# Two Topics in Applying Seasonal Adjustment Diagnostics On Large Sets of Series December 2024

# Brian C. Monsell

U.S. Bureau of Labor Statistics and NSR Solutions, Inc.

#### Abstract

This paper highlights two issues that came up in recent work on a large set of employment series from the Current Employment Statistics program in the Bureau of Labor Statistics. The first examines performing hypotheses tests on such a large set of series. As seasonal adjustment practitioners move from empirical to model-based diagnostics, care needs to be taken to mitigate the effects of multiple hypothesis testing on this set of 4700+ series. The second topic involves using the model-based seasonal F-test to determine if a series is seasonal. Since models used in seasonal adjustment production rarely include seasonal regressors, a second X-13 run is often needed to generate this test. This work examines two methods for generating the model for this second run, and what difference this makes to the results of the seasonal F-test.

**Keywords:** seasonality diagnostics, multiple hypothesis testing, seasonal diagnostics, automatic ARIMA model identification, automatic outlier identification

#### **1** Introduction

Over the last year, there has been an effort to improve the tools used by analysts at the U.S. Bureau of Labor Statistics (BLS) to evaluate seasonal adjustments for the Current Employment Statistics (CES) survey published monthly by BLS.

In the course of this work, which seeks to replace empirical seasonality diagnostics with model-based counterparts, two issues have come up which needed to be addressed.

An important issue came from generating a model-based F-statistic for stable seasonality. Lytras et al. (2007) compared a model-based F-statistic to existing seasonality diagnostics in X-12-ARIMA and showed that the model-based diagnostics was superior. The current version of X-13ARIMA-SEATS makes generating this statistic very easy when stable seasonal regressors are specified in the model. For a more detailed description of this diagnostic, see Bell et al. (2022), Section 4.1.

Unfortunately, stable seasonal regressors are not usually specified in model used in seasonal adjustment production work, so a second X-13ARIMA-SEATS run is needed to generate the model-based F-statistic. This paper will compare two methods of determining the model used for this test: automatic model identification with seasonal regressors included or revising the production model. I will give details of each method, and show results from each. Results from another model-based seasonality diagnostic will be generated to compare with these results.

Another issue is the large number of series published each month by CES - over 4,000 employment statistics for different types of employment, hours worked, etc. Performing hypothesis tests for such a large set of series poses the risk of accepting false results.

In this talk, I'll document my attempts to control the false detection rate using methods available within R (see R Core Team (2024)).

#### 2 Specifying the Model for the Model-Based F-test

As mentioned in Section 1, I'll use two methods of specifying the model for the modelbased F-test - generating a model using the automatic model identification procedure within X-13ARIMA-SEATS, and modifying the model currently used for production.

I'll give more details on each in the next few sections.

#### 2.1 Modifying production model

In this method, the ARIMA model used in production is modified for the seasonal test by removing any seasonal differencing. If there is no seasonal differencing, the model is specified as it is in production.

By removing the seasonal difference, we can then add stable seasonal regressors to the regARIMA model, leaving all other regressors specified the same.

In the example below, we show the regression and arima specs from a production X-13 spec file that uses the  $(0\ 1\ 1)(0\ 1\ 1)$  model (the airline model of Box and Jenkins):

```
regression{
    variables = (
        tc2019.10
    )
    user = (dum1 dum2 dum3 dum4 dum5 dum6
            dum7 dum8 dum9 dum10 dum11)
    start = 1986.01
    usertype = td
    file = '..\..\ces_reg\FDUM8606.dat'
    save = (td ao ls tc)
    format = "free"
}
arima{
    model = (0 1 1)(0 1 1)
}
```

I used a special version of X-13-SAM (see Lytras 2024) to change each of the production models to

- set the seasonal difference to zero,
- add seasonal regressors to the regression spec.

The resulting spec file would have these modeling specs:

This method will be labelled as the P0Q method going forward in this paper.

## 2.2 Automatic model identification

In this method, we'll use the automdl spec to generate the model used for generating the model-based F-test. The seasonal regressors will be preset in the model, which means that we should set the options for automdl carefully so that the final model will not have a seasonal difference. Having seasonal regressors in a model with seasonal differencing would cause an estimation error, and the seasonal differencing will eliminate the seasonal regressors.

