Optimal Data Analysis
Vol. 7 (March 25, 2018), 28-35

2155-0182/10/$3.00

The Australian Gun Buyback Program and Rate of Suicide by Firearm

Ariel Linden, Dr.P.H., and Paul R. Yarnold, Ph.D.
Linden Consulting Group, LLC Optimal Data Analysis LLC

In 1997, Australia implemented a gun buyback program that reduced the stock of firearms by around one-fifth, and nearly halved the number of gun-owning households. Leigh and Neill[1] evaluated if the reduction in firearms availability affected homicide and suicide rates, and reported that the buyback led to a drop in the firearm suicide rates of almost 80%, with no significant effect on non-firearm death rates. In this paper we re-evaluate the suicide rate data to assess whether any directionally-correct structural breaks in the time series could be identified prior to the buy-back program. Such a change in the time series prior to the intervention may confuse causal interpretation of the actual intervention.

When evaluating effects of an intervention or policy change on an outcome serially measured as a time series, treatment effects may seem less plausible if a parallel trend already exists in the time series prior to the actual intervention. Thus, sensitivity analyses should always be conducted to detect structural breaks in the time series that occur prior to the intervention.[2]

This paper performs a structural breaks analysis on annual suicide rates by firearms in Australia in the pre-intervention period (1968-1996), using two unrelated analytic methods to examine the data. The first analytic approach uses interrupted time-series analysis (ITSA), in which a "pseudo-intervention" effect is estimated in the data prior to the actual intervention.[3] A "pseudo-treatment effect" would be indicated by a statistically significant change in level or slope (trend) from the pre-intervention period to the "pseudo-intervention" period.

The second analytic approach uses the machine learning algorithm known as optimal discriminant (data) analysis, or ODA, to determine if (and to what degree) structural breaks can be identified in periods prior to the actual initiation of the intervention. Strengths of this methodology are that ODA provides intuitive measures of predictive accuracy, uses distribution-free permutation tests to derive *P* values, and performs cross-validation to assess generalizability of the model applied to new cases in similar settings. Therefore ODA likely reflects an approach of interest to investigators who are currently using ITSA designs, and to those more generally interested in applications of machine learning to traditional research designs.[4-18]

2155-0182/10/$3.00

Methods

Data

Data we use were originally retrieved by Leigh and Neill[1] from the Australian Bureau of Statistics, Cause of Death collection (http:// www.abs.gov.au/Causes-of-Death). We specifically focus on suicides per 100,000 population by firearms in Australia in 1968-2006, providing 29 annual observations in the pre-intervention period to assess structural breaks.

Analytic Approaches

We present brief overviews of the ITSA and ODA Structural Break analyses used herein.

Interrupted Time-Series Analysis

A single-group ITSA was estimated using the ITSA package written for Stata.[3] Two intervention periods were specified in the model—the actual intervention which occurred in 1997, and a pseudo-intervention in 1987. The year 1987 was chosen as the pseudo-intervention period based on visual inspection of the time series, which appeared to change direction (from an increasing to decreasing trend in the suicide rate) in that year. Newey-West standard errors[19] were specified to account for autocorrelation at lag 3 (see Linden[3] for a discussion of how autocorrelation is assessed for ITSA models).

ODA Structural Break Analysis

In order to systematically assess the presence or absence of structural breaks in the years prior to the intervention in 1996, a series of 28 "pseudo-interventions" was generated—one for each year commencing in 1968 and ending in 1995. For example, in the pseudo-intervention year 1968, the intervention is set to 1 for all years from 1969 onward, and 1968 represents the sole pre-intervention period and is set to 0. At the other end of the continuum all years from 1968 to 1994 represent pre-pseudo-intervention periods and are set to 0, whereas 1995 is the sole pseudo-intervention year and is set to 1.

In ODA every pseudo-intervention is treated as a two-category (pre- *vs*. post-pseudo-intervention period) class variable. The relationship between pseudo-intervention and suicide rate was ascertained using an ODA model of the form: if annual rate of suicide by firearm $\leq$ (cutpoint) then predict that the observation is from the post–pseudo-intervention period; otherwise predict the observation is from the pre-pseudo-intervention period (a non-directional two-tailed hypothesis was tested in actual analysis).[4-8]

Alpha inflation associated with multiple comparisons (28 tests of statistical significance were conducted in training analysis, and 27 in LOO analysis) was controlled by a sequentially-rejective Sidak Bonferroni-type multiple comparisons procedure to ensure an experimentwise $P < 0.05$. Permutation P values were estimated using 25,000 Monte Carlo experiments.[5,6]

