The Australian National University
Centre for Economic Policy Research
*DISCUSSION PAPER*

# Weak Tests and Strong Conclusions: A Re-Analysis of Gun Deaths and the Australian Firearms Buyback

**Christine Neill and Andrew Leigh**

**DISCUSSION PAPER NO. 555**
**June 2007**

**ISSN: 1442-8636**
**ISBN: 1 921262 26 5**

Christine Neill, Department of Economics, Wilfrid Laurier University
Email: cneill@wlu.ca, Website: http://www.wlu.ca/sbe/cneill

Andrew Leigh, Economics Program, Research School of Social Sciences, Australian National University
Email: andrew.leigh@anu.edu.au, Website: http://econrsss.anu.edu.au/~aleigh/

## ABSTRACT

Using time series analysis on data from 1979-2004, Baker and McPhedran (2006) argue that the stricter gun laws introduced in the National Firearms Agreement (NFA) post-1996 did not affect firearm homicide rates, and may not have had an impact on the rate of gun suicide or accidental death by shooting. We revisit their analysis, and find that their results are not robust to: (a) using a longer time series; or (b) using the log of the rate rather than the level (to take account of the fact that the rate cannot fall below zero). We also show that claims that the authors had allowed both for method substitution and for underlying trends in suicide or homicide rates are misleading. The high variability in the data and the fragility of the results with respect to different specifications suggest that time series analysis cannot conclusively answer the question of whether the NFA led to lower gun deaths. Drawing strong conclusions from simple time series analysis is not warranted, but to the extent that this evidence points anywhere, it is towards the firearms buyback reducing gun deaths.



*JEL Codes:* I12, K14

*Keywords:* firearms ownership, homicide, suicide

Understanding the relationship between firearms availability and gun deaths is an important policy question for many countries. Australia's 1996-97 National Firearms Agreement (NFA) – which tightened gun ownership and licencing requirements and removed 600,000 guns from a country with a population of 20 million – offers a potentially useful natural experiment. Several authors have exploited this natural experiment, including Reuter & Mouzos (2003), Ozanne-Smith *et al.* (2004), and Chapman *et al.* (2006).

In a paper that received widespread media attention in Australia, Baker & McPhedran (2006) (henceforth 'BM') seek to determine the impact of the NFA on homicide, suicide and accidental death rates. Using time series techniques to analyse Australian deaths data from 1979-2004, BM conclude 'Homicide patterns (firearm and non-firearm) were not influenced by the NFA, the conclusion being that the gun buy-back and restrictive legislative changes had no influence on firearm homicide in Australia.' The purpose of this brief comment paper is to highlight important flaws in BM.

In principle, a variety of statistical approaches could be used to model the effect of the NFA on firearm deaths. Exploring the full gamut of approaches would be an important exercise, but it is not the purpose of this paper. Rather than beginning afresh, our paper takes BM's framework as its starting point, and tests whether that paper's findings are robust to minor changes in specification – using a longer time series, and using the log of the rate rather than the level (to take account of the fact that the rate cannot fall below zero). We also discuss whether these results can properly control for changes in the underlying factors determining overall suicide and homicide rates simultaneously with testing for the possibility of method substitution.

We do, however, provide some brief comments on the methodology employed by two other recent papers on the same policy change, and include an appendix that discusses possible identification strategies that may be of use in guiding future research in this area.

## Using All Available Data

In their paper, BM show graphs of Australian homicide rates (gun and non-gun) and suicide rates (gun and non-gun) from 1915 onwards. However, despite the fact that they have 95 years of data, their econometric analysis uses only 26 years of data (1979-2004). Since the authors do not provide any justification for using less than one-third of the available data, it is useful to re-analyse their results over a longer time frame.[1] We begin by attempting to replicate BM, then use the full series. Table 1 shows how these changes in specification would affect the results as reported by BM. Details of the parameter estimates in the ARIMA(1,1,1) model are in Appendix Table B1.[2] Several facts are notable about these results.

First, predicted firearm suicides and homicides post-1997 are considerably lower when data from 1979 is used, than when the full data set is used. The main reason for the differences between the results from the shorter and longer time series is that in the late 1970s and early 1980s firearm suicide and homicide rates were at historically very high levels. Estimating an ARIMA model from a local maximum biases the estimates towards finding a long-term declining trend, resulting in predictions of firearm deaths that are relatively low. The result is that BM's estimates will tend to understate the difference between the predicted and actual decrease in firearm suicides and homicides post-1997. The models estimated from the longer time series suggest that there were on average 250 fewer firearm deaths per year after the implementation of the NFA than would have been expected based on the predictions from the ARIMA model – close to double the numbers implied by BM's estimates.

There is no discussion in BM as to the reasons for restricting their sample to the shorter time period. We can think of two possible explanations. First, there may have been a structural break in the series after 1978. BM do not attempt to make that argument, however, and we can see no obvious reason why this should have been the case. The second possible reason is that the data may have come from different sources, or have been calculated differently. The ABS provides data on deaths due to assault and deaths due to self-harm in its Cause of Death publications going back to 1915. While there have been revisions to the categories of causes of death, including a move from the use of ICD-8 to ICD-9 categorisation in 1979, this seems unlikely to have greatly affected estimates of suicide or homicide deaths. The move from ICD-9 to ICD-10 left these virtually unchanged (Kriesfeld & Harrison 2005).

