Open Journal of Statistics, 2020, 10, 31-51
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Using HAC Estimators for Intervention Analysis

Ashok K. Singh[1], **Rohan J. Dalpatadu**[2]

[1]RGG Department, Harrah College of Hospitality, University of Nevada, Las Vegas, USA
[2]Department of Mathematical Sciences, University of Nevada, Las Vegas, USA
Email: ashok.singh@unlv.edu

How to cite this paper: Singh, A.K. and Dalpatadu, R.J. (2020) Using HAC Estimators for Intervention Analysis. *Open Journal of Statistics*, **10**, 31-51.
https://doi.org/10.4236/ojs.2020.101003

Received: December 12, 2019
Accepted: January 11, 2020
Published: January 14, 2020

Abstract

The purpose of this article is to present an alternative method for intervention analysis of time series data that is simpler to use than the traditional method of fitting an explanatory Autoregressive Integrated Moving Average (ARIMA) model. Time series regression analysis is commonly used to test the effect of an event on a time series. An econometric modeling method, which uses a heteroskedasticity and autocorrelation consistent (HAC) estimator of the covariance matrix instead of fitting an ARIMA model, is proposed as an alternative. The method of parametric bootstrap is used to compare the two approaches for intervention analysis. The results of this study suggest that the time series regression method and the HAC method give very similar results for intervention analysis, and hence the proposed HAC method should be used for intervention analysis, instead of the more complicated method of ARIMA modeling. The alternative method presented here is expected to be very helpful in gaming and hospitality research.

Keywords

Time Series, ARIMA, ARMA, Autocorrelation, Partial Autocorrelation, Ljung-Box Test, Bootstrap, Simulation

1. Introduction

An intervention model or interrupted time series model [1] is an Autoregressive Integrated Moving Average (ARIMA) model of a time series in which at least one of the predictors is a dummy variable for an event, which is thought of as an interruption in a pure ARIMA process. An ARIMA model in which differencing is not used is also called an ARMA model.

The use of intervention analysis or interrupted time series analysis is very common in hospitality and tourism literature. Bonham and Gangnes [2] used intervention analysis to show that the 5% Hawaii hotel room tax started in 1987 did not significantly impact Hawaii hotel room revenues. Min's [3] intervention analysis of inbound tourism data showed that both the earthquake of September 21, 1999 and the Severe Acute Respiratory Syndrome (SARS) of 2003 had significant negative impacts on Taiwan's inbound tourism. A thorough review of time series forecasting literature is provided by De Gooijer and Hyndman [4]. Ahlgren *et al.* [5] used an ARIMA model to assess the impact of a higher gaming tax rate in the state of Illinois on gaming volume, and concluded that the gaming volume experienced a significant decrease when the tax increase took effect. Ahlgren *et al.* [6] used an ARIMA model to assess the impact of a higher gaming tax rate in the state of Illinois on marketing expenditure by a major Illinois riverboat operator. Toma *et al.* [7] used a seasonal ARIMA model to show that the book "Midnight in the Garden of Good and Evil" set in Savannah, Georgia had a significant positive effect on hotel tax receipts, while both 9/11 and hurricane Floyd had a significant negative effect. Goh and Law [8] used intervention analysis to show that relaxation of issuing out-bound visitor visas, the Asian financial crisis, the handover, and the bird flu epidemic had significant and expected impacts on Hong Kong tourism demand. Eisendrath *et al.* [9] used intervention analysis on Las Vegas Strip gaming volume to show that 9/11 had a significant negative impact lasting five months. Suh *et al.* [10] used this approach to investigate the effects of cash revenue generated from non-comped diners (CASHREV), amount spent by the casino in comped-meals (COMPREV) on gaming volume, using major holidays and Motorcycle Rally as intervention predictors; their ARIMA model showed that both CASHREV and COMPREV were significant predictors of slot coin-in, and Motorcycle Rally had a significant and negative impact on slot coin-in. Zheng *et al.* [11] used ARIMA intervention analysis to study the impact of recession on restaurant stock performance. D'Amuri and Marcucci [12] used ARIMA modeling to assess the impact of an index of Google job-search intensity on the monthly US unemployment rate. Intervention analysis has been used in other disciplines as well. Su and Deng [13] used time series regression with intervention term to predict the yield of Yu Ebao. Huang [14] has used intervention analysis to show that government intervention improved a firm's investment efficiency. The purpose of the present article is to introduce a method from econometric modeling that is simpler to use than the ARIMA method for intervention analysis.

2. Problem Statement and Methodology

The general form of an autoregressive moving average (ARMA) model is

$$Y_t = \delta + \sum_{i=1}^{p} \phi_i Y_{t-i} + a_i - \sum_{i=1}^{q} \theta_i a_{t-i}$$

where

Y_t = the response variable of interest.

δ = the intercept.

ϕ_i = the autoregressive (AR) term coefficients.

θ_i = the moving average (MA) term coefficients.

a_i = the random shocks.

The above model is referred to as ARIMA($p,d = 0,q$) model or ARMA(p,q) model [15].

Following steps are used in fitting an intervention model to a time series Y_t of the response variable as a function of predictor(s) X_t and intervention variable(s) I_t:

(1) The time series Y_t is plotted to assess the presence of trend with time. A polynomial function of time t is typically used to model the trend.

(2) A multiple linear regression (MLR) model is fitted to the data, such that the variance inflation factor (VIF) values of all predictors are not too high; values of VIF above 5 suggest the presence of multicollinearity [16], and values of VIF above 10 indicate that the MLR model suffers from serious multicollinearities [17]. One typically drops predictors with highest VIF value, one by one, in order to get a reasonable MLR model.

(3) The MLR model assumes that the errors are independent and normally distributed with 0 means and a common error variance, and the standard t-test is used to conduct significance tests for the coefficients of the predictors. In a time series data, however, the residuals are typically auto-correlated, and hence the use of the t-test is not valid. Plots of the auto-correlation function (ACF) and partial auto-correlation function (PACF) function are examined to determine the order of the ARIMA(p,d,q) × (P, D, Q) model; here p is the order of the non-seasonal auto-regressive term, d the non-seasonal differencing used, q the order of the non-seasonal moving average term, and P, D, Q are the corresponding seasonal terms.

(4) Once the ARIMA model has been identified, a time series regression model is fitted to the data with all predictors and the ARIMA terms in the model, and the residuals from this model are tested for zero auto-correlations up to h lags; the Ljung-Box test [18] is commonly used for this purpose. Hyndman and Athanasopoulos [19] recommend using h = 10 lags in the Ljung-Box test After a time-series regression model with uncorrelated residuals is found, the significance of all predictors is tested.

The last or 4th step at times proves to be quite challenging, and an alternative approach for testing the significance of all predictors and intervention variables is proposed and investigated in this study.

One of the assumptions of multiple linear regression (MLR) is that the variance of the response variable is same across the range of predictor values; when this is not the case, we say that heteroscedasticity (or heteroskedasticity) is present [20]. When the error terms (residuals) from an MLR model are autocorrelated (*i.e.*, not independent) the standard estimate of the correlation matrix needs to be corrected for both heteroscedasticity and the presence of autocorrelation. These estimators are referred to as HAC-Consistent estimators [21].

This approach uses three different heteroskedasticity and autocorrelation consistent (HAC) estimators of the covariance matrix, used in econometric modeling [22], namely HAC, Kern-HAC, and Newey-West in place of Steps (3) and (4) above. Examples from hospitality literature as well as a synthetic time series data are used to demonstrate the effectiveness of this alternative approach, and parametric bootstrap of time series will be used to compare the results from standard ARIMA-based intervention model and the results of significance testing using HAC estimators. In Step (2) above, one typically keeps only the significant predictors in the MLR model; in this paper, to keep things simple, all predictors in the model are kept as long as the VIF values are less than 5.

