLS estimates that ignore the LAC FE. Column 2 presents estimates from a FE regression that removes the LAC FE using the within transformation. As first noted by Nickell (1981), inclusion of a Lagged Dependent variable (LDV) into a model with ignored unobserved heterogeneity (as in Column 1) leads to an upward bias in the LDV estimator. Nickell (1981) also notes a downward bias in the LDV estimator when the FE is removed using the within transformation (as in Column 2). A consistent estimate therefore lies between these bounds. We can see that the coefficient on the LDV in our preferred GMM specification

Table 6. Aggregate analysis: Main results for conviction rate

	(1)	(2)	(3)
Dependent variable: Monthly conviction rate	**OLS**	**FE**	**GMM**
L. Conviction rate	0.268***	0.009	0.132
	(0.027)	(0.018)	(0.117)
DVEC Uptake rate	0.040**	0.020	-0.032
	(0.016)	(0.017)	(0.075)
Constant	0.532***	0.741***	
	(0.025)	(0.018)	
Observations	2,132	2,132	2,056
Time FE	YES	YES	YES
LAC FE	NO	YES	YES
Number of LACs	76	76	76
Instrument count			54
Hansen p-value			0.771
AR(1) p-value			< 0.001
AR(2) p-value			0.292

Note. L = First lag of associated variable, OLS = Ordinary Least Squares, FE = Fixed Effects, GMM = Generalised Method of Moments, LAC = Local Area Command, finite sample adjusted robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 7. Aggregate analysis: Main results for guilty plea rate

	(1)	(2)	(3)
Dependent variable: Monthly guilty plea rate	**OLS**	**FE**	**GMM**
L.Guilty plea rate	0.175***	-0.020	0.013
	(0.025)	(0.030)	(0.188)
DVEC uptake rate	0.0120	-0.009	-0.191*
	(0.019)	(0.025)	(0.097)
Constant	0.475***	0.601***	
	(0.023)	(0.027)	
Observations	2,132	2,132	2,056
Time FE	YES	YES	YES
LAC FE	NO	YES	YES
Number of LACs	76	76	76
Instrument count			54
Hansen p-value			0.460
AR(1) p-value			0.0117
AR(2) p-value			0.990

Note. L = First lag of associated variable, OLS = Ordinary Least Squares, FE = Fixed Effects, GMM = Generalised Method of Moments, LAC = Local Area Command, finite sample adjusted robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

(Column 3) lies between the upper (0.268) and lower (0.009) bounds provided by the OLS and FE estimates, respectively. The coefficient on the LDV in Column 3 is also below unity. Taken together these results indicate that our model is able to provide a consistent estimate and is dynamically stable.

Turning our attention to the coefficient on the DVEC uptake rate, Column 1 indicates that a one percentage point increase in the DVEC uptake rate is associated with a 0.4 percentage point increase in a LAC's expected monthly conviction rate. This effect appears to be significant at the 5% level. However, once we control for LAC FEs in Column 2, this effect reduces in size and becomes insignificant. This indicates that time invariant systematic differences between LACs are driving the size and significance behind our OLS estimate.

Column 3 presents estimates from our preferred (GMM) specification. This is our preferred specification for three reasons. First, the number of instruments (54) is below the number of LACs (76).[28] As noted in Roodman (2009), a general rule of thumb is to keep the instrument count below the number of cross-sectional units because instrument proliferation can generate biased estimates and reduce the reliability of Sargan-Hansen tests. Second, the p-value associated with the Hansen test is 0.771, indicating that our over-identifying restrictions are valid.[29] Third, Column 3 has a significant coefficient on the first lag of the residuals but an insignificant coefficient on its second lag. This indicates we can begin instrumenting from lag two.[30] Our preferred specification finds no significant relationship between a LAC's monthly DVEC uptake rate and its expected conviction rate.

