Building Better Forecasting Models with Transfer Functions

Jian Cao, PhD
JMP Division, SAS Institute

Copenhagen 03.14.2019

What are Transfer Function Models For

- *"I have collected the monthly data on sales and some predictors like spending on TV ads. How can I build a model to forecast future sales?"*

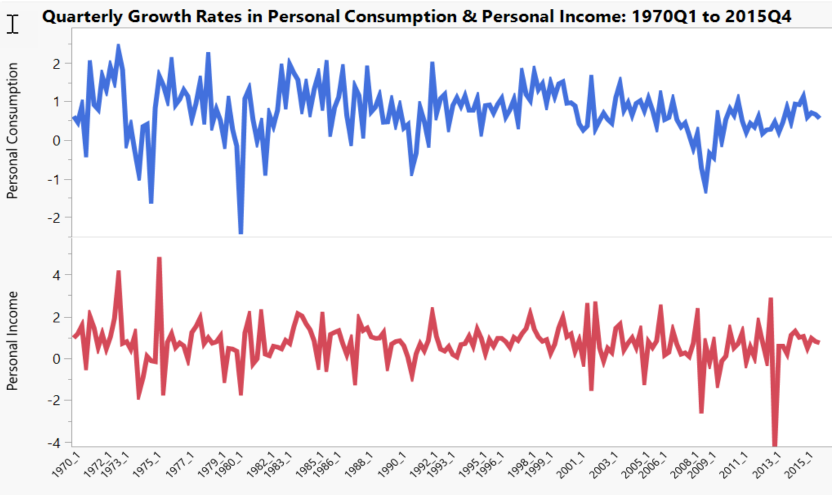
- *"A new manufacturing process has been implemented. I have the data on the defect rate before and after the change. How should I model and analyze the process data?"*

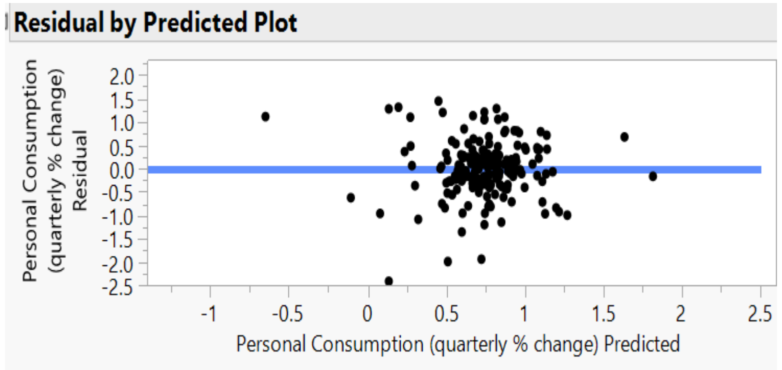
- The statistical analyses fall within the Transfer Function Models

Why Transfer Function Models

y_t: **Quarterly US Personal Consumption Growth Rate**

x_t: **Quarterly US Personal Income Growth Rate**





Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
1.692595	184	0.1531	0.0182*

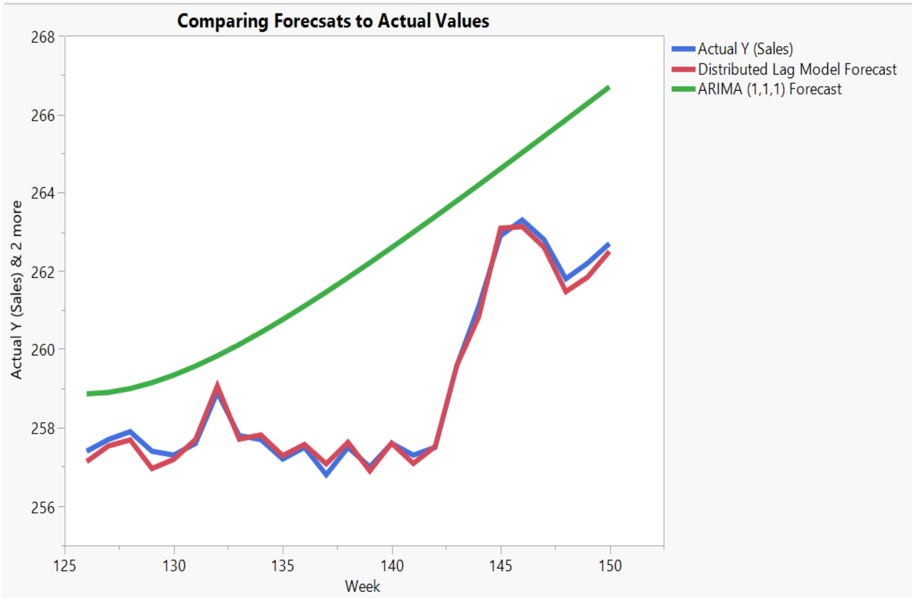
- *"Can I just run a standard least squares regression model to predict y(t)?"* ($\widehat{y_t} = \mathbf{0.55} + \mathbf{0.28} * x_t$)

  --Key assumption that the regression errors are independent is likely violated => Misleading inferences.

