Seasonal Adjustment of CPS Labor Force Series During the Great Recession[1]

Thomas D. Evans and Richard B. Tiller, Bureau of Labor Statistics
Bureau of Labor Statistics, Statistical Methods Staff,
2 Massachusetts Ave. NE, Room 4985, Washington, DC 20212

Key Words: Time Series; Level Shifts; Ramps

Introduction

The Current Population Survey (CPS) is a monthly household survey that collects information on labor force characteristics for the United States. The seasonally adjusted monthly levels of employment, unemployment and unemployment rate are key indicators of the health of the economy. Recent economic turbulence related to the Great Recession has raised public interest in how the CPS series are seasonally adjusted. The most widely used approaches to seasonal adjustment apply weighted moving averages to the original data to separate the seasonal from the nonseasonal components of the data. These weights may be derived from a model fit to each series or from a non-parametric method using pre-determined filters. These methods have been successful because of their flexibility in accommodating stochastic changes in seasonal patterns. The presence of stochastic seasonality, however, creates conceptual and statistical difficulties in separating seasonal from the nonseasonal components. The difficulties are compounded during periods of rapid economic change.

The national CPS data are adjusted using X-12-ARIMA (Findley, et al., 1998) which is a nonparametric approach. This method has a long history of development and refinement and is currently used by statistical agencies throughout the world. This paper reviews how well this method was designed to handle changes in trend-cycle and seasonality during both economic expansions and contractions, what tools are available to adjust for major economic shocks, and how the method has performed in the last recession.

The original X-11 filtering algorithm that performs seasonal adjustment remains at the core of the X-12 process (Shiskin, Young, and Musgrave, 1967). The X-11 part of X-12 has a broad range of seasonal and trend-cycle filters. The symmetric Henderson trend filter reproduces third-degree polynomials within the span of the filter which can be as short as one to two years. Thus, the trend filter is flexible enough to follow rapid changes in growth rates as well as quickly occurring turning points in the trend-cycle. Towards the end of the series less adaptive asymmetric filters must be used. Although the one-sided Henderson filter is less flexible, it can still track a local linear trend and with ARIMA forecasting so more adaptive filters are produced for the end points (Dagum 1983).

A well-known problem with seasonal adjustment is that moving averages are highly vulnerable to sudden strong atypical changes or outliers. For example, sudden shocks to the trend-cycle can be absorbed into the seasonal factor estimates and erroneously removed from the seasonally adjusted series. A second potential source of distortion is a break in the seasonal pattern which some analysts argue is likely to occur during recessions. Either type of shock may lead to distorted seasonally adjusted estimates in later years following the recession.

The best way to detect and correct for sudden disturbances is with prior information on their source, time of occurrence, magnitude, and duration. The original X-11 developers added an option for users to

[1] Disclaimer: Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

supply adjustment factors to the original data to remove the effects of shocks prior to seasonal adjustment. Users, however, had to develop these prior adjustments outside of the software. This is hardly feasible to do when there are hundreds or thousands of series to seasonally adjust. With the addition of the RegARIMA modeling capabilities in X-12 (Findley, et al., 1998), it became standard practice to utilize the automatic outlier detection routine prior to the seasonal adjustment step. RegARIMA also allows users to specify their own outlier models.

Currently, there are 144 CPS national series that are directly seasonally adjusted. One hundred and thirty-nine of these series are released monthly and the others quarterly. Many more are indirectly seasonally adjusted based on the directly-adjusted series. See Tiller and Evans (2013) for more details.

The focus of this paper is on eight CPS series that make up the national unemployment rate:

Employment (EM) ages 16-19 by gender
Employment ages 20+ by gender
Unemployment (UN) ages 16-19 by gender
Unemployment ages 20+ by gender

The eight series are directly seasonally adjusted and then used to derive the seasonally adjusted unemployment rate (UR). The official national seasonally adjusted unemployment rate is plotted over time in Figure 1. The trend for the unemployment rate (not officially published) appears in Figure 2. Both series rise sharply in the 2008-2009 period. Commentary from Wall Street analysts and other members of the press suggested that the sharp rise (fall) in UN (EM) from the fourth quarter of 2008 to the first quarter of 2009 was absorbed into the seasonal component and this caused seasonal adjustment to overestimate recent economic growth in those quarters (a decline in UN or a rise in EM) and underestimate in the second and third quarters.

Testing for Trend-Cycle and Seasonal Breaks with RegARIMA Models

To investigate the claims of distortions to the seasonal adjustment process, we use RegARIMA models to test for the presence of recession effects due to trend-cycle and seasonal breaks. For seasonal breaks we also use a graphical analysis of X-11 seasonal-irregular (SI) sub-plots by month which will be described in more detail below.

