MEMO

Date: 12/4/17

[redacted]

From: Joshua Markwell

RE: Request for 6-month forecast of sales

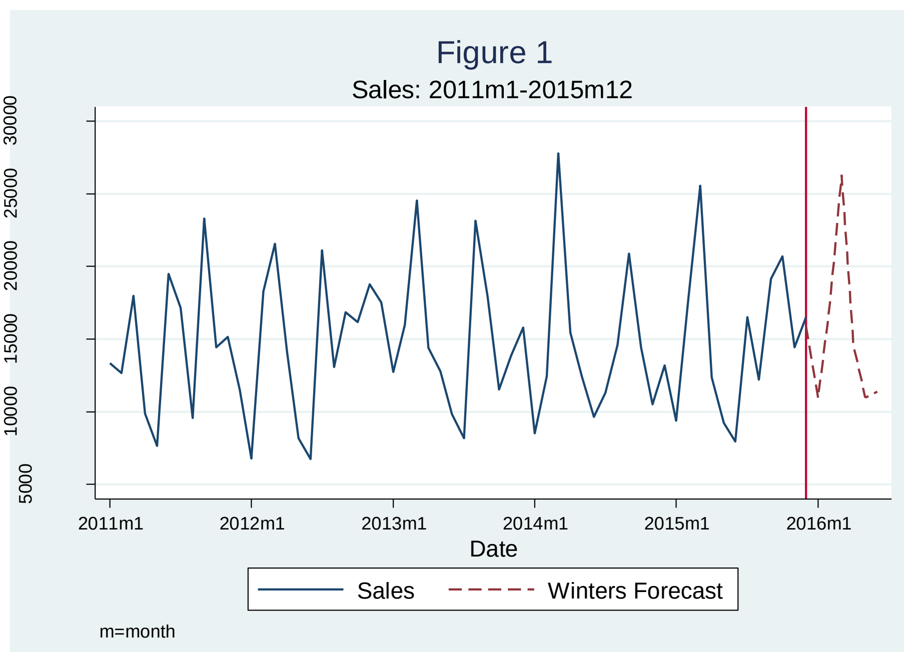
[redacted]

Per Request I have calculated a 6-month sales forecast. Below, Table 1 displays each applied models forecast results with Table 2 displaying the corresponding error.

Table 1				
Date	**Winters**	**Time Decomposition**	**AR**	**Combined**
2016m1	$10,928.60	$ 9,681.80	$ 9,455.90	$ 9,624.10
2016m2	$17,297.20	$ 16,808.60	$13,485.20	$15,510.00
2016m3	$26,308.20	$ 26,027.70	$24,831.30	$25,631.80
2016m4	$14,629.40	$ 14,755.60	$14,740.80	$14,802.40
2016m5	$10,984.70	$ 11,089.20	$11,877.40	$11,451.50
2016m6	$11,381.70	$ 8,915.60	$10,438.20	$ 9,570.10

Table 2		
Model	**MAPE**	**RMSE**
Winters	15.0%	2987.9
Time Decomposition	17.5%	3805.8
AR	17.4%	3307.4
Combined	15.1%	3092.7

Based upon the measures of error (MAPE & RMSE) the Winters model is the most accurate. The MAPE value for the Winters is only slightly lower than the Combined forecast model, however, the Winters model has the lowest RMSE value. Below is a graphic depiction of the Winters model forecast (Figure 1); a vertical line denotes the beginning of the forecasting period. A graphic depiction of every model is located in the appendix.[1]