Table 7 presents the results from the same set of specifications as in Table 6, but now the dependent variable is a LAC's monthly guilty plea rate. Overall the results from Table 7 are consistent with those of Table 6. That is, that the DVEC uptake rate has had no significant impact on the rate of guilty pleas. That said, our GMM specification now indicates that a one percentage point increase in the DVEC uptake rate results in a 0.191 percentage point decrease in the guilty plea rate. This effect is significant at the 10% level. We do not recommend taking this result too seriously for the following three reasons. First, in the robustness checks we experiment with various GMM style augmentations to Column 3 and find the coefficient to be

insignificant in all other variations. Second, GMM estimators are extremely sensitive to choice of specification, and finally, the coefficient is insignificant even when estimated using OLS.

In order to ensure the robustness of our estimates we conduct several robustness and sensitivity checks. Tables with the results from these checks are available for interested readers in Part B of the Appendix.[31, 32] These checks are consistent with our main findings and support our choice of preferred (GMM) specification. That is, these robustness checks further reinforce our finding that the DVEC uptake rate has had no significant impact on either conviction or guilty plea rates, while the sensitivity checks support our choice of GMM specification.

DISCUSSION

This study set out to assess the impact of the 2015 DVEC reforms on court outcomes in NSW. We assessed the impact using two different empirical approaches. The first approach estimated the impact of a DVEC statement on three outcomes at the individual court appearance level. These outcomes were: the probability that a defendant has at least one offence proven against them at their court appearance (which we refer to as a conviction), the probability that a defendant pleads guilty to at least one charge at their appearance, and finally, the time between charge and finalisation conditional on the outcome being a guilty plea (which we refer to as a plea delay).

We found no significant impact of DVEC on the probability of a guilty plea or on plea delay. We found mixed results when estimating the impact of DVEC on the probability of conviction. Using standard OLS methods (adjusted for observed covariates and fixed effects) our analysis indicated that the presence of a DVEC statement increased the probability of a conviction by about two percentage points. However, when estimated through IV methods, which aim to alleviate bias arising from endogenous selection into providing a DVEC statement, this effect was no longer statistically significant. Results from a distance or 'endogeneity' test indicated that DVEC can be treated as conditionally exogenous and therefore we should prefer the OLS estimate. Robustness checks showed that sign, size and (in)significance was retained within but not across estimation method. That is, our results are consistent across a variety of robustness checks when estimated through OLS or IV, but not when we switch from one estimation method to the other. We then partitioned the sample by assault severity finding that the impact of DVEC on convictions is only significant for common assaults. This part of the analysis allowed us to conclude that there is limited evidence to indicate that DVEC increased the probability of a conviction by about two percentage points.

The second approach was a dynamic panel data model with fixed effects that exploited variation in the uptake of DVEC across LACs and time. The outcome measures used in this analysis were the expected monthly DV assault conviction and guilty plea rates. We found no significant relationship between a LAC's uptake of DVEC and either outcome after controlling for time invariant systematic differences between LACs. This result is extremely robust.

Taken together these results suggest that the introduction of the DVEC reforms has had little discernible impact on court outcomes for DV assaults. There are three possible explanations for our failure to find a significant impact. First, our analysis was unable to take into account the quality of the evidence that was collected in DVEC statements. While procedures were put in place to ensure high quality DVEC statements (e.g. only sufficiently trained and experienced police officers collect DVEC statements and the statements are obtained using standardised questions), legal knowledge, experience and interviewing skills would vary enormously across officers and in most cases be inferior to that of police prosecutors who previously collected evidence-in-chief for use in DV cases. Improvements in conviction rates and guilty pleas may have been more apparent had we been able to identify matters where the DVEC evidence gathered was of higher quality and could therefore be applied more effectively to prove elements of an offence in court proceedings. As the recording of digital evidence becomes routine business for general duties police officers, the quality of DVEC statements is likely to improve.