The RegARIMA model may be represented as,

$$y_t = X_t'\beta + Z_t$$

where X_t is a vector of fixed regressors and Z_t follows a seasonal ARIMA model. The parameters of the ARIMA model implicitly determine the stochastic properties of seasonality as well as trend-cycle and other components of the series without having to specify specific models for these components. The variables of the regression part of the model provide the basis for testing for the presence of exogenous effects, such as trend and seasonal breaks that we wish to estimate and remove from the series prior to seasonal adjustment to prevent distortions to the X-11 seasonal adjustment process.

Trend breaks can be handled by modeling level shifts (LS), temporary changes (TC), and ramps which are briefly described below and the form of the regression variables for testing for each effect are shown in Table 1. These variables are part of a large set of built-in regressors available in RegARIMA.

LS: A permanent abrupt shift in level.

TC: An abrupt change in level followed by a gradual return to normal.

Ramp: A change to a new level with a user specified start (t_0) and end date (t_1) and a fixed rate of change per period $(1/(t_1 - t_0))$.

To test for the presence of seasonal breaks we use the partial change in regime test which is another predefined option available in RegARIMA. An exogenous shift in the seasonal factors occurring prior to the change point (t_0) is modeled in terms of the s-1 seasonal dummy variables ($M_{j,t}$) shown in Table 1. This effect is referred to as a partial change of regime since it is estimated for only the early span and set to zero for the complementary span (U.S. Census Bureau, 2013).

LSs and TCs are usually detected during the automatic outlier detection option of RegARIMA. This option assumes no prior knowledge of the timing or type of outliers. The procedure identifies additive outliers (AOs), level, and TCs. The setting of the critical value is always an issue since it involves multiple testing (type 1 errors are inflated) and because the outliers may have no economic explanation. To reduce the number of spurious outliers, critical values for the outlier T-values (nonstandard distributions) are adjusted for the number of observations (see Findley, et al., 1998).

Table 1: RegARIMA Outlier Variables

Regression Effect	Variable definition
Additive	$X_t^{AO} = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}$
Level shift	$X_t^{LS} = \begin{cases} -1 & t < t_0 \\ 0 & t \geq t_0 \end{cases}$
Temporary change	$X_t^{TC} = \begin{cases} 0 & t < t_0 \\ \alpha^{t-t_0} & t \geq t_0 \end{cases}$
Temporary Ramp	$X_t^{Rp} = \begin{cases} -1 & t \leq t_0 \\ (t-t_0)/(t_1-t_0)-1 & t_0 < t < t_1 \\ 0 & t \geq t_1 \end{cases}$
Partial change of regime in seasonality	$X_{j,t}^{E} = \begin{cases} M_{j,t} & t < t_0 \\ 0 & t \geq t_0 \end{cases};\ M_{j,t} = \begin{cases} 1 & t = j+ks \\ -1 & t = s \\ 0 & t \neq j+ks \end{cases}$ for $j = 1, \ldots, s\text{-}1;\ k = 0, 1, 2, 3\ldots;\ s = 12$

Since the automatic outlier identification procedure is run at the end of each year, we make no attempt to identify LSs in real time. During the recession period, we identified two LSs: May 2008 for (UN) teen males and UN teen females. Note that the original t-value for the teen females LS was originally only 3.8, even though the default critical value ≅ 4.0 (it varies by series length). Occasionally, an outlier with a slightly lower T-value may be added to a series if it is believed that doing so will be helpful or if the effect is known in advance. Should we reduce the critical value during a recession period? An examination of the “almost” critical outliers identified by the program confirmed our suspicion that lowering the default critical value would have introduced too many spurious outliers.

Table 2 shows the RegARIMA estimates for the two level shifts. Both LSs have large relative level increases. Notice that the LS for UN teen females is no longer as significant as additional data lowered the t-values further below the critical value. An obvious question is whether the LSs for UN teens affected the seasonally adjusted values? The surprising answer is no, but they did have strong effects on the trend and irregular components. If we just examined the seasonal factors, we might conclude that no recession-related LS effects occurred. This point is clearly seen in Figures 3 and 4. However, it is also important to examine the trend and irregular components for interpreting what happened. The trend in Figure 5 shows a different story in which the break is clearly noticeable. Without accounting for an LS, the trend over smoothes the recession effects, and the irregular component (Figure 6) compensates for the LS. The plot for UN teen females is in Figure 7.

Table 2: RegARIMA Estimates for Level Shifts

Level Shifts from Automatic Outlier Detection						
Series	**Date**	**Coef (exp)[2]**	**T-Value**	**AICC[3] w/LS**	**AICC w/o LS**	**AICC Difference**
UN M 16-19	May 2008	1.26	4.4	10,343	10,360	-16.5
UN F 16-19	May 2008	1.17	3.1	10,236	10,243	-6.8

If the LSs were ignored for the UN teen series, the large upward shifts in May would be interpreted as a random deviation from normal as opposed to a fundamental level shift. Data users would see no noticeable effect on the seasonally adjusted series since the seasonal patterns are not affected by the LS.

