

This time it is different! Or not?

Discounting past data when predicting the future

Philip Hans Franses

Eva Janssens

Econometric Institute

Erasmus School of Economics

EI2017-25

Abstract

We employ a simple method based on logistic weighted least squares to diagnose which past data are less or more useful for predicting the future course of a variable. A simulation experiment shows its merits. An illustration for monthly industrial production series for 17 countries suggests that earlier data are useful, for the prediction in a crisis period (2006-2011) and for the period after the crisis (2011-2016). Hence, this time, apparently it was not that different after all.

Keywords: Forecasting; Weighted Least Squares; Discounting; Logistic function; Industrial Production

This version: September 2017

Correspondence to: Philip Hans Franses, Econometric Institute, Erasmus School of Economics, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands,
franses@ese.eur.nl

1. Introduction and motivation

Recessions and expansions alternate. Where expansions can last for a long while, typical recessions last for 1 to 2 years. Forecasting of turning points is important, but it is notoriously difficult. Forecasting during recession times is important too, as one would want to know if prosperous times are near or whether the recession would continue for a while.

Time series models that allow for different regimes like recessions and expansions are non-linear time series models, and a very popular type of non-linear model is the smooth transition autoregression (STAR), see Granger and Teräsvirta (1993) and van Dijk, Teräsvirta and Franses (2002). An alternative popular model is the Markov switching model (Hamilton, 1989). A STAR model, like a Markov switching model, allows for a time series model that has varying parameters and lag lengths across the two macroeconomic regimes. An important feature of these models is that it is assumed that the models for each of the regimes are the same over time. In other words, it is assumed that expansions can be described similarly over various decades, and the same holds for recessions. More precise, and as an example, it is then effectively assumed that the 1974-1975 recession has the same characteristics as the recent 2008-2009 recession, briefly summarized as “this time it is again the same”.

In this paper, we propose a simple method that can be helpful to see if indeed earlier recessions do have similar characteristics as recent ones. The method does not require the fitting of nonlinear models with time-varying parameters or models, which makes it easy to apply in practice. The main idea of our method is to see if those older data are useful to predict newer data. The method is based on the ideas in Refenes et al. (1997), and it aims to add to the literature on detecting breaks and changes, see for example Giritatis et al. (2013) and Inoue et al. (2017) for two recent important contributions. An approximate model is set as a linear autoregression, which oftentimes quite closely fits the data. By using some version of discounted least squares, a first visual impression can be obtained as to whether, for example, earlier recession data warrant inclusion in forecasting models for future recessions, at least for many countries, as we will see in our illustration.

The outline of the paper is as follows. Section 2 presents the main idea in more detail, and examines the performance of the method using a small-scale simulation experiment. Section 3 illustrates the method for monthly industrial production data for 17 countries. A key empirical result is that the 2008-2009 recession indeed seems not to be that different from earlier ones. Section 4 concludes with a discussion.

2. Main idea and simulations

It is not unlikely that data generating processes of time series experience changes over time. These changes can be permanent or recurrent. Furthermore, these changes may happen at once or gradually. For example, there may be a smooth transition from one data generating process to another one but there may also be a sudden break. Consider for example the improvement in the measurement of inflation and Gross Domestic Product (GDP) over the recent decades by national statistics' agencies, which may induce a smooth transition.

Otherwise, older observations of GDP and inflation could be biased or suffer from measurement error, and one may want to take this into account when modeling the data. Yet another example of a potential source of changes in the data generating process is that central banks change the way they decide on setting interest rates. Different central bank presidents may have different views on the economy, so it could well be that interest rates changes during the presidency of Paul Volcker were influenced by different variables or by different impact of the same variables than the interest rate decisions are of Janet Yellen (Judd and Rudebusch, 1998).

A popular model to deal with possible changes in the data generating process is the Smooth Transition Autoregressive (STAR) model, which for first order dependence looks as follows:

$$y_t = (\phi_{1,0} + \phi_{1,1}y_{t-1})(1 - G(q_t; \gamma, c)) + (\phi_{2,0} + \phi_{2,1}y_{t-1})G(q_t; \gamma, c) + \varepsilon_t$$

where $G(q_t; \gamma, c)$ is a continuous function that changes smoothly from 0 to 1 as the values of some variable q_t increase, see Granger and Teräsvirta (1993) and Van Dijk et al. (2002). Note that the commonly applied version of this model assumes a model in each of the two regimes, where the regime-specific model is the same within each regime. In other words, when $G(q_t; \gamma, c)$ associates with expansions and recessions, then the parameters in the AR(1) models in each of the regimes are the same for all expansions and recessions. In certain settings, this assumption might seem too strong and in this paper we try to see if there is a simple visual method that can help to indicate whether this assumption is plausible from the onset or not.

Our method relies on the idea that we can analyze which past observations are actually helpful and relevant to predict future observations, and at the same time, which past observations could well be discarded. As an approximate model, we consider a linear

autoregression. To infer which data can be discarded, we use a weight function to weight the observations as done in weighted least squares, where now data before a certain time α will get a weight approaching zero. The weight function that we will employ is the logistic weight function

$$w(t) = \frac{1}{1 + e^{-\gamma(t-\alpha)}}, \text{ with } \gamma > 0$$

As with weighted least squares (WLS), each variable is multiplied with $\sqrt{w(t)}$, and hence the variable to be modeled, y_t , becomes $y_t\sqrt{w(t)}$. After this transformation, ordinary least squares (OLS) estimation is used to estimate the parameters in the autoregression, now for the transformed variables. For convenience, we call our method LWLS, short for logistic weighted least squares. The positive-valued parameter γ in $w(t)$ measures the speed with which the function runs from 0 to 1. A larger value of γ gives the function a larger-valued derivative when $t = \alpha$. The parameter α moves the weight function along the x-axis. In the experiments and illustrations below, we will set $\gamma = 10$ to allow for some smoothness. Figure 1 depicts a few variants of the weight function.

To evaluate the effect of α in the weighting function, the parameters in the approximate autoregression are estimated for many values of α . Next, we will look at how the sum of squared residuals (SSR) (or one-step-ahead forecast errors each time after re-estimating the parameters) for the last 20% of the observations (the hold-out sample to evaluate the forecasts) evolves for different values of α . Other choices than 20% can also be made, of course, as we will do in our empirical illustrations.

Figures 2 and 3 are informative on how LWLS works and how the results can be interpreted. Figure 2 displays the SSR for the last 20% of the observations for 800 different values of α for the following data generating process (DGP)

$$\begin{aligned} \text{DGP 1:} \quad t = 1, \dots, 300: \quad & y_t = 0.1 + 0.2y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \\ t = 301, \dots, 1000: \quad & y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \end{aligned}$$

In this DGP, the data before the 300th observation would not be beneficial to predict the last 200 of the 1000 observations, that is, observations 801 to 1000.

In Figure 2, we notice that there is a saddle point in the SSR function at $\alpha = 300$, where we applied LWLS to an approximate AR(1) model. In this saddle point, the derivative of the SSR function is zero, but the function is not at an extreme value point. This observation is confirmed by Figure 3, which displays an approximation of the first derivative of the SSR function, by computing for each integer value of α the following score:

$$\frac{(SSR(\alpha) - SSR(\alpha - 1))}{1}.$$

Looking for the saddle point in the SSR function, or finding the point where the (approximate) derivative of the SSR function is zero, can thus lead to a useful choice for α , in practice. The SSR function seems to be able to detect the location of the change in the DGP.

To see the contrast between the SSR functions for data with or without a break, consider Figure 4, where we consider a DGP without a break. For the no-break DGP:

$$DGP\ 2: \quad y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1, 2, \dots, 1000,$$

we notice that SSR is again a decreasing function in α , but now there is no saddle point. The downward trend in Figure 4 arises from the fact that the sum of squared residuals is increasing in the number of observations, so giving older variables a lower weight will generally decrease the SSR.

To see what one can expect to observe if there is a temporary change in the data, like for a DGP as

$$\begin{aligned} DGP\ 3: \quad t = 1, 2, \dots, 300: & \quad y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, & \quad \varepsilon_t \sim N(0,1) \\ t = 301, \dots, 500: & \quad y_t = 0.3 + 0.5y_{t-1} + \varepsilon_t, & \quad \varepsilon_t \sim N(0,1) \\ t = 501, \dots, 1000: & \quad y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, & \quad \varepsilon_t \sim N(0,1) \end{aligned}$$

consider Figure 5. In the first part SSR increases slightly, then there clearly is a turning point around observation 300, with after that an increase in SSR until around observations 500, after which the SSR starts to decline again. This pattern suggests that earlier data can be relevant to predict the future, and Figure 5 is a clear example of a case where “this time is the same again”.

A similar but slightly different example of the case “this time is the same again” emerges for the following DGP:

$$\begin{aligned}
DGP\ 4: \quad t = 1, 2, \dots, 300: \quad & y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1) \\
t = 300, \dots, 310: \quad & y_t = \varepsilon_t, \quad \varepsilon_t \sim N(0, 25) \\
t = 311, \dots, 1000: \quad & y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1)
\end{aligned}$$

In this case, we observe that the SSR experiences a very steep decline for α between 300 and 310. This is visualized in Figure 6. This pattern suggests that it is harmful for the fit of the model to include the period of outliers around $t = 300, \dots, 310$, but removing period $t=1, \dots, 300$ on the other hand harms the regression estimates as well. The reason for considering this specific DGP is because Figure 6 bears close resemblance to the patterns we shall observe in the actual data in the next section.

At last, we consider a fifth data generating process:

$$\begin{aligned}
DGP\ 5: \quad t = 1, 2, \dots, 550: \quad & y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,49) \\
t = 550, \dots, 650: \quad & y_t = 0.4 + 0.5y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, 81) \\
t = 650, \dots, 1000: \quad & y_t = 0.3 + 0.8y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1)
\end{aligned}$$

The idea behind this DGP is that a large part of the estimation sample ($t=1, \dots, 800$) is generated by a DGP with the same parameters (0.3 and 0.8) as in the evaluation sample, but with a very large variance. As such, the influence of the structural break ($t=550, \dots, 650$) is limited as long as α is small (far from 550). However, when α is close to 550-650, the parameter estimates become biased, deteriorating the model fit and leading to a large increase in the SSR. This has been visualized in Figure 7.

