7. IMPACT OF SEASONAL LEVEL SHIFT (SLS) ON TIME SERIES FORECASTING

Hayat SHAHID[1]
Amena UROOJ[2]
Zahid ASGHAR[3]

Abstract

The effect of not treating Seasonal Level Shift (SLS) outliers on forecast accuracy, and prediction intervals is the focus of this study. We examine the impact of SLS on point and interval forecasts using simulation experiment for time series models including SAR (1) and SMA (1) for different parameter values, sample sizes and time of occurrences. We extend the strategy suggested by Asghar and Urooj (2017) to forecasting in the presence of SLS by looking at forecast accuracy and prediction interval. We demonstrate that SLS significantly increases the inaccuracy of the SARIMA models, increases the bias in the SARIMA estimates, and significantly affects the prediction intervals. However, after detection and adjustment of SLS, SARIMA estimates become less biased, and forecast accuracy measure and prediction interval significantly improve. The difference of location of SLS from forecast origin has similar effect on bias and forecast accuracy in SAR (1) model. While, in SMA (1) model, the SLS occurring at the beginning of the series has greater adverse effect than that occurring at the middle or end of the series. Three monthly time series data from Pakistan are used to explore the issue.

Keywords: Seasonal Level Shift (SLS), SARIMA, forecast accuracy, point forecasts, interval forecasts

JEL Classification: C15, C18, C63, C32, C87

[1] *Pakistan Institute of Development Economics, Islamabad. Email: hayat.pide20@gmail.com.*

[2] *Pakistan Institute of Development Economics, Islamabad. Corresponding Author. Email: amna@pide.org.pk.*

[3] *Quaid-i-Azam University, Islamabad. Email: g.zahid@gmail.com.*

Appendix

Table A1. Sampling Distribution of SMA-Hat $\widehat{\theta_1}$ and Forecast Accuracy

n=50, W=5sigma, C.V=3.5 BP=7							
Series with SLS							
SMA(1)	Coef	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.6912	0.1253	0.0242	2.0084	1.5820	-0.2054	0.2873
0.4	0.7564	0.1191	0.0241	1.9266	1.5168	-0.2063	0.2659
0.6	0.8290	0.1132	0.0311	1.8707	1.5019	-0.1671	0.2448
0.8	0.9299	0.1077	0.0240	1.7714	1.3875	-0.1968	0.2214
Series adjusted for SLS							
SMA(1)	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.3752	0.0869	-0.004	1.269	1.010	0.0705	0.698
0.4	0.5083	0.0846	-0.005	1.233	0.981	0.0566	0.6574
0.6	0.6584	0.0832	-0.005	1.213	0.965	0.0450	0.6052
0.8	0.9573	0.0784	-0.004	1.155	0.917	0.0269	0.5100
Series without SLS							
SMA(1)	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.1941	0.0685	3.5E-06	0.9829	0.7785	-0.0056	0.6898
0.4	0.3924	0.0688	7.6E-06	0.9850	0.7806	-0.0069	0.6540
0.6	0.5943	0.0688	6.0E-06	0.9876	0.7824	-0.0044	0.6063
0.8	0.7976	0.0690	-7.6E-05	0.9899	0.7840	-0.0047	0.5533

Note: n is the number of samples have four quarter so total observation 50*4. C.v* is the critical value for detecting of SLS. BP* is the break point where SLS occur. All measure have mean value in the table getting from 5000 number of iteration*

Table A2. SAR-Hat $\widehat{\emptyset_4}$, Forecast Accuracy with Increase in Sample size

n=50 , W=5sigma,3.5 BP=7							
Series with SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.880	0.088	0.024	1.317	1.021	-0.015	0.287
0.4	0.890	0.082	0.024	1.238	0.956	-0.012	0.265
0.6	0.908	0.077	0.051	1.149	0.912	0.017	0.244
0.8	0.924	0.073	0.024	1.120	0.858	-0.002	0.221
Series Adjusted for SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.550	0.081	-0.009	1.150	0.919	0.034	0.572
0.4	0.641	0.077	-0.012	1.099	0.878	0.032	0.528
0.6	0.734	0.074	-0.017	1.045	0.835	0.013	0.463
0.8	0.825	0.072	-0.018	1.032	0.821	0.033	0.384
n=100 , W=5sigma,3.5 BP=7							
Series with SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.895	0.062	0.012	1.290	1.014	-0.014	0.287
0.4	0.910	0.058	0.012	1.205	0.945	-0.009	0.265
0.6	0.922	0.055	0.008	1.125	0.882	-0.095	0.244
0.8	0.936	0.052	0.012	1.082	0.844	-0.005	0.221
Series Adjusted for SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.560	0.057	-0.003	1.148	0.915	0.0130	0.533
0.4	0.653	0.054	-0.005	1.098	0.875	0.0131	0.493
0.6	0.742	0.052	-0.007	1.057	0.841	0.0140	0.443
0.8	0.841	0.050	-0.010	1.026	0.816	0.0163	0.361

Note: n* is the number of samples have four quarter so total observation 50*4, And 100*4. C.v* is the critical value for detecting of SLS. BP* is the break point where SLS occur. All measure have mean value in the table getting from 5000 number of iteration.