Why Transfer Function Models (Cont'd)

- *"How about ARIMA Models?"*

--A pure ARIMA model uses the past values of an variable to predict its future values. If there is a good predictor bring it in.



Blue line: Actual Sales
Green Line: Forecast from Pure ARIMA
Red Line: Forecast from Transfer Function Model

Transfer Function Models

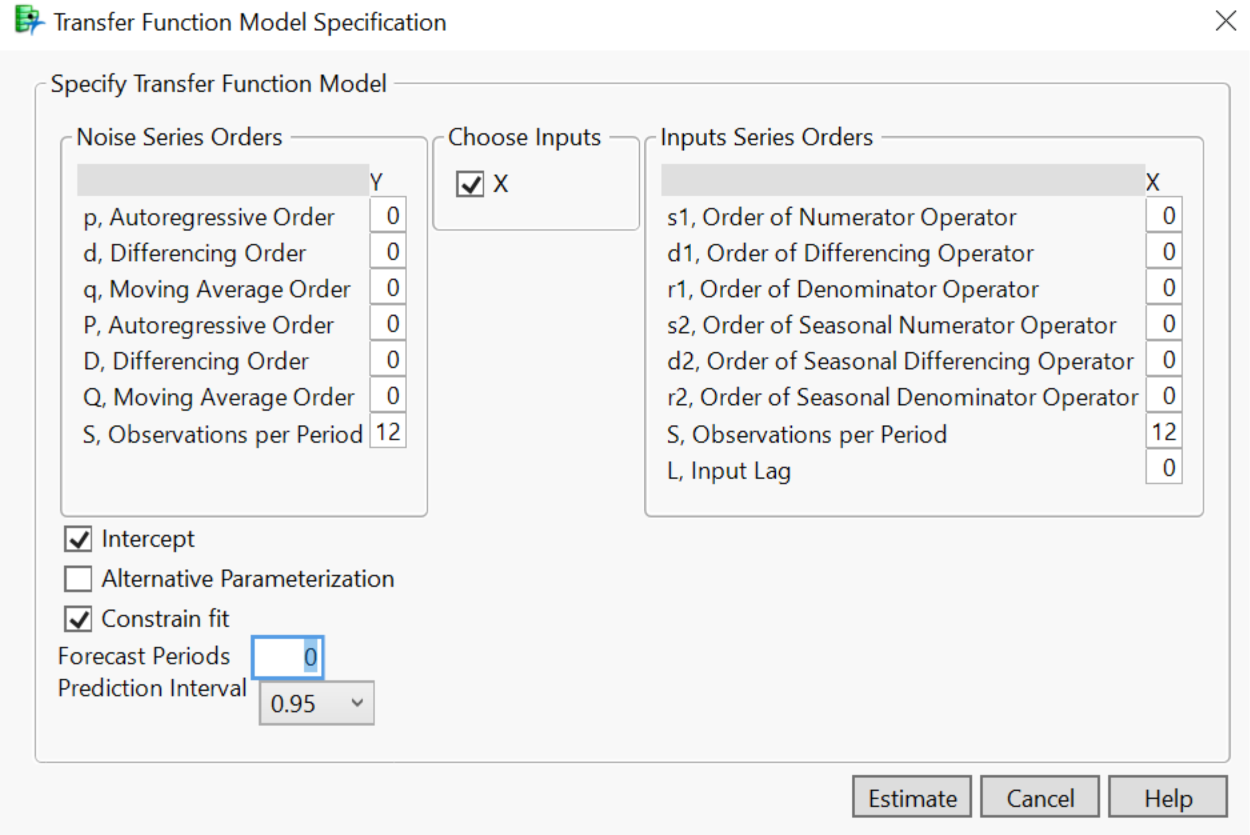
$$y_t = \mu + x_t\beta + e_t \qquad \text{(OLS)}$$

$$y_t = \mu + \frac{\theta(B)}{\varphi(B)} e_t \qquad \text{(Pure ARIMA)}$$

Transfer Function ARIMA Error Term

$$y_t = \mu + \frac{\omega(B)}{\delta(B)} x_{t-l} + \frac{\theta(B)}{\varphi(B)} e_t$$

Launch Window for Transfer Function Model



A very versatile transfer function model!