After testing for LSs, we moved to experimenting with ramps. Ramps have not been used for CPS series and are not commonly used in other applications of seasonal adjustment. Examples of the use of ramps are given by Buszuwski and Scott (1993), Maravall and Perez (2011), and (Lytras and Bell 2013).

The specification of a ramp is not automated and requires the user to specify beginning and end points for the adjustment period. This leads to the problem of how to select t_0 and t_1. Since there is not much guidance without prior information, we can visually look for segments of series where the change appears to be relatively constant. Another problem is that in this process it can be fairly easy to achieve "better" fits but they may be spurious. We try to determine this by utilizing the AICC goodness-of-fit criterion.

Test results appear in Tables 3 and 4 below. In Table 3, the exponentiated coefficients show small effects even though the T-values appear significant. Revisions between the concurrent and most recent seasonally adjusted values with and without ramps (Table 3) show little differences, yet the AICC tests seem to indicate a preference for including the ramps as regressors in the model. As the testing is not conclusive, we examine plots. The seasonally adjusted results for EM 20+ with and without ramps are plotted in Figure 8. As the coefficient value close to one implies, the addition of the ramp makes almost no difference. Using the three series with tested ramps to derive the national UR in Figure 9 also shows only very small differences.

Table3: Ramp Results

Series	**Ramp Start**	**Ramp End**	**Ramp Coef (exp)**	**T-Value**
EM M 20+	Nov 2008	Mar 2009	0.99	-5.66
UN M 20+	Apr 2008	Feb 2009	1.06	4.92
UN F 20+	Apr 2008	Feb 2009	1.04	4.79

[2] Coefficients in the tables are exponentiated from logs to levels.

[3] See U.S. Census Bureau (2013) for a detailed explanation of the AICC and model-selection criteria in RegARIMA. Models with minimum AICC are usually preferred.

Table 4: Seasonal Adjustment Revisions due to Ramps and AICC Comparisons

Series	SA Revision Medians, Average Absolute %		AICC	AICC w/Ramp	AICC Difference
	Official	With Ramp			
EM M 20+	0.07	0.08	11,511.45	11,484.39	-27.06
UN M 20+	0.78	0.81	6,822.91	6,805.60	-17.32
UN F 20+	0.83	0.81	6.732.70	6,713.66	-19.04

The partial change of regime test for a break in the seasonal pattern is shown in Table 5. While we tested other months, October 2008 makes sense as a potential breakpoint, yet the AICC criterion rejects the presence of a deterministic break for all 8 series.

Table 5: AICC Results for Seasonal Breaks for October 2008

	AICC		
Series	**No Break**	**With Break**	**Difference**
EM M 16-19	10,698.7	10,711.3	12.7
EM F 16-19	10,672.0	10,685.4	13.3
EM M 20+	11,511.5	11,517.7	6.2
EM F 20+	11,525.6	11,527.9	2.2
UN M 16-19	10,343.3	10,348.6	5.3
UN F 16-19	10,235.7	10,257.7	22.0
UN M 20+	6,822.9	6,829.9	7.0
UN F 20+	6,732.7	6,748.4	15.7

We further explore the possibility of a seasonal break with a visual examination of X-11 SI sub-plots for each series. The SI is the de-trended series produced in the X-11 part of the procedure. The trend-cycle is estimated by the Henderson filter and removed from the original series. The seasonal factors are later derived by smoothing the SIs by month, usually with a 3x5 moving average. These SI sub-plots by months are useful for identifying potential breaks in seasonality. However, our examination found no apparent breaks for any of the eight series. An example of an SI sub-plot is shown in Figure 10.

Performance of Post-Recession Seasonal Adjustment

The NBER dates the Great Recession as lasting from December 2007 to June 2009. We use the period January 2008-December 2009 as a more relevant recession period for the labor market since the unemployment rate doubled and then began to decline around January 2010. Largely because NBER dating procedures depend heavily on real product and income measures, it is not unusual for labor market recovery to lag the end of NBER designated recessions.

In order to examine the effect of the recession data (January 2008-December 2009) on the January 2010-March 2012 period for the national UR, we treated the recession data as missing. When observations are set to zero, RegARIMA automatically adds AO regressors to produce forecasts based on pre-recession data. The original series with forecasted values for the recession period was seasonally adjusted with a January 2010 level shift added to the model. Since unemployment is still much higher in 2012 compared to pre-recession levels, we use the term “post-recession” advisedly.

The officially seasonally adjusted national UR series is plotted against the seasonally adjusted series with recession data replaced with forecasts in Figure 12. Note how the seasonally adjusted series with the recession data forecasted ignores the actual rise due to the recession. Starting in 2010, the two series become very close again which indicates that the recession data does not affect the post-recession seasonal adjustment much.

