Forecasting Pieces of Introduced Legislation in the United States Congress

By: Daston Arman (020597DAA1)

Submitted: Dec 3rd, 2018

Class: BA-BHAAV6008U Forecasting in Business and Economics

Teacher: Natalia Khorunzhina

1 Introduction

The legislative body of the United States federal government is the U.S. Congress. The 538 total members have, at minimum, a two-year term to follow through on campaign promises and enact change through legislation. An important measure of the "productivity" of a congress is the amount of new legislation that is introduced by congressional members. This paper will forecast monthly pieces of introduced legislation by the U.S. Congress.

The ability for a country to legislate effectively has consequences for the country's citizens, its economy, and the outside world. As times change, regulations must change as well. Old government programs like Social Security constantly need to be re-visited while new problems like the Opioid drug epidemic have to be fought in original ways. Effective legislation is critical for democracy to work and is of *prima facie* importance to a nation.

While the real goal for a legislative body is effective governance, that goal can be split up into two parts, writing legislation and making sure that it was the right legislation. The latter goal is massively subjective and political, but the former is much easier to quantity.

The U.S. Congress is split up into two houses, the Senate and the House of Representatives. The Senate has less members and six-year terms while the House has more members with two-year terms. With elections every two years, the whole House is always up for re-election while about 1/3 of the Senate is on the ballot. Because of this, every two years results in a new U.S. Congress. The last Congress was the 115$^{\text{th}}$ Congress which met from January 2017 to December 2019.

The essential parts of the legislative process are not very complex. A congressional member from either the Senate or the House "introduces" a piece of legislation by simply submitting the bill to everybody else. This is a piece of introduced legislation. The magic of law-making pursues, many different votes occur, and if the proper majorities are reached, then the bill becomes an "enacted" piece of legislation.

This paper will focus on introduced pieces of legislation rather than enacted pieces for a few reasons. First, partisanship is at an all-time high in the U.S. which limits the amount of legislation that can pass through any recent Congress for various political reasons. Second, divided governments can't always pass legislation, but any congressional member can always introduce legislation no matter their party. Third, introduced legislation is more a measure of the ideas and work that congressional members are putting in than enacted legisalation is, which is muddied by many other factors.

This paper will model the trend, seasonality, and cyclicality in the time series of pieces of introduced legislation. The trend component will be quadratic, the seasonality will be handled with dummy variables, and the cyclicality will be modelled with an ARMA(3, 1) process. Finally, an in-sample and out-of-sample forecast will be constructed.

2 Data

2.1 Data Source

Going back to the 93rd Congress, which started serving in January 1973, data for pieces of legislation has been available online through government websites. Data before this has either not been kept at all, been stored in physical documents, or is low quality. Only since 1973 has all the data needed for this series been available.

Unfortunately, this data has not been made available in easy-to-use formats such as a csv of all pieces of legislation with their date of introduction. Because of this, the time series data for this paper was scraped from government websites using a Python script. The script simply recorded the pieces of legislation and their introduction date. This was then connected to the correct Congress and aggregated on a monthly basis.

In all, data is available for 23 different Congressional meetings (93rd Congress to the 115th Congress) ranging from 1973 to 2018. Each Congress is technically in power for 24 full months, but with elections occurring in November of their last year, there is effectively only 22 months a Congress has to introduce legislation.

As a result, this dataset contains 506 data points with a frequency of 22 data points per season. Each data point represents the number of introduced pieces of legislation during a single calendar month starting from January and ending on October of the following year. The dataset is available for download at https://github.com/dastonarman.

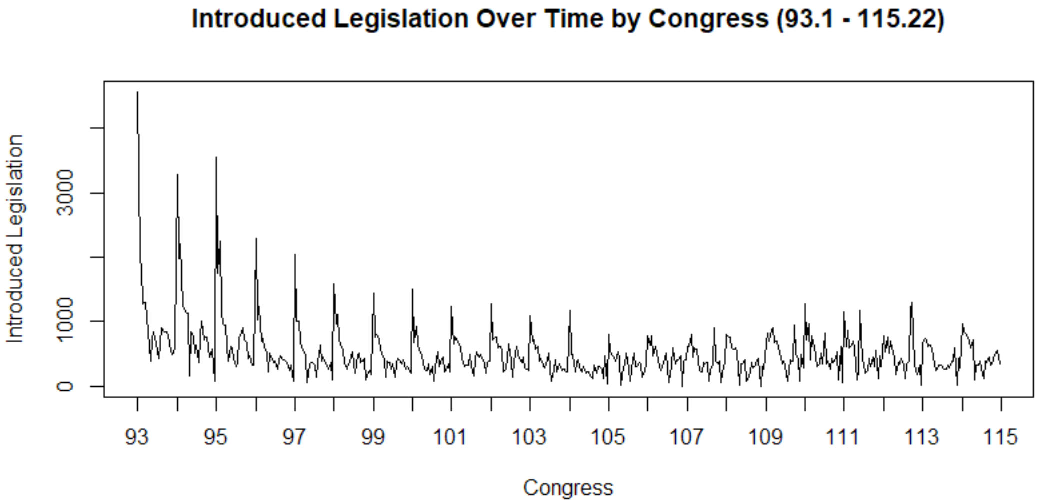
2.2 Data Summary

Below is a table of summary statistics for the dataset which is followed by a time series plot.