The maxdiff argument of the automdl spec is used to set the maximum order of seasonal differencing to zero.

```
automdl{
    maxorder = (3 1)
    maxdiff = (1 0)
    savelog = amd
}
```

We then add the seasonal regressors to whatever user-defined regressors are used in the production model for the CES series. An example of this appears below:

```
regression{
   variables = seasonal
   user = (dum1 dum2 dum3 dum4 dum5 dum6 dum7 dum8 dum9
        dum10 dum11)
   start = 1986.01
   usertype = td
   file = '..\..\ces_reg\FDUM8606.dat'
   save = (td ao ls tc)
   format = "free"
}
```

At first, I ran the automatic outlier identification procedure as well as the automatic modeling procedure to generate the model used to get the F-statistic for the seasonal

regressors. In doing this, we encountered fatal model estimation errors for a number of series, though they may be small in number when considering the large number of CES series analyzed (over 4,000).

The main estimation errors I found were:

- The covariance matrix of the ARMA parameters is singular.
- Estimation failed to converge maximum iterations reached.
- Cannot compute outlier t-statistic for outlier backward deletion the residual root mean square error is zero.
- Adding an outlier exceeds the number of regression effects allowed in the model (80).

I attempted to mitigate this in two ways - first, I tried using the order of differencing from the production models in conjunction with automatic model identification. To implement this, I used the diff argument to set the order of nonseasonal differencing in the model to match that of the production model; the order of seasonal differencing was set to zero.

An example of this appears below:

```
automdl{
    maxorder = (3 1)
    savelog = amd
    diff = (1 0)
}
```

The second change incorporates the outliers from the production spec file. If there are user-defined regressors in the production model, they will be kept in the spec file for generating the model-based F-test.

I used different combinations of these options in this study. Table 1 shows these set of options along with the total number of estimation errors generated. A code for each set of options is also provided.

Table 1: Options used in automatic model identification, with associated estimation errors

Code	Outlier Option	Differencing Option	Fatal Errors
auto0	automatic identification	automatic identification	72
auto1	automatic identification	set from production model	70
auto2	set from production model	set from production model	65
auto3	set from production model	automatic identification	67

Using outliers from the production model seems to reduce the number of fatal estimation errors generated from the X-13ARIMA-SEATS automatic modeling run, with a smaller reduction when the nonseasonal differencing order is set. In both cases, the differences are small.

Note there are no estimation errors when running the P0Q method.

## **3 Methodology**

Production spec files from the CES series used in 2023 were edited using the X-13-SAM program to create the spec files needed to generate the model-based F-test. These spec files were the same provided by CES on the BLS website during 2023. They were broken up between different types of statistics collected and the timing with which the series were released.

CES series may be seasonally adjusted directly (by applying seasonal factors directly to the not seasonally adjusted series) or indirectly (through the aggregation process). There are two sets of series published by CES:

- The lowest level seasonally adjusted series published with first preliminary estimates are used for aggregating to higher levels. These series will be referred to as the First Closing in this paper.
- The series published after the release of first preliminary estimates are seasonally adjusted directly but are not used in aggregation. These series are noted as independently seasonally adjusted, because they are not used in aggregation. These series will be referred to as the Second Closing in this paper.

Note that in this study, I will only be examining series that are directly seasonally adjusted.

The majority of series released by CES are Second Closing series - there are 298 series in the First Closing set versus 3,730 series in the Second Closing.

For more information on CES Seasonal Adjustment, access the CES National Calculations page (see Bureau of Labor Statistics 2024).

Once the spec files are generated, they are read into R using the seasonal package (Sax 2018; Sax and Eddelbuettel 2018), which is an interface between R and the X-13ARIMA-SEATS seasonal adjustment program (see Time Series Software Group 2023). I created scripts that generated seas objects for each series for each method or set of options. The model-based F-tests can then be extracted from the seas objects for those series for whatever series where X-13ARIMA-SEATS runs without incident, including the p-values for the F-tests. The FDR adjustment is applied to the set of p-values once they all are generated, and summaries of the results for the different methods and options can be generated within R.