In November 2010, an upward blip occurs in the seasonally adjusted national UR series (see Figure 12). Various sources have attributed this movement to seasonal bias (for example, see Zentner, et al., 2011). Actually, the blip is due to variation in the irregular component that was eliminated by the trend filter. In short, we see no evidence of seasonal bias during the post-recession period in question.

Summary

Our overall analysis of the performance of the seasonal adjustment for the CPS national unemployment rate is that there is no evidence of trend or seasonal breaks that biased the adjustments during the post-recession period. The major findings are:

- Six of the eight series analyzed need no outlier adjustments during the recession.
- Recessionary level shifts were identified in the normal way for UN male and female teenagers. This reduced the irregular variation but had little effect on the seasonally adjusted series.
- Fitting ramps models had little effect.
- Removing the recession data had little effect.
- The standard X-11 symmetric filters appear to perform well during and after the 2008-2009 recession.

References

Buszuwski, J., and Scott, S. (1993), "Some Issues in Seasonal Adjustment when Modeling Interventions," in JSM Proceedings, Business and Economics Section, 208-213.

Dagum, E. (1983), *The X-11-ARIMA Seasonal Adjustment Method*, Ottawa: Statistics Canada.

Dagum, E., and Morry, M. (1985), "Seasonal Adjustment of Labour Force Series during Recession and Non-Recession Periods," *Survey Methodology*, XI, no. 2.

Findley, D., Monsell, B., Bell, W., Otto, M., and Chen, B. (1998), "New capabilities and methods of the X-12-ARIMA seasonal adjustment program," *Journal of Business and Economic Statistics* **16**, 127--177 (with discussion). Available online at http://www.census.gov/ts/papers/jbes98.pdf.

Lytras, D., and Bell, W. (2013) "Modeling Recession Effects and the Consequences on Seasonal Adjustment," presented at the 2013 Joint Statistical Meetings.

Maravall, A., and Perez, D. (2011), "Applying and Interpreting Model-Based Seasonal Adjustment. The Euro-Area Industrial Production Series," Bank of Spain. Available online at http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/11/Fich/dt1116e.pdf.

Shiskin, J, Young, A., and Musgrave, J. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Bureau of the Census, Technical Paper 15. Available online at http://www.census.gov/ts/papers/ShiskinYoungMusgrave1967.pdf.

Tiller, R., and Evans, T. (2013), "Methodology for Seasonally Adjusting National Household Labor Force Series with Revisions for 2013," Bureau of Labor Statistics CPS Technical Documentation. Available online at http://www.bls.gov/cps/cpsrs2013.pdf.

U.S. Census Bureau (2013*)*, *X-13ARIMA-SEATS Reference Manual* (Version 1.0), Washington, DC: Author. Available online at http://www.census.gov/ts/x13as/docX13AS.pdf.

Zentner, E., Amemiya, A., and Greenberg, J. (2011), "Stronger Data Ahead: Explanation and Implication," Nomura Global Weekly Economic Monitor. Available online at https://www.bea.gov/about/pdf/Stronger%20Data%20Ahead.pdf.

Figure 1: Official National Seasonally Adjusted Unemployment Rate

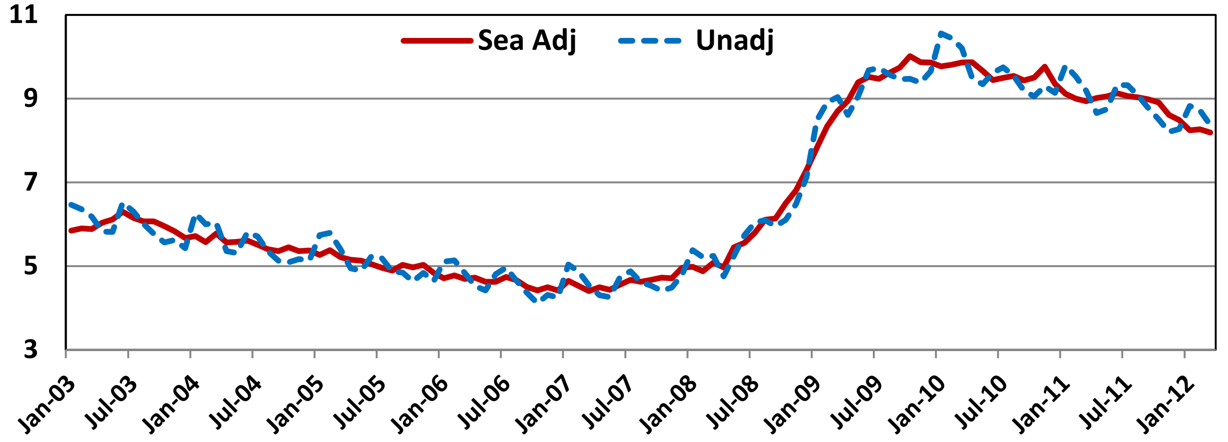

Figure 2: National Unemployment Rate Trend (not published)