3. Illustration for monthly industrial production

In this section, we apply our simple method to data on Industrial Production, which we obtained from: <https://data.oecd.org/industry/industrial-production.htm>. The specific focus is on obtaining visual impressions whether older recessions' data would still be useful to forecast future recession data, and more in general, whether past data anyway seem to be useful. We use monthly seasonally adjusted data and we address the total index. To be able to draw some generalizing conclusions, we look at 17 different countries from 1965 January (1965M01) to 2016 July (2016M07). The goal is to analyse the forecast quality for the period 2006-2016, and

that of 2011-2016, based on an approximate AR(1) model for the first differences of the natural log transformed Industrial Production Index. We use one-step-ahead forecasts based on the actually observed variables.

Below, for each country, the sum of squared residuals is presented for the last 20% of the observations, which is the period 2006-2016 for all relevant values of α . On the x-axis, we do not see the value of α itself, but the date to which it corresponds. The data before that date get approximately weight zero, as we again set γ equal to 10.

In the bottom panel in each set of graphs, we do the same, but now the sum of squared residuals is presented for the period 2011-2016, so the last 10% of the observations, the period considered to be after the great recession.

In the center panel, we consider the SSR for the period 2006-2011, which includes the recent recession around 2008-2009, and its immediate aftermath.

For some of the countries, we present an interpretation of the graphs, which can be seen as an example as how to interpret the other graphs. Especially the pattern observed in Germany in Figure 8 is quite common across countries. The most remarkable countries are the Netherlands and Norway, as their graphs are very different from the patterns we observe across the other countries.

Individual countries

Let us first have a look at the results for Germany in Figure 8. When we include the financial crisis in our evaluation sample, which is presented in the top and center panel, we see that older data is of little added value, and we especially can observe a large improvement in the fit around 1985. The pattern seems to suggest that the period 1983-1985 has similarities to the evaluation sample. However, when we evaluate the fit after 2011, presented in the bottom panel of Figure 8, we see that down-weighting older data strongly deteriorates the fit, especially around 1985. Much older data has little added value in explaining the data during the financial crisis, but removing old data strongly deteriorates the fit for the period after the crisis, unless we also down-weight all data during the crisis. The observations in the bottom panel are very similar to what is observed in Figure 6 and Figure 7. The Germany data in the top and center panel have similarities to the simulated data in Figure 5. Taking everything together, it indicates that the crisis can be seen as a temporary change in DGP and the periods before and after can be described quite well by the same data generating process. The data before 1985 likely had a

larger variance, leading to observations with a high leverage and as such a good model fit when included in the estimation window, despite the period around the recent crisis which had a very different DGP. The date 1985 might well correspond to the ‘Great Moderation’, explaining the similarities to Figure 7.

For the Netherlands, in Figure 9, we observe very little differences between the three different evaluation windows. In the first and second graph we see that down weighting data up till 1985 strongly improves the fit, both for the larger sample 2006-2016 and the smaller sample 2011-2016. The pattern of the SSR bears close resemblance to the pattern in Figure 6, indicating that around 1985, there was a DGP different than the DGP in the evaluation windows. The second graph in Figure 9 suggests that the period 1970-1983 bears resemblance to the period 2006-2011, corresponding to the periods of the oil crisis (1973) and the early 1980s recession. The same holds for the period 1999-2001. On the other hand, the data around 1984, 1993 and 2002-2003 onward have a DGP suggestedly very different from the DGP around 2006-2011.

For Norway, in Figure 10, we obtain three graphs which are in some way comparable to those for the simulated data in Figure 7, where we also observed a sudden strong increase in the SSR followed by a steep decrease. This suggests that this period with a very high peak in SSR has a very different DGP than the periods around it and that the period before the peak has similar parameter values as the evaluation window, but a larger standard deviation. In contrast to the other countries we consider, this odd looking period concerns the years 2002 to 2004. The difference between the first/second and third graph can be interpreted as the difference between Figure 6 and Figure 7. The evaluation sample in the third graph of Figure 10 likely has a lower variability than in the first and second graph, such that the older data has relatively high variability and as such is important for the fit of the sample of 2011-2016, but less for the fit of samples containing the financial crisis.

For Sweden, we can observe in Figure 11 a pattern very similar to the Netherlands, where the data between 1975 and 1978 suggest to obey the same data generating process as in the period 2006-2011, as the SSR goes up in that period, but down weighting the data from 1978-1985 is very beneficial for the fit of the periods 2006-2016 and 2006-2011. So, for Sweden we see that the recent recession does look like the recession at the end of the nineteen seventies. In a similar sense, down weighting the data around the 1980s is bad for the fit of the model for the period 2011-2016. The period 2007-2009 is also very different from 2011-2016.

For Finland, Figure 12, we can observe that the period 1989-1992 seems to bear resemblance to the period after the financial crisis (because down weighting this data

deteriorates the fit), but would better be excluded when interested in forecasting the financial crisis itself. On the other hand, the financial crisis bears resemblance to the years 1988-1989. A similar effect is observable for the period 2002-2006; down weighting this period deteriorates the fit for 2011-2016, but improves the fit of 2006-2011.

Figure 13 presents the results for Luxemburg. We see that data between 1985-1997 are very influential for the fit of the last 10% sample 2011-2016, as down weighting these data leads to a much higher SSR.

For Italy, in Figure 14, we notice two remarkable increases, around 1968 and 1974. These periods apparently bear similarities to the DGP in the financial crisis. However, after 1980s, there is a convincing decline very similar to Figure 2, which would suggest that the data in 2006-2011 follows a different DGP than in 1980-1995, which is confirmed by the bottom panel, that indicates that the DGP in 1980-1995 was similar to the DGP in 2010-2016. The pattern in the bottom graph is similar to Figure 7, suggesting that the DGP around 2007-2010 was very different from 2011-2016.

The results for Belgium in Figure 15 show increases in SSR around 1974, 1991 and 2000. This suggests that the DGP in the periods 1974-1983 and 1991-1993 was similar to that of 2006-2011 and that periods 1983-1985 and 1994-1995 were very different from 2006-2011.. Once this data leaves the sample again, the SSR decreases. On the other hand, 1994-1995 was very similar to 2010-2016, as was period 2002-2004. This can be concluded by comparing these graphs to Figures 5, 6 and 7.

For the United Kingdom in Figure 16, the center graph demonstrates that the data around 1972-1974 and around 1979 has close resemblance to the DGP of 2006-2011. On the other hand, the periods 1975-1978 and 1981-1990 are noticeably different; downweighting them in the sample strongly decreases the SSR. These dates correspond to the 1973-1975 recession and to the early 1980s recession. Hence, deleting the recession data in those years reduces forecast performance leading to large increases in SSR for the recent financial crisis.

Figure 17 displays the graphs for France. To some extent these are similar to Germany, but a remarkable date around 1968, where removing data before this date strongly improves the fit for the sample 2006-2016, but decreases the fit for 2011-2016. We observe that when optimizing the fit for 2011-2016, there is great resemblance to Figure 7, suggesting that the data between 2007 and 2009 follows a very different DGP than the data between 2011-2016 and the data before 2007. The data before 2007 was most likely very variable, especially compared to the data after 2009, causing the pattern as was also observed in Figure 7.

For Spain, Figure 18, there is a slight upward trend in SSR in the period 1970-1973, followed by steep decrease. In the period 1977-1981, something similar happens. This is somewhat similar to the pattern in Figure 6. The data around the steep decrease then clearly follows a different DGP than the evaluation sample.

For Japan in Figure 19, we notice an increase around 1974, suggesting that data from 1974-1975 followed a DGP similar as in 2006-2011. So, here the crisis in those two periods seems similar with respect to the DGP of Industrial Production. The pattern in the bottom graph again is similar to Figure 7,

For the United States, in Figure 20, there are many increases in SSR, somewhat similar to Figure 5. More precisely, we can observe peaks around 1974, 1981 and 1992. These dates coincide with the 1973-1975 recession, early 1980s recession and the early 1990s recession in the United States. Comparing this with Figure 7, the pattern indicates that at the point where the SSR experiences a steep increase, data is removed that was important for the model fit and now data with a different DGP than in the evaluation sample becomes more important. Therefore, we can conclude that the DGP during these recessions must be somehow similar to the DGP around the crisis in the period of 2006-2011 and the periods right after the recessions are convincingly different. So, here we say that “this time things are similar”, that is, Industrial Production in previous crises in the US has a similar DGP as in the recent financial crisis.

For Portugal in Figure 21, we observe upswings in SSR around 1975, 1993, and 2004. Especially the increase around 1993 is large, suggesting that these periods before the upswing bear similarities to the DGP in the evaluation sample, while the period corresponding to the temporarily larger level of SSR and the steep decrease has a different DGP. The bottom panel can be interpreted in a similar way as Figure 7..

The results for Greece in Figure 22 can be interpreted as follows. When optimizing the fit for the sample 2006-2016, we should only use data after 2000, as leaving out the data before that year strongly improves the fit. When optimizing the fit for 2011-2016, we observe great resemblance to Figure 7, suggesting that the data around 2003 and 2008-2009 follow a DGP very different from the data generating process before 2003 and during 2011-2016.

Figure 23 displays the results for Austria, and there the first two graphs suggest that 1973-1975, 1979-1982 and 1988-1993 have a DGP very similar to the evaluation sample in these graphs. On the other hand, for the evaluation window 2010-2016 we observe a pattern as in Figure 7, suggesting a large difference between the DGP from 1996-2000, 2003-2010 and the evaluation sample.

Finally, the bottom graph for Canada in Figure 24 is quite similar to the series simulated in Figure 7, suggesting that the data between 2003 and 2009 follows a very different DGP. The most notable increase occurs for 2003-2009. This strongly suggests that these data have a very similar DGP as the data in 2010-2016. On the other hand, the data from 1999-2002 have a DGP similar to the period 2006-2010.