**Table 1. Comparison of predicted and observed rates of firearm homicide and suicide, ARIMA(1,1,1) model (dependent variable is the number of deaths per 100,000 people)**

| | BM | Replication | Full Sample (1915-2004) | Extend by a Decade (1969-2004) |
|---|---|---|---|---|
| | *Panel A: Dep Var is Firearm Homicide* | | | |
| Predicted rate | 0.28 | 0.303 | 0.450 | 0.400 |
| Observed rate | 0.27 | 0.269 | 0.269 | 0.269 |
| Difference | -0.01 | -0.034 | -0.181 | -0.131 |
| P (Predicted≤Observed) | 0.14 | 0.0248 | 0.0002 | 0.0011 |
| Lives Saved per Year | 2 | 7 | 36 | 26 |
| | *Panel B: Dep Var is Firearm Suicide* | | | |
| Predicted rate | 1.85 | 1.854 | 2.465 | 2.020 |
| Observed rate | 1.22 | 1.230 | 1.230 | 1.230 |
| Difference | -0.630 | -0.624 | -1.235 | -0.790 |
| P (Predicted≤Observed) | 0.001 | 0.0000 | 0.0000 | 0.0000 |
| Lives Saved per Year | 126 | 125 | 247 | 158 |

Note: Figures in the column headed “BM” are taken directly from the text in Baker and McPhedran (2006). Replication shows our best attempts to produce the same results, using the same time period (1979-1996). Predicted rate is based on estimating an ARIMA model using data up to 1996, and forecasting out to 2004. Observed rate is the average in the data from 1996 to 2004. Lives saved per year is calculated by multiplying the difference (change in the rate per 100,000 people) by 200 (since the population of Australia is 20 million). The model estimates and predictions are from R.

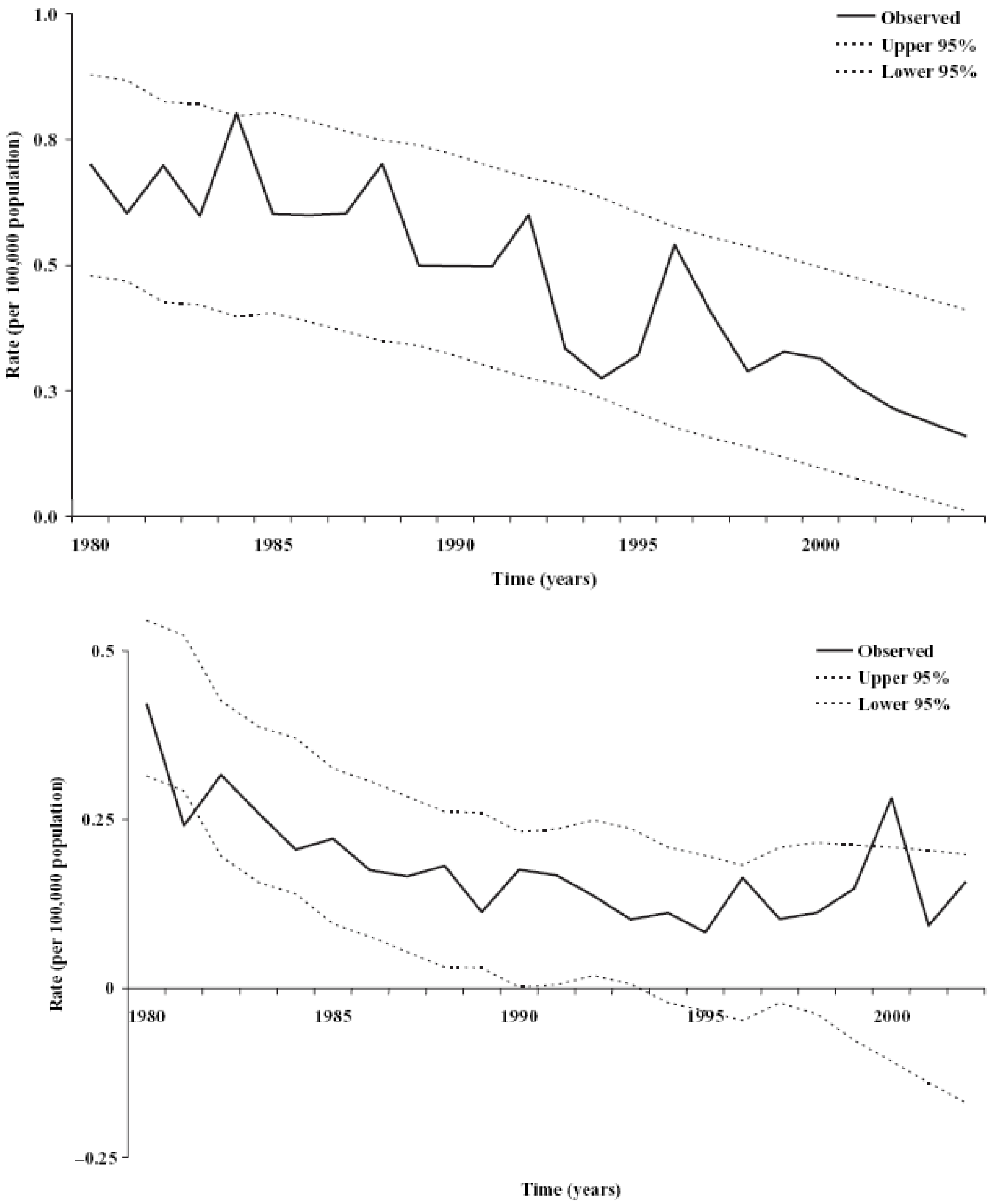
A second key concern with the time period selected by BM is that estimating an ARIMA(1,1,1) model using data only from 1979-1996 is very dubious, especially for firearm homicides. The point estimates over the shorter time period are very sensitive to the specification and even the statistical package used. Indeed, using the ARIMA routine in STATA and exactly the data shown in BM, we could not achieve convergence in this model for non-firearm homicides. This partly reflects the well-known difficulties associated with estimating an ARIMA model with such a short time series. However, it also reflects the fact that the ARIMA(1,1,1) model appears to be particularly inappropriate in this case – statistical tests consistently reject the hypothesis that homicide rates follow a non-stationary process. BM present no statistical tests to show whether these models are appropriate. Simple augmented Dickey-Fuller tests, which are a simple way to determine whether a series is integrated (non-stationary) are shown in the Appendix Table B2. They strongly reject the null hypothesis that the series are integrated in the case of firearm and non-firearm homicides.