Table 1

Min.	1st Q.	Median	3rd Q.	Max.	Mean	S.D.
1	306	412	522	4556	522	425

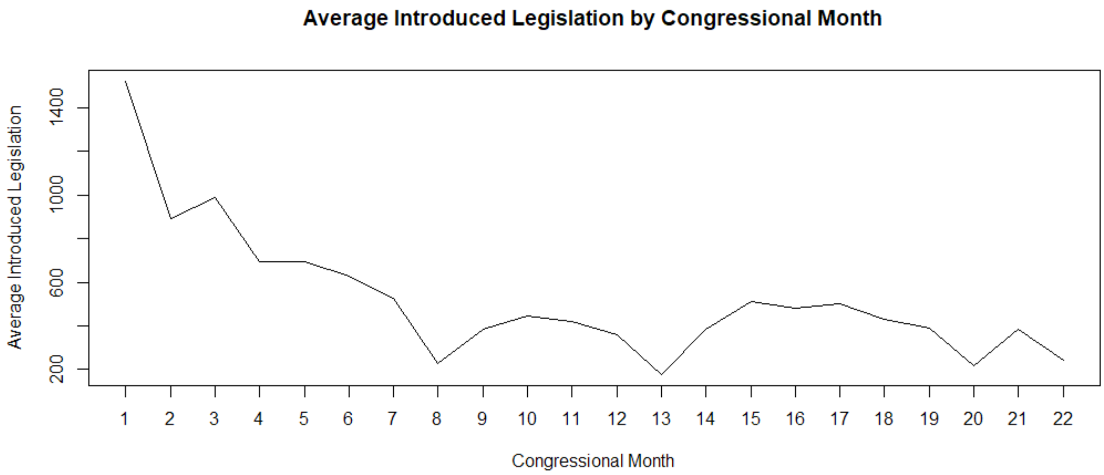
Figure 1



There are a few things to note from this simple plot. First, it seems that Congress has become less productive over time. The most productive Congress in the dataset is also the first, the 93rd Congress starting in 1973. There is a clear decline in introduced legislation until around the 105th Congress starting in 1997. This fits with popular beliefs that Congressional members have been less productive and less effective in recent times.

Second, seasonality is very evident. The 22-month cycle of a Congress is a heavily calculated political object where politicians think critically about when to introduce certain pieces of legislation so it is natural to have seasonality. Below is a plot of the average introduced legislation by month.

Figure 2



The first month is by far the most productive until decreasing to a stable minimum after about eight months.

Finally, variance is very uneven. More productive Congress' have higher variance than less productive Congress'. Because of this, a log conversion of the data is necessary.

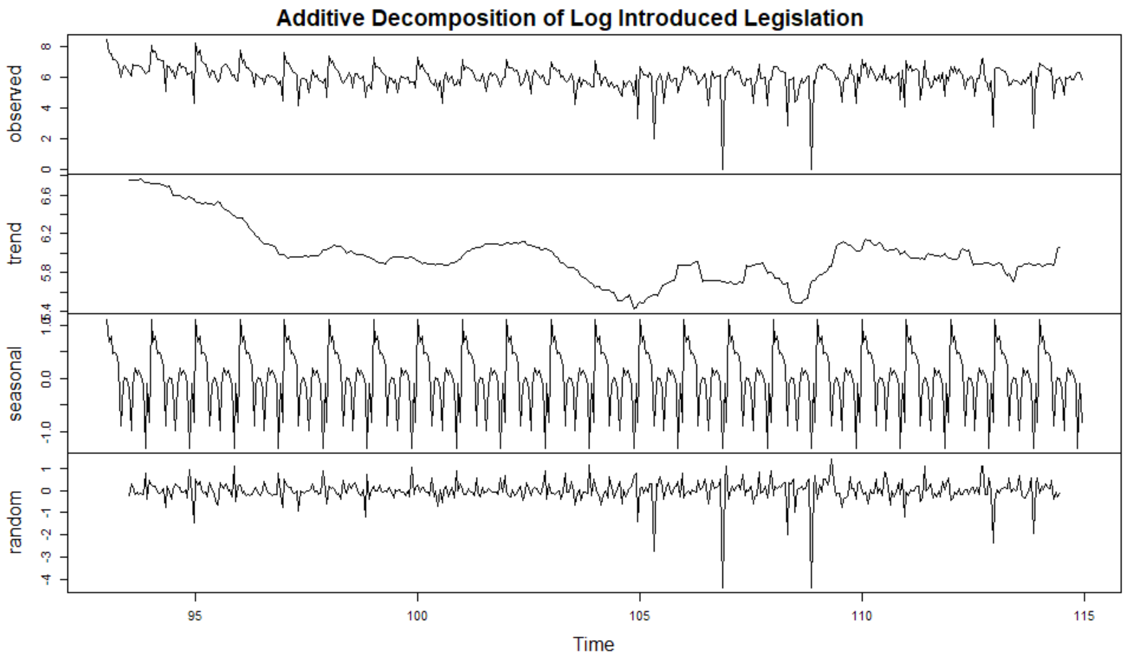
3 Model

In this section, a quadratic trend and seasonality model with an ARMA(3, 1) process for cyclicality is proposed to forecast the time series. The model structure is shown below.

$$\ln y_t = T_t + S_t + C_t + \varepsilon_t$$

The trend, seasonality, and cyclicality will each be investigated seperately in this section. Below is an additive decomposition of the log time series. Unless mentioned, the data will be in logs from here on out.