The upper-bound of the expected cross-generalizability of ODA findings across time was assessed by "leave-one-out" (LOO) single-sample jackknife analysis.[5,6,20,21] Obtaining identical ESS in training and validity analysis suggests the training model may cross-generalize to the following year with comparable reliability and strength. But, if ESS is lower in LOO than in training analysis, this suggests the cutpoint that maximizes predictive accuracy in training analysis may not cross-generalize to the following year with comparable reliability and strength. LOO requires at least two observations in both the pseudo-pre and post-intervention periods, and thus LOO findings not reported for pseudo-interventions with one observation per class category (1968 and 1995).

This analysis performed for data prior to 1996 (the year of the actual intervention) conducts a systematic search for intervention-like effects (e.g., decreased rate of suicide by firearm) occurring *prior to the actual intervention*.

A second ODA Structural Break analysis was conducted for the entire time series (1968 to

2155-0182/10/$3.00

2005), involving 38 pseudo-interventions (and 75 tests of statistical hypotheses), to systematically identify any/all intervention-like effects which occurred *within the time series*.

Results

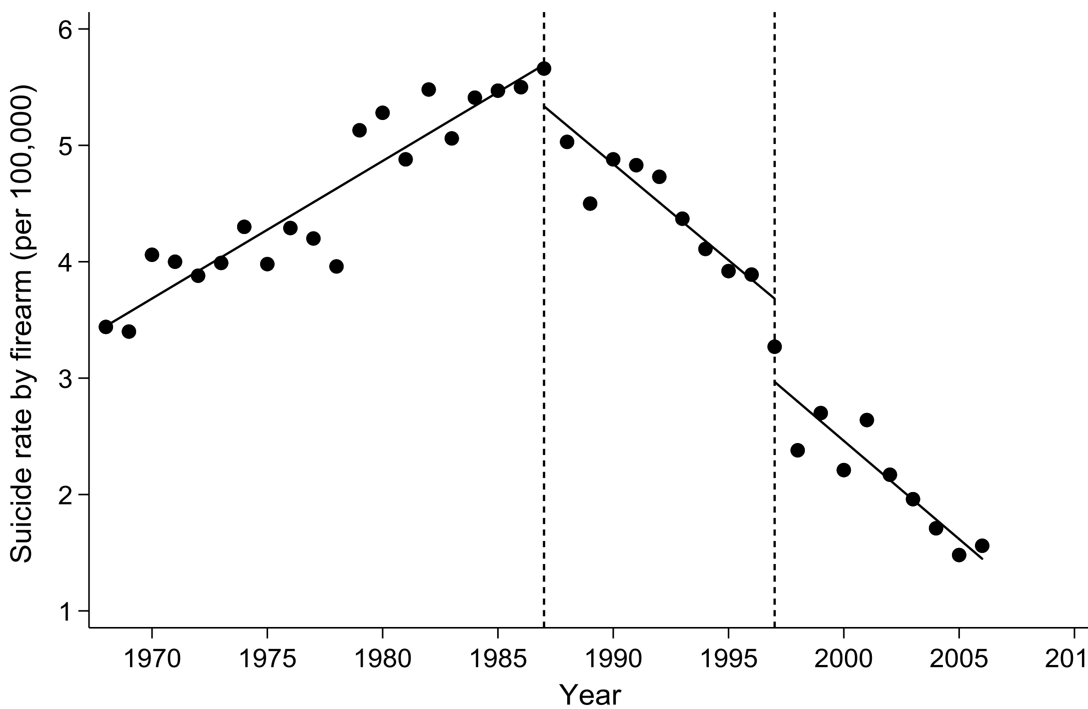
Findings obtained using ITSA are presented first, followed by findings obtained using ODA Structural Break analysis.

Interrupted Time-Series Analysis

As seen in Figure 1, ITSA concurs (as shown by the linear trend lines) with the visual assessment that a structural break (pseudo-treatment effect) did occur in 1987. The difference in trends between the pre-intervention and pseudo-intervention periods was -.284 suicides per 100,000 population [95% CI: -.323, -.244] ($P < 0.001$).

One can also see from Figure 1 that there was a level change in suicide rates in the year immediately following introduction of the buy-back plan (1997), but the trend thereafter is no different than the trend in the pseudo-intervention period. This is verified statistically, where the difference in trends between the pseudo-intervention period (1987-1996) and the actual intervention period (>1997) is -.004 suicides per 100,000 per year [95% CI: -.050, .043] ($P < 0.879$).