Taking all data together

Finally, we will turn to an analysis of all 17 countries into a single model. The data are the same as described before, but now we consider Seemingly Unrelated Regression (SUR). To the data in this SUR model, we will again apply the logistic weighting function. To estimate the SUR model, we will use Feasible Generalized Least Squares. The estimation of FGLS combined with the logistic weighted least squares function consists of the following steps.

1. First, we standardize all 17 time series. Next, we weight all data using the logistic function described above, for a specific choice of α .
2. Estimate the following multi-equation model, where m denotes the number of countries (17):

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_m \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} + u$$

3. Next, we use the residuals from this regression to estimate the elements of matrix Σ as follows:

$$\hat{\sigma}_{ij} = \frac{1}{n} \hat{u}_i \hat{u}_j$$

4. Then, we run the generalized least squares regression:

$$\hat{\beta} = (X'(\hat{\Sigma}^{-1} \otimes I_R)X)^{-1} X'(\hat{\Sigma}^{-1} \otimes I_R)y.$$

5. Next, to evaluate the SSR for the SUR model for the last part of the data, we compute the fitted values and look at the differences with the actual data. Next, for every country, we sum the squared residuals for the last part of the observations.

We perform these steps for multiple values of α of the weighting function and we consider both an approximate AR(1) and AR(2) model for all countries. Similar to the previous application, we look at the SSR for the last 20% of the observations (2006-2016), the period 2006-2011 and the period 2011-2016, for both model specifications. To limit the number of SUR regressions to perform, we only test a limited amount of α values (here corresponding to a frequency of once every half a year). The results are displayed in the graphs in Figures 25 and 26.

As can be seen, the graphs of the SUR model bear very close resemblance to most of the graphs for the individual countries. In general, we observe that when the goal is to optimize the fit between 2006-2011 and/or 2006-2016, we actually observe a pattern very similar to the pattern in Figure 4, with a normal DGP without breaks or changes. The most remarkable periods are around 1973-1975, 1979-1981 and 1989-1994, where we observe actual increases in the SSR, something that was observed in the simulations performed in Figure 5. This observation suggests that the periods from 1973-1975, 1979-1981 and 1989-1994 were very similar to the evaluation period 2006-2016, but the period directly followed by it was not. This supports the idea that the financial crisis period that falls into the evaluation sample was not that different after all. In fact, its DGP for Industrial Production is closely related to the three periods mentioned above.

When the goal is to optimize the fit between 2011 and 2016, the period after the recent crisis, we observe that the graphs suggest that we should only use data after 1980 and set the parameter for α in the weighting function correspondingly. We can easily see the similarities with Figure 7. This implies that the DGP of Industrial Production for these 17 countries was very different between 2008-2010 than it was in the period 2010-2016. This is the case both in the AR(1) model and AR(2) model.

4. Conclusion and discussion

This paper presented a simple method to diagnose whether one can do without older data when predicting the recent past (and the future). Easy to interpret graphs can be based on the outcomes of logistic weighted least squares. These graphs can also be used to analyze whether periods in time share the same data generating process or not.

For almost all countries, we noticed that it pays off the trouble to apply our method, as improvement in forecast quality is almost always to be gained, and, on the other hand, loss of forecast quality can be observed if some older data are discarded. For some countries we saw that we can learn from past crises about the most recent one, and hence that this time things are not different. In such a case, one can safely try to fit for example STAR models, where the models in the regimes are the same.

Figures

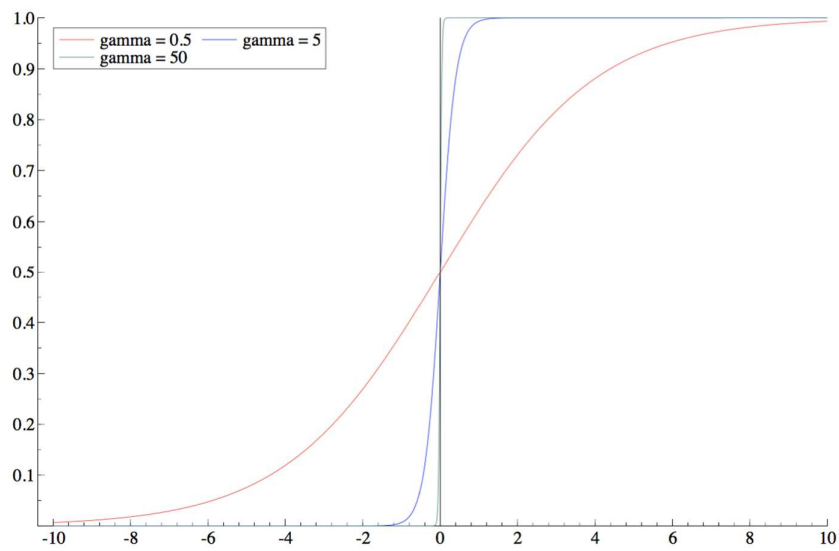


Figure 1: Logistic weight function for $\gamma = 0.5, 5$ and 50 . The parameter $\alpha = 0$.

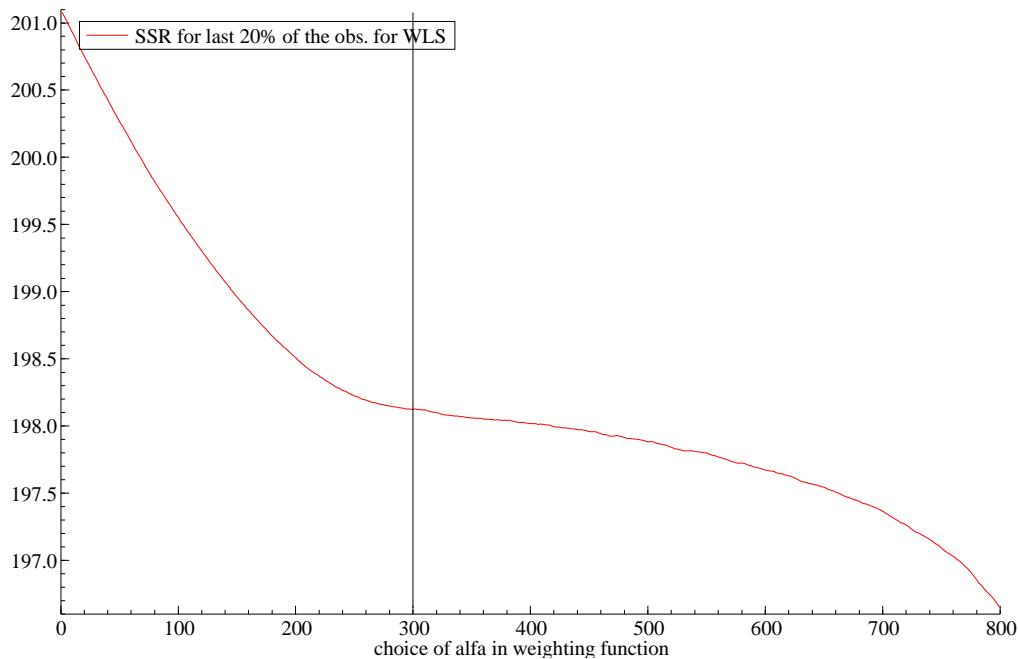


Figure 2. Sum of squared residuals for different values of α (see weight function for interpretation of α), α on x-axis. Average over 1000 simulation runs for DGP 1. Gamma in weight function is set equal to 10.

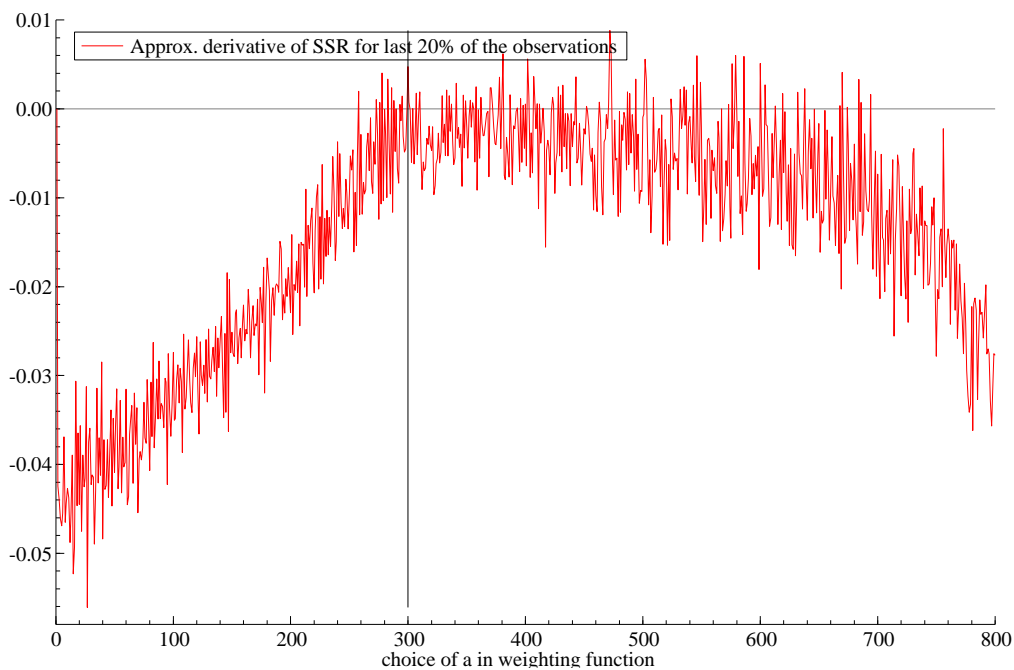


Figure 3. Approximation of the derivative of the function in Figure 2. Based on 1000 simulation runs for DGP 1.