Figure 3



By taking the log of the time series, variance has reduced considerably and become much more stationary. More formal tests for stationarity will be conducted in a later section. One problem with working in logs is that months with only a handful of pieces of introduced legislation are stretched way down in the graph to near zero. Several of these downward spikes are seen in the random component above. Other than reducing measures like R^2 by a small amount, these spikes will not be too harmful to our model.

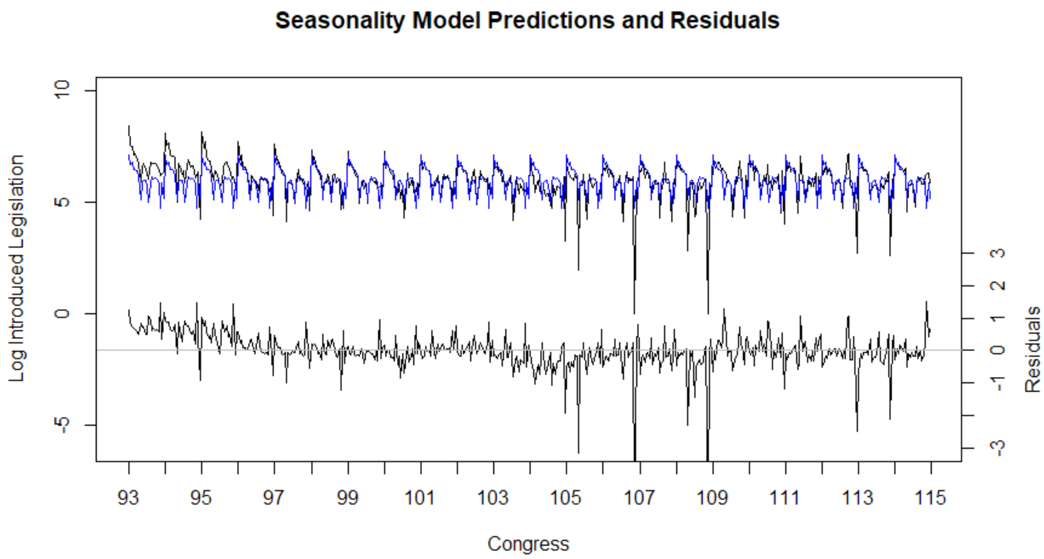
3.1 Seasonality

The first step in building our model is to account for seasonality. As Figure 2 showed, seasonality is strong and needs to be accounted for linearly in logs. The proposed seasonality component of the model is written below.

$$S_t = \beta_0 + \sum_{j=1}^{21} \delta_j D_{jt}$$

The full model output is provided in Appendix 1 and a plot of model performance is show below.

Figure 4



All but 2 of the 22 estimated parameters are significant in the model. With an R^2 of 48.9%, the fit with just seasonality is fairly adequate, but there is still much left over in the residuals. A downward curve in the residuals is seen which implies a quadratic trend component would be appropriate.

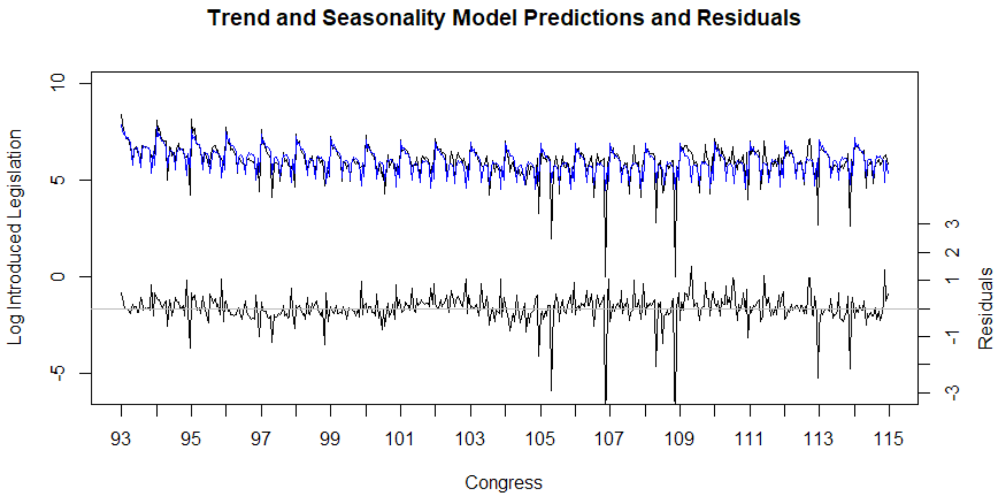
3.2 Trend

With seasonality adjusted for, the trend component of the model will be modelled as below.

$$T_t = \beta_1 TIME_t + \beta_2 TIME_t^2$$

An addition of just the linear trend component produces a significant coefficient of β_1 and an increase of R^2 to 53.2%. But, an addition of the quadratic component produces a significant β_2 and an even higher increase of R^2 to 58.9%. Because of this, the quadratic trend was chosen. No higher-order polynomials were considered for parsimony. The estimation of this model is included in Appendix 2.