3. Comparison of P-Values from ARIMA Model and HAC Estimators

Three different time series data sets are used to compare the ARIMA method for intervention analysis and significance tests using HAC estimators of the covariance matrix of the estimated coefficients of the MLR. The time series in the first two examples are real data sets from gaming literature, the first one modeled by an ARIMA(3,0,0) process with a cubic trend, and the second by an ARIMA(0,0,2) process with a cubic trend. For the third example, a synthetic monthly time series of length 84 was used.

Following steps are used for this comparison:

(a) The time series Y_t is plotted to assess the presence of a trend.

(b) An MLR model is fitted to Y_t as a function of the predictors including a dummy column for the intervention term, and a polynomial trend function; predictors with high VIF values are removed.

(c) ACF and PACF of the residuals from MLR are plotted for identification of ARIMA terms p and q.

4An ARIMA(p,0,q) time series regression model is fitted, with p and q determined in Step (c) above. The P-values for each predictor term in the ARIMA model are calculated.

(d) ACF of the residuals from the ARIMA model of Step (d) is plotted to show that these residuals are not auto-correlated, which is followed by running the L-B test for 10 lags. If the P-values from the L-B test exceed 0.05 at each lag, these residuals are deemed not auto-correlated, which validates the correctness of the P-values computed in Step (d).

(e) P-values for each term in the MLR model of Step (b) are computed using the HAC estimators of the covariance matrix of the estimated coefficients.

(f) P-values computed in Steps (d) and (f) are compared.

The package Sandwich of the statistical software environment R [23] was used to compute the P-Values using the HAC estimators.

In addition, for each time series data set, 1000 bootstrap samples are generated, and Steps (a)-(f) are repeated for each bootstrap sample. Histograms of the 1000 P-values computed in Steps (d) and (f) are plotted to compare the ARIMA method to the HAC-method of intervention analysis.

4. Examples from Tourism and Hospitality Literature

Example 1: Impact of 9/11 on Las Vegas Strip Gaming Revenue

Eisendrath *et al.* [9] used Las Vegas Strip slot coin-in data, complied from the Nevada Gaming Control Reports for the period January 1990 through November 2004. In this study, data for the period July 1997 to November 2005 are used. The potential predictors of Las Vegas Strip coin-in are dummy variables for months February, March, …, December, a cubic trend, and a dummy column D9_11 for 9/11 that was set to 1 for 6 months starting from September 2011, and 0 for all other months. A time-series plot of Las Vegas Strip coin-in for the sampling period, with all six points labeled as 1 corresponding to the interrupted time-series (Figure 3, top left plot) shows two bends, suggesting the presence of a cubic term in the time series. In order to keep VIF values small, the variable t (indicating month number) was standardized first, and then squared and cubed: (Table 1)

$$Zt = \left(t - mean(t)\right) / sd(t),\ t = 1, 2, \cdots, 101$$

$$Zt2 = \left(Zt\right)^2$$

$$Zt3 = \left(Zt\right)^3$$

Figure 1 is a plot of the ACF and PACF values for lags 1 to 20; this graph suggests using ARIMA(3,0,0) model for the residuals [22].

The final Time Series Regression model is shown in Table 2; it can be seen from Table 2 that

Table 1. MLR model fitted to Las Vegas Strip Slot Coin-in Data.

Term	Coeff	SE	t-stat	P-Value
(Intercept)	178,074	4253.1	41.87	0.00
Zt	5321	2823.7	1.88	0.06
Zt2	3371	1334.7	2.53	0.01
Zt3	12,251	1455	8.42	0.00
DFeb	−15,827	5626.7	−2.81	0.01
DMar	10,343	5663.9	1.83	0.07
DApr	−7969	5665.2	−1.41	0.16
DMay	949	5666.9	0.17	0.87
DJun	−10,794	5669.2	−1.90	0.06
DJul	−521	5504.4	−0.10	0.92
DAug	−3477	5503.7	−0.63	0.53
DSep	−1103	5477.2	−0.20	0.84
DOct	2524	5479.8	0.46	0.65
DNov	−6164	5483.9	−1.12	0.26
DDec	−24,787	5626.7	−4.41	0.00
D9_11	−11,474	5126	−2.24	0.03

Table 2. ARIMA(3,0,0) regression model fitted to Las Vegas Strip Slot Coin-in Data.

Term	Coeff	SE	t-stat	P-Value
ar1	−0.16	0.10	−1.64	0.10
ar2	0.20	0.10	2.14	0.03
ar3	0.35	0.10	3.55	0.00
intercept	178,540	4103.90	43.50	0.00
Zt	6052	3663.60	1.65	0.10
Zt2	3763	1663.70	2.26	0.02
Zt3	11,731	1772.80	6.62	0.00
DFeb	−15,770	5431.90	−2.90	0.00
DMar	9228	4707.30	1.96	0.05
DApr	−7972	4362.70	−1.83	0.07
DMay	793	5191.00	0.15	0.88
DJun	−11,732	4862.10	−2.41	0.02
DJul	−1369	4739.50	−0.29	0.77
DAug	−4392	4763.40	−0.92	0.36
DSep	−1814	5019.50	−0.36	0.72
DOct	1863	4253.80	0.44	0.66
DNov	−6755	4559.50	−1.48	0.14
DDec	−26,473	5427.20	−4.88	0.00
D9_11	−12,800	4694.60	−2.73	0.01

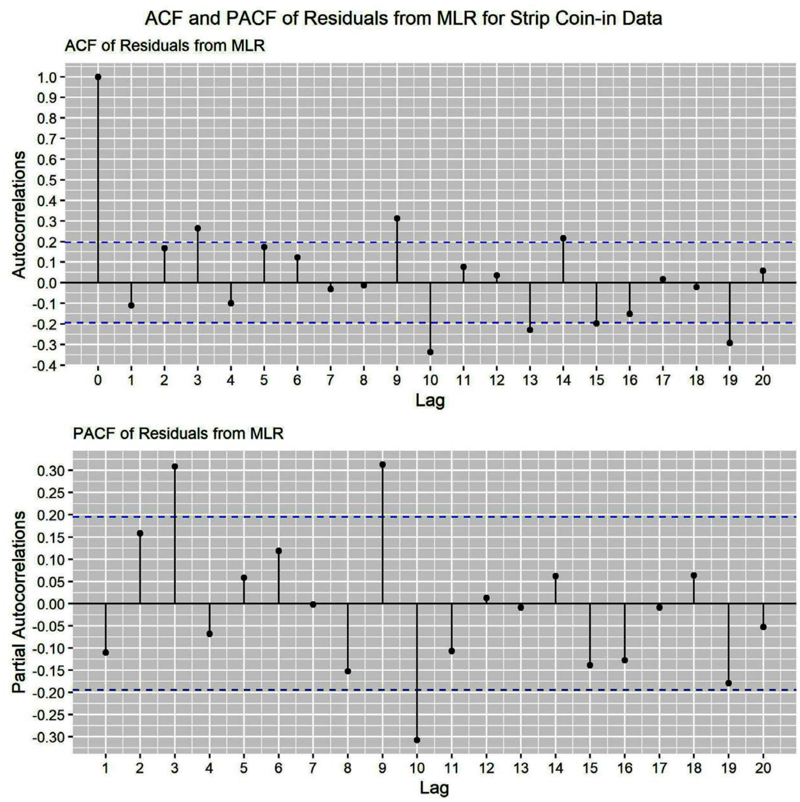

Figure 1. ACF and PACF plots of residuals from the MLR model for Strip Coin-in.

1) the months February, June and December are statistically significant, each with lower average slot coin-in than the other ten months, 2) the terrorist attack of September 2011 (predictor D9_11) had a significant and negative impact on slot coin-in, and 3) the ARIMA terms AR2 and AR3 are statistically significant.