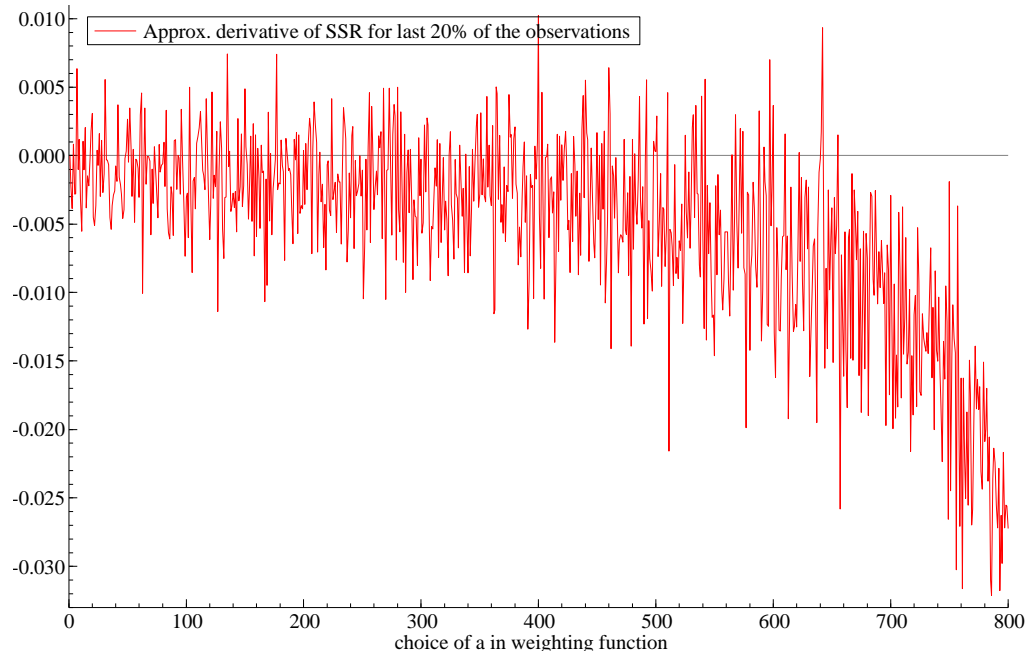
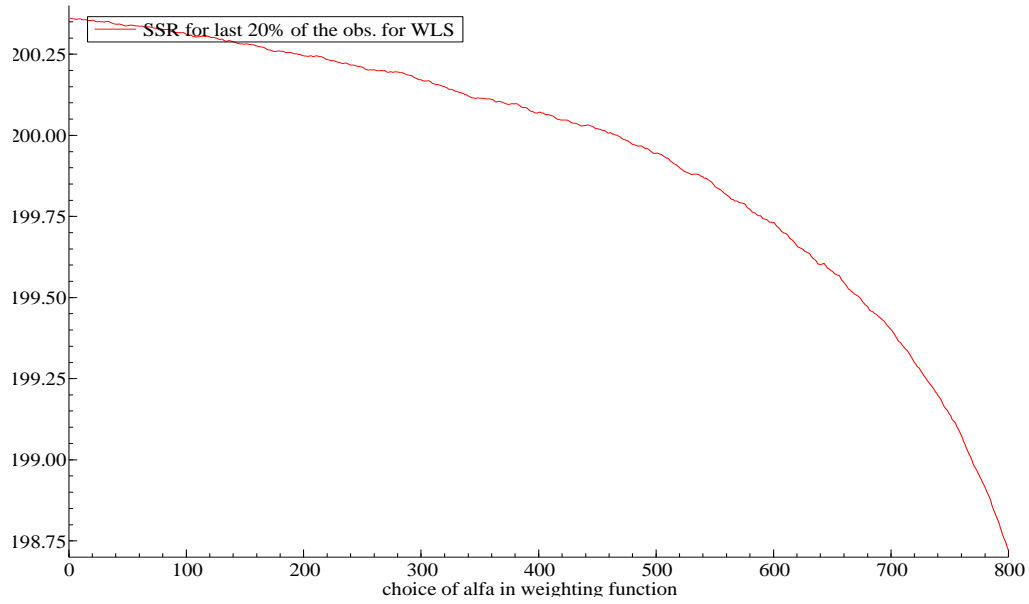


Figure 4. SSR for different values of α in the weighting function and approximation of the derivative of the SSR. Based on 1000 simulation runs for DGP 2.

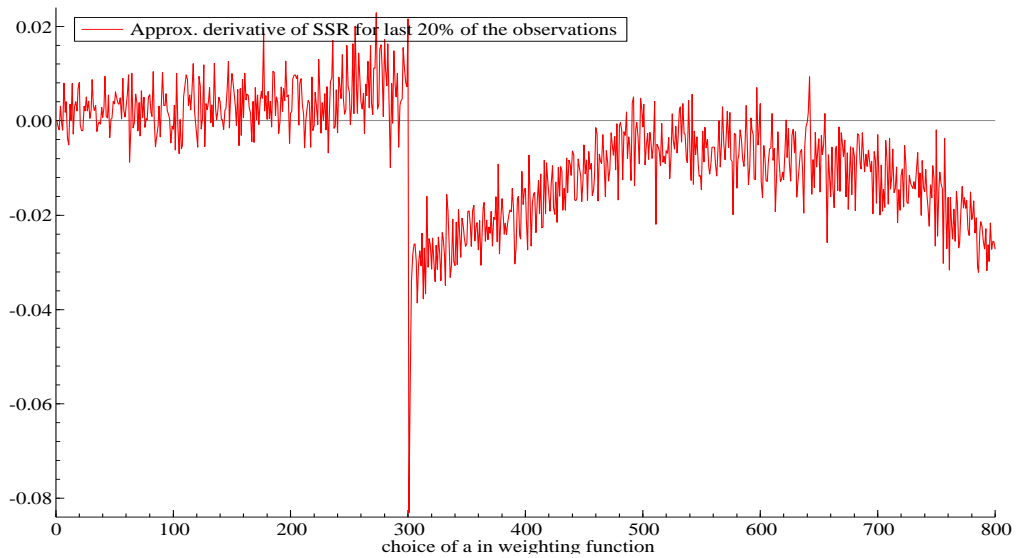
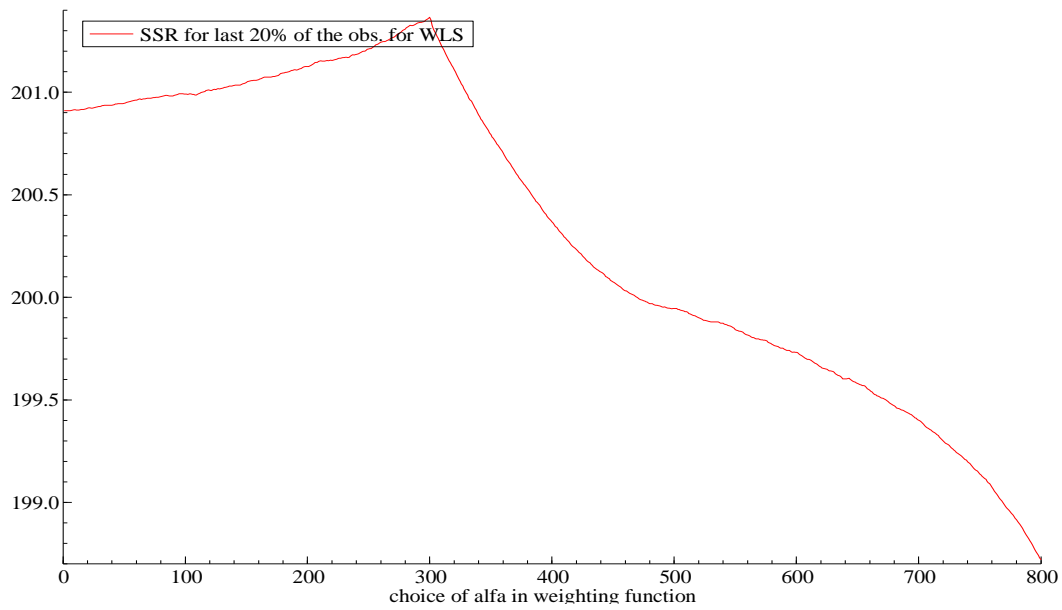


Figure 5. SSR for different values of α in the weighting function and approximation of the derivative of the SSR. Based on 1000 simulation runs for DGP 3.