## 4 Multiple Hypothesis Testing

As mentioned in Section 1, there was a concern for how to perform hypothesis testing on such a large group of series. Repeatedly performing hypothesis tests at the  $\alpha$  level would increase the  $\alpha$  level of the entire data set without some adjustment to control for false detection.

Using the classic Bonferroni (testing at the  $\alpha/n$  level) correction was thought to be too conservative and would reject more series than necessary.

I examined two methods of adjusting for false detection: the FDR (False Detection Rate) adjustment of Benjamini and Hochberg (see Benjamini and Hochberg 1995), and q-values, which measures the proportion of false positives incurred (called the false discovery rate) when that test is called significant.

The basic FDR method takes all the p-values and sorts them by size, from lowest to highest, for all N series. Then where P(i) is the ith p-value in the ordered set, we check to see if the following is satisfied:

$$P(i) \le \alpha \times i/m$$
, where  $m = N - i + 1$ 

If this is true, the test is significant.

I used the FDR adjustment implemented in the p.adjust function in base R. This function returns an adjusted version of the p-value such that

$$PAdj(i) = P(i) \times m/i$$

where m is as defined in the first equation.

As for q-values, Storey (2018) contains a brief introduction to local FDRS and q-values, and a summary of the steps needed to generate q-values are given in Appendix B of Storey and Tibshirani (2003).

To generate the q-values, I used the R package qvalue, which is part of the BioConductor set of open-source packages for bioinformatics (Storey et al. (2024)). This package takes a list of p-values resulting from the simultaneous testing of many hypotheses and estimates their q-values and local FDR values. Storey and Bass (2024) is a vignette of the package and shows some of the plots and output that are given below.

Applying these methods to the p-values of the seasonal adjustment diagnostic used in this study, I found that the FDR adjustments for the p-values of the model-based F-statistic were very close to the actual p-values, and the q-values showed many more of the series to have significant seasonality.

However, to apply these tests correctly would require the p-values to be continuous and uniformly distributed; this required an examination of the distribution of the generated p-values which appears in Figure  $1_{-}$ 

These histograms show the distributions of the p-values taken from two implementations of the model-based F and the QS diagnostic. Note the skewness in the distributions for the different diagnostics, and the uptick in the distribution at the end of the distribution for the QS statistic, which we will discuss in more detail in Section 6.

Given that I could not use the FDR or q-values, I decided to pause consideration of multiple hypothesis testing and use the p-values as is standard practice.

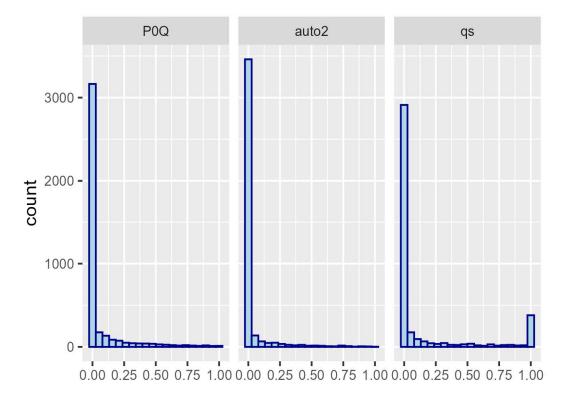


Figure 1: P-value histograms for different diagnostics

#### **5 Model-Based F-tests Failed**

Table 2 shows the percentage of series that fail the model-based F-test using each of the methods described in Section 2. The codes used for the automatic model identification methods are the same as in Table 1.

Here, we are testing at the one percent level, so a failure in this case means that the FDR adjusted p-value is greater than 0.01.

	auto0	auto 1	auto2	auto3	P0Q
First Closing	4.79	4.45	5.14	5.80	11.74
Second Closing	14.85	14.57	15.94	16.28	26.11
All CES	14.11	13.82	15.14	15.50	25.05

Table 2: Percentage of Failed Model-Based Seasonal F-tests in CES Series

Some observations:

- First Closing series fail far less frequently than the Second Closing series.
- Failure rates from the automatic models are very close to one another, as one would expect.
- The failure rate from the P0Q method is consistently higher than those of any of the automatic methods. This is true even for the auto2 set of options, which take their outliers and differencing orders from the same production model as the P0Q method.