Figure 5



The trend and seaonality visually fits the time series even more well, but there is still a systematic component left in the residuals. Visually, there are wave-like patterns through the residuals series that are forecastable. With a Durbin-Watson statistic of 1.77 and a p-value of 0.6%, it is true that there is some sort of ARMA process that would help make the leftover residuals into unforecastable white noise. Test results are in Appendix 3.

Figure 6

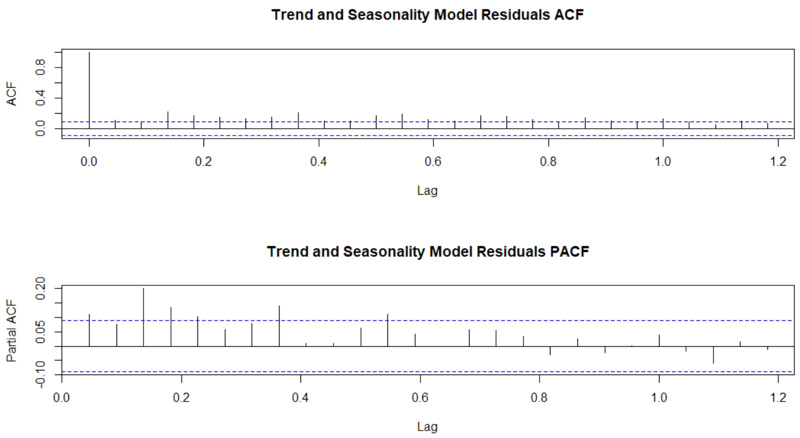


Figure 6 above shows the ACF and PACF of the residuals from the trend and seasonality model. Both the ACF and the PACF have a few significant lags, but it is not clear which ARMA process would fit the residuals the best.

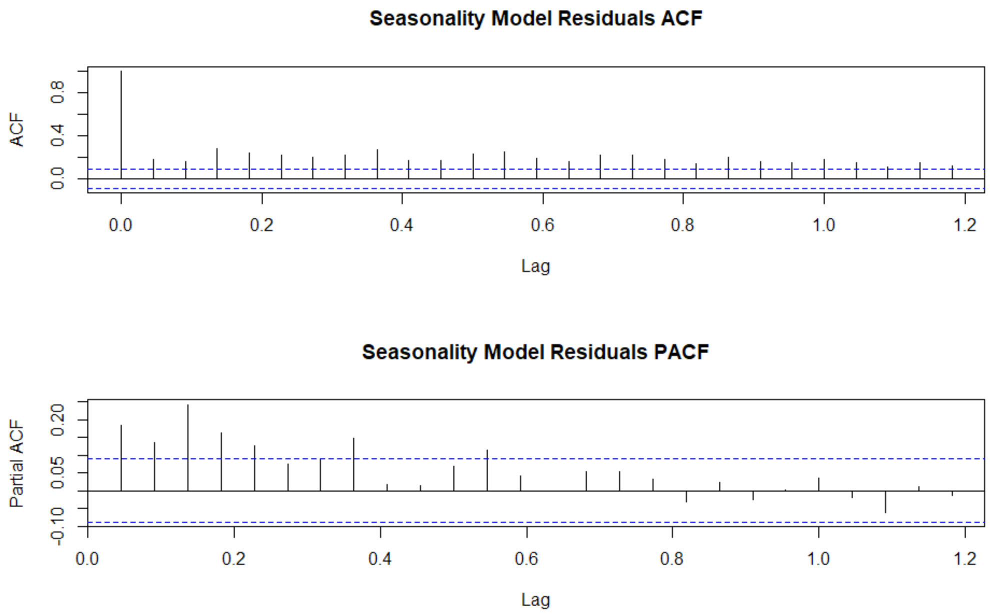
3.3 Unit Roots

Before adding the cyclical component of the model, this section will take a moment to look for stochastic trend in the time series. The previous section assumed trend was deterministic, but in reality the trend could be better modelled in first differences if the process is a random walk.

Visually, the common signs of stochastic trend are permanent shocks to the time series having a permanent effect. As seen in Figure 3, this does not seem to be the case. The de-seasonalized time series is mostly smooth with no large, permanent shocks.

Another sign of stochastic trend is an ACF that fails to damped with a PACF that instantly dampens after the first lag. Figure 7 below shows the ACF and PACF of the deseasonalized model *before* quadratic trend was added to the model.

Figure 7



The ACF dampens fairly quickly while the PACF does not instantly dampen after the first lag.

The final test for stochastic trend is the Dickey-Fuller test against the null hypothesis that there is a unit root. The test produces a test statistic of -4.0 which is significant to beyond the 0.1% percentile. This test is evidence against the null hypothesis that there is a unit root. Test results are in Appendix 4.

Because of looking at the time series visually, looking at the ACF and PACF, and because of the Dickey-Fuller test, the trend in this time series is concluded to be deterministic. The quadratic trend component will not change.