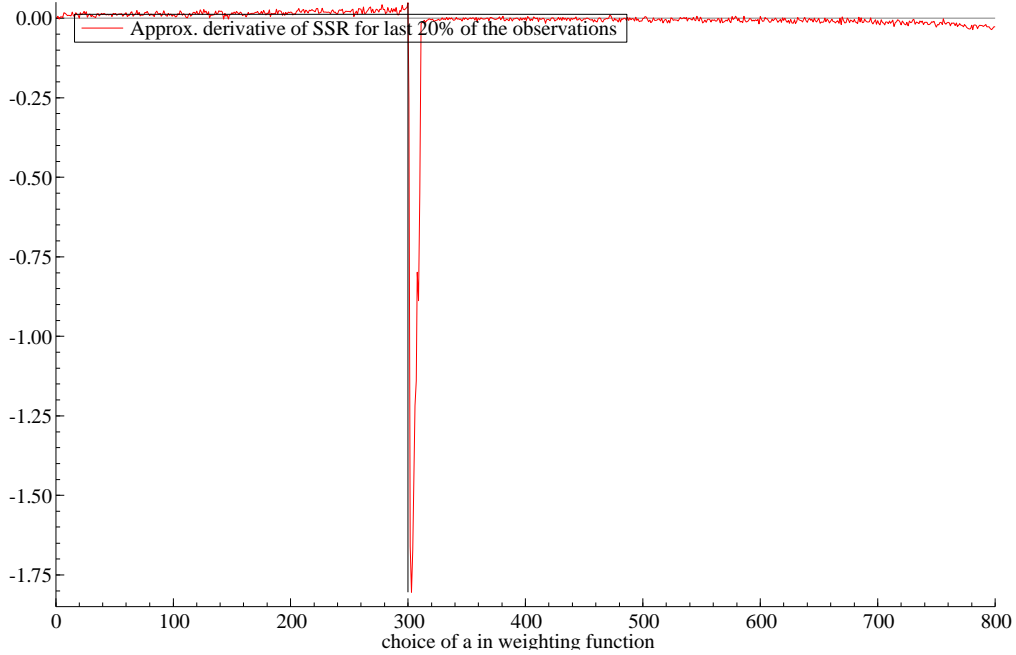
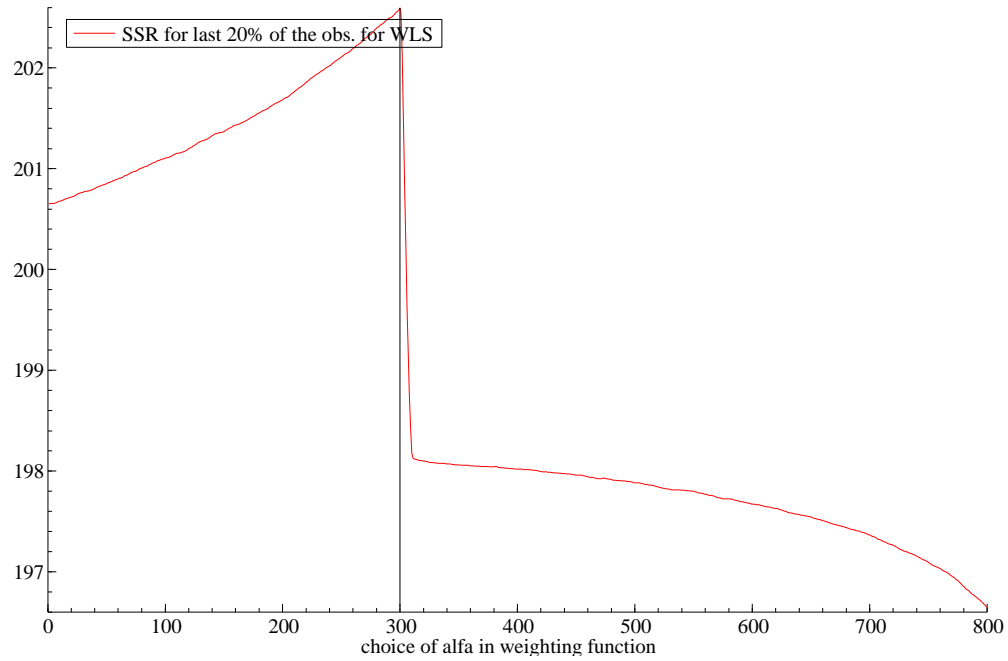


Figure 6. SSR for different values of α in the weighting function and approximation of the derivative of the SSR. Based on 1000 simulation runs for DGP 4.

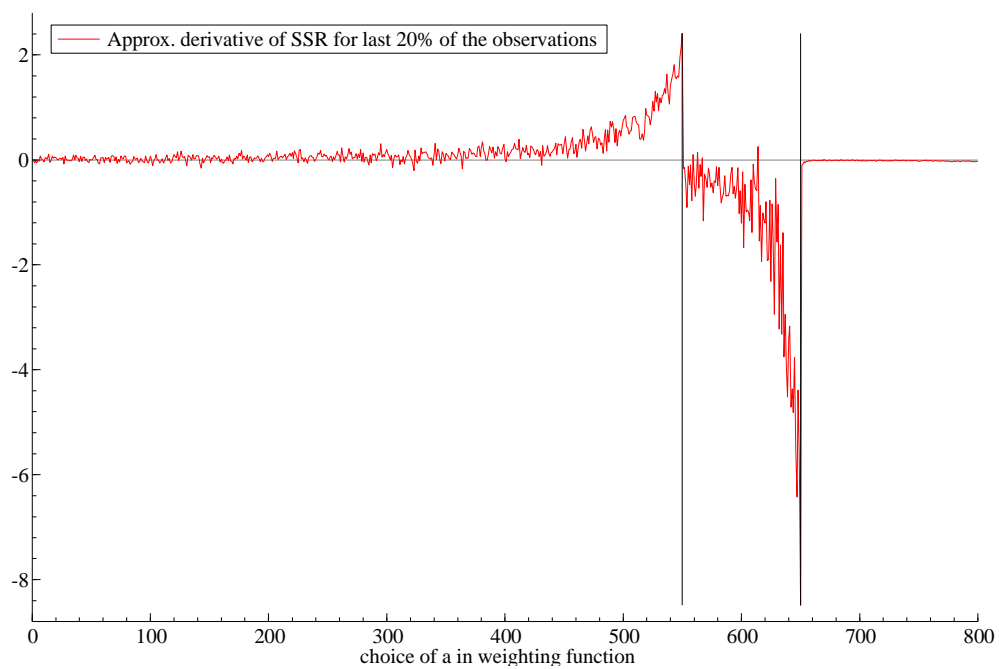
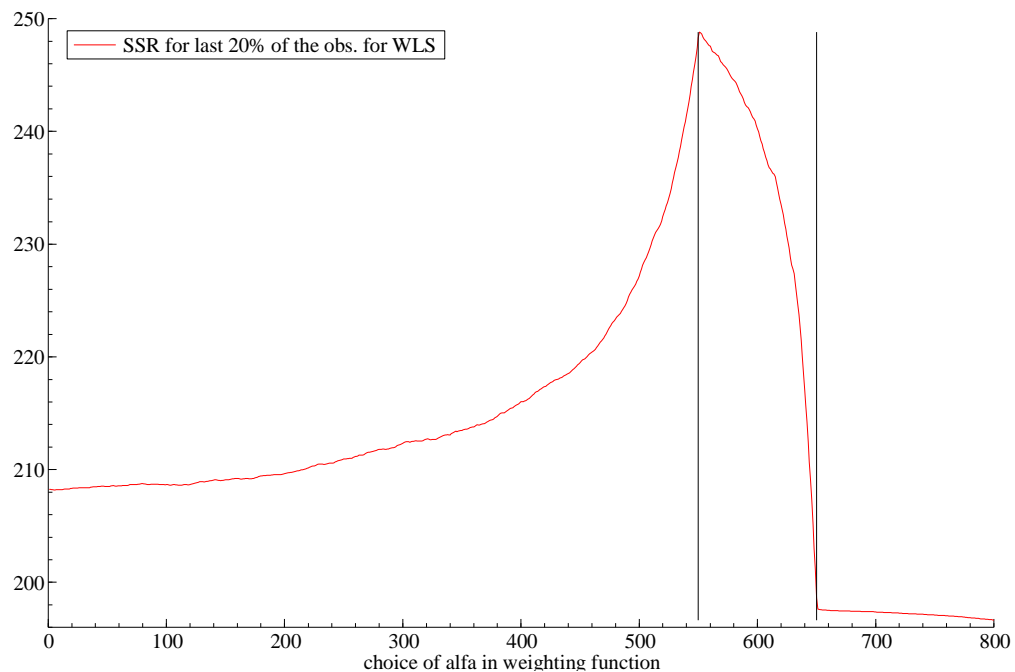


Figure 7. SSR for different values of α in the weighting function and approximation of the derivative of the SSR. Based on 1000 simulation runs for DGP 5.

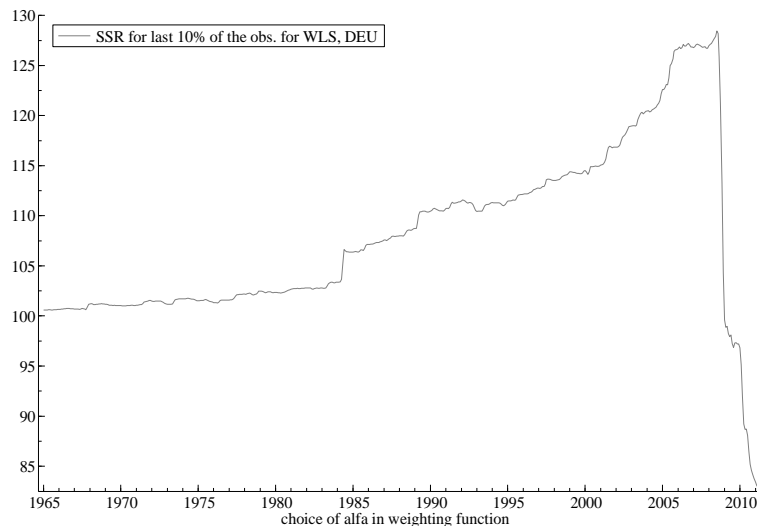
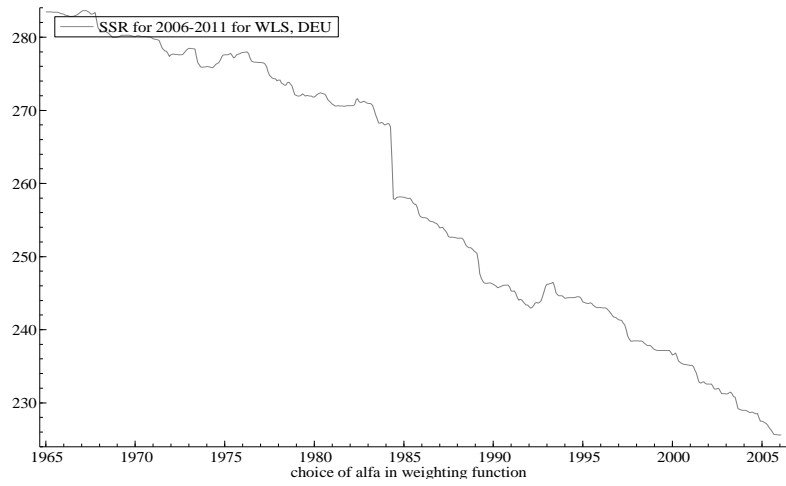
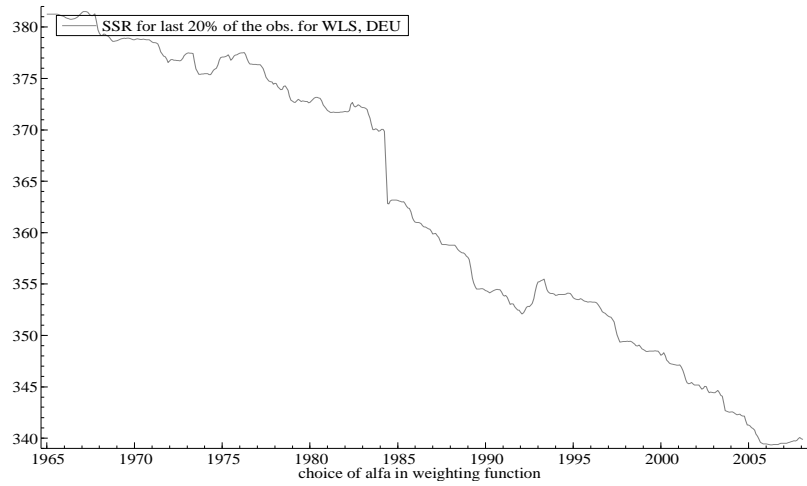