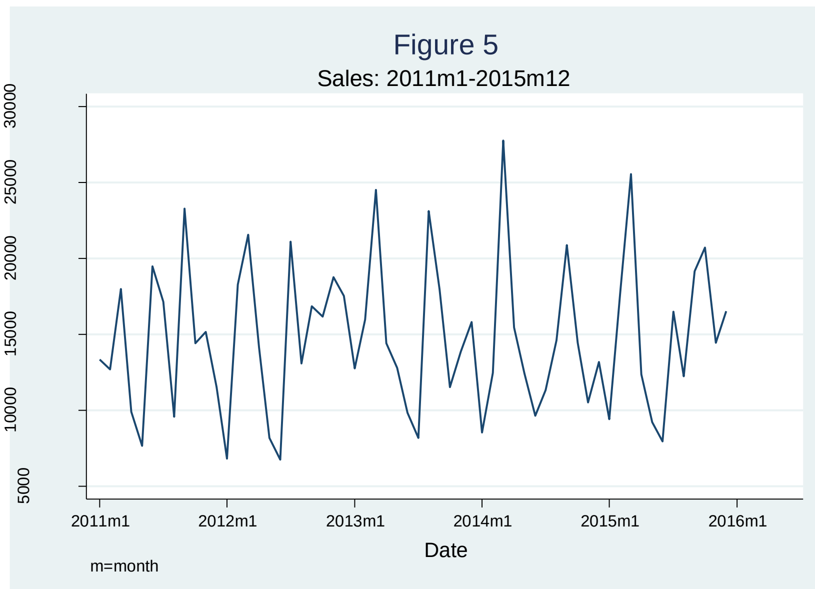
[1] Figures 1-4.

Methodology & Model Selection

This section will discuss why each of the following four models were chosen and how the forecast for each model was calculated.

1. Winters Smoother
2. Time Series Decomposition
3. Auto Regressive
4. Combined
    a. Using Time Decomposition and Auto Regressive

These models were chosen because each handles the seasonal and stationary nature of the data. This can be seen in Figure 5 below. The stationary nature of the data can be confirmed using an Augmented Dickey Fuller Test.[2]



Winters Smoother:

The sales data is seasonal, as seen in Figure 5. Therefore, the Winters model is an appropriate candidate to perform a forecast of the series. The Winters model includes a seasonal estimate which accounts for the repeating behavior within each year. However, the Winters model also incorporates a trend estimate, but there is no trend in our data due to its stationary nature. The lack of trend does not necessarily adversely affect the model or its forecasts; but it is worth noting that the Winters model attempts to account for something that is not there. According to the error measurements in Table 2 (Page 1), the Winters model attempting to account for trend is not causing it to be less accurate than the other three forecasts.

[2] Test 1 in appendix

Variables:

W_{t+m}: Winters' forecast for m periods into the future

F_t: Smoothed value for period t

m: Number of periods in the forecast lead period

T_t: Trend value

S_t: Seasonality estimate

p: Number of periods in the seasonal cycle

Model 1: *sales_forecast*= $W_{t+m} = (F_t + mT_t) \bullet S_{t+m-p}$

Time Series Decomposition Model:

To use a Time Series Decomposition model four steps must be followed:

1. De-seasonalize the data
2. Calculate the trend estimates
3. Calculate the cyclical estimates
4. Calculate forecast

Step 1:

To de-seasonalize the data a 6-period moving average and corresponding centered moving average is found for each month (that is mathematically possible). The first three and last two data points for the series are eliminated after the moving averages are found. The centered moving average for each month is found by adding the moving average of a particular month, the next month, and then dividing the value by two. This eliminates another data point (the very last one). A seasonal coefficient (*sf*) is found for every individual month. The *sf* values are found by dividing the *sales* value for each month by the corresponding centered moving average. The *sf* values for every month (1-12) are averaged over every year creating the *S* averages. The *S* values for months 1 through 6 will be used in the forecast since the forecast period ranges over months 1 through 6. Table 3 below displays the monthly estimates for *S*.

Table 3

Month	*S*- averages
1	0.577
2	0.634
3	0.720
4	0.951
5	0.960
6	0.973
7	0.977
8	0.987
9	1.032
10	1.099
11	1.343
12	1.698