### 6 The QS statistic

For each of these runs, we can compare the results from the QS seasonality diagnostic developed by Maravall with these results. It is described in detail in Bell et al. (2022), Section 4.2 - it uses the sample autocorrelation of the first two seasonal lags to construct a measure of the strength of the seasonal autocorrelation and it attempts to test the null hypothesis that these first two seasonal autocorrelations are zero. The resulting statistic is assumed to be distributed as a chi-squared statistic with two degrees of freedom.

The QS diagnostic is a function of the first two seasonal lag autocorrelations – those for lags 12 and 24 for monthly data, and for lags 4 and 8 for quarterly data – and it attempts to test the null hypothesis that these first two seasonal autocorrelations are zero. The rationale behind QS is that a series with seasonality, or residual seasonality, should exhibit substantial positive autocorrelation at these lags. Seasonal autocorrelation may extend beyond lags 24 or 8, but such higher lags are not used by QS. Note that if the autocorrelation at the first seasonal lag is zero or negative, then QS is set to 0 and the p-value is set to 1.

Appendix B of Bell et al. (2022) gives a formula for computing the QS statistic.

Rather than generate a separate QS statistic for each of the methods listed in Table 2, we'll take our QS statistic from the production runs. The spec files will be processed in R, and we'll extract the QS statistic for the original series adjusted for extremes for the last 8 years of data.

#### Table 3: Percentage of QS Statistics in CES Series

	Failed	Number	Percent
First Closing	40	298	13.42
Second Closing	1281	3730	34.34
All CES	1321	4028	32.80

In general, the QS statistic appears to fail more series than either the automatic or P0Q methods, though for most categories it appears to be closer to the P0Q method than any of the automatic methods. For First Closing, the results are very similar between QS and P0Q.

It is natural to apply these diagnostics in tandem - using them together to classify which series are clearly seasonal, which series are clearly not seasonal, and which series require more work. How much would the method of computing the seasonal F-test affect this type of classification?

There's another element to this - for the automatic modeling method, there are a group of series where there is not an F-test for stable seasonality due to fatal estimation errors. In this case, we'll only use the results from the QS statistic.

Table 4 gives the percentage of CES series that fail both the QS and model-based seasonal F-test. Again, the results for the automatic model methods are quite consistent and using the P0Q method fail a larger percentage of series, particularly for the Second Closing series.

	auto0	auto1	auto2	auto3	P0Q
First Closing	3.69	3.69	4.36	5.03	6.38
Second Closing	12.60	12.17	13.83	14.24	20.59
All CES	11.94	11.54	13.13	13.56	19.54

**Table 4:** Percentage of CES Series That Fail Both Diagnostics

Table 5 gives the percentage of CES series that pass both the QS and model-based seasonal F-test. Again, the results for the automatic model methods are quite consistent, and using the P0Q method would pass a smaller percentage of series, particularly for the First Closing series.

Table 5: Percentage of CES Series That Pass Both Diagnostics

	auto0	auto l	auto2	auto3	POQ
First Closing	85.23	85.57	85.23	85.23	81.21
Second Closing	63.14	63.08	63.22	63.22	60.13
All CES	64.77	64.75	64.85	64.85	61.69

Finally, Table 6 shows the series that pass only one of the diagnostics. Here, there is still a difference between the PØQ method and the automatic model methods, but it is different for the First Closing series than for the Second Closing series.

Table 6: Percentage of CES Series That Pass Only One Diagnostics

	auto0	auto 1	auto2	auto3	P0Q
First Closing	11.08	10.74	10.41	9.74	12.41
Second Closing	24.26	24.75	22.95	22.54	19.28
All CES	23.29	23.71	22.02	21.59	18.77

# 7 Choosing a Method for the Model-Based F-test

The automatic options seem to give very similar results, while the P0Q method give consistently higher failure rates than the automatic options. Which should one choose?

There are good reasons for selecting the P0Q method:

- using this option rather than a method that automatically identifies the model will be faster;
- even though there are a relative few series for the automatic options that have fatal execution errors, the POQ method is unlikely to generate such an error; and
- the results from the QS statistic seem to align more with the P0Q method.