Third, these results are not driven by some distant historical episode. The final column of Table 1 shows that even if we extend the sample period by only a decade, the probability that gun deaths were not lower than would be expected based on a simple ARIMA(1,1,1) model is well below 1 per cent.

## The Resurrection Problem

As the dependent variable, the authors use the number of deaths of a particular type per 100,000 people. If the firearm death rate were high and stable, this might not present a problem. However, since the firearm death rate is low and volatile, the estimates in BM place a non-zero probability on the death rate for accidental firearm deaths and firearm homicides falling below zero. As is shown graphically in Figure 1, which merely repeats figures 4A and 5 from BM, it cannot be rejected at the 95 per cent confidence level that there will be negative deaths after 1994 for accidental firearm deaths, and after 2004 for firearm homicides. Projecting out further, the models predict that by 2010 deaths attributable to assault with a firearm or an accidental firearm incident would be negative. This is clearly concerning: an effective modelling strategy should place a zero probability on the occurrence of a logically impossible event. It points to a clear flaw in the specification of the model.

**Figure 1: Death Rates Cannot Fall Below Zero**




Panels are from BM. First panel shows their Figure 4A (firearm homicide rate per 100,000 people) and Figure 5 (accidental firearm death rate per 100,000 people).

In the case of death rates, there are straightforward solutions to this problem – in order to ensure that predicted death rates are always positive, researchers can use a Poisson model, or take log of the rate, rather than use the rate itself. Such approaches are common in studies of the impact of gun laws on deaths (see eg Ozanne-Smith *et al.* 2004; Duggan 2002; Beautrais, Fergusson & Horwood 2006, Chapman *et al.* 2006). Indeed, we have been unable to find another study examining firearm-related deaths that simply uses the death rate as a dependent variable with no further specification checks.

Table 2 shows estimates of the post-1996 reduction in deaths, taken from ARIMA(1,1,1) models with the dependent variable as the log of the death rate, both for the period 1979-2004, and for the full period 1915-2004. The estimates are converted into level terms to facilitate comparison. The results show an increase in the predicted death rate in the log model compared with the levels models shown in Table 1. More importantly, though, there is a very clear reduction in the probability that the observed series is greater than the predicted series for firearm homicides in the model estimated from 1979 to 1996. The estimated probability that the predicted series is smaller than the actual series drops from 2.6 per cent to 0.2 per cent. Thus, there is extremely strong evidence in this model that the observed firearm homicide rate was lower after 1997 than would have been expected based on an ARIMA(1,1,1) log model.

**Table 2. Comparison of predicted and observed rates of firearm homicide and suicide, ARIMA(1,1,1) model (dependent variable is the log of the death rate)**

| | BM 1979-1996, depvar is rate | Replication 1979-1996, depvar is log | Full Sample 1915-1996, depvar is log |
|---|---|---|---|
| | *Panel A: Dep Var is Firearm Homicide* | | |
| Predicted rate | 0.28 | 0.338 | 0.446 |
| Observed rate | 0.27 | 0.269 | 0.269 |
| Difference | -0.01 | -0.069 | -0.177 |
| P (Predicted≤Observed) | 0.14 | 0.0026 | 0.0002 |
| Lives Saved per Year | 2 | 14 | 35 |
| | *Panel B: Dep Var is Firearm Suicide* | | |
| Predicted rate | 1.85 | 1.941 | 2.396 |
| Observed rate | 1.22 | 1.230 | 1.230 |
| Difference | -0.63 | -0.712 | -1.166 |
| P (Predicted≤Observed) | 0.001 | 0.0000 | 0.0000 |
| Lives Saved per Year | 126 | 142 | 233 |

Note: Figures in the column headed "BM" are taken directly from the text in Baker and McPhedran (2006). The other columns use the log of the death rate, rather than the death rate itself, as the dependent variable in the model estimated to obtain forecasts. These forecasts are then converted back to a death rate, to ensure comparability with Table 1. Model estimates and predictions are from R.