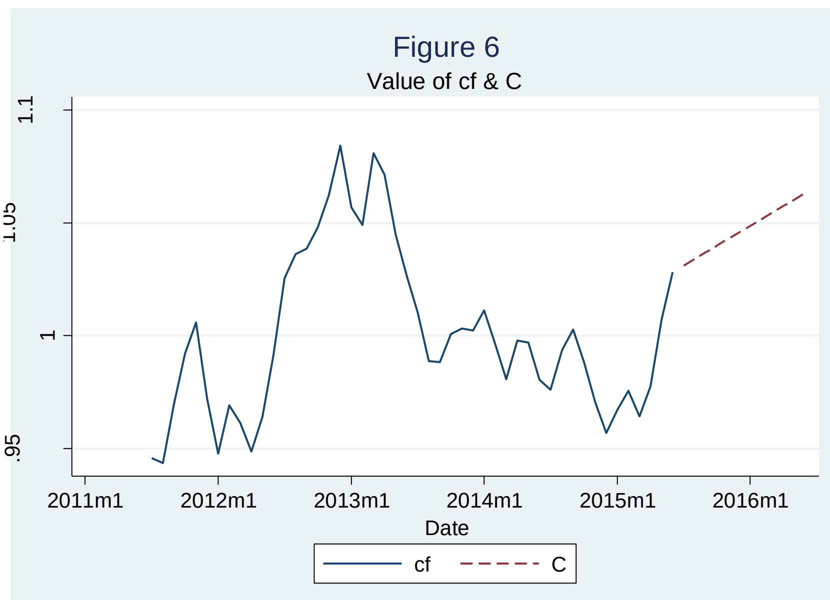
For example: The value of 0.577 for month 1 indicates that on average, in month 1 of every year, the sales values are approximately 42% lower because of seasonality.

Step 2:

To calculate the trend estimates, denoted as *T,* a time trend model is used. The centered moving averages are regressed using time as the independent variable.[3] This model creates a trend estimate for every month, including the six forecasting periods.

Step 3:

Finding the cyclical estimates, denoted as *C*, requires using the trend estimates *T,* from step 2, and the centered moving averages found in step 1. The cyclical coefficients, denoted *cf*, for every month, are equal to the centered moving averages divided by the trend estimates *T*. I forecasted the values of *cf* to find the values of *C*. I used a Holts smoother to forecast the values of *cf,* creating the six values for *C*. I chose the Holts smoother because the *cf* values display an upward trend at the end of its series. Observe Figure 6 below, the upward trend is noticed after period, 2015 month 3, as denoted by the vertical line.



The *cf* values' upward trend and lack of seasonality make the Holts model appropriate to use. The series ***could*** turn downward; however, it could also continue to rise as seen in Figure 6. Table 4 below displays the *C* estimates used for the forecast values.

Table 4	
Date	***C*- estimates**
2016m1	1.048689
2016m2	1.051614
2016m3	1.054538
2016m4	1.057463
2016m5	1.060388
2016m6	1.063313

Step 4:

[3] Equation 1 in appendix

The *S, T, C* and *I* (assume I=1) estimates are multiplied together to calculate the forecasted values. I assume *I,* the 'irregular', coefficient is equal to one because there is no indication of an irregular influence.

Model 2: $sales\_forecast = S \bullet T \bullet C \bullet I$

Auto Regressive Model:

The Auto Regressive model allows me to treat the sales data as a function of itself. To use the auto regressive approach the data must be stationary. The data is stationary based upon the Augmented Dickey Fuller test shown on page 1. I began with an AR(5); considering the series is only 60 observations long, 5 lags terms is a reasonable place to begin. Table 6 below displays the initial results for the potential AR models.

Table 6			
Model	**Significant Lag terms**	**AIC**	**BIC**
AR(5)	lag 2, lag 4	1084.8	1096.9
AR(4)	lag 2, lag 4	1064.7	1074.7
AR(3)	lag 2	1069.2	1077.2

The criterion I used to judge the AR models is based upon significant terms[4] and the information criterion (AIC, BIC). As shown in Table 6, the AR(4) has two significant lag terms and the lowest information criterion, therefore it is the best model out of those three. The insignificant lag terms are removed leaving an AR(2, 4). However, there is no aspect of the AR model that controls for seasonality. Thus, I added monthly dummy variables. Observed Equation 2 below.