Figure 1: ITSA of Annual Rate of Suicide by Firearm



ODA Structural Break Analysis

For analysis of data occurring *prior to the intervention in 1996*, Table 1 presents the annual suicide by firearm rate; the ODA-identified cutpoint and direction (direction indicates if rates are larger or smaller over time) for comparing suicide rate between the pre- and post-pseudo-intervention periods; and statistical reliability (*P* value) and accuracy (ESS) in both training and LOO analysis.

2155-0182/10/$3.00

Table 1: ODA Structural Break Analysis for Pre-Intervention Time-Series

Year	Rate	Predict Intervention If rate is	Training *P* value	Training ESS	LOO *P* value	LOO ESS
1968	3.44	> 3.660	0.1411	96.43	---	---
1969	3.40	> 3.660	0.0051	100.00	0.0025	100.00
1970	4.06	> 4.085	0.0564	73.08	0.2667	35.90
1971	4.00	> 4.085	0.0133	76.00	0.1049	47.00
1972	3.88	> 4.085	0.0038	79.17	0.0358	55.00
1973	3.99	> 4.085	0.0007	82.61	0.0106	61.59
1974	4.30	> 4.335	0.0063	68.18	0.0190	53.90
1975	3.98	> 4.085	0.0014	73.21	0.0326	46.43
1976	4.29	> 4.335	0.0004	75.00	0.0005	70.00
1977	4.20	> 4.335	0.0002	78.95	0.0002	73.68
1978	3.96	> 4.335	0.0001	83.33	0.0001	77.78
1979	5.13	> 4.335	0.0006	74.02	0.0004	68.14
1980	5.28	> 4.335	0.0019	65.87	0.0020	59.62
1981	4.88	> 4.335	0.0086	58.57	0.0070	51.90
1982	5.48	> 4.335	0.0237	51.90	0.0199	44.76
1983	5.06	> 4.335	0.0692	45.67	0.0476	37.98
1984	5.41	> 4.335	0.1548	39.71	0.0987	31.37
1985	5.47	> 4.335	0.3292	33.84	0.1815	24.75
1986	5.50	> 4.335	0.5858	27.89	0.3000	17.89
1987	5.66	≤ 5.045	0.0207	40.00	0.1919	23.89
1988	5.03	≤ 4.855	0.3770	35.12	1.0000	-25.00
1989	4.50	≤ 4.855	0.5596	31.17	0.4532	12.34
1990	4.88	≤ 4.855	0.1763	47.83	0.1824	31.16
1991	4.83	≤ 4.780	0.1950	50.00	0.2361	30.00
1992	4.73	≤ 4.435	0.1540	56.00	0.3257	27.00
1993	4.37	≤ 4.155	0.0886	69.23	0.3161	32.05
1994	4.11	≤ 3.940	0.0502	88.89	0.0247	88.89
1995	3.92	≤ 3.905	0.2740	89.29	---	---

There was only one perfect structural break: 1968 and 1969 had a *lower* firearm suicide rate than all subsequent years in training and LOO analysis. However, due to the small number (two) of pre-pseudo-intervention years this effect was only statistically significant at the per-comparison criterion.[6]

Findings consist of two parts: 1968 to 1986—when the ODA model direction term indicated suicide rate was *greater* post-pseudo-intervention *vs*. pre-pseudo-intervention (suicide rate was increasing over time); and 1987 to 1995—when the direction term indicated suicide rate was *lower* post-pseudo-intervention (suicide rate was decreasing over time).

In the initial (early) part of the series, when firearm suicide rates were increasing, the training models for 1973 and 1976-1979 met the experimentwise criterion for statistical significance, as did LOO models for 1976-1979. The ESS values of these models varied between relatively strong (>50) and strong (>75), but within a given year the ESS was always lower in LOO *vs*. training analysis. More broadly, training and

2155-0182/10/$3.00

LOO effects were statistically significant at the per-comparison criterion in 1972-1982. In contrast, there were no significant effects in 1983-1986, indicating that there was no statistically reliable direction of firearm suicide rate over time during this transitional period.

In the second (later) part of the series, 1987-1995, all of the ODA models switched to the ≤ direction—indicating that firearm suicide rates were decreasing over time. During this period, only the moderately strong 1987 training model met the per-comparison criterion for statistical significance, but this was not confirmed in LOO analysis. The strong 1994 training model was statistically marginal, and was the only model in Table 2 with stable ESS in LOO analysis—which met the per-comparison criterion for statistical significance.