A second possible explanation for our negative results may be that DVEC statements are now being taken in matters where the evidence available would not usually be considered sufficient to proceed to prosecution (e.g. in matters where there are no visible injuries and there are conflicting testimonies). If in these matters, DVEC induces police to proceed with charges anyway (without sufficient evidence to ensure a conviction), then this may offset any positive impact DVEC has had on convictions. That is, while DVEC may increase probability of a conviction through higher quality evidence-in-chief in some instances, it may simultaneously, decrease the likelihood of a conviction in other instances, as police may have become more likely to proceed to prosecution in matters where they are less likely to obtain a conviction, irrespective of DVEC. In principle these effects could offset each other thus preventing us from detecting a significant impact of the policy.

Finally, DV assault matters for which DVEC is likely to prove most effective in increasing convictions are those involving reluctant or unreliable complainants. DVEC was designed to reduce the trauma for victims associated with re-telling the events and make it easier for the victim to recall these events during cross-examination. It also sought to reduce the capacity for the defendant to intimidate the complainant to change or recant their evidence. However, even with a DVEC statement, a victim must be available for cross-examination if the defendant

pleads not guilty and the matter proceeds to a hearing. This can be traumatic; particularly in matters where defendants are self-represented. Victims may also be intimidated by the defendant before their matter reaches court and pressured to withdraw their cooperation with the prosecution. In these situations, DVEC statements can be submitted as evidence in court proceedings (i.e. in the absence of complainant consent) but police are reluctant to do so as it would involve subpoenaing the victim and treating them as a hostile witness. Rather than risk re-traumatisation police will typically elect to withdraw the charges or proceed without the complainant's evidence. If the presence of a DVEC statement does not substantially affect a victim's willingness to proceed with criminal charges its impact on proven outcomes or guilty pleas is likely to be diminished.

There are other potential benefits arising from the implementation of the DVEC reforms that have not been considered in this evaluation and should be pursued in future research. As noted above, the DVEC reforms were also intended to reduce the trauma for victims when providing evidence in DV proceedings, reduce the difficulty for victims to recall events during cross examination and reduce the capacity of the defendant to intimidate victims into changing or recanting their evidence. The victim experience of providing DVEC statements and the use of DVEC in criminal proceedings was beyond the scope of this evaluation but is certainly worthy of further investigation. The impact of the new technology on police efficiency should also be considered. Prior to DVEC, police would attend a DV incident, and either take notes at the time of the incident and ask victims to return to the station to give a formal statement, or take a notebook statement and then type up a formal statement back at the station. With the new digital cameras, police can now take electronic statements at the time of the incident and then need only upload the file into COPS when entering other relevant details about the incident. These time savings and the associated costs would be significant given the volume of DV offences that police attend and are likely to grow as these processes become embedded in police practice.

The impact of DVEC on court outcomes should continue to be monitored. NSW courts were the first in Australia to accept video statements as evidence-in-chief for DV matters. It may take time for complainants to begin to trust the new process, be willing to consent to video statements and pursue criminal charges. Police expertise in collecting evidence of sufficient quality will also continually improve. If further enhancements are made to procedural safeguards for victims in criminal proceedings, then over the longer term we may see the DVEC reforms achieve their ultimate aim; to enhance victim safety and reduce domestic and family violence in our community.

ACKNOWLEDGEMENTS

We would like to thank Wai-Yin Wan, Don Weatherburn and the research team at BOCSAR more broadly for their advice and feedback. We would also like to thank Inspector Sean McDermott from the NSW police for providing the data used in this report and for his constructive input into the evaluation. We would also like to thank our anonymous referees for their valuable suggestions, Florence Sin for desktop publishing this report, and finally, Clare Ringland for her extraordinarily detailed feedback.

NOTES

1 In the case of an Aboriginal or Torres Strait Islander; is or has been part of the extended family or kin of the other person according to the Indigenous kinship system of the person's culture.

2 Hoyle (1998) also outlines not wanting to leave their partner for emotional, social and economic reasons as another driver for under-reporting.

3 DVEC also allows these recordings to be used as evidence-in-chief for concurrent or related proceedings for applications for Apprehended Domestic Violence Orders under the *Crimes (Domestic and Personal Violence) Act 2007*.