Equation 2: $sales_t = \beta_0 + \beta_1\,\text{sales}_{t-2} + \beta_2\,\text{sales}_{t-4} + \beta_3\,\text{m1} + \beta_4\,\text{m2} + \ldots + \beta_{13}\,\text{m11} + \varepsilon_t$

- This equation uses month 12 (December as a reference variable)

After running the model with monthly dummy variables, I checked the information criterion and the significance of each monthly variable. I then removed insignificant monthly variables and saw that the information criterion had improved.[5] Model 3 below displays the best model.

Model 3: $sales_t = \delta_0 + \delta_1\,\text{sales}_{t-2} + \delta_2\,\text{sales}_{t-4} + \delta_3\,\text{m1} + \delta_4\,\text{m3} + \delta_5\,\text{m6} + \delta_6\,\text{m9} + \varepsilon_t$

Calculating the forecast values requires finding the fitted values for the regression and then recursively building for the forecast for each forecast period. The process is as follows: Run Model 3, predict the fitted values, include the fitted values in the *sales* values of the regression, and then repeat the process for each forecasting period.[6]

[4] Significant at the 5% level
[5] See appendix for full results (Table 7)
[6] See appendix for STATA code example

Combined Forecast:

The combined forecast model uses the statistical power of multiple forecasters to find the forecasted values. The process for creating a combined forecast is as follows:

Step 1:

Use the chosen forecasters as independent variables in a regression where sales serves as the dependent variable. Observe Equation 3 below and Table 8 with regression results[7].

Variables:

Sales: Sales data

TD: Time decomposition model forecast

Winters: Winters smother model forecast

AR: Auto Regressive model forecast

Equation 3: $sales_t = \beta_o + \beta_1 TD_t + \beta_2 Winters_t + \beta_3 AR_t + \varepsilon_t$

Table 8		
Model & βo	**p-value**	**Significant**
TD	0.031	Yes
Winters	0.834	No
AR	0.120	Yes
βo	0.907	No

The model is evaluated based upon the significance of the constant term (β_o). If the constant term is insignificant, then our underlying forecasters are unbiased. The results in Table 8 confirm that the constant term is insignificant.

Step 2:

Once the underlying forecasters are determined to be unbiased, the model is run again but without the constant term, observe Equation 4. Whichever forecasters remain insignificant are removed from the model, observe Table 9.

Equation 4: $sales_t = \beta_1 TD_t + \beta_2 Winters_t + \beta_3 AR_t + \varepsilon_t$

Table 9	
Model	**p-value**
TD	0.027
Winters	0.764
AR	0.111

I evaluated this model at the 15% level rather than 5%. Therefore, the Winters model is removed, but the Time Decomposition and Auto Regressive models are kept. The AR is significant at the 15% level; it was not removed because while it is not significant at the 5% level, the fact it is close to being significant at the 10%

[7] Results significant at the 5% level

level indicates it has statistical power. Eliminating the AR model would eliminate useful information for my combined forecast.

Step 3:

The significant forecasters (based upon Table 9), that will be included in the model are the Time Decomposition forecaster and the Auto Regressive forecaster. Observe Model 4.

Model 4: $sales_t = \beta_1 TD_t + \beta_2 AR_t + \varepsilon_t$

The predicted fitted values of Model 4 serve as the forecasted values for the combined forecast.

Other Contributing Factors to Selection:

The seasonal nature of the data immediately eliminated using an ARIMA model. I could have implemented a SARIMA model, but because I have never implemented a SARIMA before I decided to use a different model. While I did use regression in the AR and combined forecast methods, other regression techniques such as an ADL model or Causal Regression model were not used because of my lack of independent variables. I only had one series of data; therefore, I chose models that could handle seasonal and stationary data, but only required one series of information. Of the models we have studied this year, the three that fit the criterion are the Winters Smoother, Time Decomposition, and Auto Regressive.