3.4 Cyclicality

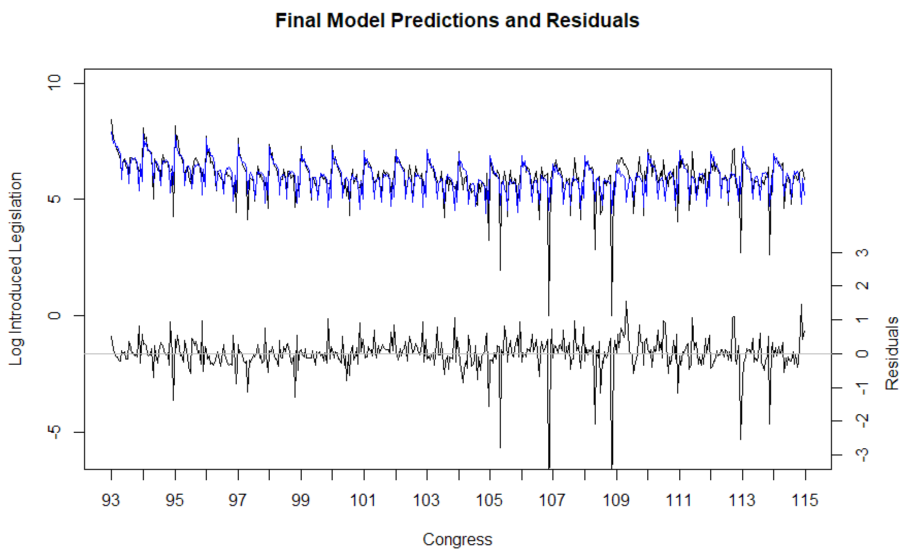
The final aspect of the model to include is the cyclical component. As previously stated, an ARMA process of some order will be chosen with the form as below, where ỹ is detreneded and deseasonalized y.

$$C_t = \sum_{i=1}^{p} \rho_i \tilde{y}_{t-i} + \sum_{k=1}^{q} \theta_k \varepsilon_{t-k}$$

The procedure to determine the best order ARMA process is to fit all ARMA model with order of p and q less than 10, and to then choose the model that minimizes the Akaike Information Criterion (AIC). AIC is a measure of goodness of fit that automatically accounts for parsimony. This procedure suggests that an ARMA(3, 1) process is the correct model for the residuals from the trend and seasonality model since its AIC of 800.2 is the lowest of all tested models.

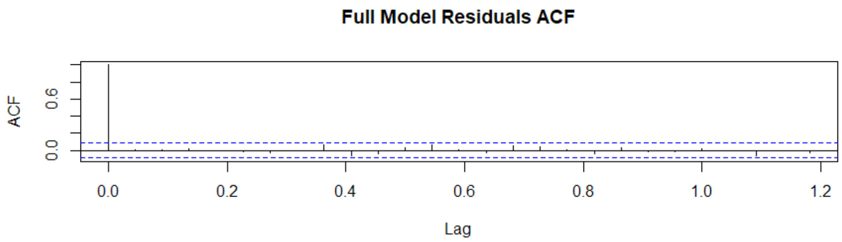
The residuals now have a Durbin-Watson statistic of 1.99 which is highly insignificant. This means that the residuals are not serially-correlated and unforecastable. The figure below shows the final model's performance.

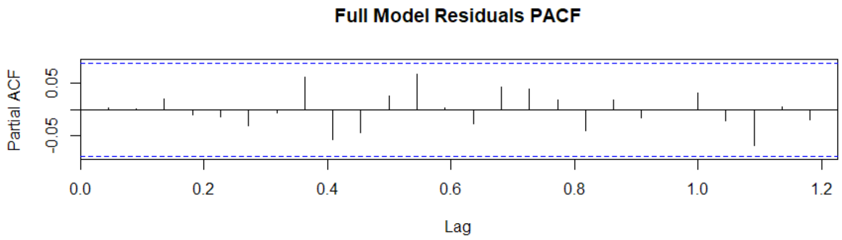
Figure 8



Visually, the residuals have lost much of the wavelike pattern seen in the trend and seasonality model. Besides the few large dips, which are a product of the log scale, the residuals look much more like white noise.

Figure 9





The final check is for the ACF and PACF of the full model. As compared to Figure 7, the ACF and PACF both show no correlations. In fact, none of the first 26 lags are significant for either the ACF or the PACF. This is further proof that the ARMA(3, 1) model for the remaining residuals was a useful addition to the full forecasting model.

4 Results

Full model specifications are reported in Appendix 5, but key results are included here.

4.1 Model Coefficients

Table 2

Coefficient	Estimate	S.E.	Significance
AR1	0.7948	0.0945	***
AR2	-0.0229	0.0577	
AR3	0.1245	0.0465	***
MA1	-0.8243	0.0857	***
TIME	-0.0069	0.0009	***
$TIME^2$	0.0012	0.0003	***

The largest AR coefficient is AR1 which has a positive sign. AR2 is not significant, but is included so AR3 can be included which is also positive. This reflects the fact that strucutural fluctuations in congress persist over time. This could mean that a certain congress is consistently different than just what trend and seasonality could predict. As well, this could be showing the effects that certain long-serving congressional members can have on the decades they are in power.

The sign on the MA1 coefficient is negative which means that the error factor after adjusting for structural factors tends to alternate. If it so happens that one month is a little lower on introduced legislation after accounting for relevant factors, then it is likely that the next month will have more legislation introduced. This makes sense as sometimes bills just don't get submitted in time.

The signs on the trend coefficients suggest a downward curing parabola which is indeed what was seen in Figure 4. The signs and weights on the coefficients make sense when compared with prior knowledge.

4.2 Goodness of Fit

Various goodness of fit measures are included below.