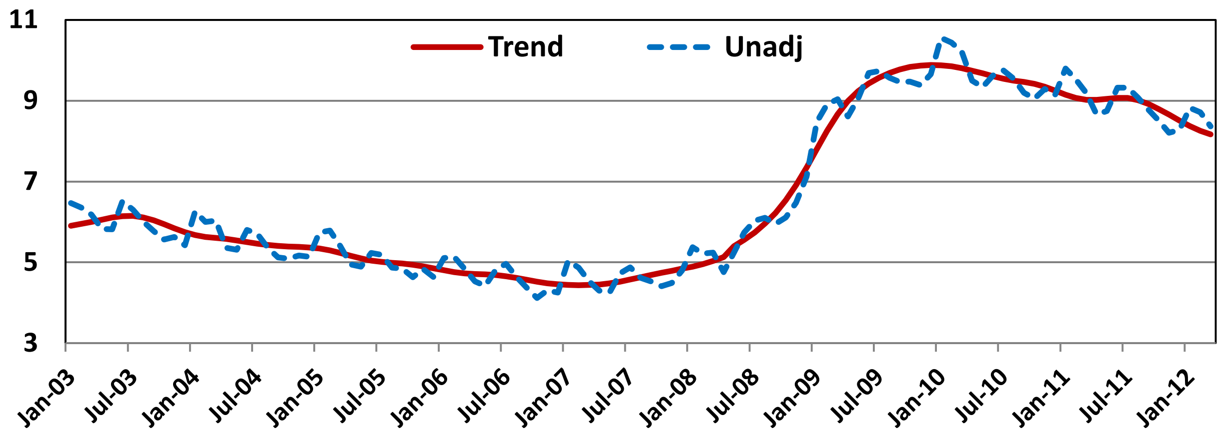

Figure 3: Level Shift Effect on UN Male Teen Multiplicative Seasonal Factors

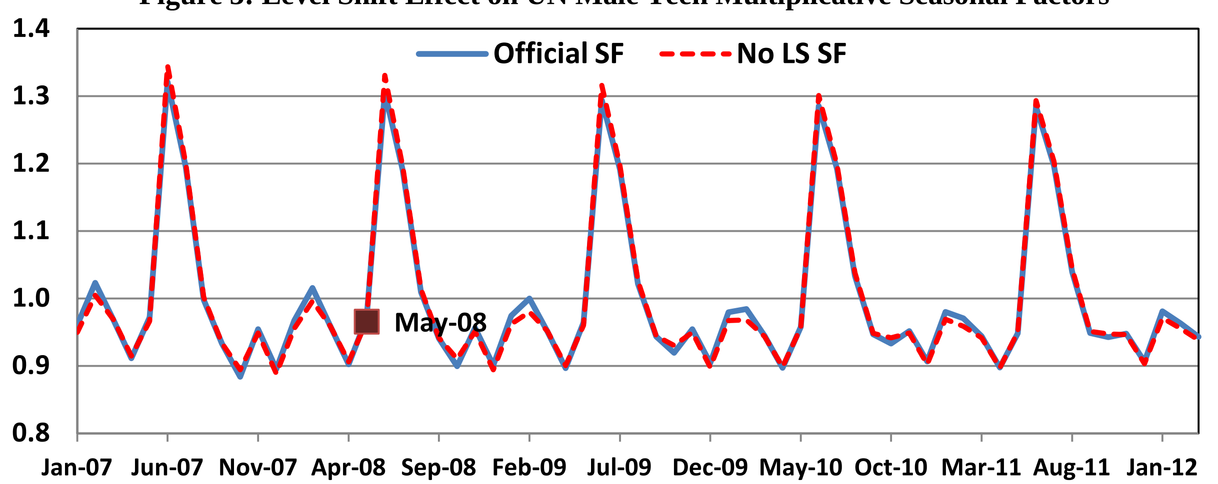

Figure 4: Level Shift Effect on UN Male Teen Seasonally Adjusted Series

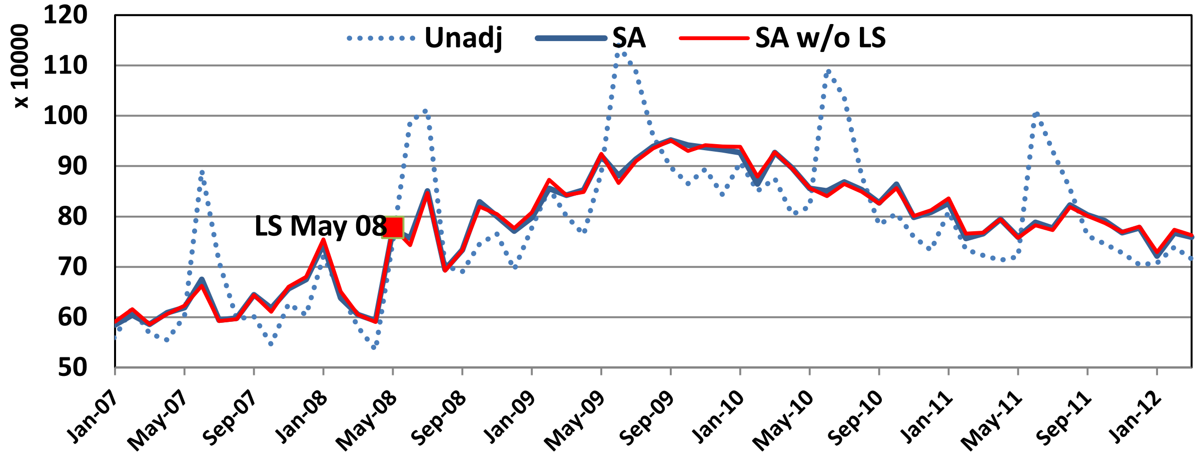