Use Cases

- Regression with ARMA Errors

  Ex1: Quarterly Growth Rates of Personal Consumption and Income in US

- Distributed Lag Models

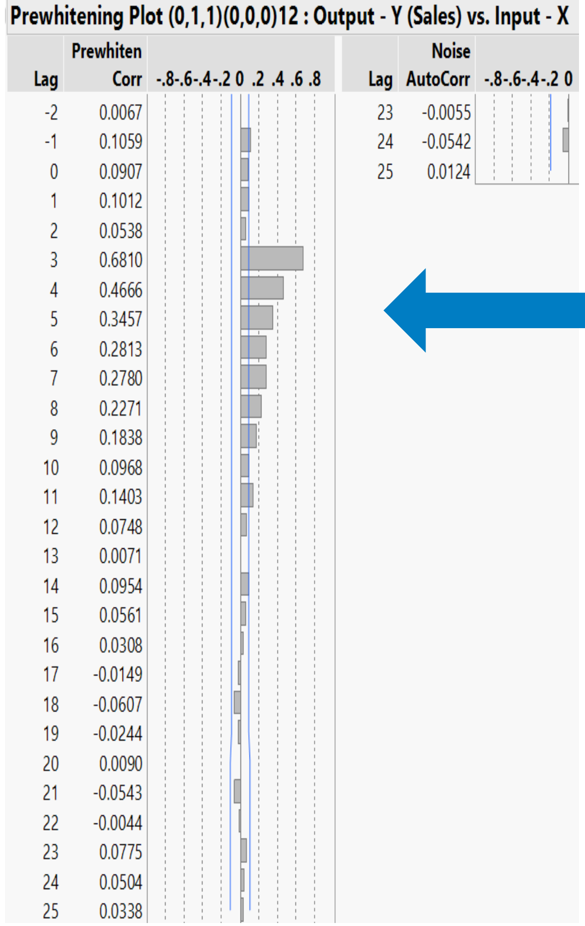
  Ex2: Forecasting Sales using Lagged Predictors

- Intervention Models

  Ex3: Modeling Manufacturing Process Change

Prewhitening

Identifying *when* the input effect takes place, *how long* it lasts and in *what* shape it decays.



Prewhitening Plot (0,1,1)(0,0,0)12 : Output - Y (Sales) vs. Input - X

Lag	Prewhiten Corr
-2	0.0067
-1	0.1059
0	0.0907
1	0.1012
2	0.0538
3	0.6810
4	0.4666
5	0.3457
6	0.2813
7	0.2780
8	0.2271
9	0.1838
10	0.0968
11	0.1403
12	0.0748
13	0.0071
14	0.0954
15	0.0561
16	0.0308
17	-0.0149
18	-0.0607
19	-0.0244
20	0.0090
21	-0.0543
22	-0.0044
23	0.0775
24	0.0504
25	0.0338

Lag	Noise AutoCorr
23	-0.0055
24	-0.0542
25	0.0124

Ex2: Forecasting Sales using Lagged Predictors

Input series x_t is delayed by 3 lags and then exponentially decreasing:

$$\omega_0(x_{t-3} + \delta x_{t-4} + \delta^2 x_{t-5} + \delta^3 x_{t-6} + \delta^4 x_{t-7} + \cdots)$$

try the transfer function $\frac{\omega_0}{1-\delta B} x_{t-3}$ to approximate the input-output relationships

References

George Box, Gwilym Jenkins, Gregory Reinsel and Greta Ljung. 2016. *Time Series: Forecasting and Control* (5th ed.). Hoboken, NJ: John Wiley & Sons.

Rob J. Hyndman and George Athanasopoulos. 2018. *Forecasting: Principles and Practices* (2nd ed.). Melbourne, Australia: OTexts.org (*https://otexts.com/fpp2/*)

SAS Institute Inc. 2018. *JMP® 14 Predictive and Specialized Modeling.* Cary, NC: SAS Institute Inc. (*https://www.jmp.com/support/help/14-2/time-series-analysis.shtml#*)