Table 3

AIC	S.E.	Training RMSE	Training MAE
800.2	0.52	0.52	0.31

A majority of values on the log scale fall between 4 and 8 so a residual standand error of 0.52 is a significant reduction in variance. As well, the training RMSE is almost equal to the in-sample residual standard error which is one sign that the model has not overfit the sample too much. If more random noise was memorized by the model, then one would expect training RMSE to be significantly lower than residual standard error.

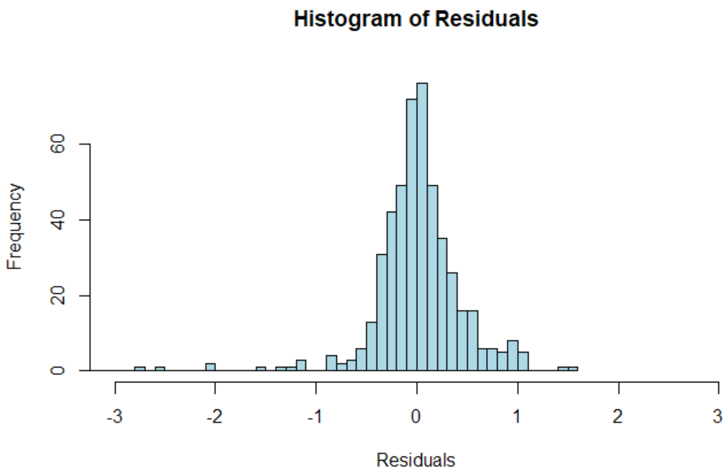
5 Forecast

With the full forecasting model developed, the final step is to perform an in-sample forecast to check for validity and then to extrapolate to an out-of-sample forecast for future use.

5.1 In-Sample Forecast

For simplicity in confidence intervals, residuals will need to be normally-distributed. The below figure is histogram of residuals from the full model.

Figure 10



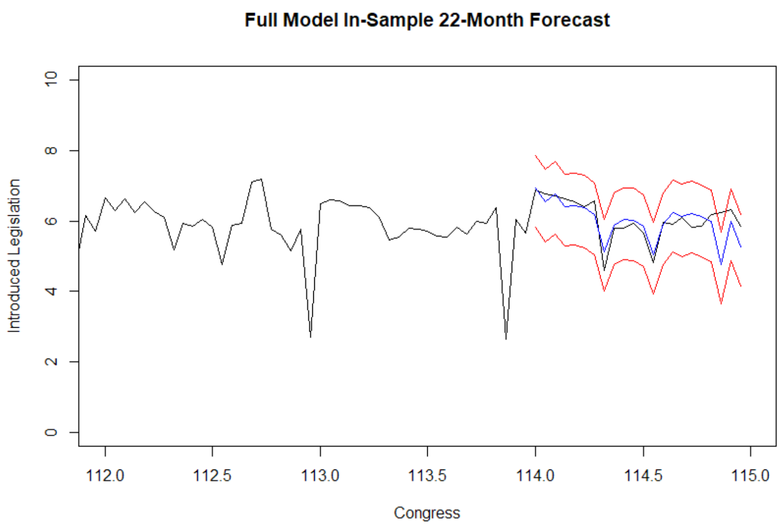
Visually, the histogram looks approximately normally distributed. For more assurance, the often-used Kolmogorov-Smirnov test is applied.

Test Statistic (D)	P-Value
0.27034	2.2e-16

Since the test concludes that the residuals are not likely to be distributed normally, normality assumptions for confidence intervals cannot be made. Instead, 95% confidence intervals will be made through a 95% coverage interval of the residuals (-1.1, 0.9).

The below figure is a 22-month in-sample forecast.

Figure 11



The forecast fits very well with the real time series. There is only one month that falls ourside the 95% confidence interval. Still, Figure 11 is on the log scale while we really care about the forecast in levels. Figure 12 is the same graph in levels instead.

Figure 12

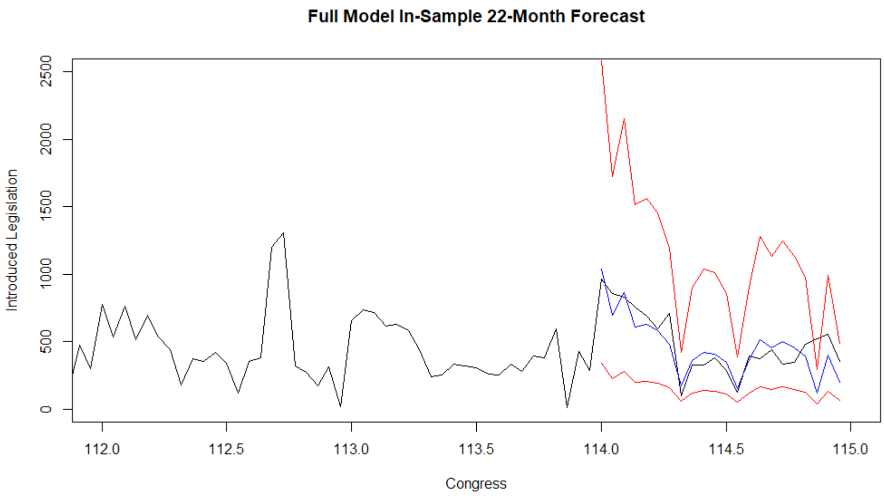


Figure 12 reveals how changing to levels makes the forecast unsymmetric. The uncertainty on the high end of predictions is much more than the lower end. The most prominent example of this is the first month of legislation which normally has a huge spike. This is an un-avoidable part of working in logs to build a model.