Table A3.SMA-Hat $\widehat{\emptyset_4}$, Forecast Accuracy with Increase in Sample size

n=50 , W=5sigma,3.5 BP=7 Series with SLS							
SMA(1)	Coef	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.6912	0.1253	0.0242	2.0084	1.5820	-0.2054	0.2873
0.4	0.7564	0.1191	0.0241	1.9266	1.5168	-0.2063	0.2659
0.6	0.8290	0.1132	0.0311	1.8707	1.5019	-0.1671	0.2448
0.8	0.9299	0.1077	0.0240	1.7714	1.3875	-0.1968	0.2214
Series Free of SLS							
SMA(1)	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.1941	0.0685	3.5E-06	0.9829	0.7785	-0.0056	0.6898
0.4	0.3924	0.0688	7.6E-06	0.9850	0.7806	-0.0069	0.6540
0.6	0.5943	0.0688	6.0E-06	0.9876	0.7824	-0.0044	0.6063
0.8	0.7976	0.0690	-7.6E-05	0.9899	0.7840	-0.0047	0.5533
n=100 , W=5sigma,3.5 BP=7 Series with SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.6973	0.0913	0.0113	2.0521	1.6383	-0.2126	0.2873
0.4	0.7630	0.0863	0.0115	1.9626	1.5663	-0.2134	0.2659
0.6	0.8329	0.0821	0.0440	1.8971	1.4814	-0.2483	0.2448
0.8	0.9294	0.0780	0.0122	1.7973	1.4307	-0.2109	0.2214
Series Adjusted for SLS							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.2	0.1964	0.0493	1.4E-06	0.9922	0.7888	-0.0031	0.6915
0.4	0.3972	0.0493	-1.1E-05	0.9933	0.7894	-0.0013	0.6555
0.6	0.5963	0.0494	-7.1E-06	0.9939	0.7900	-0.0025	0.6063
0.8	0.7977	0.0495	-3.0E-05	0.9953	0.7914	-0.0031	

Note: n is the number of samples have four quarter so total observation 50*4, And 100*4. C.v* is the critical value for detecting of SLS. BP* is the break point where SLS occur. All measure have mean value in the table getting from 5000 number of iteration.*

Table A4. SAR-Hat $\widehat{\emptyset_4}$, Forecast Accuracy, with Different Location of SLS

n=50 , W=5sigma, C.V= 3.5 BP= 7							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.1	0.871	0.092	0.0247	1.364	1.060	-0.018	0.299
0.3	0.888	0.085	0.0248	1.275	0.987	-0.013	0.276
0.5	0.902	0.080	0.0252	1.203	0.927	-0.008	0.256
0.7	0.915	0.075	0.0247	1.145	0.879	-0.005	0.234
0.9	0.936	0.072	0.0238	1.099	0.840	0.000	0.203
n=50 , W=5sigma, BP= 25							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.1	0.812	0.090	0.0044	1.332	1.033	-0.040	0.409
0.3	0.840	0.084	0.0037	1.247	0.964	-0.035	0.277
0.5	0.863	0.078	0.0036	1.177	0.906	-0.031	0.256
0.7	0.887	0.074	0.0027	1.121	0.861	-0.027	0.234
0.9	0.923	0.071	0.0001	1.077	0.823	-0.022	0.204
n=50 , W=5sigma, BP= 38							
SAR1	Coeff	SE	ME	RMSE	MAE	ACF	MASE
0.1	0.687	0.086	0.0075	1.289	0.995	-0.053	0.553
0.3	0.738	0.081	0.0063	1.215	0.936	-0.045	0.277
0.5	0.783	0.076	0.0058	1.154	0.886	-0.038	0.256
0.7	0.833	0.073	0.0041	1.108	0.849	-0.030	0.234
0.9	0.903	0.070	0.0040	1.074	0.821	-0.022	0.204

Note: n is the number of samples have four quarter so total observation 50*4, And 100*4. C.v* is the critical value for detecting of SLS. BP* is the break point where SLS occur. All measure have mean value in the table getting from 5000 number of iteration*