## Control or Substitute?

In analysing the impact of the NFA on gun deaths, BM attempt to address two questions:

(a) did the NFA lead to method substitution in homicide or suicide?

(b) were there other social factors apart from the NFA that affected death rates after 1997?

To answer the first of these questions, it is necessary to see how non-firearm death rates changed after 1997. To answer the second of these questions, it is necessary to identify a control group, which is affected by the same social factors, but unaffected by the NFA.

BM state: ‘The inclusion of suicide and homicide by methods other than firearm provided a control against which the political, social and economic culture into which additional legislative requirements for civilian firearm ownership occurred could be evaluated, as well as determining the level of method substitution within homicide and suicide.’ It is clear, however, that the potential for method substitution invalidates the use of the non-firearm death rate as a control for the firearm death rate. A good control group must be unaffected by the policy change. If the gun buyback caused an increase in non-gun homicides, the non-gun homicide rate cannot be a good control group for the gun homicide rate. For a more formal discussion of the problems that arise when attempting to use non-firearm deaths to look at both substitution effects and as a control group, see the Appendix.

A key problem with the use of non-firearm death rates as both a control group and a test for method substitution is that, even with a finding of a statistically significant decrease in firearm homicides or suicides post-NFA, BM could draw virtually any conclusion they wish. In the event that non-firearm deaths increased after the buyback, they could claim that this is due to method substitution, showing the NFA too have been ineffective. In the event that non-firearm deaths fell after the buyback, they could claim that non-firearm deaths are a control for firearm deaths, and therefore that some factor other than the buyback must have caused the reduction in firearm deaths.

In our view, the lack of any marked, sustained increase in non-firearm suicide or homicide rates after 1997, as shown in BM, suggests there is little reason to suspect any strong method substitution effect.[3] It is difficult to be certain of this without first having in place a model that satisfactorily explains movements in total deaths, and simple time series models are not sufficient in this regard. Further, it would likely be difficult to identify method substitution if it occurred, given that firearm deaths are small relative to total numbers of deaths. Despite these difficulties, and although we cannot rule out the possibility that non-firearm suicide and homicide would have fallen faster in the absence of the NFA, the fact that overall violent deaths have fallen since 1996 – even accounting for the Port Arthur effect – suggests there has not been substantial method substitution.

## Other Recent Approaches to Analysing the NFA

Two other recent papers have also used aggregate-level data beginning in 1979 to explore Australian gun law reforms – Chapman *et al.* (2006) and Ozanne-Smith *et al.* (2004). As we have argued above, taking a starting point of 1979 biases both of these two papers against finding any statistically significant reduction in death rates. Nonetheless, both find significant reductions in firearms deaths in Australia after 1997.

Of the two, Chapman *et al.* (2006) takes an approach closer to that of BM. Unlike BM, they use a negative binomial functional form, which corrects for the fact that death rates cannot fall below zero. They estimate the model allowing for a break in both the level and the trend in the series at 1997.[4] The authors attempt to deal with the issue of preexisting trends by allowing for such a trend, and testing for a break in that trend. They find what appear to be statistically significant downward movements in both firearm suicides and homicides, and a faster rate of decrease in those series after 1997 (although in the case of firearm homicides, this is not statistically significant). They also recognise the possibility of method substitution, and conclude that the fact that non-firearm deaths also decreased post-1997 suggests that method substitution did not occur.

Ozanne-Smith *et al.* (2004) takes a somewhat different approach. The authors note that Victoria's firearm laws were tightened in 1988, and then in 1997 the rest of the states essentially 'caught up' with Victoria's tougher legislation. They then use this quasi-experiment to compare the rates of decline of firearm deaths in Victoria relative to the rest of Australia after 1988 and then again after 1997. Because they use sub-national variation in policy, they are able to control for any national-level changes in firearm death rates by including a full set of year dummy variables, rather than relying on time trends. This avoids most of the problems associated with beginning the analysis in 1979, and mitigates the problems of potentially predicting negative death rates. The latter is in any case not a concern since they use a log specification. To be clear, Ozanne-Smith *et al.* (2004) is subject to none of the concerns that we have raised in relation to BM's paper. They do not, however, consider the possibility of method substitution in any detail.

## Conclusion

A common axiom in social science research is that a test that fails to reject the null hypothesis should not automatically lead the researcher to accept the null hypothesis. If the sample size is small or the dependent variable is mis-specified, the test is commonly described as being ‘weak’. In such an instance, there may well be a true effect, but the test lacks the statistical power to identify the effect at conventional levels of statistical significance. Indeed, we surmise that most papers in social science would fail to find an effect if we restricted the sample size to just 26 datapoints.