Finally, a 44-month out-of-sample forecast is presented both in logs (Figure 13) and in levels (Figure 14).

Figure 13

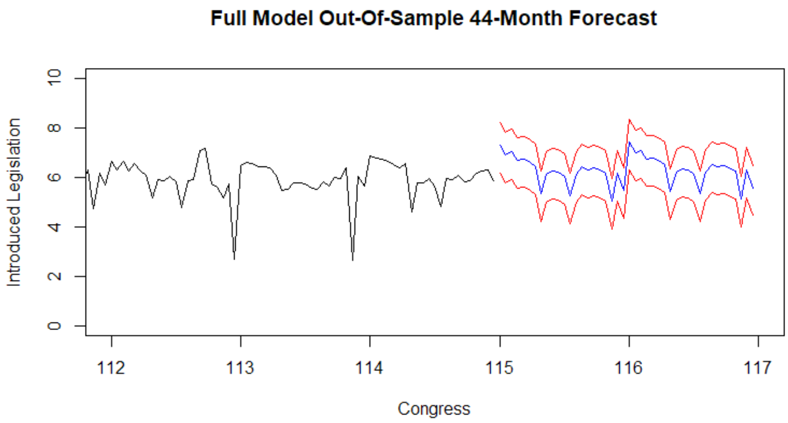
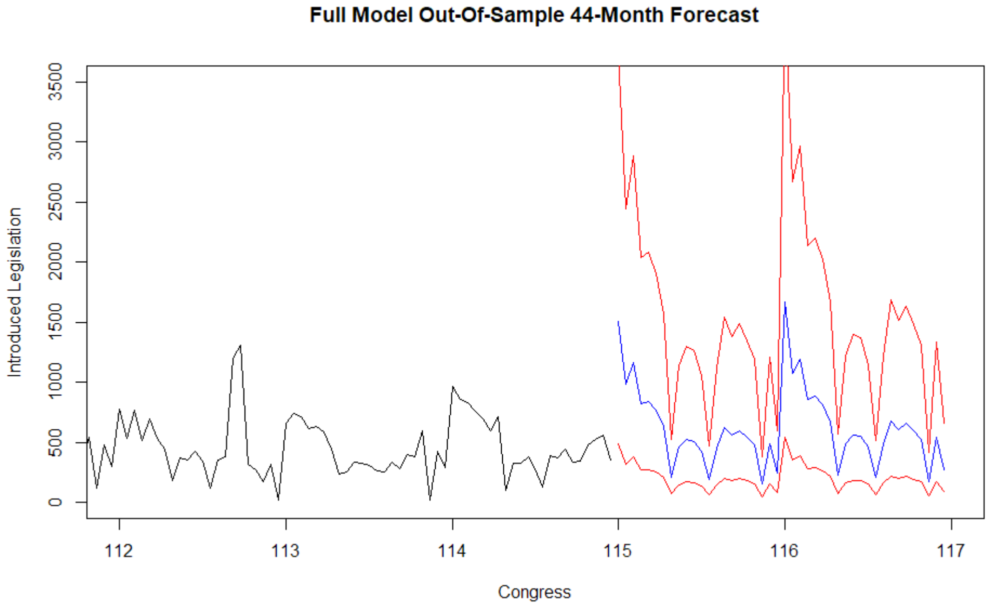


Figure 14



6 References

Data was scraped from https://www.congress.gov/. All other work is original with help from lecture notes and teacher-provided sample code.

R-code and the dataset can be found at https://github.com/dastonarman.

Appendix

Appendix 1: Seasonal Model Output

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.19781    0.12874  40.375  < 2e-16 ***
MS1          1.95987    0.18206  10.765  < 2e-16 ***
MS2          1.50946    0.18206   8.291 1.24e-15 ***
MS3          1.60936    0.18206   8.840  < 2e-16 ***
MS4          1.27547    0.18206   7.006 8.74e-12 ***
MS5          1.30130    0.18206   7.147 3.47e-12 ***
MS6          1.20917    0.18206   6.641 8.75e-11 ***
MS7          1.01282    0.18206   5.563 4.50e-08 ***
MS8         -0.07859    0.18206  -0.432 0.666204    
MS9          0.68509    0.18206   3.763 0.000190 ***
MS10         0.81597    0.18206   4.482 9.35e-06 ***
MS11         0.79110    0.18206   4.345 1.71e-05 ***
MS12         0.60458    0.18206   3.321 0.000969 ***
MS13        -0.20203    0.18206  -1.110 0.267720    
MS14         0.65570    0.18206   3.601 0.000351 ***
MS15         0.97455    0.18206   5.353 1.37e-07 ***
MS16         0.86039    0.18206   4.726 3.05e-06 ***
MS17         0.93040    0.18206   5.110 4.72e-07 ***
MS18         0.82883    0.18206   4.552 6.79e-06 ***
MS19         0.69881    0.18206   3.838 0.000141 ***
MS20        -0.45385    0.18206  -2.493 0.013024 *  
MS21         0.70854    0.18206   3.892 0.000114 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 
1