Figure 2 shows the ACF of the residuals from the ARIMA(3,0,0) Time Series Regression Model of Table 2, and also the P-values of the Ljung-Box (LB) test for 10 lags. The ACF plot shows that the residuals from the ARIMA(3,0,0) Time Series Regression Model are not auto-correlated, which is confirmed by the LB test since P-values at lags 1 through 10 are all above 0.05.

Table 3 shows the results of t-tests using the HAC estimator of the covariance matrix; these results are very similar to the ones obtained from ARIMA model (Table 2).

Figure 3 shows plots of the Las Vegas Strip Coin-in and five bootstrap samples generated from it using the estimated ARIMA(3,0,0) model given in Table 2. The bootstrap samples are seen to have the same general pattern as the Las Vegas Strip Coin-in time series. Figure 4 shows the histogram of 1000 P-values for D9_11 obtained from 1000 bootstrap runs. Since D911 term is highly significant (see Table 2 and Table 3), the P-values from bootstrap samples are expected to be small. The term D11 (dummy column for November) is not significant (see Table 2 and Table 3), and hence the 1000 P-values for D11 are expected to exceed 0.05. Figure 4 and Figure 5 clearly show that all of the intervention analysis methods yield similar results.

Table 3. Results of significance tests using the three HAC estimators for Las Vegas Strip Slot Coin-in Data.

		P-Value		
Term	Coeff	HAC	Kern-HAC	Newey-West
(Intercept)	178,074	0.00	0.00	0.00
Zt	5321	0.06	0.04	0.05
Zt2	3371	0.01	0.00	0.00
Zt3	12,251	0.00	0.00	0.00
DFeb	−15,827	0.00	0.00	0.00
DMar	10,343	0.05	0.04	0.04
DApr	−7969	0.07	0.08	0.04
DMay	949	0.84	0.84	0.82
DJun	−10,794	0.02	0.02	0.01
DJul	−521	0.91	0.91	0.90
DAug	−3477	0.50	0.49	0.45
DSep	−1103	0.82	0.83	0.81
DOct	2524	0.43	0.40	0.31
DNov	−6164	0.26	0.29	0.24
DDec	−24,787	0.00	0.00	0.00
D9_11	−11474	0.00	0.00	0.00

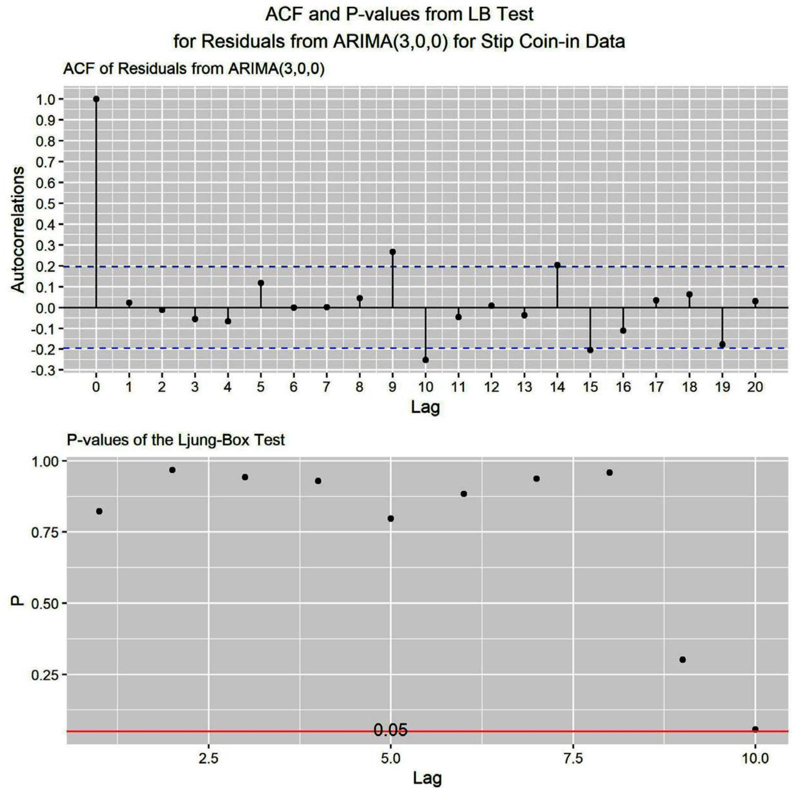

Figure 2. Plots of ACF and P-values of L-B test for residuals from the ARIMA(3,0,0) model for Strip Coin-in.

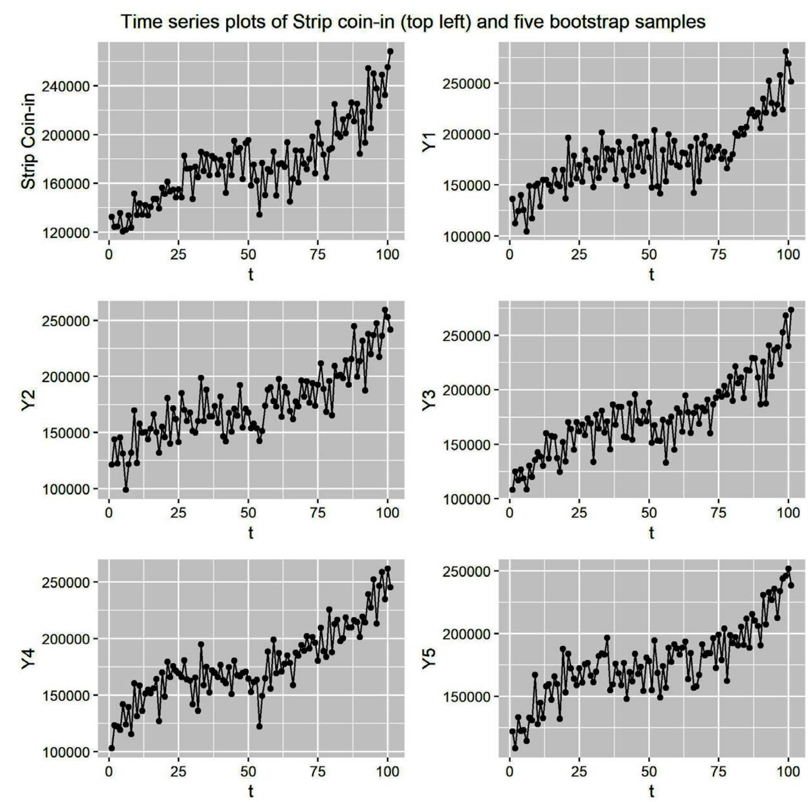

Figure 3. Time series plots of Strip Coin-in (top left) and five bootstrap samples.

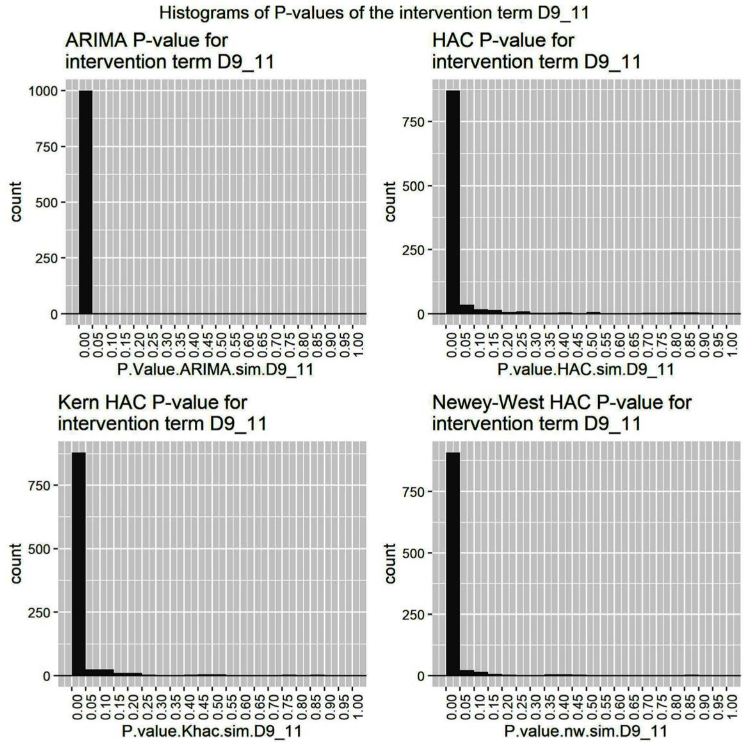

Figure 4. Histograms of P-values of the intervention term D9_11 for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the Strip Coin-in data.

