Appendix for "Forecasting New Product Life Cycle Curves: Practical Approach and Empirical Analysis"

July 19, 2017

Appendix

A. Launch Dates

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Number launched	13	44	3	2	2	26	21	3	17	11	24	4

Table A1 Distribution of number of new products' launches across different calendar months. Launches are distributed across different months. Reproduced from Acimovic et al. (2017)

B. Adjusting for Seasonality

Here, we outline an approach to adjust for seasonality. We applied these steps on the data from Dell, and reported the forecast accuracy results in Table 10. As shown in that table, adjusting for seasonality in this way did not lead to better forecasts. For the ACME data (which we report on in Section 7), we followed these same steps with a minor adjustment because the company provided us with their proprietary seasonality factors, which are based on years of observations across many products.

Steps to adjust for seasonality:

1. Normalize data so that cumulative volume of each product equals 1. In this way, the seasonal effect will not be disproportionately affected by high volume products.

2. Apply an additive or multiplicative seasonal effect model to the normalized data. This model would estimate customer orders based on the following independent variables:
   (a) Seasonal effect (for example, quarterly/monthly/weekly)
   (b) A generalized PLC that the model would estimate
   (c) Fixed effects for year and other attributes

The seasonally-adjusted data is then $D_t^{C[n+1],i} = f_{Season}^{-1}\left(D_t^{C[n],i}\right)$, where $f_{Season}^{-1}(\cdot)$ denotes the deseasonalization function derived from the seasonal effects model, n represents the most recent iteration of cleaning and processing before adjusting for seasonality, and $C[n+1]$ is the data after adjusting for seasonality. The approach for computing the seasonally-adjusted data is the same if seasonal factors are provided externally rather than estimated. Adjusting for seasonality would occur after the Detecting cancellations step and before the Excluding configure-to-order customer orders (Outlier correction) step. This is because what may appear as outliers in a given week may actually be known to be a very high week of demand, for instance, for toys the week after Thanksgiving.

C. Time series clustering

The time series distance metric proposed by Chouakria and Nagabhushan (2007) measures proximity by two dimensions: behavior and value. The proximity of *behaviors* between two series X and Y is evaluated by means of the first order temporal correlation coefficient, which is defined by

$$CORT(X_t, Y_t) = \frac{\sum_{t=1}^{T-1}(X_{t+1} - X_t)(Y_{t+1} - Y_t)}{\sqrt{\sum_{t=1}^{T}(X_{t+1} - X_t)^2}\sqrt{\sum_{t=1}^{T-1}(Y_{t+1} - Y_t)^2}} \tag{A1}$$

$CORT(X_t, Y_t)$ falls into $[-1, 1]$ with 1 meaning that two series behave similarly, i.e. their increase or decrease at any instant of time are similar in direction and rate, -1 meaning that the two series have similar rate of change but opposite in direction, and 0 meaning that the two series are stochastically linearly independent.

The proximity on *values* are measured as the conventional Euclidean distance with $d(X_t, Y_t) = \sqrt{\sum_{t=1}^{T}(X_t - Y_t)^2}$.

The dissimilarity index to measure the proximity between series X_t and Y_t is proposed as

$$d_{CORT}(X_t, Y_t) = \phi_m[CORT(X_t, Y_t)]d(X_t, Y_t) \tag{A2}$$

where $\phi_m(\cdot)$ is an adaptive tuning function to adapt the distance metrics $d(X_t, Y_t)$ to the temporal correlation $CORT(X_t, Y_t)$. With m to be the tuning parameter, the function $\phi_m(u)$ is written as

$$\phi_m(u) = \frac{2}{1 + e^{m\mu}}, m \geq 0.$$

In our case, we use m equal to 2, which is the default choice recommended by Chouakria and Nagabhushan (2007). When using other values for m, the results change very little. Note the dissimilarity measure $d_{CORT}(X_t, Y_t)$ is model-free, it allows us to cluster the fitted PLC curves based on their features in terms of temporal structure and scales.

D. Scree plot

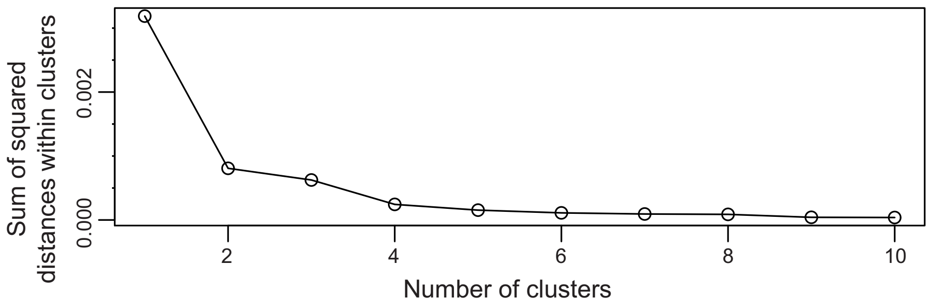

Figure A1 **Sum of squared distances ($\delta_{CORT}(X_t, Y_t)$) within clusters versus number of clusters. We chose four clusters because there is little reduction in sum of squared distances for values above 4.**

E. Stitched forecasts figure

Figure A2 **Illustration of stitched forecasts for two SKUs. The top figure shows** $ScaleActualStitch$**, the middle shows** $ScaleActualSegment$**, while the bottom shows** $ScaleForecast$**. The vertical lines show the stitch points (the epochs at which one forecast horizon ends and we adopt the next quarter's forecasted values).**