Figure 8: Germany

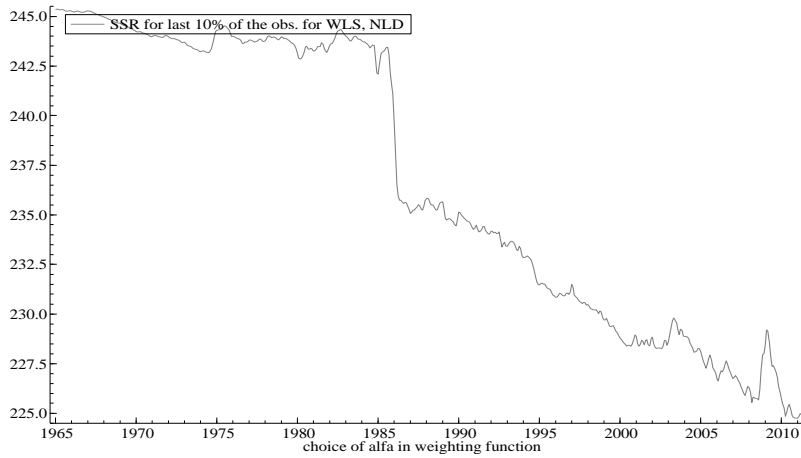
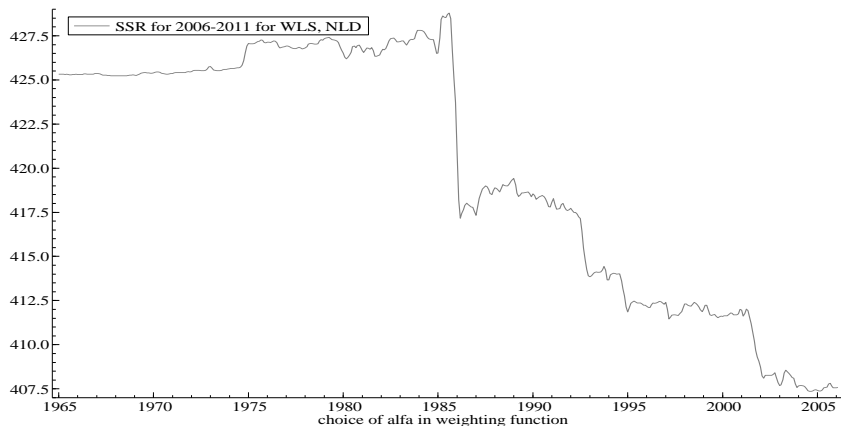
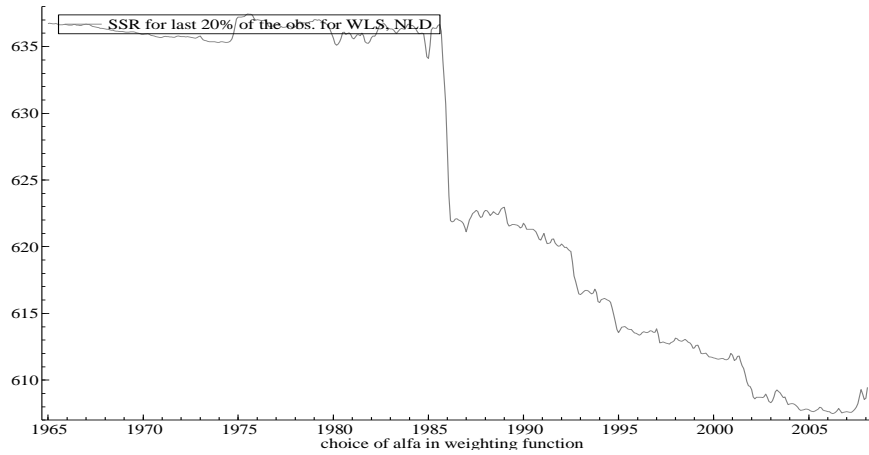


Figure 9: The Netherlands

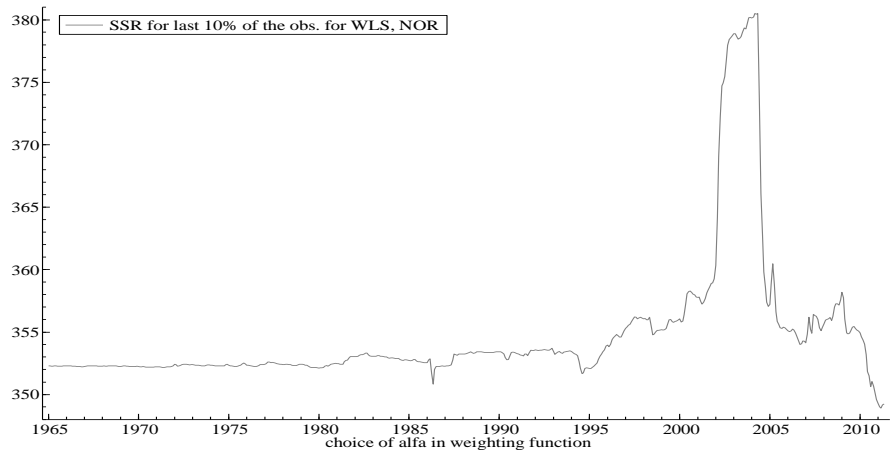
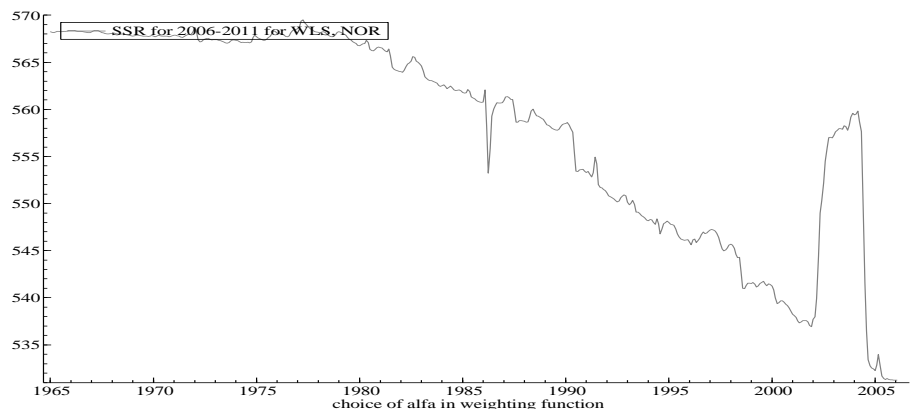
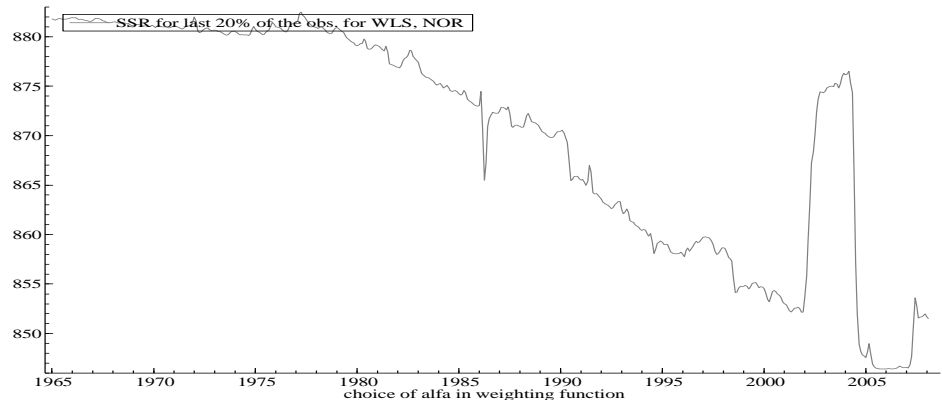