To look at this further, I compared the AICCs (for the Corrected Akaike Information Criterion) generated for the models used for the auto2 and P0Q methods. I chose auto2 because it has the same outliers and differencing orders as the models specified for P0Q, making this a fairer comparison than for the other automatic methods.

The AICC is defined as

$$AICC_N = -2L_N + 2n_p \left(1 - \frac{n_p + 1}{N}\right)^{-1}$$

where the number of estimated parameters in the model, including the white noise variance, is  $n_p$ , N is the number of observations after applying the model's differencing and seasonal differencing operations, and  $L_N$  is the estimated maximum value of the exact log likelihood function of the model for the untransformed data.

I'll define the differences in AICC by diffAICC = AICC(P0Q) - AICC(auto2), where AICC(P0Q) is the AICC computed from a model formed by the P0Q method and AICC(auto2) if the AICC computed from a model identified with options specified for auto2 as specified in Table 1.

We will use the criteria from Symonds and Moussalli (2011). We consider differences larger than two in absolute value to be significant, meaning that if diffAICC > 2.0, we'd prefer the model from auto2, and if diffAICC < -2.0, we'd prefer the model from P0Q.

Table 7 shows that the auto2 model is the overwhelming favorite. This suggests using one of the automatic modeling options. The fully automatic method may be easier to implement (auto0 from Table 1), but there is a small reduction in fatal estimation errors if one incorporates the outlier regressors from the production model (auto2).

Table 7: Model Preference Due to AICC Difference

	Number of series
Preference for auto2 model	2,715
No preference for either model	826
Preference for P0Q model	422

## **8** Series That Pass One Diagnostic

If a given series passes or fails both the model-based F-test and QS diagnostics, the interpretation is straightforward.

What should we do when the series only passes one diagnostic, either by passing only one of the diagnostics, or in the case of estimation errors where the model-based F-test is not available?

There have been systems developed by Agustin Maravall and at the Bundesbank that used additional diagnostics in a sequential way to determine seasonality or the quality of the seasonal adjustment. One of those diagnostics is the spectrum of the original series.

X-13ARIMA-SEATS generates spectrum plots for the original series adjusted for outlier and extreme values and uses a visual significance criteria to determine if there are seasonal peaks in a given spectral plot. As described in Bell et al. (2022), the program uses the "star" as a unit of measure, which is 1/52 of the range of the spectral values. This is based on the ASCII text representation of the graph of the spectrum. To be considered a peak, the spectrum at a given frequency must have a height that is at least six (6) stars above the taller of the two nearest-neighbor frequencies on the plot and above the median height of all the plotted frequencies.

A sample spectrum plot from the X-13ARIMA-SEATS output of the All Employees, Oil and Gas Extraction, First Closing series is given in the Appendix at the end of this paper.

You can see the second, third and fourth seasonal frequencies (labelled "S" on the plot) are much stronger than the other seasonal frequencies and are more than six "stars" higher than the nearest spectral frequencies.

Figure 2 shows a high-resolution spectrum plot for the same First Closing CES series. While you cannot see the "stars" as you could on the ASCII plot from the printout, the seasonal frequencies are marked by dotted lines on the plot, and the visually significant peaks are marked by stars on the second, third, and fourth seasonal frequencies.

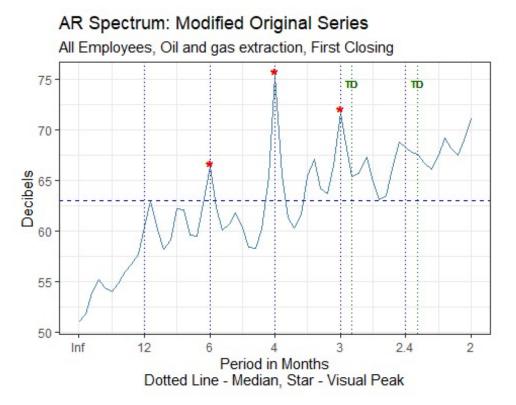


Figure 2: Spectrum Plot Produced by X-13ARIMA-SEATS

In this situation, the number of visually significant seasonal peaks will be collected for each of the CES series that passed only one of the diagnostics. If none of the seasonal peaks are visually significant, the series fails this test.