F. Additional results for comparing clustering approaches

Limit to SKUs for which we know:			**Count of SKUs**	**Reduction in MASE over no clustering**			
Features	Dell forecasts	Product category		Dell's own forecasts	Our forecasts' clustering methods: *Features*	*Category*	*Time series*
✓		✓	86		1.71%	2.15%	4.83%
✓	✓	✓	34	-0.13%	0.02%	0.54%	2.38%
	✓	✓	45	-1.93%		0.45%	1.97%
		✓	161			3.44%	5.00%

Table A2 **This table has the same structure as Table 8 except it is measuring improvement in MASE.**

G. Additional data and figures for ACME forecasting

Config	Bass	poly2	poly3	poly4	triangle	trapezoid
RMSE						
- SKU01-p	0.01877	0.01699	0.01471	0.00960	0.01233	0.01095
- SKU02-p	0.04083	0.04189	0.04080	0.03855	0.03369	0.03369
- SKU03-p	0.03374	0.03498	0.03495	0.03305	0.03330	0.03298
Log likelihood (sum)	103.01	103.57	106.28	114.98	112.85	114.93
AIC	-188.02	-189.14	-188.56	-199.96	-221.7	-199.86

Table A3 **Summary statistics of PLC fits to the deseasonalized ACME data. Boxes with a double (single) outline denote the best (second best) value in each row. Consistent with Dell PLC fitting results (see Table 2), trapezoid and poly4 curves most effectively fit the data.**

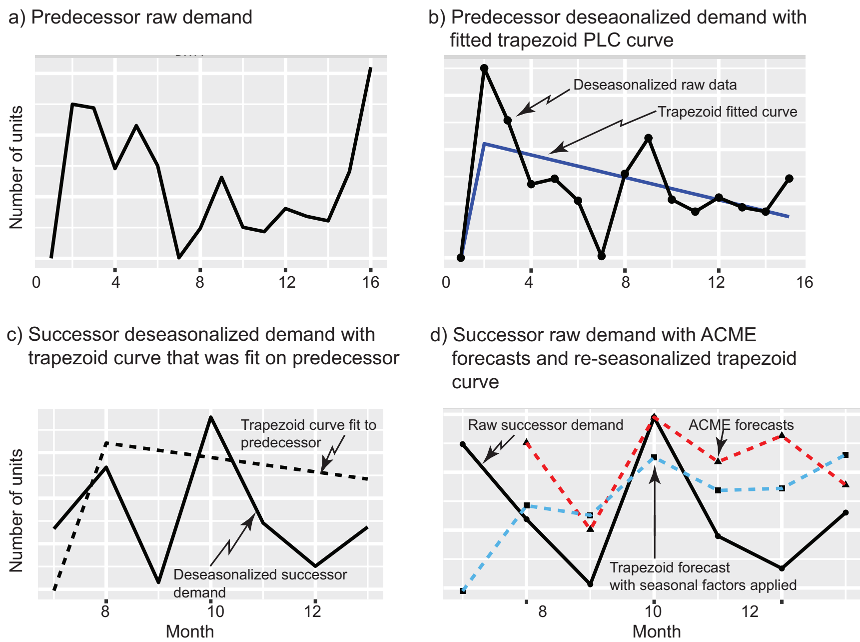

Figure A3 **PLC fitting and forecasting for ACME SKU02**

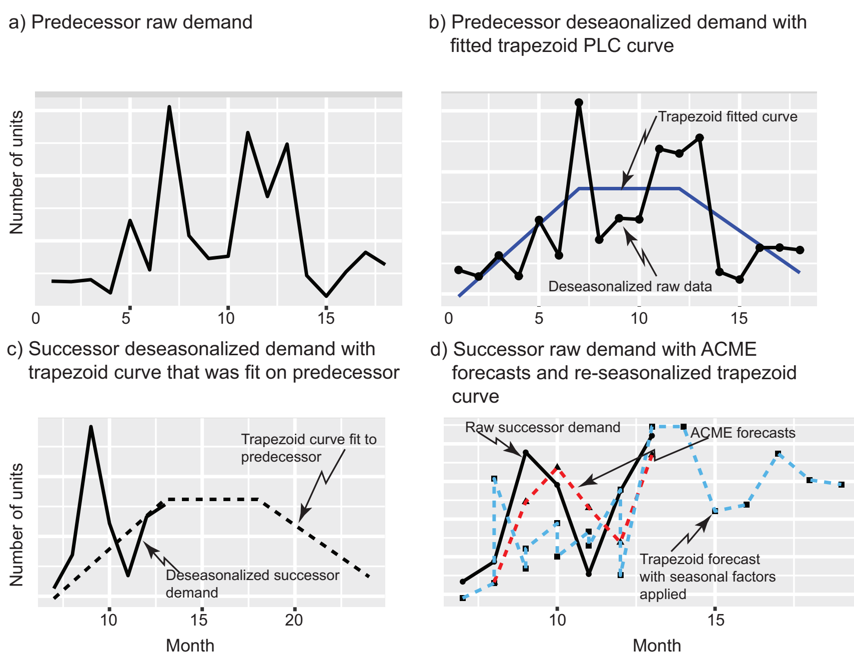

Figure A4 **PLC fitting and forecasting for ACME SKU03**

References

Acimovic, Jason, Francisco Erize, , Kejia Hu, Douglas J. Thomas, Jan A. Van Mieghem. 2017. Product life cycle data-set: Raw and cleaned data of weekly orders for personal computers. *Working paper* .

Chouakria, Ahlame Douzal, Panduranga Naidu Nagabhushan. 2007. Adaptive dissimilarity index for measuring time series proximity. *Advances in Data Analysis and Classification* **1**(1) 5–21.