4 In both QLD and WA these protections are extended further to include a variety of adult witnesses the court considers to be vulnerable. See *Evidence Act 1977 (QLD) s 21A(1); Evidence Act 1906 (WA) s 106R*.

5 Pre-recorded evidence-in-chief is also available in these jurisdictions for victims of sexual assault. In some jurisdictions pre-recorded evidence is restricted to children and in others it extends to adult victims courts deemed to be at risk of disadvantage.

6 The children must be under the age of 16 at the time of the offence and under the age of 18 when they give evidence.

7 DV incidents recorded by officers from specialist jurisdictions are removed from the data. These jurisdictions do not represent geographical areas and have extremely low counts of recorded DV incidents. For example, some of these specialist jurisdictions include; major events and incidents group, counter-terrorism operations and transport commands.

8 Information on the number of incidents was only available for the period January 2014 - June 2016. After recording an incident police can then choose whether or not to proceed with charges. In this sense charges are a subset of incidents.

9 The LPM in Equation 1 refers to our first two outcome measures. For the plea delay regressions we estimate Equation 1 as: $y_{icjt} = \beta * D_{icjt} + \mathbf{X}_{icjt} + \alpha_c + \delta_j + \lambda_t + \epsilon_{icjt}$. Similar reasoning applies to Equation 3.

10 As first noted by Bloom (1984), one-sided non-compliance with no always-takers results in the Local Average Treatment Effect being approximately equal to Average Treatment Effect on the Treated.

11 As outlined in Angrist & Krueger (2001) 2SLS estimators have a causal interpretation that is robust to non-linarites' induced by binary dependent variables. The main advantages presented by competing non-linear models such as Probit or Logit specifications are incurred when the objective is prediction not causal inference.

12 There are instances (about 7% of our total sample) in which the same person appears more than once in our sample. In order to safeguard against this possible violation of the independence assumption, in our robustness checks we restrict the analysis to include only an individual's first court appearance in our sample. We find no meaningful change to the main results.

13 Angrist & Pischke (2008) show how re-coding a continuous instrument into a binary variable provides a parsimonious non-parametric model for the first relation E(Di|zi). For example, see Angrist, Graddy & Imbens (2000).

14 In the appendix we also provide robustness checks where we use the full continuous form of the instrument with no meaningful change to the main results.

15 Strictly speaking valid identification in an IV set-up also requires monotonicity of the instrument(s) as outlined in Angrist & Imbens (1994). Our instrument meets this assumption.

16 We look at DV assault charges finalised within 6 months in NSW Local Courts only.

17 FOD involves subtracting the average of all available future observations from each variable. Therefore, not only does FOD remove unobserved heterogeneity, but it also allows us to implement the transformation over all observations except the last period for each LAC. This mitigates the impact of missing observations and preserves the size of our sample. For concreteness, let w denote any of the variables in Equation 4. The FOD transform of w is then given by $w_{j,t+1} = c_{jt}\left(w_{jt} - \frac{1}{T_{jt}}\sum_{s>t} w_{js}\right)$, where the sum is over available future observations, T_{jt}, and the scale factor is calculated as $c_{jt} = \sqrt{T_{jt}/(T_{jt+1})}$.

18 The FOD transformation is implemented instead of FD for two reasons. First, FD GMM estimators are sensitive to missing observations in unbalanced panels. The sample over which the analysis is conducted contains several missing values due to the method in which the dependent variables are constructed. Second, Hayakawa (2009) shows that when T ≥ 10 the combination of FOD transform on the equation of interest and the backward orthogonal deviation transform on the instruments exhibits a smaller bias relative to difference GMM.

19 In this bulletin we do not investigate the potential long-run effect of an increase in the DVEC uptake rate as we do not have enough post-policy periods to reasonably support this type of analysis. An investigation of the long-run effect of the policy is left for a follow up study.

20 We look at the FD instead of FOD residuals because all FOD residuals are mathematically interrelated as they depend on forward lags.

21 In part A of the Appendix we remove clustering, cluster at both the court and LAC level, as well as the LAC level alone with no meaningful change to the results.