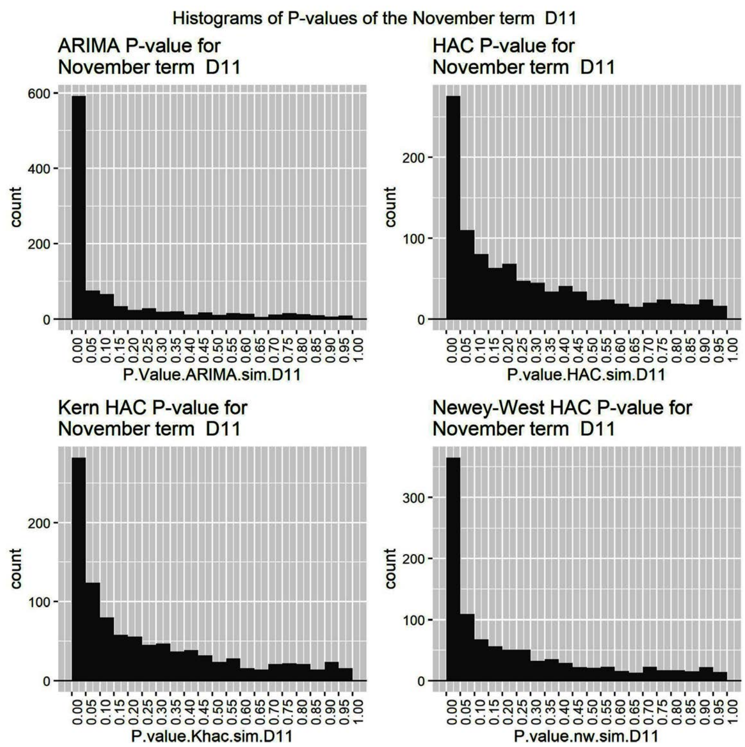

Figure 5. Histograms of P-values of the dummy variable for November for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the Strip Coin-in data.

Example 2: Impact of tax rate increase on Marketing Expenditure

Ahlgren *et al.* [6] used secondary data for the period January 2000 to December 2006, provided by a major Illinois riverboat operator, to assess the impact of a gaming tax rate increase in the state of Illinois on marketing expenditure by the riverboat operator. A time series regression model was fitted to marketing expenditure; the predictors were a cubic trend, eleven dummy columns for the months of February through December (see Example 1), a dummy column DTax for Illinois tax increase which was 1 for months 43 through 67 and 0 for all other months, and an interaction term between the linear term and the DTax column.

Table 4 shows the MLR model fitted to the Marketing Expenditure data. The cubic trend is significant, along with the month of July, the intervention event DTax and the interaction term.

Figure 6 shows plots of ACF and PACF of the residuals from the MLR model of Table 4. The behavior of the autocorrelation functions suggests an ARIMA(0,0,1) process but the residuals turned out to be auto-correlated. An ARIMA(0,0,2) process provided good fit to the Marketing Expenditure data, as can be seen from Figure 7.

Table 5 shows the fitted ARIMA model. The intervention term DTax is highly significant, and the quadratic trend component Zt2 is not significant (P-value = 0.91). The t-tests using HAC estimators yield similar results (see Table 6).

Table 4. MLR model fitted to Marketing Expenditure Data.

Term	Coeff	SE	t-stat	P-Value
(Intercept)	2,401,290	85,168	28.20	0.00
DFeb	50,976	102,531	0.50	0.62
DMar	186,509	102,626	1.82	0.07
DApr	193,334	102,780	1.88	0.06
DMay	156,254	102,991	1.52	0.13
DJun	139,804	103,258	1.35	0.18
DJul	270,399	103,991	2.60	0.01
DAug	170,899	103,471	1.65	0.10
DSep	91,241	103,587	0.88	0.38
DOct	116,540	103,771	1.12	0.27
DNov	162,265	104,030	1.56	0.12
DDec	174,118	104,369	1.67	0.10
DTax	−872,334	99,982	−8.73	0.00
Zt	−79,785	79,378	−1.01	0.32
Zt2	−10,461	30,231	−0.35	0.73
Zt3	107,666	38,177	2.82	0.01
Zt × DTax	574,225	140,373	4.09	0.00

Table 5. ARIMA(0,0,2) regression model fitted to Marketing Expenditure Data.

Term	Coeff	SE	t-stat	P-Value
ma1	0.79	0.13	6.04	0.00
ma2	0.40	0.13	3.08	0.00
intercept	2,336,700	96,680	24.17	0.00
DFeb	53,826	60,783	0.89	0.38
DMar	190,250	86,570	2.20	0.03
DApr	191,670	98,658	1.94	0.05
DMay	149,160	98,846	1.51	0.13
DJun	127,250	99,162	1.28	0.20
DJul	258,060	99,678	2.59	0.01
DAug	199,570	99,452	2.01	0.04
DSep	114,340	99,455	1.15	0.25
DOct	133,910	99,579	1.34	0.18
DNov	169,570	88,794	1.91	0.06
DDec	174,830	64,935	2.69	0.01
DTax	−919,370	125,930	−7.30	0.00
Zt	−211,180	113,700	−1.86	0.06
Zt2	5014	43,350	0.12	0.91
Zt3	162,250	53,238	3.05	0.00
Zt × DTax	945,590	190,680	4.96	0.00

Table 6. t-test results from the three HAC estimators for Marketing Expenditure Data.

				P-Value		
Term	Coeff	SE	t-stat	HAC	Kern-HAC	Newey-West
(Intercept)	2,401,290	86,223	27.85	0.00	0.00	0.00
DFeb	50,976	52,603	0.97	0.34	0.35	0.30
DMar	186,509	69,202	2.70	0.01	0.01	0.00
DApr	193,334	101,366	1.91	0.06	0.08	0.04
DMay	156,254	99,852	1.56	0.12	0.12	0.08
DJun	139,804	109,682	1.27	0.21	0.14	0.10
DJul	270,399	107,795	2.51	0.01	0.01	0.00
DAug	170,899	120,226	1.42	0.16	0.13	0.08
DSep	91,241	130,436	0.70	0.49	0.47	0.41
DOct	116,540	113,469	1.03	0.31	0.27	0.21
DNov	162,265	98,990	1.64	0.11	0.05	0.03
DDec	174,118	50,850	3.42	0.00	0.00	0.00
DTax	−872,334	134,931	−6.47	0.00	0.00	0.00
Zt	−79,785	157,052	−0.51	0.61	0.74	0.70
Zt2	−10,461	48,646	−0.22	0.83	0.86	0.84
Zt3	107,666	68,712	1.57	0.12	0.32	0.25
Zt × DTax	574,225	109,356	5.25	0.00	0.00	0.00