Appendix

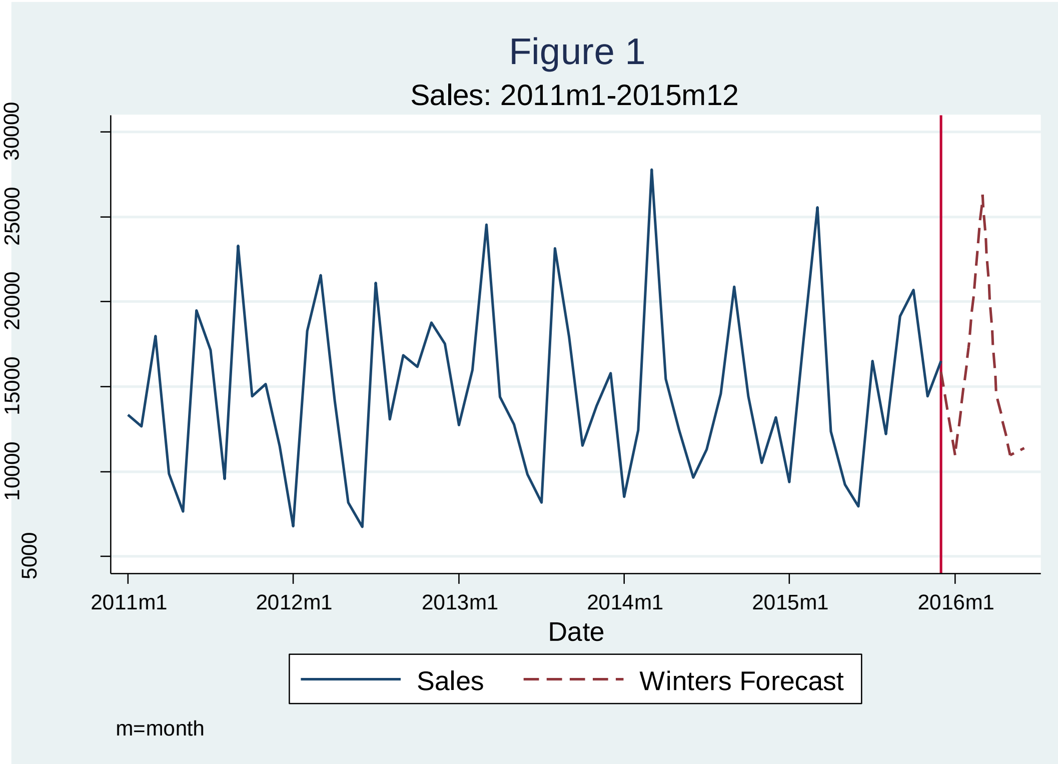


Figure 1
Sales: 2011m1-2015m12
30000
25000
20000
15000
10000
5000
2011m1
2012m1
2013m1
2014m1
2015m1
2016m1
Date
Sales
Winters Forecast
m=month