22 The estimate for the DVEC indicator presented in Column 2 is an average marginal effect, not the raw Probit coefficient. This holds true for all Tables in this paper and it's Appendix that present DVEC estimates from a non-linear model. Further note that the number of courts in this Column is 123 instead of 137. These 14 courts are removed from our sample because there is no variation in the dependent variable within these clusters. That is, appearances at these courts either always or never result in a conviction.

23 The insignificance of this test may reflect low power of the test. However, as far as we are aware, no such literature relating to the power of this test exists.

24 Note that in Part A of the Appendix we rotate between combinations of time, court and LAC FEs with finding that the LAC and/or court FEs are important for capturing the impact of this experience.

25 In the first battery of robustness checks we check the robustness of our results to a variety of sample restrictions. Specifically we exclude all control variables that aren't significant in an OLS regression on Equation 1, restrict the sample to appearances in which the defendant has no prior court appearances (in our sample), and finally, restrict the sample to include only appearances in which the defendant has no prior convictions. Overall we find no significant change to our main results.

26 In the second set of robustness checks, we experiment with varying levels of Fixed Effects (FEs) by iteratively removing and replacing the time, LAC and court FEs. Overall we find no significant change to our main results.

27 In the third set of robustness checks we iteratively rotate between standard errors that have no clustering but are robust to arbitrary forms of heteroskedasticity, have no clustering and are not robust, as well as robust standard errors that are clustered at the LAC and court, and just the LAC level. The size of our standard errors does not significantly vary with the different levels of clustering and therefore we find no meaningful change to our inference.

28 In this (and all GMM specifications reported in Tables 6 and 7) we restrict the number of lags for the endogenous variables to two in order to avoid instrument proliferation. In the robustness checks we relax this restriction and also present results from several other augmentations to the number/type of lags used as instruments with no meaningful change to the overall results.

29 We omit results from a Sargan test because these results are not robust to heteroskedasticity present in our model.

30 These three points justify our choice of GMM specification over the competing GMM alternatives outlined in Part B of the Appendix.

31 Broadly speaking the robustness checks include re-estimating Equation 4 using a variety of GMM style augmentations to our preferred specification, investigating the possibility that DVEC may have had a lagged impact on conviction or guilty plea rates, rotating between various levels of fixed effects, various sample restrictions, and switching the unit of analysis from monthly to quarterly. We find no meaningful change to our main results in any of these checks.

32 The sensitivity checks include testing the (joint) significance of the fixed effects, checking the sensitivity of the significance of the estimates to various types of standard errors, conducting a Choi (2001) unit root test in order to determine whether or not our dependent variables are stationary, and finally, testing for autocorrelation in the residuals using the Wooldridge (2010) test. Results from these checks support our choice of specification.

REFERENCES

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

Angrist, J., & Krueger, A. B. (2001). Instrumental variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives, 15*(4), 69-85.

Angrist, J. D., Graddy, K., & Imbens, G. W. (2000). The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish. *The Review of Economic Studies, 67*(3), 499-527.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricists Companion.* Princeton University Press.

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies, 58*(2), 277-297.

Australian Bureau of Statistics. (2006). Personal Safety Survey. (Cat No. 4906.0). Retrieved from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4906.0Main+Features12005%20(Reissue)?OpenDocument

Birdsey, E, M., & Smith, N. (2012). *The Domestic Violence Intervention Court Model: A follow-up study* (Crime and Justice Bulletin No. 155). Retrieved from NSW Bureau of Crime Statistics and Research website:

http://www.justice.nsw.gov.au/Documents/Media%20Releases/2012/060312_BOCSAR_CJB155.pdf

Birdsey, E, M., & Snowball, L. (2013). *Reporting Violence to Police: A survey of victims attending domestic violence services* (Bureau Brief No. 91). Retrieved from NSW Bureau of Crime Statistics and Research website: http://www.bocsar.nsw.gov.au/Documents/BB/bb91.pdf

Bloom, H. S. (1984). Accounting for No-Shows in Experimental Evaluation Designs. *Evaluation Review, 8*(2), 225-246.