Table A5. Forecast Interval or Prediction Interval In Case of SAR (1) Model

n=50 , W=5sigma, BP=7				
	95 % Forecast Interval in Case of SLS		95% Forecast Interval in Case of without SLS	
SAR= 0.1	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	6.9808	1.8302	1.9574	-1.9514
Q2	2.5892	-2.5613	1.9546	-1.9542
Q3	2.6101	-2.5403	1.9555	-1.9532
Q4	2.6101	-2.5679	1.9555	-1.9558
SAR= 0.5	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	7.1284	1.9255	2.1049	-2.1084
Q2	2.6204	-2.5825	2.1079	-2.1054
Q3	2.6147	-2.5882	2.1043	-2.1090
Q4	2.6186	-2.5842	2.1069	-2.1064
SAR= 0.8	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	6.8780	2.4896	1.9996	-1.9320
Q2	2.1912	-2.1971	1.9501	-1.9815
Q3	2.2303	-2.1579	1.9835	-1.94813
Q4	2.2007	-2.1875	1.9585	-1.9731

n* is the number of samples have four quarter so total observation 50*4. BP* is the break point where SLS occur. All measures have mean value in the table getting from 5000 number of iteration.

Table A6. Prediction Interval In Case of SMA (1) Model

n=50 , W=5sigma, BP=7				
	95 % Forecast Interval in Case of SLS		95% Forecast Interval in Case of without SLS	
SMA = 0.1	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	5.9069	-2.4936	1.9543	-1.9486
Q2	4.3284	-3.4310	1.9518	-1.9511
Q3	4.3335	-3.4258	1.9520	-1.9508
Q4	4.3335	-3.4357	1.9520	-1.9530
SMA = 0.5	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	6.1913	-1.7550	2.1117	-2.1193
Q2	4.5795	-3.3667	2.1156	-2.1154
Q3	4.5892	-3.3570	2.1179	-2.1130
Q4	4.5791	-3.3671	2.1161	-2.1148
SMA = 0.8	Upper Limit	Lower limit	Upper Limit	Lower Limit
Q1	5.8773	-1.1034	2.0003	-1.9902
Q2	4.2270	-2.7538	1.9911	-1.9994
Q3	4.2420	-2.7388	2.0068	-1.9837
Q4	4.2431	-2.7377	2.0019	-1.9886

Table A7. Models Results of Monthly FBR Tax collection

Model for Series with Outlier					
Model	Estimates(S.E)			AIC	sigma^2
	MA(1)	SAR(1)	SAR(2)		
(0,1,1)*(2,1,0)	-1.255(0.077)	-0.206(0.096)	-0.109(0.112)	1167.3	173.8
Model for series Adjusted for Outlier					
Model	Estimates (S.E)			AIC	sigma^2
	MA(1)	SAR(1)	SAR(1)		
(0,1,1)*(2,1,0)	-0.702(0.059)	-0.288(0.094)	0.191(0.099)	1074.28	137.2

Table A8. Models Results of Money in Circulation

Model With SLS				
Model	Estimates (S.E)		AIC	sigma^2
	AR(1)	SAR(1)		
(1,1,0)*(1,0,0)	-0.295(0.073)	0.596(0.066)	1887.15	2209
Model Adjusted for SLS				
Model	Estimates (S.E)		AIC	sigma^2
	AR(1)	SAR(1)	1591.94	390.9
(1,1,0)*(1,0,0)	0.035(0.077)	0.885(0.032)		

Table A9. Models Results of Monthly Broad Money

Model With SLS				
Model	Estimates (S.E)		AIC	sigma^2
	AR(1)	AR(2)		
(2,1,0)*(0,0,0)	0.033(0.084)	0.367(0.083)	1595.69	26634
Model Adjusted for SLS				
Model	Estimates (S.E)		AIC	sigma^2
	AR(1)	AR(2)	1494.78	11595
(2,1,0)*(0,0,0)	0.158(0.075)	0.547(0.075)		

Table A10. Forecast Accuracy of Different SARIMA Models

ME	RMSE	MAE	MASE	ACF
Series:1 FBR Monthly Tax Collection SARIMA $(0,1,1)*(2,1,0)_{12}$				
1.594	12.597	8.442	0.217	-0.095
1.342	11.195	8.133	0.229	-0.057
Series 2: Monthly Money Circulation Model SARIMA $(1,1,0)*(1,0,0)_{12}$				
11.634	46.869	25.919	0.739	-0.079
4.559	19.716	13.785	0.498	-0.056
Series 3: Monthly Broad Money SARIMA $(2,1,0)*(0,0,0)_{12}$				
58.424	162.534	119.635	0.937	-0.308
29.467	107.241	81.503	0.752	-0.361