In addition, we will keep track of whether the fifth seasonal frequency is the only visually significant frequency, as this translates to an effect that has a period of 2.4 months, which is an odd frequency.

Table 8 shows the results for series that pass only one of the QS or model-base F statistic (shows as MBF in the table) generated from the auto2 method as shown in Table 1.

#### Table 8: Spectral peak results for auto2 method

QS	MBF	Number of series	No Visual Peaks	S5 Only Visual Peak
fail	pass	782	239	21
pass	fail	108	25	5
pass	run failed	41	0	0

These results show several additional series (about 28 percent) have no visual spectral peaks even though one of the seasonal diagnostics of significance passed at the one percent level.

# 9 Conclusions and Future Work

I would recommend using automatic model selection to select the model used to generate the Model-Based F-test for stable seasonality.

I would prefer outliers from production models, and would use automatic difference identification, but the differences between the different options are small.

For future work, I will redo this analysis with updated series to see how robust the selections are, though I may limit the number of automatic modeling options used. Also, I might investigate other options for reducing the number of estimation errors.

Also, Bell et al. (2022) points out that the QS and spectral diagnostics "may detect moderate or even mild seasonal autocorrelation that would not necessarily produce discernible seasonal patterns in the data". If the QS diagnostic is the only significant seasonality diagnostic (perhaps in addition to the spectral diagnostic), it would be good to examine sample autocorrelations of the series to determine if there is a peak discernible at the seasonal lags of the ACF.

## **10 Acknowledgements**

I want to acknowledge the contributions of several people:

- John Stewart, Steve Mance, Victoria Battista, and others from the Bureau of Labor Statistics working with CES series for their help getting data and working through issues with groups of series,
- Richard Tiller, Thomas Evans, and Jennifer Oh of the Bureau of Labor Statistics for conversations on technical issues and suggestions that helped form this work,
- Richard Tiller and Lindsey Platt of the Bureau of Labor Statistics for comments on drafts of this paper,
- Demetra Lytras of the US Census Bureau for creating new versions of X-13-SAM to address issues related to this study. Without these updated versions of X-13-SAM, this paper would not have been possible.

The plots in the paper were produced using the ggplot2 package (see Wickham 2016).

#### **11 References**

- Bell, W., McDonald-Johnson, K., McElroy, T., Pang, O., Monsell, B., and Chen, B. (2022), <u>"Identifying seasonality,"</u> BLS Research Reports, U.S. Bureau of Labor Statistics.
- Benjamini, Y., and Hochberg, Y. (1995), "Controlling the false discovery rate—a practical and powerful approach to multiple testing," Journal of Royal Statistical Society B, 57.
- Bureau of Labor Statistics (2024), <u>"Current Employment Statistics National:</u> <u>Calculation,"</u> U.S. Bureau of Labor Statistics, Department of Labor [Online; Last Modified Date: February 02, 2024].
- Lytras, D. (2024), <u>X-13-SAM (Spec-file Automated Modifications)</u>, Version 1.1, Washington, DC: U.S. Census Bureau, U.S. Department of Commerce.
- Lytras, D., Feldpausch, R. M., and Bell, W. R. (2007), <u>"Determining Seasonality: A</u> <u>Comparison of Diagnostics from X-12-ARIMA."</u>
- R Core Team (2024), <u>R: A language and environment for statistical computing</u>, Vienna, Austria: R Foundation for Statistical Computing.
- Sax, C. (2018), Seasonal: R Interface to X-13ARIMA-SEATS. R package version 1.7.
- Sax, C., and Eddelbuettel, D. (2018), <u>"Seasonal adjustment by X-13ARIMA-SEATS in</u> <u>R,"</u> Journal of Statistical Software, 87, 1–17. <u>https://doi.org/10.18637/jss.v087.i11</u>.
- Storey, J. D. (2018), <u>"False discovery rates,"</u> in International Encyclopedia of Statistical Science, Berlin, Heidelberg: Springer.
- Storey, J. D., and Bass, A. J. (2024), "Bioconductor's qvalue package version 2.36.0," Princeton University.
- Storey, J. D., Bass, A. J., Dabney, A., and Robinson, D. (2024), qvalue: <u>Q-value</u> estimation for false discovery rate control.
- Storey, J. D., and Tibshirani, R. (2003), <u>"Statistical significance for genome-wide experiments,"</u> Proceedings of the National Academy of Sciences, 100, 9440–9445.
- Symonds, M., and Moussalli, A. (2011), <u>"A brief guide to model selection, multimodel</u> <u>inference and model averaging in behavioral ecology using Akaike's information</u> <u>criterion,"</u> Behavioral Ecology and Sociobiology, 65, 13–21.
- Time Series Software Group (2023), <u>X-13ARIMA-SEATS Reference Manual, Version</u> <u>1.1</u>, U.S. Census Bureau, U.S. Department of Commerce.
- Wickham, H. (2016), ggplot2: <u>Elegant graphics for data analysis</u>, Springer-Verlag New York.