When statistical tests are weak, the proper approach is to focus attention on the point estimates and their associated standard errors (McCloskey and Ziliak; 1996). BM report results suggesting that the NFA prevented 2 deaths per year from homicide (statistically insignificant), and prevented 126 deaths per year from suicide (statistically significant). Their results showed no evidence that the gun buyback led to method substitution, with the non-firearm homicide rate staying stable, and the non-firearm suicide rate falling.

Re-analysing these results either with a longer time series or using the log of the death rate, however, strengthens the evidence against the null hypothesis that the NFA had no effect on firearm suicides or homicides, and more than doubles the estimated number of lives ‘saved’, from 128 to 282 lives per year. In particular, our re-analysis shows a statistically significant reduction in deaths due to both firearm homicides and suicides.

As we point out, the high degree of variability in the underlying data and the fragility of the estimated results with respect to different specifications and even statistical packages used suggest that time series analysis alone cannot conclusively answer the question of whether the NFA cut gun deaths. However, contrary to Baker and McPhedran's summary of their own findings,[5] to the extent that the time series evidence points anywhere, it is towards the conclusion that the NFA reduced gun deaths.

## References

Baker, J & S McPhedran (2006) ‘Gun Laws and Sudden Death: Did the Australian Firearms Legislation of 1996 Make a Difference?’ *British Journal of Criminology*, Advance Access published on October 18, 2006. (doi:10.1093/bjc/azl084)

Bertrand, M, E Duflo & S Mullainathan (2004) ‘How much should we trust difference-in-differences estimates?’ *Quarterly Journal of Economics*, vol 119, no 1, pp 249-75

Chapman, S, P Alpers, K Agho & M Jones (2006) ‘Australia’s 1996 gun law reforms: faster fall in firearm deaths, firearm suicides and a decade without mass shootings’, *Injury Prevention*, vol 12, no 6, pp 365-372

Beautrais, AL, DM Ferguson & LJ Horwood (2006) ‘Firearms legislation and reductions in firearm-related suicide deaths in New Zealand’, *Australian and New Zealand Journal of Psychiatry*, vol 40, pp 253-259

Duggan, M (2001) ‘More Guns, More Crime’, *Journal of Political Economy*, vol 109, no 5, pp 1086-1114

Kriesfeld, R & J Harrison (2005) ‘Injury Deaths, Australia, 1999’, *Injury Research and Statistics Series*, Number 24 (AIHW cat. no. INJCAT 67). Adelaide: AIHW

McCloskey, DN & ST Ziliak (1996) ‘The Standard Error of Regressions’, *Journal of Economic Literature*, vol 34, no 1, pp 97-114

Ozanne-Smith, J, K Ashby, S Newstead, V Z Stathakis & A Clapperton (2004) ‘Firearm related deaths: the impact of regulatory reform’, *Injury Prevention*, vol 10, pp 280-286

Reuter, P & J. Mouzos (2003) ‘Australia: A Massive Buyback of Low-Risk Guns’, In J Ludwig & PJ Cook (eds) *Evaluating Gun Policy: Effects on Crime and Violence*, Washington, DC, pp 121-156

## Appendix A. A Simple Model of Controls Versus Substitutes

The following discussion is framed in terms of suicide, where method of choice has been more commonly studied, but could be applied equally to murder. It is described in terms of firearm and hanging, simply to make the model more concrete and relevant to the issues discussed in this comment. Suppose these are the only two possible forms of suicide, and that numbers of suicides per year in a particular country are determined by:

$$Firearm_t = \alpha_1 + \beta_1 X_t + \rho_1 guns_t + \varepsilon_t \qquad (1)$$
$$Hanging_t = \alpha_2 + \beta_2 X_t + \rho_2 guns_t + \upsilon_t \qquad (2)$$

We assume that there are sufficient numbers of deaths that the errors are approximately normally distributed, and that there is no serial correlation in the error terms. We also assume that $\mathrm{cov}(\varepsilon_t \upsilon_t) = 0$.

In this simple model, there are thus two reasons why firearm and hanging deaths may be related to each other. First, movements in $X_t$ cause changes in deaths. This variable is intended as a simple representation of what BM refer to as the "political, social and economic culture", which may affect the decision to suicide. Suppose, for instance, that increases in the unemployment rate increase the numbers of suicides – then $X_t$ is the unemployment rate and $\beta_1$ and $\beta_2$ would be positive. For the purposes of evaluating the effect of changes in the gun laws, we are not concerned about these coefficients, except that we want to ensure that we do not mis-attribute falls in firearm-related suicides to general social changes.

The variable $guns_t$ represents the legal environment around gun ownership. The model above allows that changes in gun laws may have led to changes in both firearm and hanging suicides. The simple intuition that drives tightening of gun laws is typically that $\rho_1$ would be negative: a tightening of gun laws would reduce firearm-related suicides. However, there is a possibility that method substitution may occur, so that changes in gun laws that reduce firearm-related suicides may actually *increase* suicides by hanging.