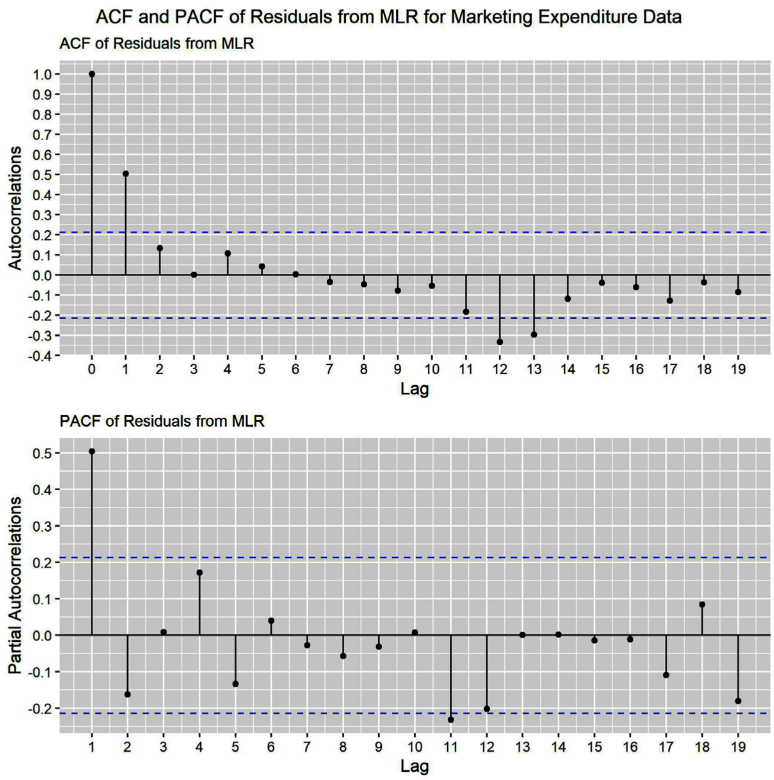

Figure 6. ACF and PACF plots of residuals from the MLR model for Marketing Expenditure.

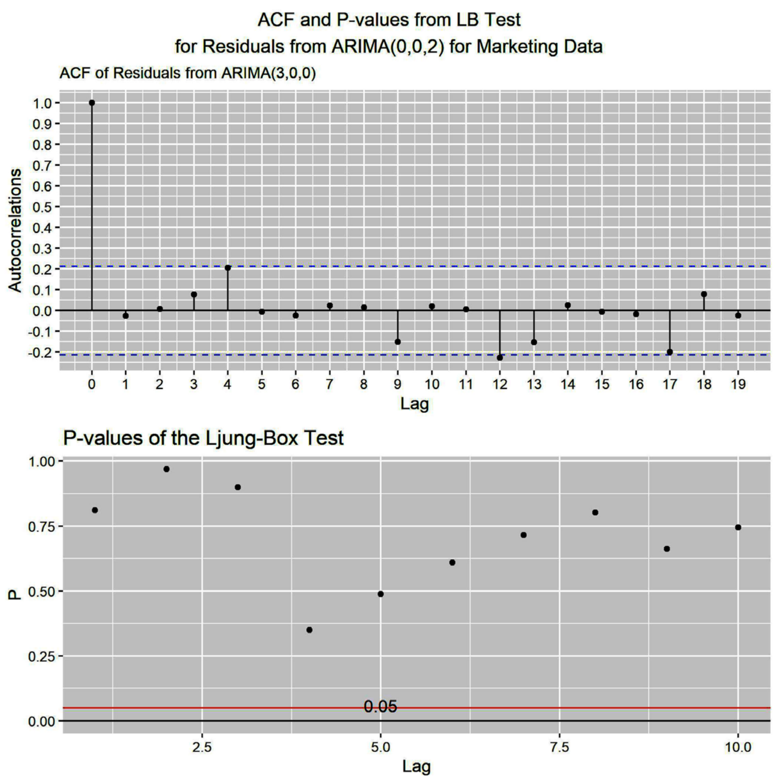

Figure 7. Plots of ACF and P-values of L-B test for residuals from the ARIMA(3,0,0) model for Marketing Expenditure.

Figure 8 shows the Marketing Expenditure time series (top left) and five bootstrap samples generated from it using the estimated ARIMA(0,0,2) model given in Table 5.

Figure 9 shows the histograms of 1000 bootstrap P-values for ARIMA and HAC methods for the statistically significant term DTax; Figure 10 shows the same for the insignificant term Zt2. Both of these figures show that the ARIMA and HAC methods provide similar results.

Example 3: Synthetic time series

The advantages of working with a simulated (synthetic) time series are that the truth is known and hence the estimates can be compared to the corresponding true parameter values.

The synthetic time series was generated from the following model:

$$Y_{0t} = 5000 + 20 \times t + 3000 \times \text{DJun} + 3200 \times \text{DJul} + 2500 \times \text{DAug} + 1000 \times \text{DSep} + 4500 \times \text{DE} + e$$

where

$t = 1, 2, \cdots, 84$ with 1 representing January month of Year 1 and 84 representing December of Year 7.

DE = dummy variable for the intervention event E.

DE = 1 for $t = 43, 44, \cdots, 67$; 0 otherwise.

e = ARIMA(2,0,2) error process with $\phi_1 = 0.8897$, $\phi_2 = -0.4858$, $\theta_1 = -0.2279$, $\theta_2 = 0.2488$ and sd $\sigma = 1000$.

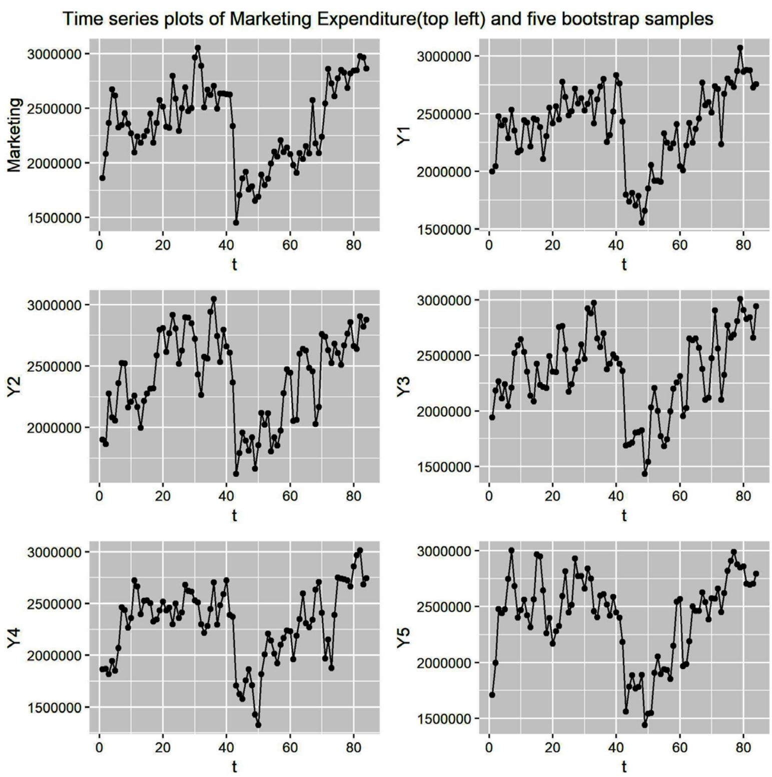

Figure 8. Time series plots of Marketing Expenditure (top left) and five bootstrap samples.