Figure 10: Norway

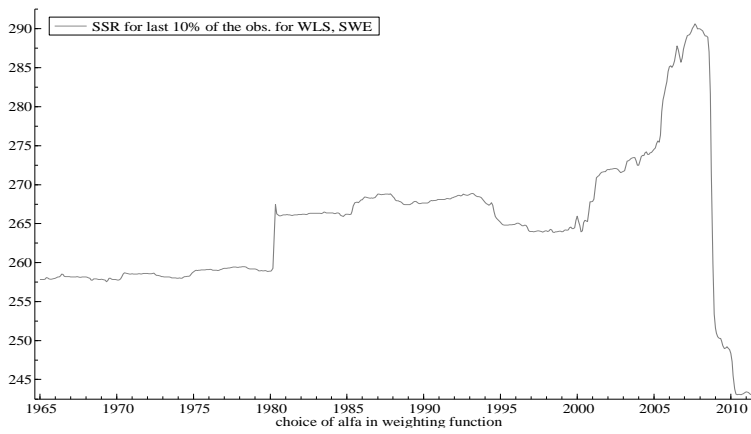
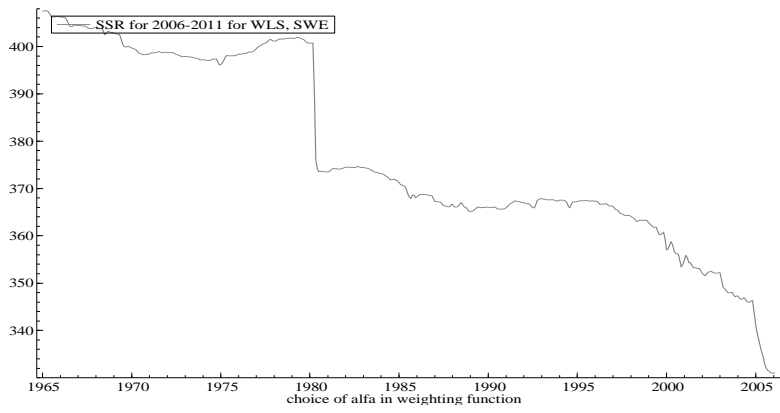
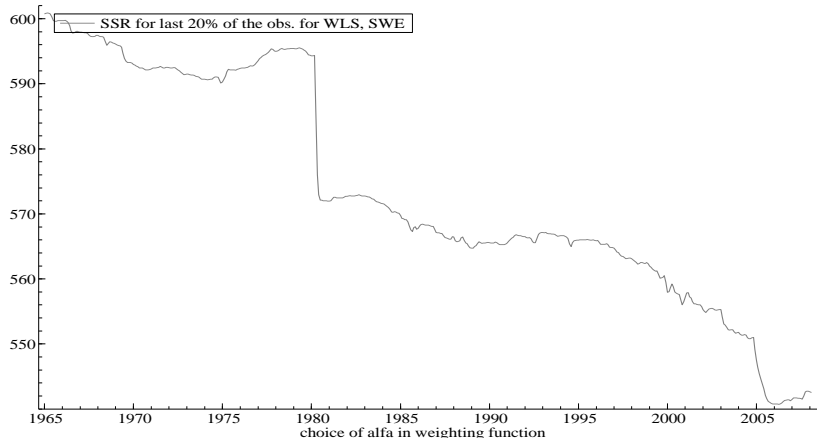


Figure 11: Sweden

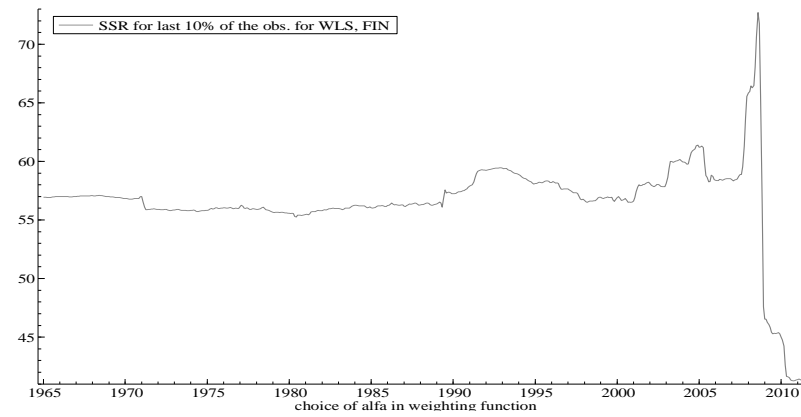
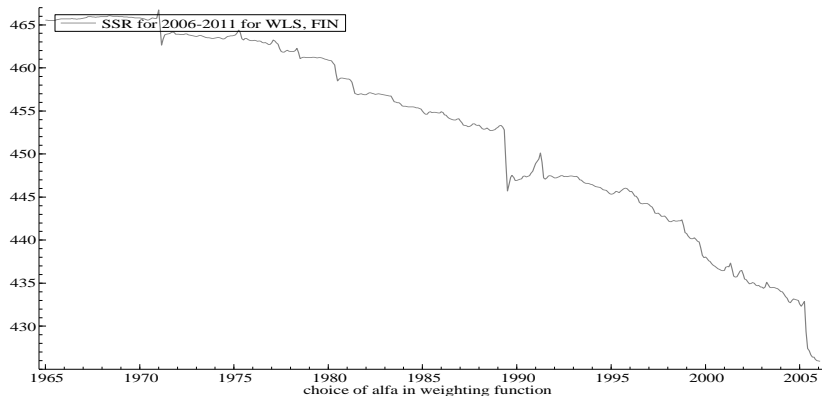
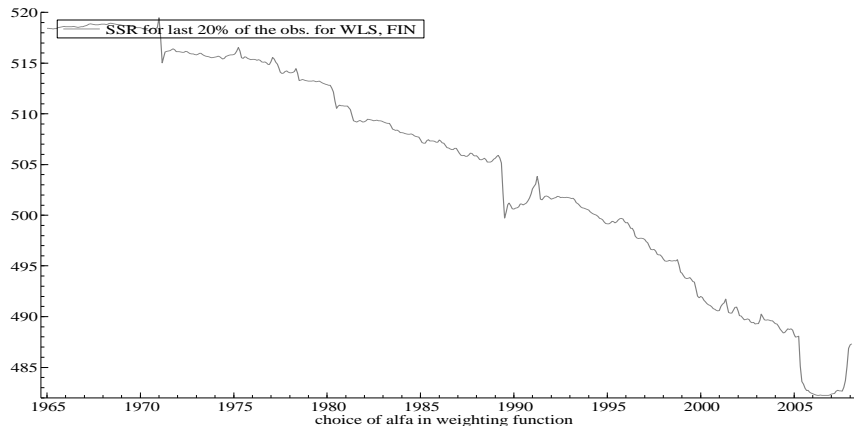


Figure 12: Finland

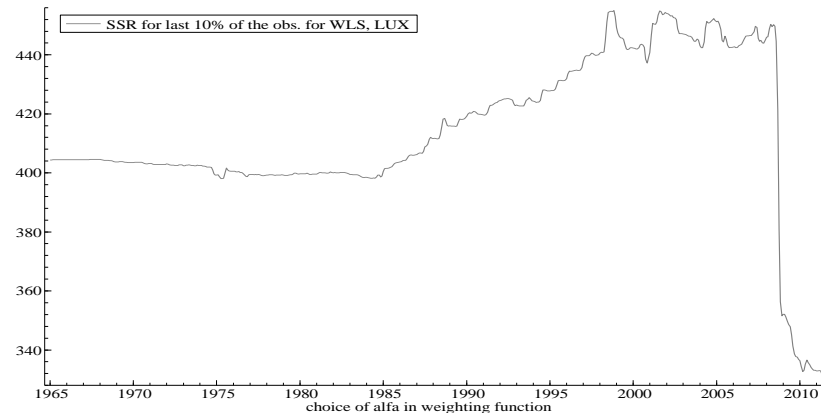
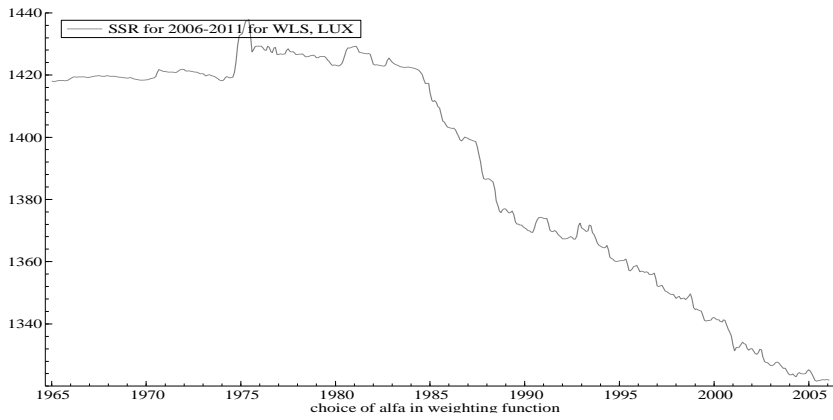
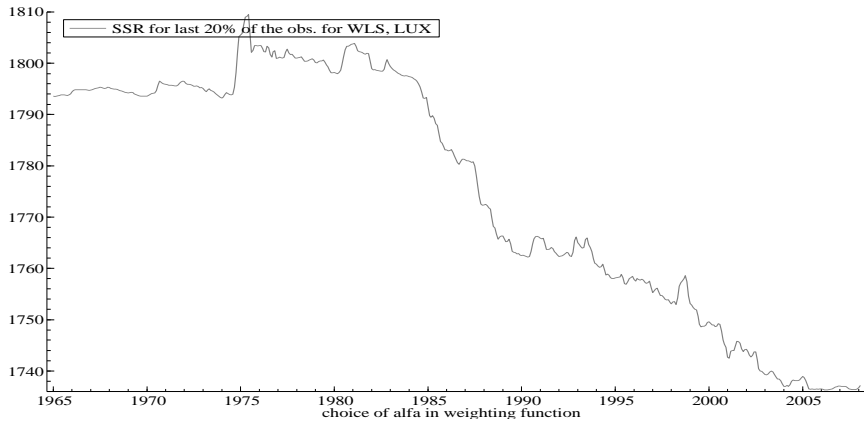