Given this model, total suicides are:

$$\begin{aligned} Suicide_t &= Firearm_t + Hanging_t \\ &= (\alpha_1+\alpha_2) + (\beta_1+\beta_2)X_t + (\rho_1+\rho_2)guns_t + (\varepsilon_t + \upsilon_t) \\ &= (\alpha_1+\alpha_2) + (\beta_1+\beta_2)X_t + (\rho_1+\rho_2)guns_t + \mu_t \qquad (3) \end{aligned}$$

The central questions for this paper are:

(1) is $\rho_1$ negative? (did firearm-related deaths fall after the change in laws?);
(2) is $\rho_2$ positive? (was there method substitution?); and
(3) is $(\rho_1 + \rho_2)$ negative? (did overall suicides fall after the change in laws?)

A way of determining this would be to estimate a stacked model:

$$\begin{matrix} Firearm_t \\ \\ Hanging_t \end{matrix} = \alpha_1 + \beta_1 X_t + \rho_1 guns_t + \alpha_2 H_t + \beta_2 (H_t X_t) + \rho_2 (H_t guns_t) + \mu_t \qquad (3)$$

Suppose, however, that we have no information on X available to us. Can we answer these questions? There are two possibilities. First, if $\beta_1 = \beta_2$, then there is a potentially simple solution: subtract Hangings from Firearm deaths, so that we have:

$$\begin{aligned} Firearm_t - Hanging_t &= (\alpha_1 - \alpha_2) + (\beta_1 - \beta_2) X_t + (\rho_1 - \rho_2) guns_t + (\varepsilon_t + \upsilon_t) \\ &= (\alpha_1 - \alpha_2) + (\rho_1 - \rho_2) guns_t + (\varepsilon_t + \upsilon_t) \qquad (4) \end{aligned}$$

In this case, however, we will only be able to test the third of the propositions – we will be able to say nothing about the impact on firearm related suicides, or on method substitution – and that only implicitly by assuming that $\rho_1 <0$ and that $\rho_2 >=0$. The assumption that firearm and hanging deaths respond in the same way to the same stimuli is also unsatisfactory. For example, if unemployment rates affect suicides in low-income families more than high-income families, but firearms tend to be owned by high-income families, then $\beta_1 \neq \beta_2$.

What if we simply omit $X_t$ and perform the stacked regression analysis described in (3)? Then we have:

$$\begin{matrix} Firearm_t \\ \\ Hanging_t \end{matrix} = \alpha_1 + \rho_1 guns_t + \alpha_2 H_t + \rho_2 (H_t guns_t) + \zeta_t \qquad (5)$$

where $\zeta_t = \beta_1 X_t + \beta_2 (H_t X_t) + \mu_t$. Estimates of $\rho_1$ and $\rho_2$ will be biased and inconsistent if there is any correlation between $X_t$ and $guns_t$. Given the particular nature of this model, where there is only time series variation, and where $guns_t$ is a dummy equal to one after 1997, this is inevitable if there is any time-series variation in $X_t$. This is roughly similar to the problem of using a difference-in-difference estimator with time series variation and serial correlation described in Bertrand, Duflo & Mullanaithan (2004), except that in this case the lack of a reasonable control group exacerbates the difficulties. Suppose, for instance, that a drought tends to lead to higher firearm suicides, but not higher suicides by hanging ($X_t$ is drought, $\beta_1 > 0$, $\beta_2 = 0$). Then the drought of the mid-2000s should have led to more firearm suicides, but not hangings. Failing to account for this in estimating (5) would see all those excess suicides attributed to the gun laws of 1997: that is, the estimate of $\rho_1$ would be biased in a positive direction.

In short, time series variation in and of itself is unable to recover plausible estimates of the effect of the Australian gun laws on deaths, except in the presence of extremely restrictive (and probably inaccurate) assumptions on the determinants of the numbers of deaths. Without having some credible control group – or at a minimum a model that introduces time-varying factors that may affect homicide and suicide rates, such as personal income growth, unemployment rates, or health and social program changes –

we can draw no conclusions on either the extent of method substitution, or on the underlying direction of overall homicide and suicide rates.