Figure 5: Level Shift Effect on UN Male Teen Trend Component

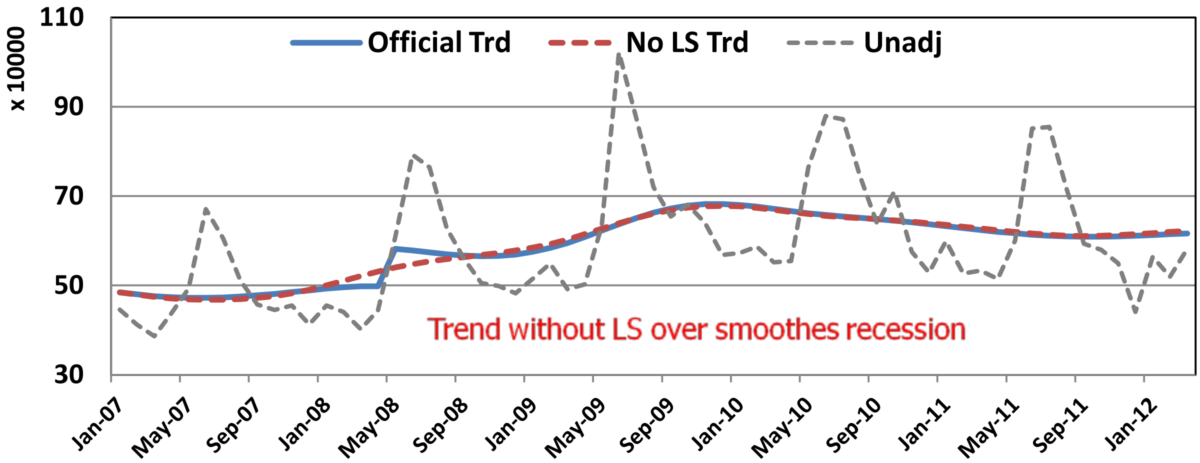

Figure 6: Level Shift Effect on UN Male Teen Multiplicative Irregular Component

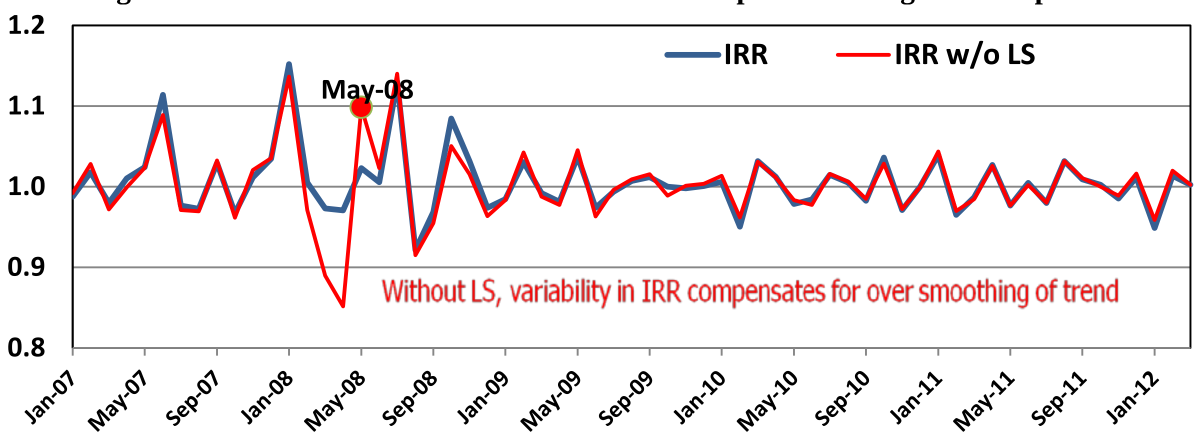