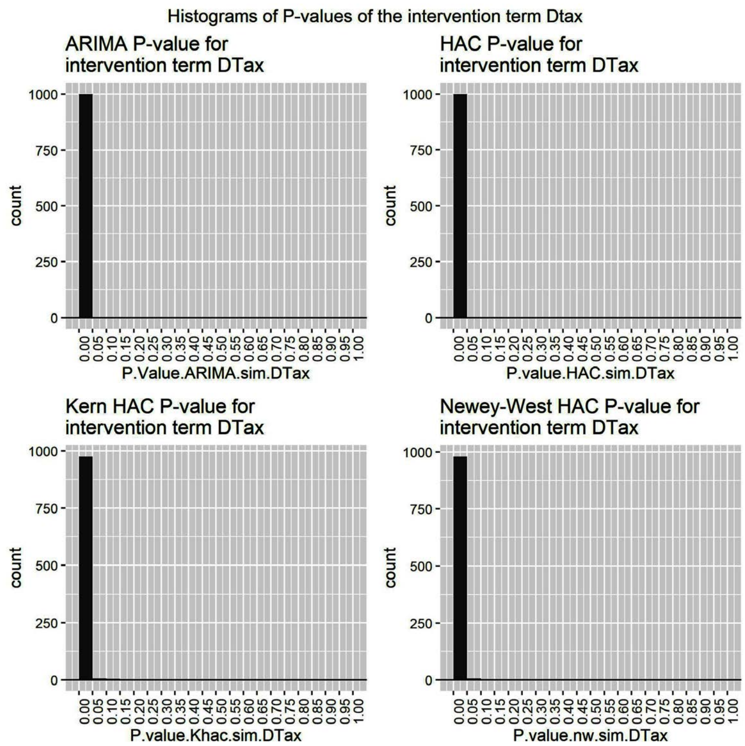

Figure 9. Histograms of P-values of the intervention term DTax for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the Marketing Expenditure data.

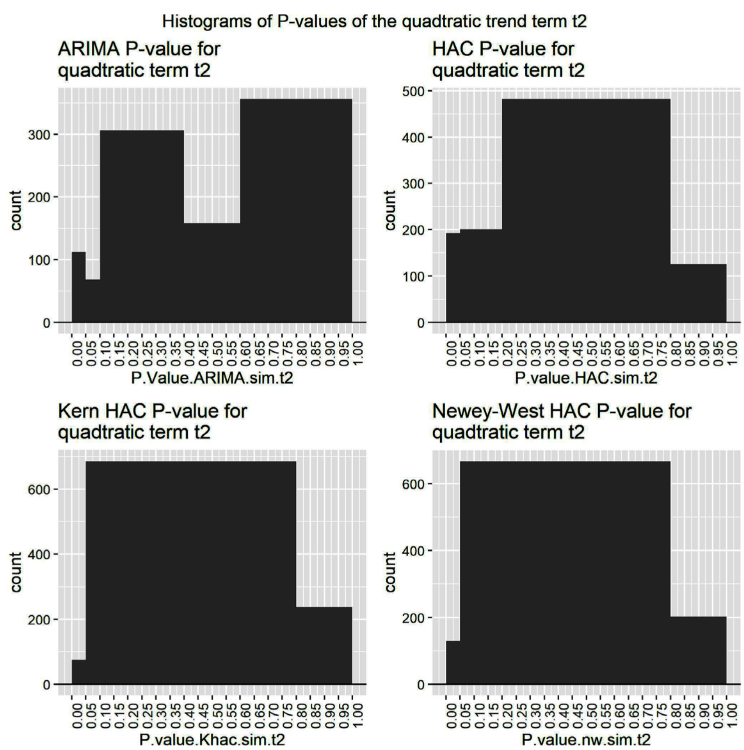

Figure 10. Histograms of P-values of the dummy variable for the quadtratic term Zt2 for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the Marketing Expenditure data.

Figure 11, plots of the ACF and PACF functions, suggest an ARIMA(2,0,2) model, and Figure 12 shows that ARIMA(2,0,2) model yields uncorrelated residuals. Note that the synthetic time series was generated from an ARIMA(2,0,2) process.

Tables 7-9 show the fitted MLR model, the ARIMA(2,0,2) model, and the significance test results from the HAC estimators, respectively. The true values used to generate the synthetic time series are also shown in these three tables. It can be seen from this table that all of the estimated coefficients are close to their corresponding true values. The intervention term DE is seen to be highly significant, and DNov is not significant.

Figure 13 shows the synthetic time series and five bootstrap samples generated from the synthetic series and the estimated ARIMA(2,0,2) model of Table 8. The histograms of 1000 bootstrap P-values for the intervention term DE and the dummy variable for November (DNov) are shown in Figure 14 and Figure 15, respectively. Both of these figures again show that the ARIMA and HAC methods provide similar results.

5. Discussion

For each of the three examples presented in this paper, the ARIMA method of intervention analysis and HAC methods yielded similar results. The four methods (ARIMA, HAC, Kern-HAC, and Newey-West) also yielded similar results for 1000 bootstrap samples from the original time series in each case. These results demonstrate that the simpler HAC method can be used for intervention

Table 7. MLR model fitted to the synthetic data.

Term	True value	Estimated Coeff	SE	t-stat	P-Value
(Intercept)	5000	5052.46	474.93	10.64	0.00
t	20	22.69	5.55	4.09	0.00
DFeb	0	782.41	613.40	1.28	0.21
DMar	0	−210.10	613.47	−0.34	0.73
DApr	0	−418.36	613.60	−0.68	0.50
DMay	0	−202.43	613.78	−0.33	0.74
DJun	3000	3325.94	614.00	5.42	0.00
DJul	3200	3690.96	614.93	6.00	0.00
DAug	2500	3304.80	614.60	5.38	0.00
DSep	1000	1466.00	614.98	2.38	0.02
DOct	0	302.53	615.40	0.49	0.62
DNov	0	−300.60	615.88	−0.49	0.63
DDec	0	−734.41	616.40	−1.19	0.24
DE	4500	4054.31	292.26	13.87	0.00

Table 8. ARIMA(0,0,2) regression model fitted to the synthetic data.

Term	True value	Estimated Coeff	SE	t-stat	P-Value
ar1	0.8897	0.86	0.26	3.29	0.00
ar2	−0.4858	−0.54	0.17	−3.08	0.00
ma1	−0.2279	−0.20	0.29	−0.67	0.50
ma2	0.2488	0.22	0.18	1.20	0.23
intercept	5000	5196.30	451.88	11.50	0.00
t	20	20.52	6.14	3.34	0.00
DFeb	0	659.54	383.31	1.72	0.09
DMar	0	−369.36	543.55	−0.68	0.50
DApr	0	−533.84	619.73	−0.86	0.39
DMay	0	−255.27	614.44	−0.42	0.68
DJun	3000	3296.66	566.10	5.82	0.00
DJul	3200	3619.11	534.03	6.78	0.00
DAug	2500	3186.39	556.57	5.73	0.00
DSep	1000	1327.86	605.77	2.19	0.03
DOct	0	228.50	617.62	0.37	0.71
DNov	0	−247.24	552.31	−0.45	0.65
DDec	0	−599.44	405.99	−1.48	0.14
DE	4500	4125.18	319.54	12.91	0.00

Table 9. Results of significance tests using the three HAC estimators for the synthetic data.