## Appendix B. Details of Regression Results

**Table B1. Estimates from ARIMA(1,1,1) models of firearm and non-firearm suicides and homicides**

| | Firearm suicide | | | Firearm homicide | | | Non-firearm suicide | | | Non-firearm homicide | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1979-96 | 1969-96 | 1915-96 | 1979-96 | 1969-96 | 1915-96 | 1979-96 | 1969-96 | 1915-96 | 1979-96 | 1969-96 | 1915-96 |
| **Level** | | | | | | | | | | | | |
| I(1) | **-0.087** | -0.024 | -0.006 | -0.020 | -0.004 | -0.001 | **0.196** | 0.058 | 0.023 | 0.009 | 0.015 | 0.000 |
| | (0.0154) | (0.0363) | (0.0063) | (0.0044) | (0.0086) | (0.0043) | (0.0445) | (0.1278) | (0.0193) | (0.0175) | (0.0138) | (0.0064) |
| AR | **0.562** | 0.156 | **0.771** | 0.001 | -0.149 | 0.138 | **-0.444** | **-1.000** | **0.847** | -0.551 | -0.340 | -0.183 |
| | (0.2338) | (0.5928) | (0.0827) | (0.2703) | (0.2782) | (0.1813) | (0.2345) | (0.0050) | (0.0662) | (0.3001) | (0.3611) | (0.1580) |
| MA | -1.000 | -0.319 | -1.000 | -1.000 | -0.554 | -0.736 | -1.000 | 0.990 | -1.000 | -0.192 | -0.375 | -0.548 |
| | (0.1985) | (0.5388) | (0.0404) | (0.1832) | (0.1966) | (0.1296) | (0.1634) | (0.0474) | (0.0338) | (0.3163) | (0.4111) | (0.1294) |
| AIC | -0.08 | 5.5 | 62.87 | -21.09 | -35.07 | -102.01 | 41.67 | 63.8 | 196.68 | -11.7 | -18.01 | -70.61 |
| *Observed mean* | | 1.230 | | | 0.269 | | | 11.312 | | | 1.327 | |
| *Predicted mean* | 1.854 | 2.020 | 2.465 | 0.303 | 0.400 | 0.450 | 11.819 | 11.253 | 10.409 | 1.393 | 1.393 | 1.326 |
| *P (Predicted $\leq$ Observed)* | 0.0000 | 0.0000 | 0.0000 | 0.0248 | 0.0011 | 0.0002 | 0.2146 | 0.5464 | 0.9768 | 0.2092 | 0.0692 | 0.5060 |
| | | | | | | | | | 0.0232 | | | |
| **Log** | | | | | | | | | | | | |
| I(1) | **-0.031** | -0.010 | -0.002 | **-0.038** | -0.008 | -0.001 | **0.021** | 0.006 | -0.001 | 0.007 | **0.012** | 0.000 |
| | (0.0064) | (0.0128) | (0.0025) | (0.0115) | (0.0174) | (0.0090) | (0.0051) | (0.0135) | (0.0090) | (0.0129) | (0.0032) | (0.0064) |
| AR | **0.668** | 0.158 | **0.799** | 0.213 | 0.119 | 0.227 | **0.487** | **-1.000** | 0.227 | -0.548 | 0.146 | -0.217 |
| | (0.2321) | (0.6782) | (0.0778) | (0.2777) | (0.3536) | (0.1790) | (0.2315) | (0.0030) | (0.1790) | (0.3081) | (0.2475) | (0.1677) |
| MA | -1.000 | -0.285 | -1.000 | -1.000 | -0.662 | -0.782 | -1.000 | 0.991 | -0.782 | -0.197 | -1.000 | -0.513 |
| | (0.1900) | (0.6297) | (0.0379) | (0.1712) | (0.2401) | (0.1218) | (0.1623) | (0.0383) | (0.1218) | (0.3281) | (0.1552) | (0.1499) |
| AIC | -36.46 | -53.69 | -105.53 | 3.89 | 2.23 | 28.34 | -34.55 | -57.34 | 28.34 | -21.85 | -28.25 | -77.62 |
| *Observed mean* | | 1.230 | | | 0.269 | | | 11.312 | | | 1.327 | |
| *Predicted mean** | 1.941 | 2.031 | 2.396 | 0.338 | 0.405 | 0.446 | 12.125 | 11.296 | 10.396 | 1.351 | 1.556 | 1.315 |
| *P (Predicted $\leq$ Observed)* | 0.0000 | 0.0000 | 0.0000 | 0.003 | 0.0005 | 0.0002 | 0.1294 | 0.5119 | 0.9785 | 0.2590 | 0.0007 | 0.6212 |

* For ease of interpretation, predictions for the log models were converted to levels for calculating the means and conducting the t-tests.
Coefficients in bold are significant at the 5% level.
Note: the estimates show the unsuitability of using the ARIMA(1,1,1) model for this data analysis. We nonetheless present only ARIMA(1,1,1) estimates in this paper to highlight the fragility of BM's results. Simple OLS estimates using the log of the death rate as a dependent variable model and allowing for linear time trends and a structural break in either levels or time trend in 1997 typically find a negative shift in firearm suicides after 1997, with firearm homicides shifting negatively, but not significantly unless a longer time series is used. That type of analysis is very similar to Chapman *et al.* (2006), which we discuss briefly. Because the purpose of this paper is not to provide our own estimates, and because a similar analysis is already in publication, we do not provide the full results here. Results are available on request.