Spectrum esti	ECTRUM) of the differenced mated from 2014.Nov to 2022	.Oct.					
	********		+++++++++++++++++++++++++++++++++++++++	++++++	++++++++		
75.49I		S				I	75.49
I		S				I	
I		S S S				I	
I		S				I	
73.60I		S				I	73.60
I		Š				Ī	
I		S S S				I	
I		S C					
		2	6			I	71 70
71.72I		S	S			I	71.72
I		S S S S S S	S S S			I	
I		S	S			SI	
I		S	S			SI	
69.83I		S	S			SI	69.83
I		ŝ	S			SI	
Ī		ç	S			* *SI	
I		ر د	S	*	*	* *SI	
		2				51	67.04
67.941		S S	S		*S	** *SI	67.94
I		S	S			****SI	
I		S	* S	* *	*S*T *	****SI	
I		S	* *S	* *	*S*T* *	****SI	
66.06I	S	S	* *S	* *	**S*T***	****ST	66.06
I	S	S	-	* **	**S*T***		00100
I	S	*S*		*T**	**S*T***		
			5	-			
I	S	*S*	-	*T***	**S*T***		
64.17I	S	*S*	*** *S	-	**S*T***		64.17
I	S	*S*	*** *S	*T***	**S*T***	****SI	
I	S	*S*	****S	*T*** *	***S*T***	****SI	
I	* S	*S*	****S	*T****	***S*T***	****SI	
62.29I	* *S	*S*	****	*T****	***S*T***	****ST	62.29
I	* ** *S* *	*S*	0	•	***S*T***	01	02125
I	* ** *S* *	*S*	-		***S*T***		
	5						
I		-	-	-	***S*T***		co 10
60.40I	S* ** *S* **				***S*T***		60.40
I	S** ** *S****	-	-	-	***S*T***		
I	S** ****S****	**S**	******S	*T****	***S*T***	****SI	
I	S** *****S****	**S**	******S	*T****	***S*T***	****SI	
58.52I	S** *****S****	**S**	******S	*T****	***S*T***	****SI	58.52
I	S*******						
Ī	*5************	-	-	-			
I	*S*******	-	-	-			
		-	-	-			FC (2)
56.63I	**S***********************************						56.63
I	**S*******						
I	***S*******S*****						
I	***S******S*****	**S**	******S	*T****	***S*T***	****SI	
54.74I	* ***S******S*****	**S**	******S	*T****	***S*T***	****SI	54.74
I	** ****S*******S*****	**S**	******S	*T****	***S*T***	****SI	
Ī	*****						
I	*****						
	**************************************						F2 96
52.861		-	-	-			52.86
I	**************************************						
I	*******S******	-	-	-			
I	*********S*******S******	**S**	*****S	*T****	***S*T***	****SI	
50.971*	********S*****************************	**S**	******S	*T****	***S*T***	****SI	50.97
	+++++++++++++++++++++++++++++++++++++++						
	=SEASONAL FREQUENCIES, T=TR						
5	-SENSONAL TREQUENCIES, I-IN	DTHO	DAT TAL	QUENCII			

# Appendix: Sample Spectral Plot from X-13ARIMA-SEATS Output