Figure 7: Level Shift Effect on UN Female Teen Seasonally Adjusted Series

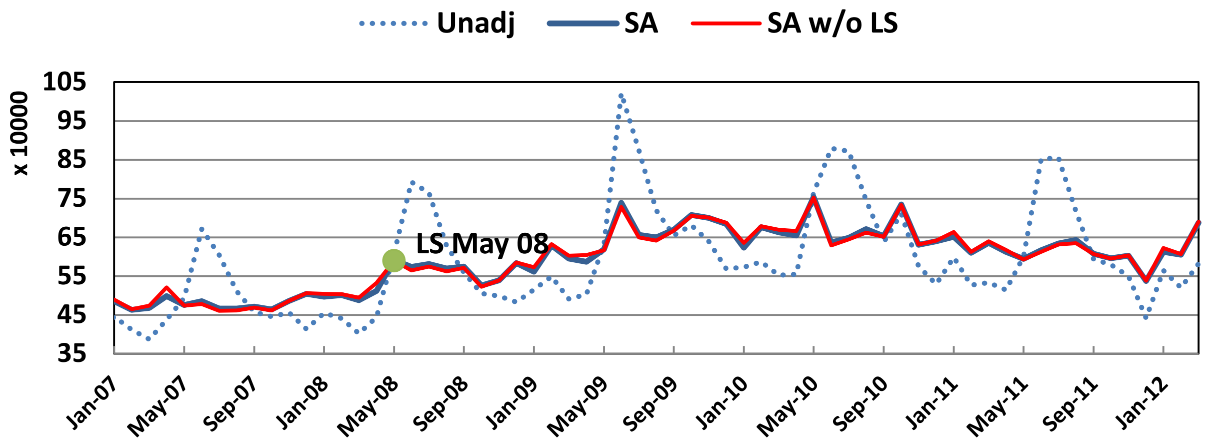

Figure 8: Seasonally Adjusted EM M 20+ with and without Ramps

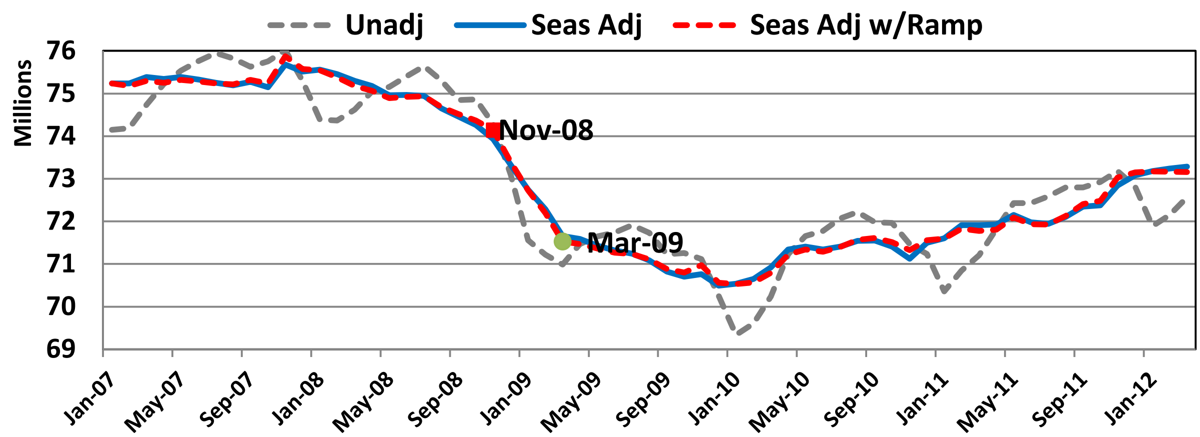

Figure 9: Effects of Modeling Ramps on National UR

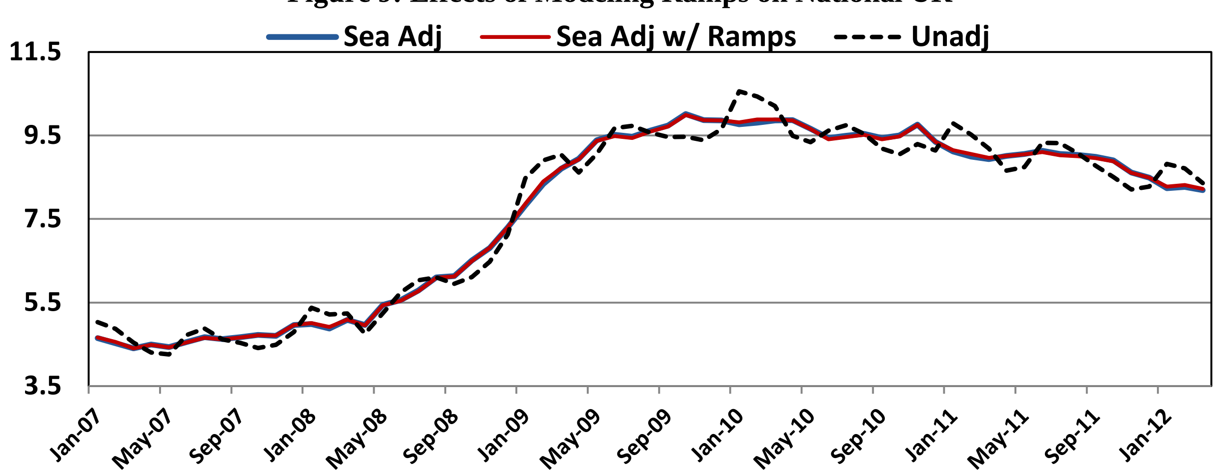