					P-Value		
Term	True value	Estimated Coeff	SE	t-stat	HAC	Kern-HAC	Newey-West
(Intercept)	5000	5052.46	524.70	9.63	0.00	0.00	0.00
t	20	22.69	6.48	3.50	0.00	0.13	0.06
DFeb	0	782.41	673.07	1.16	0.25	0.17	0.18
DMar	0	−210.10	799.27	−0.26	0.79	0.77	0.75
DApr	0	−418.36	554.21	−0.75	0.45	0.40	0.36
DMay	0	−202.43	504.41	−0.40	0.69	0.76	0.74
DJun	3000	3325.94	595.24	5.59	0.00	0.00	0.00
DJul	3200	3690.96	542.29	6.81	0.00	0.00	0.00
DAug	2500	3304.80	641.68	5.15	0.00	0.00	0.00
DSep	1000	1466.00	474.09	3.09	0.00	0.01	0.01
DOct	0	302.53	516.78	0.59	0.56	0.57	0.53
DNov	0	−300.60	404.99	−0.74	0.46	0.19	0.26
DDec	0	−734.41	389.84	−1.88	0.06	0.00	0.00
DE	4500	4054.31	239.75	16.91	0.00	0.00	0.00

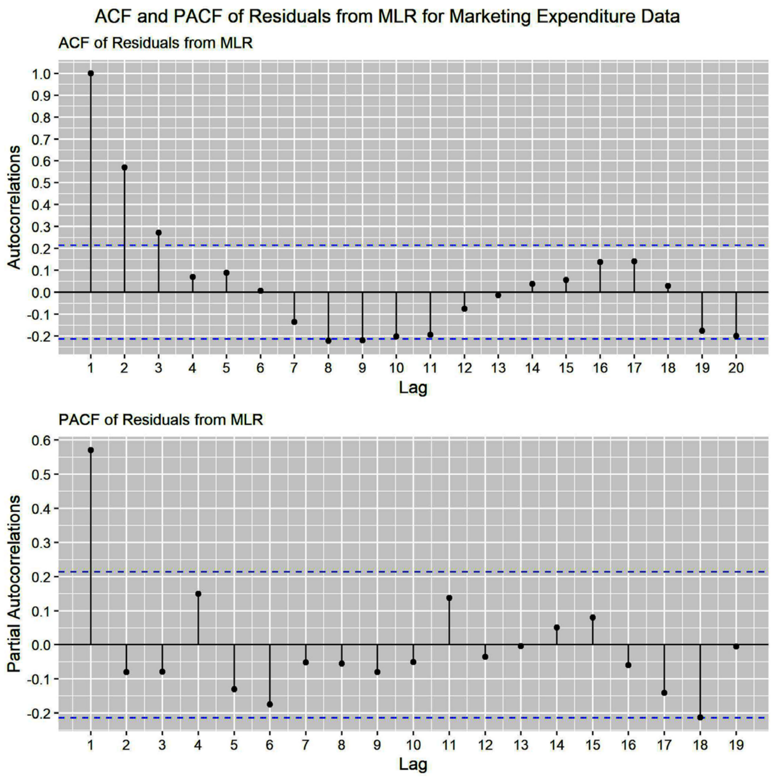

Figure 11. ACF and PACF plots of residuals from the MLR model for the synthetic data.

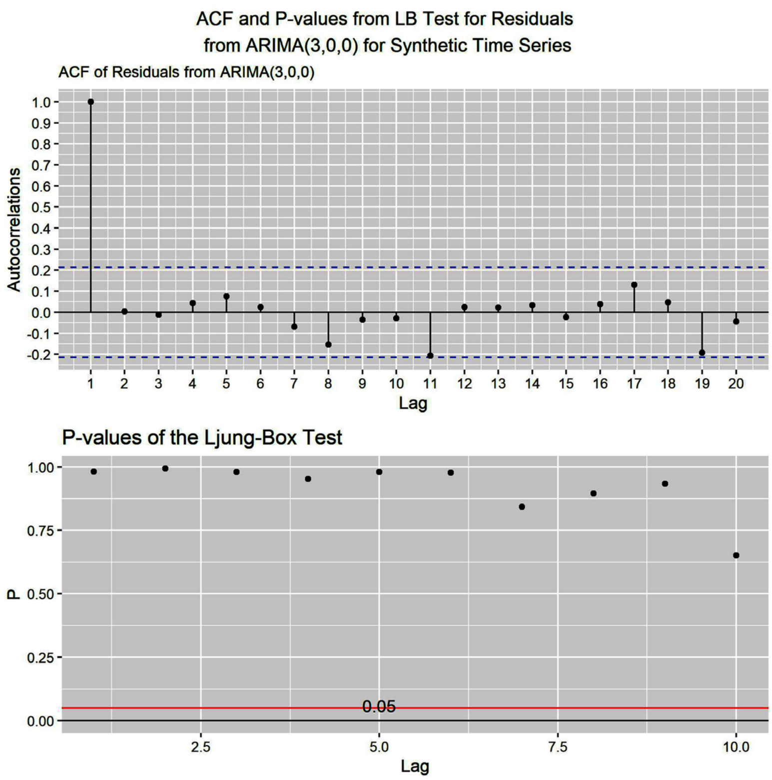

Figure 12. Plots of ACF and P-values of L-B test for residuals from the ARIMA(3,0,0) model for the synthetic data.

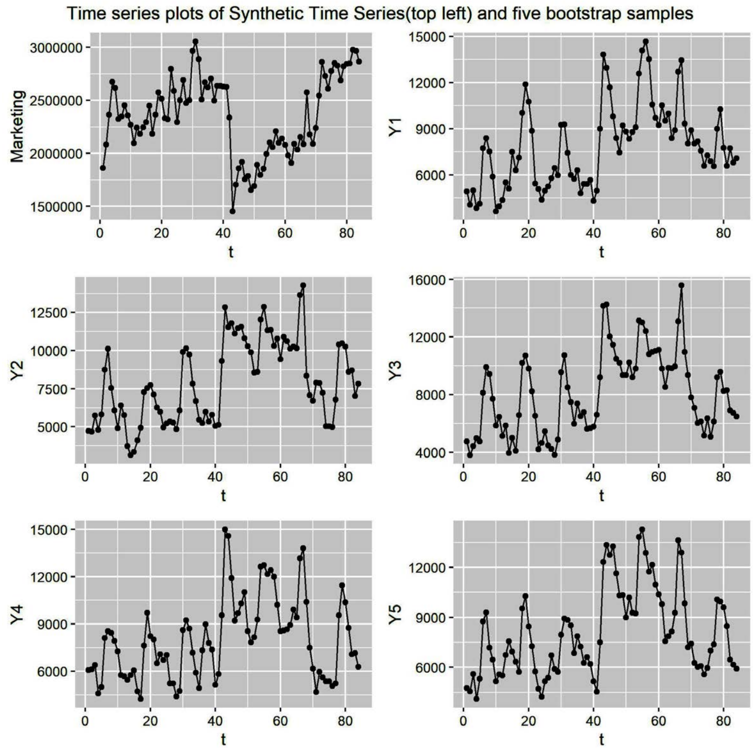

Figure 13. Time series plots of the synthetic data Y0 (top left) and five bootstrap samples.

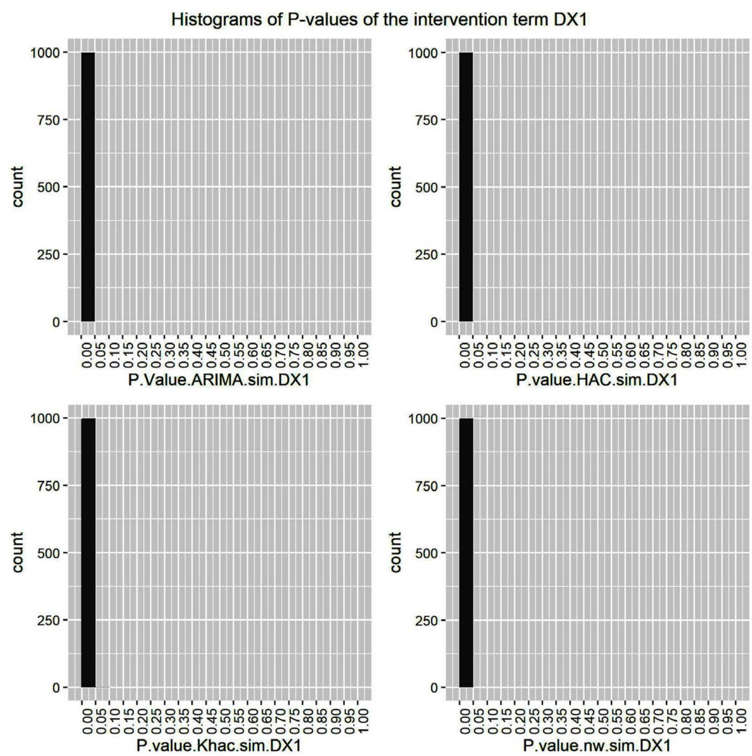

Figure 14. Histograms of P-values of the intervention term DE for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the synthetic data.