Table A11. Prediction interval of FBR Tax Collection

	95 % Forecast Interval in Case of SLS			95% Forecast Interval in Case of series Adjusted for SLS		
	Point forecast	Upper limit	Lower limit	Point forecast	Upper limit	Lower limit
Jul	208.61	241.69	175.53	202.60	226.03	179.17
Aug	242.91	276.67	209.16	248.16	272.61	223.71
Sep	321.35	355.76	286.93	330.97	356.39	305.54
Oct	276.85	311.91	241.79	279.22	305.58	252.85
Nov	278.56	314.26	242.86	277.45	304.73	250.17
Dec	375.71	412.04	339.39	382.81	410.97	354.65
Jan	260.07	297.01	223.13	255.98	284.98	226.97
Feb	272.60	310.15	235.05	275.56	305.40	245.72
Mar	350.27	388.41	312.13	347.90	378.54	317.26
Apr	295.84	334.57	257.12	296.65	328.07	265.22
May	349.30	388.60	309.99	347.71	379.90	315.51
Jun	504.27	544.15	464.40	401.97	434.91	369.02

Table A12. Prediction interval of Broad Money

	95 % Forecast Interval with SLS			95% Forecast Interval Adjusted for SLS		
	Point forecast	Upper limit	Lower limit	Point forecast	Upper limit	Lower limit
Jan	16245.79	16572.19	15919.39	16280.76	16496.12	16065.40
Feb	16299.57	16768.79	15830.35	16370.68	16700.15	16041.21
Mar	16315.91	16971.13	15660.69	16425.69	16922.96	15928.42
Apr	16336.2	17139.77	15532.63	16483.55	17128.35	15838.75
May	16342.87	17294.41	15391.33	16522.76	17328.88	15716.64
Jun	16350.54	17432.01	15269.07	16560.58	17517.79	15603.37
Jul	16353.24	17557.78	15148.70	16587.99	17697.66	15478.32
Aug	16356.14	17673.15	15039.13	16612.99	17867.96	15358.01

Table A13. Prediction interval of Money Circulation

	95 % Forecast Interval with SLS			95% Forecast Interval for series Adjusted for SLS		
	Point forecast	Upper limit	Lower limit	Point forecast	Upper limit	Lower limit
Dec	3606.06	3700.06	3512.05	3606.91	3646.45	3567.37
Jan	3643.71	3758.71	3528.70	3714.96	3771.86	3658.06
Feb	3672.24	3809.24	3535.24	3758.71	3828.82	3688.60
Mar	3687.90	3842.68	3533.12	3781.61	3862.81	3700.40
Apr	3692.26	3863.28	3521.24	3788.21	3879.16	3697.26
May	3767.11	3952.88	3581.35	3899.50	3999.24	3799.75
Jun	3917.97	4117.42	3718.52	3978.41	4086.25	3870.58
Jul	3861.18	4073.43	3648.94	3992.08	4107.43	3876.73
Aug	3844.18	4068.49	3619.87	4014.12	4136.53	3891.71
Sep	3906.84	4142.60	3671.08	4107.31	4236.39	3978.23
Oct	3924.85	4171.53	3678.17	4134.09	4269.51	3998.66
Nov	3953.25	4210.39	3696.12	4125.06	4266.54	3983.57

Figure 2: FBR Tax Collection

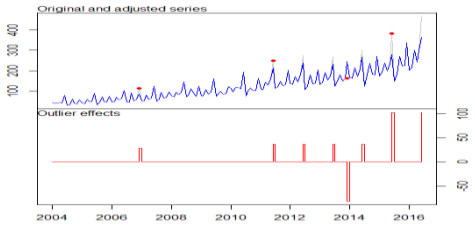

Figure 3: Broad Money

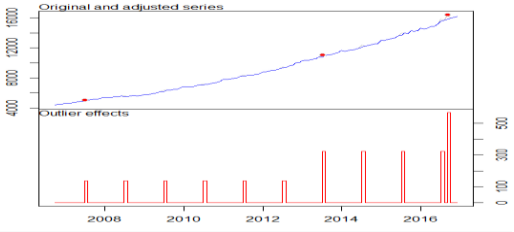

Figure 4: Money Circulation

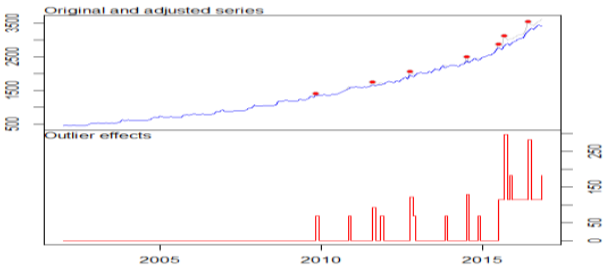

Figure 5: Interval forecasts of the series