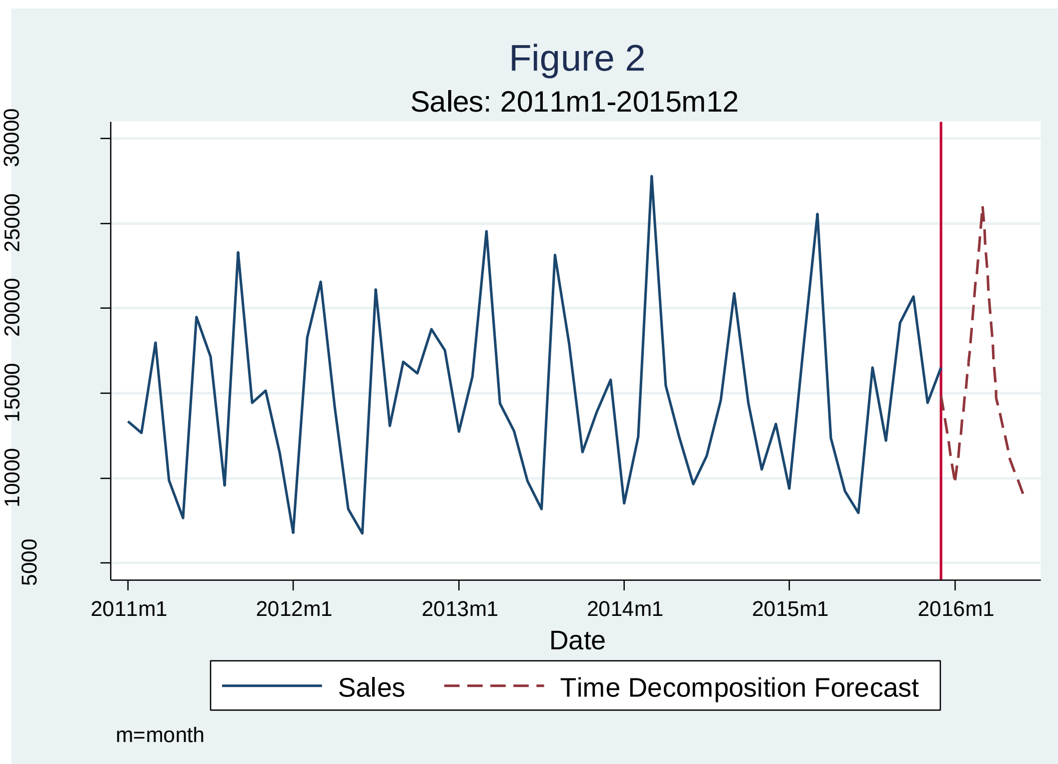


Figure 2
Sales: 2011m1-2015m12
30000
25000
20000
15000
10000
5000
2011m1
2012m1
2013m1
2014m1
2015m1
2016m1
Date
Sales
Time Decomposition Forecast
m=month

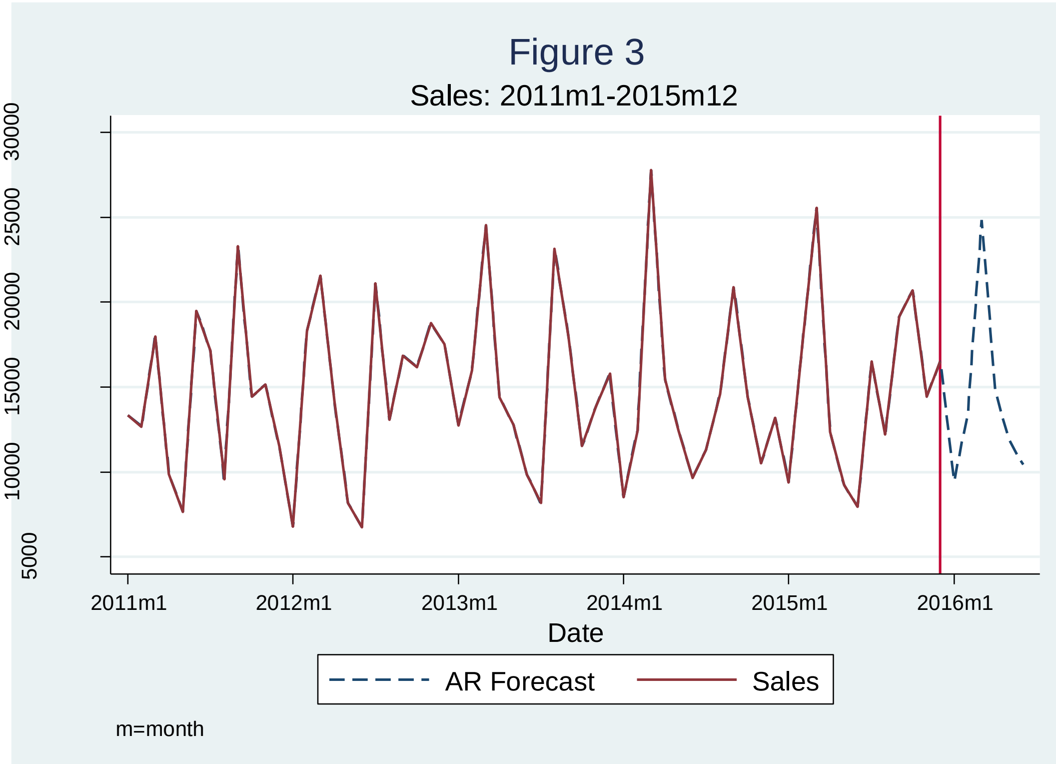


Figure 3
Sales: 2011m1-2015m12
30000
25000
20000
15000
10000
5000
2011m1
2012m1
2013m1
2014m1
2015m1
2016m1
Date
AR Forecast
Sales
m=month