**Table B2. Augmented Dickey Fuller test results**

| | | Level | | Log | |
|---|---|---|---|---|---|
| | | ADF statistic | p-value | ADF statistic | p-value |
| **1979-2004** | | | | | |
| No trend: | | | | | |
| | Firearm suicides | 0.44 | 0.9817 | 1.16 | 0.9933 |
| | Firearm homicides | -1.30 | 0.6296 | -0.47 | 0.8990 |
| | Non-firearm suicides | -1.62 | 0.4712 | -1.60 | 0.4845 |
| | Non-firearm homicides | **-3.76** | **0.0033** | **-3.75** | **0.0035** |
| With trend: | | | | | |
| | Firearm suicides | -3.05 | 0.1200 | -1.79 | 0.7106 |
| | Firearm homicides | **-5.44** | **0.0000** | **-3.66** | **0.0253** |
| | Non-firearm suicides | -1.43 | 0.8510 | -1.35 | 0.8734 |
| | Non-firearm homicides | **-3.69** | **0.0230** | **-3.70** | **0.0227** |
| **1915-2004** | | | | | |
| No trend: | | | | | |
| | Firearm suicides | -1.98 | 0.2963 | -0.45 | 0.9037 |
| | Firearm homicides | **-4.54** | **0.0002** | **-3.92** | **0.0019** |
| | Non-firearm suicides | -2.38 | 0.1471 | -2.28 | 0.1778 |
| | Non-firearm homicides | **-3.86** | **0.0024** | **-3.92** | **0.0019** |
| With trend: | | | | | |
| | Firearm suicides | -2.29 | 0.4389 | -0.94 | 0.9516 |
| | Firearm homicides | **-4.53** | **0.0014** | **-3.94** | **0.0109** |
| | Non-firearm suicides | -2.97 | 0.1398 | -2.88 | 0.1696 |
| | Non-firearm homicides | **-4.33** | **0.0028** | **-4.37** | **0.0025** |

Note: Coefficients in bold are significant at the 5% level. This indicates that there is substantial evidence against the null hypothesis that the series is integrated. That is, it indicates that an ARIMA(1,1,1) model is inappropriate for the particular series. Typically, the preferred ADF test statistic is one based on a model that allows for a time trend. Models use data to 2004, because the modelling strategy employed by BM requires stability in the underlying structural model for the entire period 1979 to 2004.

## Acknowledgements

We thank Jenny Mouzos for providing the data on firearm and non-firearm death rates used in the empirical analysis, and Don Weatherburn and Justin Wolfers for helpful comments on a draft of this paper.

---

[1] We could not obtain the full set of data underlying Figures 1 and 2 in BM from the corresponding author. Our results are therefore based upon a dataset which is made up of the numbers of deaths provided in Table 1 of BM, supplemented with data kindly provided by Jenny Mouzos of the Australian Institute of Criminology. We also use the most recent population data from the Australian Bureau of Statistics, which differ slightly from those in BM. The dataset from 1979 to 2004 is therefore precisely that used by BM except for the population statistics. Results vary little if we instead use the population statistics in BM.

[2] Estimates shown in this paper are those from R, which came closer to replicating the results in BM than did the estimates from STATA. All data used in this paper and the programs used to obtain the statistical results (in STATA and R), as well as details of those results in numerical and graphical form, are available at http://www.wlu.ca/sbe/cneill.

[3] As in BM, our re-analysis of the statistics on non-firearm death rates suggests there was little change in these death rates post-NFA. Using a log rather than a level specification increases predicted non-firearm suicide rates slightly, but as in BM, the t-tests still reveal no statistically significant reduction in suicide rates post-1996. Using the full time series, on the other hand, suggests that suicide rates were statistically significantly higher post-1996, largely due to the jump in non-firearm suicides in 1997 and 1998. Given that the non-firearm suicide rate increased by around 2.2 per 100,000 between 1996 and 1998, while the firearm suicide rate in 1996 was 2.1 per 100,000, this is clearly not attributable to method substitution alone. We do not discuss our findings on non-firearm death rates in detail in this paper, because they have little bearing on our key arguments. All results are available on request.

[4] Our replications of their work suggest this makes it more difficult to find drop in firearms deaths after 1997. Results available on request.

[5] In statements to the Australian media following the publication of the paper, the authors were even less circumspect. In a summary of the research released at the Sporting Shooters' Association of Australia website, the conclusions are:

- 'The reforms did not affect rates of firearm homicide in Australia.
- The reforms could not be shown to alter rates of firearm suicide, because rates of suicide using other methods also began to decline in the late 1990s.
- …
- It must be concluded that the gun buyback and restrictive legislative changes had no influence on firearm homicide in Australia.
- The lack of effect of a massive buyback and associated legislative changes in the requirements for obtaining a firearm licence or legally possessing a firearm has significant implications for public and justice policy, not only for Australia, but internationally.' http://www.ssaacms.com/pdfs/pdf453c38e7ab1e0.pdf

Referring to an interview with Jeanine Baker, a newspaper article says 'The findings were clear, she said: The policy has made no difference. There was a trend of declining deaths that has continued.' (Sydney Morning Herald http://www.smh.com.au/news/national/buyback-has-no-effect-on-murder-rate/2006/10/23/1161455665717.html).

A press release by the International Coalition for Women in Shooting and Hunting (WiSH), which listed Samara McPhedran as the contact, described the findings of the research thus: 'Published in the influential British Journal of Criminology, the research showed that the legislative reforms, bans, and buybacks of 1996 did not impact on already declining rates of firearm homicide in Australia. It could not be concluded that the tough laws had any impact on firearm suicide rates, because suicides using other methods also started to decline in the late 1990s.' http://www.ic-wish.org/WiSH%20Gun%20Study%20Response%20Overwhelmingly%20Positive%20PR%20241006.pdf
All documents available at 5 February 2007.