Residual standard error: 0.6038 on 462 degrees of freedom
Multiple R-squared:  0.4889,	Adjusted R-squared:  0.4657 
F-statistic: 21.04 on 21 and 462 DF,  p-value: < 2.2e-16
```

Appendix 2: Trend and Seasonal Model Output

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.951e+00  1.358e-01  43.804  < 2e-16 ***
t           -6.682e-03  7.076e-04  -9.443  < 2e-16 ***
t2           1.123e-05  1.413e-06   7.950 1.46e-14 ***
MS1          1.934e+00  1.637e-01  11.817  < 2e-16 ***
MS2          1.485e+00  1.637e-01   9.074  < 2e-16 ***
MS3          1.586e+00  1.637e-01   9.693  < 2e-16 ***
MS4          1.254e+00  1.637e-01   7.662 1.10e-13 ***
MS5          1.281e+00  1.636e-01   7.828 3.44e-14 ***
MS6          1.190e+00  1.636e-01   7.274 1.52e-12 ***
MS7          9.953e-01  1.636e-01   6.082 2.50e-09 ***
MS8         -9.478e-02  1.636e-01  -0.579 0.562755    
MS9          6.702e-01  1.636e-01   4.096 4.97e-05 ***
MS10         8.024e-01  1.636e-01   4.903 1.31e-06 ***
MS11         7.787e-01  1.636e-01   4.759 2.61e-06 ***
MS12         5.935e-01  1.636e-01   3.627 0.000319 ***
MS13        -2.119e-01  1.636e-01  -1.295 0.195897    
MS14         6.470e-01  1.636e-01   3.954 8.90e-05 ***
MS15         9.670e-01  1.636e-01   5.910 6.67e-09 ***
MS16         8.540e-01  1.636e-01   5.219 2.73e-07 ***
MS17         9.251e-01  1.636e-01   5.654 2.76e-08 ***
MS18         8.247e-01  1.636e-01   5.040 6.71e-07 ***
MS19         6.957e-01  1.636e-01   4.252 2.57e-05 ***
MS20        -4.559e-01  1.636e-01  -2.786 0.005552 ** 
MS21         7.075e-01  1.636e-01   4.324 1.88e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 
1

Residual standard error: 0.5427 on 460 degrees of freedom
Multiple R-squared:  0.589,	Adjusted R-squared:  0.5684 
F-statistic: 28.66 on 23 and 460 DF,  p-value: < 2.2e-16
```

Appendix 3

```
        Durbin-Watson test

data:  trend_seasonal_model
DW = 1.7685, p-value = 0.006295
alternative hypothesis: true autocorrelation is greater than
0
```

Appendix 4

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-4.6119 -0.1974  0.0062  0.2402  1.6766

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0656858  0.0577049   1.138 0.255582
z.lag.1     -0.3650146  0.0910970  -4.007 7.17e-05 ***
tt          -0.0002967  0.0002139  -1.387 0.165967
z.diff.lag1 -0.6215890  0.0925012  -6.720 5.36e-11 ***
z.diff.lag2 -0.6211134  0.0926175  -6.706 5.83e-11 ***
z.diff.lag3 -0.4699385  0.0918868  -5.114 4.62e-07 ***
z.diff.lag4 -0.3666953  0.0876866  -4.182 3.46e-05 ***
z.diff.lag5 -0.2817785  0.0797933  -3.531 0.000455 ***
z.diff.lag6 -0.2222897  0.0654530  -3.396 0.000742 ***
z.diff.lag7 -0.1439871  0.0466312  -3.088 0.002138 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
1

Residual standard error: 0.539 on 463 degrees of freedom
Multiple R-squared:  0.5058,   Adjusted R-squared:  0.4961
F-statistic: 52.64 on 9 and 463 DF,  p-value: < 2.2e-16


Value of test-statistic is: -4.0069 5.5972 8.3926

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

Appendix 5

```
Coefficients:
         ar1      ar2     ar3      ma1  (Intercept)        t    t2
      0.7948  -0.0229  0.1245  -0.8243       7.8990  -0.0069 0.0012
s.e.  0.0945   0.0577  0.0465   0.0857       0.1536   0.0009 0.0001
           S2       S3       S4       S5       S6       S7       S8
      -0.4482  -0.3443  -0.6776  -0.6509  -0.7419  -0.9370  -2.0272
s.e.   0.1601   0.1618   0.1511   0.1528   0.1546   0.1545   0.1546
           S9      S10      S11      S12      S13      S14      S15
      -1.2624  -1.1303  -1.1541  -1.3395  -2.1449  -1.2860  -0.9661
s.e.   0.1549   0.1551   0.1551   0.1552   0.1551   0.1551   0.1549
          S16      S17      S18      S19      S20      S21      S22
      -1.0791  -1.0080  -1.1086  -1.2371  -2.3880  -1.2279  -1.9361
s.e.   0.1546   0.1545   0.1545   0.1527   0.1509   0.1612   0.1595

sigma^2 estimated as 0.2713:  log likelihood = -371.1,  aic = 800.2
1

Training set error measures:
                       ME    RMSE       MAE       MASE       ACF1
Training set -0.000622767 0.52082 0.3022285  0.5240371  0.0028757
```

Appendix 6

```
        One-sample Kolmogorov-Smirnov test

data:  residuals
D = 0.27034, p-value < 2.2e-16
alternative hypothesis: two-sided
```