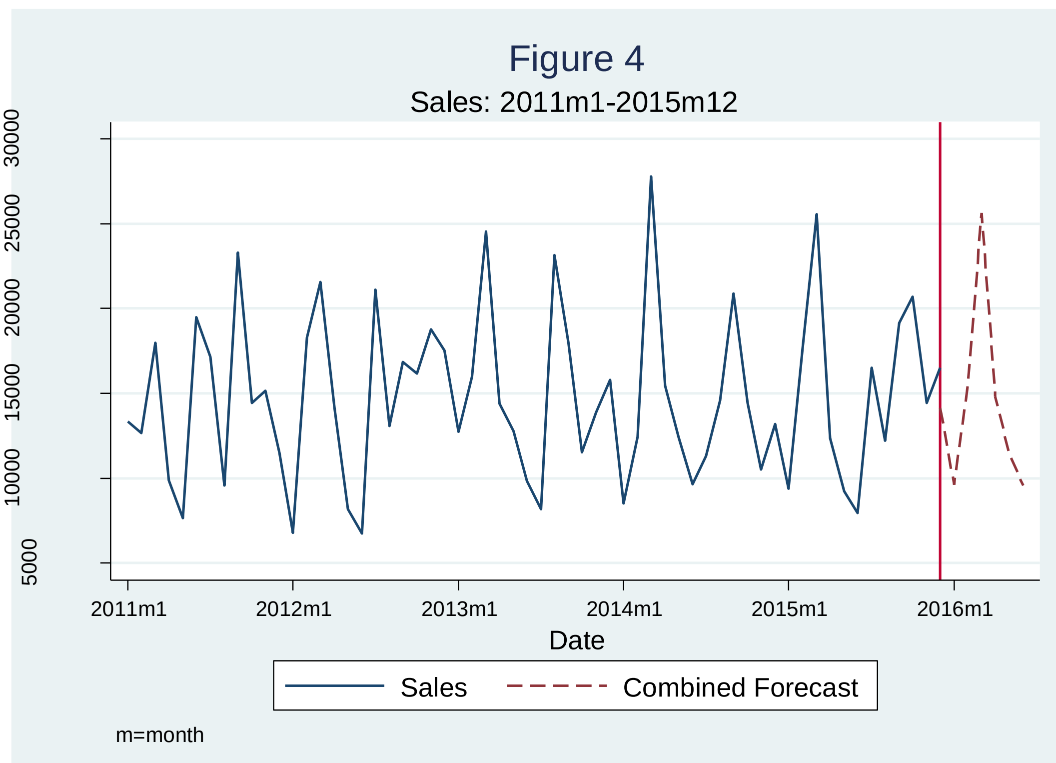


Figure 4
Sales: 2011m1-2015m12
30000
25000
20000
15000
10000
5000
2011m1
2012m1
2013m1
2014m1
2015m1
2016m1
Date
Sales
Combined Forecast
m=month

Test 1:

Augment Dickey-Fuller:

H_o: δ=0 , stochastic trend: non-stationary

H_a: δ<0 , no stochastic trend: stationary

Equation 1: $cma_t = \beta_o + \beta_1 \, Time_t + \varepsilon_t$

Table 7			
Model	**Significant Terms**	**AIC**	**BIC**
AR(2, 4) m1-m11	lag 2, lag 4, m1, m3, m6, m9	1063.6	1069.6
AR(2, 4) m1 m3 m6 m9	lag 2, lag 4, m1, m3, m6, m9	1035.1	1063.3

AR Forecast STATA code example

```
/*Forecast 1*/
reg sales l2.sales l4.sales m1 m3 m6 m9
predict ar_f1
gen ar_forecast=sales
replace ar_forecast=ar_f1 if date==m(2016m1)

/*Forecast 2*/
reg sales l2.ar_forecast l4.ar_forecast m1 m3 m6 m9 if date<m(2016m1)
predict ar_f2
replace ar_forecast=ar_f2 if date==m(2016m2)
```

*process if repeated for all six forecast periods