Burton, M., Evans, R., & Sanders, A. (2006). Are special measures for vulnerable and intimidated witnesses working? Evidence from the Criminal Justice Agencies, *Home Office Online Report, 01/06.*

Cashmore, J., & Trimboli, L. (2005). *An Evaluation of the NSW Child Sexual Assault Specialist Jurisdiction Pilot*. Retrieved from NSW Bureau of Crime Statistics and Research website: http://www.bocsar.nsw.gov.au/Documents/r57.pdf

Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance, 20*(2), 249-272.

Davies, E., & Hanna, K. (2013). Pre-Recording Testimony in New Zealand: Lawyers and Victim Advisors Experiences in Nine Cases. *Australian & New Zealand Journal of Criminology, 46*(2), 289-305.

Dutton, D. G. (1995). The Domestic Assault of Women: Psychological and Criminal Justice Perspectives. Vancouver: University of British Columbia Press.

Elliott, D. S. (1989). Criminal Justice Procedures in Family Violence Crimes. *Crime and Justice, 11,* 427-480.

Fagan, J. (1989). Cessation of Family Violence: Deterrence and Dissuasion. *Crime and Justice, 11,* 377-425.

Ford, D., & Regoli, M, J. (1993). The Indianapolis Domestic Violence Prosecution Experiment, National Institute of Justice.

Grech, K., & Burgess, M. (2011). Trends and Patterns in Domestic Violence Assaults: 2001 to 2010 (Bureau Brief, 61). Retrieved from NSW Bureau of Crime Statistics and Research website: http://www.bocsar.nsw.gov.au/Documents/BB/bb61.pdf

Hamlyn, B., Phelps, A., Turtle, J., & Sattar, G. (2004). Are Special measures Working? Evidence from Surveys of Vulnerable and Intimidated Witnesses, Home *Office Research, Development and Statistics Directorate.*

Hayakawa, K. (2009). A Simple Efficient Instrumental Variable Estimator for Panel AR(p) Models when both N and T are large. *Econometric Theory, 25*(3), 873-890.

Hoyle, C. (1998). Negotiating Domestic Violence: Police, Criminal Justice, and Victims, Clarendon Press.

Kleibergen, F., & Paap, R. (2006). Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics, 133*(1), 97-126.

McConachy, D. (2002). An Evaluation of the Electronic Recording of Children's Evidence: Final Report, *New South Wales Police,* Sydney.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica, 49*(6), 1417-1426.

Roodman, D. (2009). How to do xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal, 9*(1), 86-136.

Schmidt, J., & Steury, E. H. (1989). Prosecutorial Discretion in Filing Charges in Domestic Violence Cases. *Criminology, 27*(3), 487-510.

Staiger, D. O., & Stock, J. H. (1994). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65*(3),557-586.

Stock, J.H. and Yogo, M. 2005. Testing for Weak Instruments in Linear IV Regression. In D.W.K. Andrews and J.H. Stock (Eds). *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. 80-108. New York: Cambridge University Press.

Westera, N. J., Kebbell, M. R., & Milne, R. (2011). Interviewing Rape Complainants: Police Officers Perceptions of Interview Format and Quality of Evidence. *Applied Cognitive Psychology, 25*(6), 917-926.

Windmeijer, F. (2005). A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators. *Journal of econometrics, 126*(1), 25-51.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, MIT press.

APPENDIX

Further robustness, sensitivity and diagnostic checks are available electronically at www.bocsar.nsw.gov.au

NSW Bureau of Crime Statistics and Research - Level 1, Henry Deane Building, 20 Lee Street, Sydney 2000
bcsr@justice.nsw.gov.au • www.bocsar.nsw.gov.au • Ph: (02) 8346 1100 • Fax: (02) 8346 1298
ISSN 1030-1046 (Print) ISSN 2204-5538 (Online) • ISBN 978-1-925343-35-9