In sum, ODA Structural Break analysis of temporal firearm suicide rate *prior to the Gun Buyback intervention* indicates firearm suicide rate increased in 1968-1982; was directionless in 1983-1986; and decreased in 1987-1994. The only perfect structural break occurred in 1969, indicating firearm suicide rate was *increasing*.

For data analysis *over the entire time series* (1968-2005), Table 2 presents the annual firearm suicide rate. Two perfect structural breaks were identified: 1997 (one year after the Gun Buyback intervention began) and 2003 (the first time in the series that firearm suicide rate fell beneath 2.00 suicide deaths by firearms per 100,000) had a *lower* firearm suicide rate than all prior years in training and LOO analysis.

Paralleling the preceding analysis, the findings consist of two parts: 1968 to 1979—when the ODA model direction term indicated suicide rate was *greater* post-pseudo-intervention (suicide rate was increasing over time); and 1980 to 2005—when the direction term indicated suicide rate was *lower* post-pseudo-intervention (suicide rate was decreasing over time).

In the initial (early) part of the series when firearm suicide rates were increasing, no models met the experimentwise criterion for statistical significance. Four training models were statistically significant if evaluated by the per-comparison criterion (1973, 1976-1978), but ESS achieved in LOO analysis only met this criterion in 1977 and 1978. ESS for these effects indicated moderate to relatively strong effects.

In the second (later) part of the series, 1980-2005, all of the ODA models switched to the ≤ direction—indicating that firearm suicide rates were decreasing over time. The ESS values of these models varied between moderate and perfect, but except for three exceptions (1997, 1998, and 2003), within a given year the ESS was always lower in LOO *vs*. training analysis. More broadly, effects were statistically significant at the per-comparison criterion in 1983-2004 for training analysis, and in 1981-2004 for LOO analysis. And, effects were statistically significant at the experimentwise criterion in 1990-2003 for training analysis, and in 1987 and 1992-2003 for LOO analysis.

In sum, ODA Structural Break analysis of temporal firearm suicide rate *over the entire time series* indicates firearm suicide rate increased in 1968-1978; was directionless in 1979 and 1980; and decreased in 1981-2004. Two perfect structural breaks occurred—in 1997 and 2003—both indicating that the firearm suicide rate was *decreasing*.

Discussion

Both analytic approaches identified a structural break in the time series prior to the actual initiation of the gun buy-back plan. Given that Leigh and Neill[1] reported that the buyback led to a drop in the firearm suicide rates of almost 80%, one may question whether this is a true effect, or an artifact of the directionally-correct decrease in firearm suicide rates that began in 1987. Given that these are observational data, the true result cannot be known. However, this study highlights the importance of routinely performing structural break analyses when using the single-group ITSA framework as a way to test the sensitivity of treatment effect estimates.

2155-0182/10/$3.00

Table 2: ODA Structural Break Analysis for Entire Time-Series

Year	Rate	Predict Intervention If rate is	Training *P* value	Training ESS	LOO *P* value	LOO ESS
1968	3.44	>3.660	0.6135	71.05	---	---
1969	3.40	>3.660	0.1802	72.97	1.0000	0
1970	4.06	>4.085	0.3357	52.78	1.0000	-33.33
1971	4.00	>4.085	0.1656	54.29	0.3223	26.43
1972	3.88	>4.085	0.0941	55.88	0.1869	32.94
1973	3.99	>4.085	0.0442	57.58	0.1022	37.88
1974	4.30	>4.335	0.1111	46.88	0.1538	29.46
1975	3.98	>4.335	0.0778	48.39	0.0964	32.66
1976	4.29	>4.335	0.0414	50.00	1.0000	0
1977	4.20	>4.335	0.0248	51.72	0.0052	48.28
1978	3.96	>4.335	0.0118	53.57	0.0027	50.00
1979	5.13	>4.335	0.0641	43.52	1.0000	-8.33
1980	5.28	≤3.335	0.1397	38.46	0.1913	19.23
1981	4.88	≤3.335	0.0802	40.00	0.0097	36.00
1982	5.48	≤3.335	0.0579	41.67	0.0062	37.50
1983	5.06	≤3.335	0.0394	43.48	0.0039	39.13
1984	5.41	≤3.335	0.0239	45.45	0.0024	40.91
1985	5.47	≤3.335	0.0154	47.62	0.0014	42.86
1986	5.50	≤3.335	0.0067	50.00	0.0008	45.00
1987	5.66	≤3.335	0.0064	52.63	0.0005	47.37
1988	5.03	≤3.335	0.0027	55.56	0.0191	35.71
1989	4.50	≤3.335	0.0013	58.82	0.0108	39.30
1990	4.88	≤3.335	0.0005	62.50	0.0019	49.46
1991	4.83	≤3.940	0.0001	67.50	0.0008	54.17
1992	4.73	≤3.940	0.0001	73.71	0.0003	59.43
1993	4.37	≤3.940	0.0001	80.77	0.0001	65.38
1994	4.11	≤3.940	0.0001	88.89	0.0001	72.22
1995	3.92	≤3.335	0.0001	90.91	0.0001	71.10
1996	3.89	≤3.335	0.0001	100.00	0.0001	90.00
1997	3.27	≤2.985	0.0001	100.00	0.0001	100.00
1998	2.38	≤2.985	0.0001	96.77	0.0001	96.77
1999	2.70	≤2.670	0.0001	96.87	0.0001	79.46
2000	2.21	≤2.670	0.0001	93.94	0.0006	74.24
2001	2.64	≤2.190	0.0001	100.00	0.0003	77.06
2002	2.17	≤2.065	0.0001	100.00	0.0005	75.00
2003	1.96	≤1.835	0.0004	100.00	0.0002	100.00
2004	1.71	≤1.635	0.0030	100.00	0.0041	97.30
2005	1.48	≤1.635	0.1020	97.37	---	---