Figure 10: Seasonal-Irregular Sub-Plot by Month Example

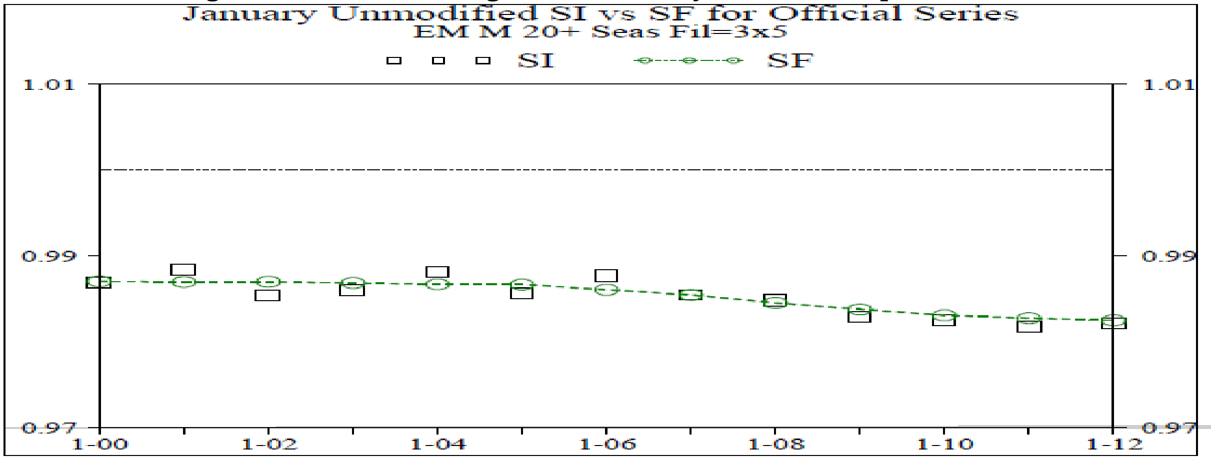

Figure 11: Seas Adj National UR with Recession Data (2008-2009) Removed

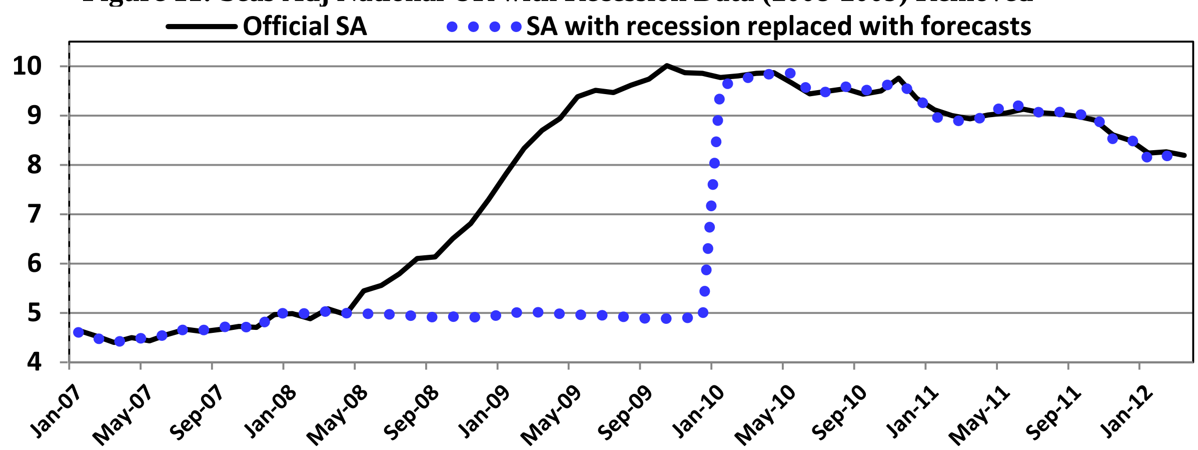

Figure 12: November 2010 UR Blip due to Irregular