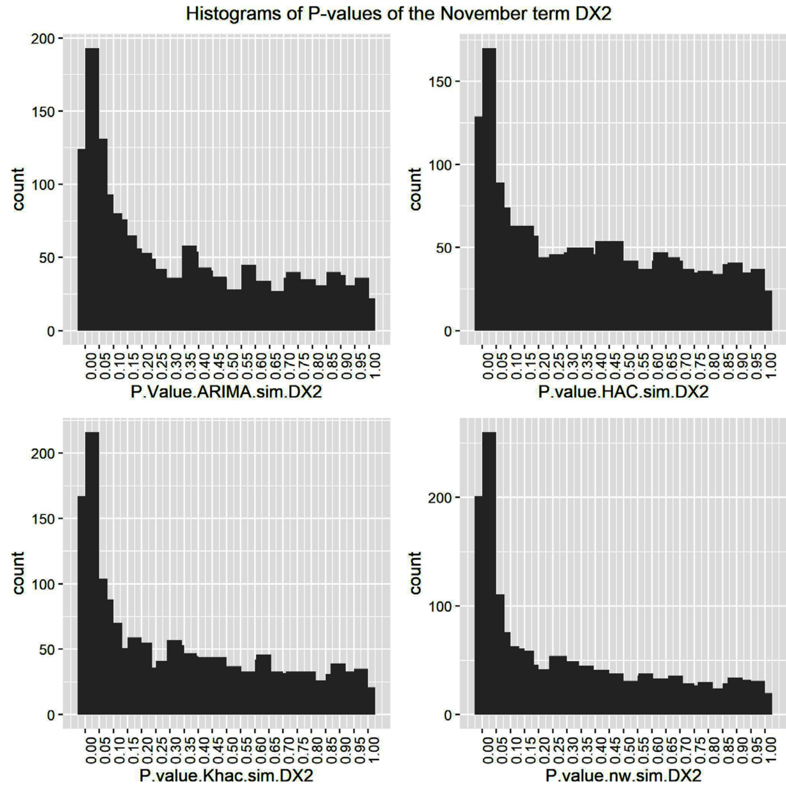

Figure 15. Histograms of P-values of the dummy variable for DDec for the four methods considered in this paper from 1000 parametric bootstrap samples generated from the synthetic data.

analysis instead of the ARIMA model, especially in situations where finding the right ARIMA model turns out to be challenging.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] McDowall, D. (1980) Interrupted Time Series Analysis. Sage University Papers Series: Quantitative Applications in the Social Sciences. Sage Publications, Thousand Oaks. https://doi.org/10.4135/9781412984607

[2] Bonham, C. and Gangnes, B. (1996) Intervention Analysis with Cointegrated Time Series: The Case of the Hawai'i Hotel Room Tax. *Applied Economics*, **28**, 1281-1293. https://doi.org/10.1080/000368496327831

[3] Min, J.C.H. (2008) Intervention Analysis of Inbound Tourism: A Case Study of Taiwan. In: Chen, J.S., Ed., *Advances in Hospitality and Leisure*, Advances in Hospitality and Leisure, Volume 4, Emerald Group Publishing Limited, Bingley, 53-74. https://doi.org/10.1016/S1745-3542(08)00003-9

[4] De Gooijer, J. and Hyndman, R. (2006) 25 Years of Time Series Forecasting. *International Journal of Forecasting*, **22**, 443-473. https://doi.org/10.1016/j.ijforecast.2006.01.001

[5] Ahlgren, M., Dalbor, M. and Singh, A.K. (2009) Estimating the Effects of the 2003 Gaming Tax Restructuring on Riverboat Gaming Volume. *UNLV Gaming Research & Review Journal*, **13**, 45-58.

[6] Ahlgren, M., Tanford, S. and Singh, A.K. (2013) Impact of the 2003 Illinois Gaming Tax Rate Increase on Marketing Spending. *UNLV Gaming Research & Review Journal*, **17**, 1-15.

[7] Toma, M., McGrath, R. and Payne, J.E. (2009) Hotel Tax Receipts and the "Midnight in the Garden of Good and Evil": A Time Series Intervention Seasonal ARIMA Model with Time-Varying Variance. *Applied Economics Letters*, **16**, 653-656. https://doi.org/10.1080/13504850701221808

[8] Goh, C. and Law, R. (2002) Modeling and Forecasting Tourism Demand for Arrivals with Stochastic Nonstationary Seasonality and Intervention. *Tourism Management*, **23**, 499-510. https://doi.org/10.1016/S0261-5177(02)00009-2

[9] Eisendrath, D., Bernhard, D., Lucas, A.F. and Murphy, D.J. (2008) Fear and Managing in Las Vegas: An Analysis of the Effects of September 11, 2001, on Las Vegas Strip Gaming Volume. *Cornell Hospitality Quarterly*, **49**, 145-162. https://doi.org/10.1177/1938965508315369

[10] Suh, E., Tanford, S. and Singh, A.K. (2012) The Indirect Gaming Contributions of Cash and Comped Casino Dining: Does Providing Complimentary Meals Pay off at the Slots? *International Journal of Hospitality Management*, **31**, 1303-1310. https://doi.org/10.1016/j.ijhm.2012.03.013

[11] Zheng, T.S., Farrish, J. and Wang, X.F. (2013) How Did Different Restaurant Segments Perform Differently through the Recession? An ARIMA with Intervention Analysis on US Restaurant Stock Indices. *Journal of Hospitality Financial Management*, **20**, Article 1. https://scholarworks.umass.edu/jhfm/vol20/iss2/1

[12] D'Amuri, F. and Marcucci, J. (2017) The Predictive Power of Google Searches in Forecasting US Unemployment. *International Journal of Forecasting*, **33**, 801-816. https://doi.org/10.1016/j.ijforecast.2017.03.004

[13] Su, J. and Deng, G.M. (2014) Application of Intervention Analysis Model in Yu Ebao Yield Prediction. *Modern Economy*, **5**, 864-868. https://doi.org/10.4236/me.2014.58079

[14] Huang, Y. (2019) Government Intervention and Corporate Investment Efficiency: Evidence from China. *Journal of Service Science and Management*, **12**, 267-276. https://doi.org/10.4236/jssm.2019.123018

[15] Bowerman, B., O'Connell, R.T. and Koehler, A.B. (2004) Forecasting, Time Series, and Regression. 4th Edition, Thomson Brooks/Cole, Belmont, 435-436.

[16] Weil, R.L., Frank, P.B., Hughes, C.W. and Wagner, M.J. (2007) Litigation Services Handbook: The Role of the Financial Expert. John Wiley & Sons, New York, 6.

[17] Kennedy, P. (2003) A Guide to Econometrics. MIT Press, Cambridge, 213.

[18] Ljung, G.M. and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika*, **65**, 297-303. https://doi.org/10.1093/biomet/65.2.297

[19] Hyndman, R.J. and Athanasopoulos, G. (2014) Forecasting: Principles and Practice. OTexts. https://play.google.com/store/books/details?id=gDuRBAAAQBAJ

[20] Zeileis, A. (2006) Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software*, **16**, 1-16. http://www.jstatsoft.org/v16/i09 https://doi.org/10.18637/jss.v016.i09

[21] Kutner, M.H., Nachtsheim, C. and Neter, J. (2004) Applied Linear Regression Models. McGraw-Hill/Irwin, New York, 217-218.

[22] Andrews, D. (1991) Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, **59**, 817-858. https://doi.org/10.2307/2938229

[23] R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org