References

[1]Leigh A, Neill C (2010). Do Gun Buybacks Save Lives? Evidence from Panel Data? *American Law and Economics Review*, *12*, 509–557.

[2]Linden A, Yarnold PR (2016). Using machine learning to identify structural breaks in single-group interrupted time series designs. *Journal of Evaluation in Clinical Practice*, *6*, 855-859.

[3]Linden A (2015). Conducting interrupted time-series analysis for single- and multiple-group comparisons. *The Stata Journal*, *15*, 480–500.

[4]Yarnold PR, Soltysik RC (1991). Theoretical distributions of optima for univariate discrimination of random data. *Decision Sciences*, *22*, 739-752

[5]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: Guidebook with software for Windows*. Washington, D.C.: APA Books.

[6]Yarnold PR, Soltysik RC (2016). *Maximizing predictive accuracy*. Chicago, IL: ODA Books. DOI: 10.13140/RG.2.1.1368.3286

[7]Yarnold PR (2017). What is optimal data analysis? *Optimal Data Analysis*, *6*, 26-42.

[8]Linden A, Yarnold PR (2016). Using data mining techniques to characterize participation in observational studies. *Journal of Evaluation in Clinical Practice*, *6*, 839-847.

[9]Linden A, Yarnold PR (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*, *6*, 848-854.

[10]Linden A, Yarnold PR, Nallamothu BK (2016). Using machine learning to model dose-response relationships. *Journal of Evaluation in Clinical Practice*, *6*, 860-867.

[11]Linden A, Yarnold PR (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice*, *6*, 868-874.

[12]Linden A, Yarnold PR (2016). Combining machine learning and propensity score weighting to estimate causal effects in multivalued treatments. *Journal of Evaluation in Clinical Practice*, *6*, 875-885.

[13]Linden A, Yarnold PR. Using classification tree analysis to generate propensity score weights. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12744

[14]Linden A, Yarnold PR. Modeling time-to-event (survival) data using classification tree analysis. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12779

[15]Linden A, Yarnold PR. Minimizing imbalances on patient characteristics between treatment groups in randomized trials using classification tree analysis. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12792

[16]Linden A, Yarnold PR. Identifying causal mechanisms in health care interventions using classification tree analysis. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12848

[17]Linden A, Yarnold PR. Estimating causal effects for survival (time-to-event) outcomes by combining classification tree analysis and propensity score weighting. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12859

[18]Yarnold PR, Linden A (2017). Computing propensity score weights for CTA models involving perfectly predicted endpoints. *Optimal Data Analysis*, *6*, 43-46.

2155-0182/10/$3.00

[19]Newey WK, West KD (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703-708.

[20]Yarnold PR (2016). Using UniODA to determine the ESS of a CTA model in LOO analysis. *Optimal Data Analysis*, *5*, 3-10.

[21]Yarnold PR (2016). Determining jackknife ESS for a CTA model with chaotic instability. *Optimal Data Analysis*, *5*, 11-14.

Author Notes

This article uses publically-available data and is exempt from IRB review. No conflict of interest was reported.