Figure 13: Luxemburg

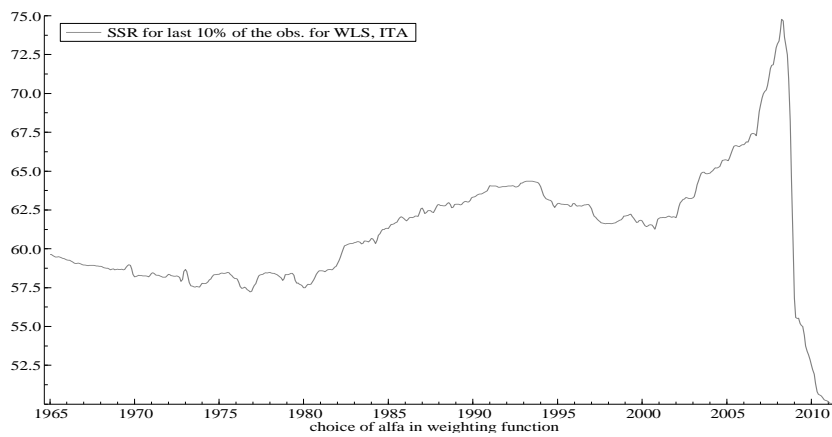
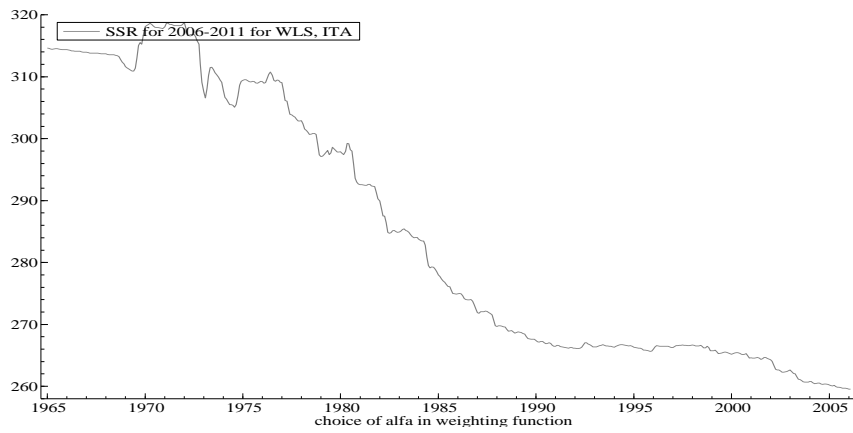
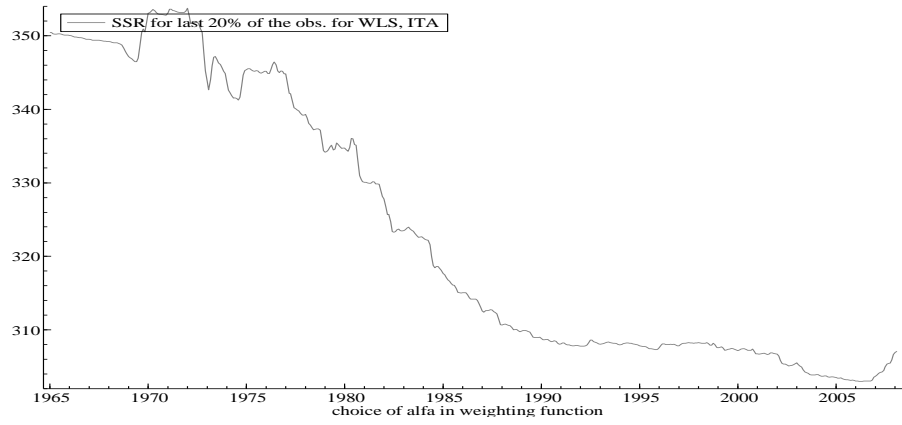


Figure 14: Italy

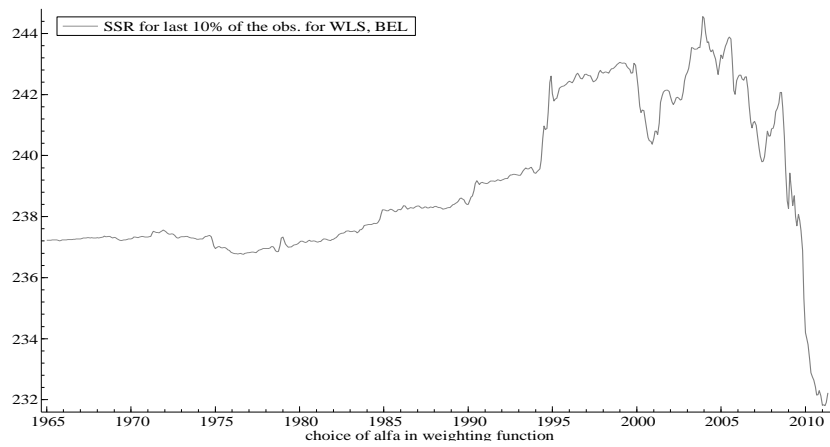
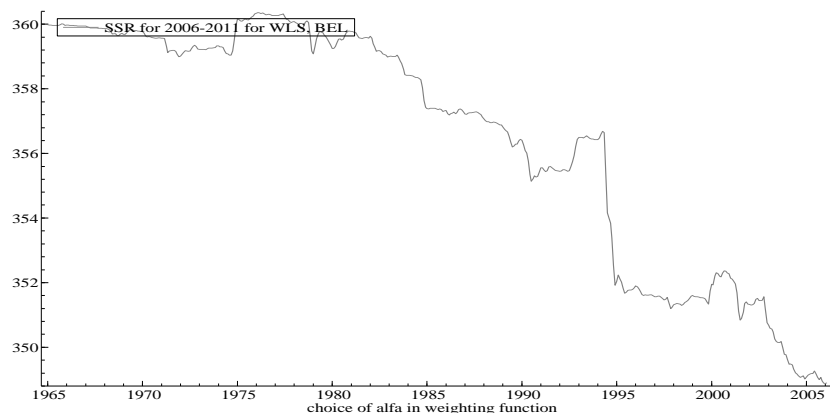
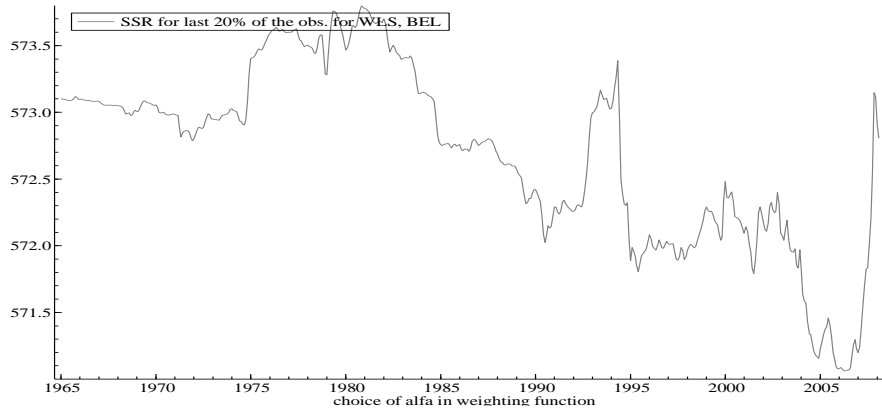


Figure 15: Belgium

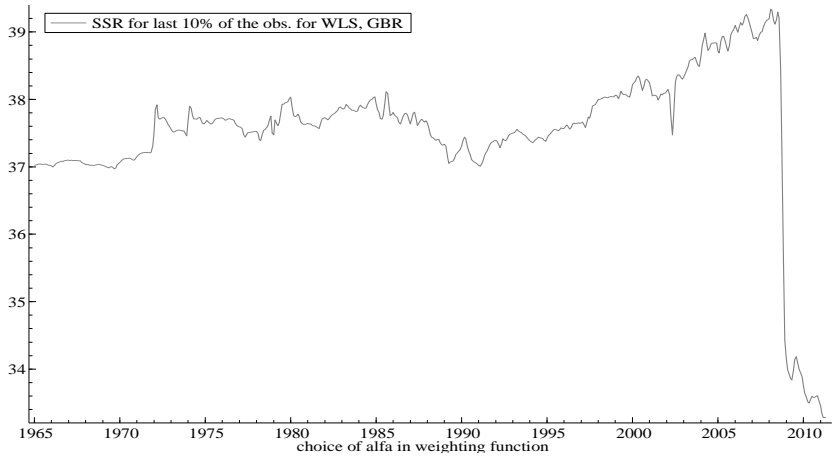
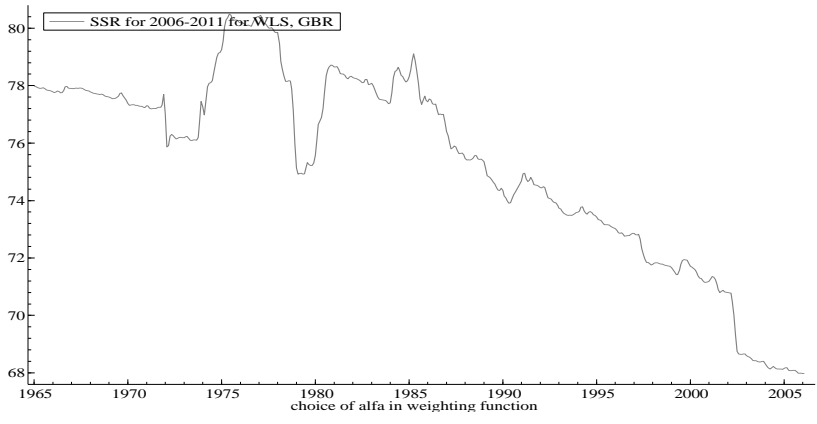
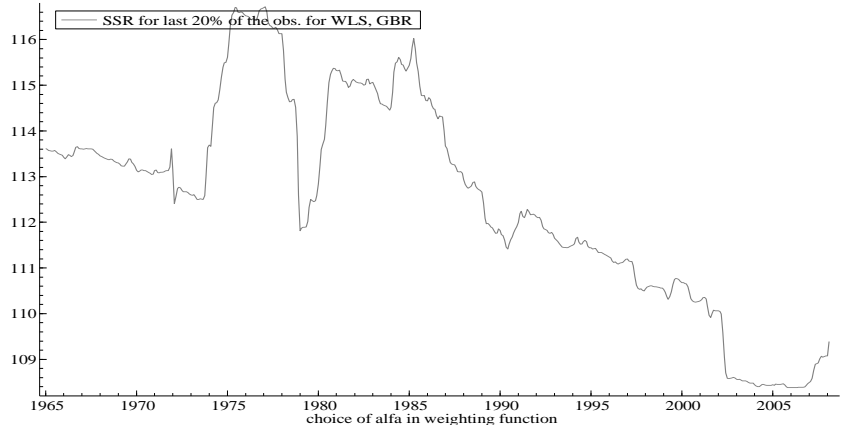


Figure 16: United Kingdom

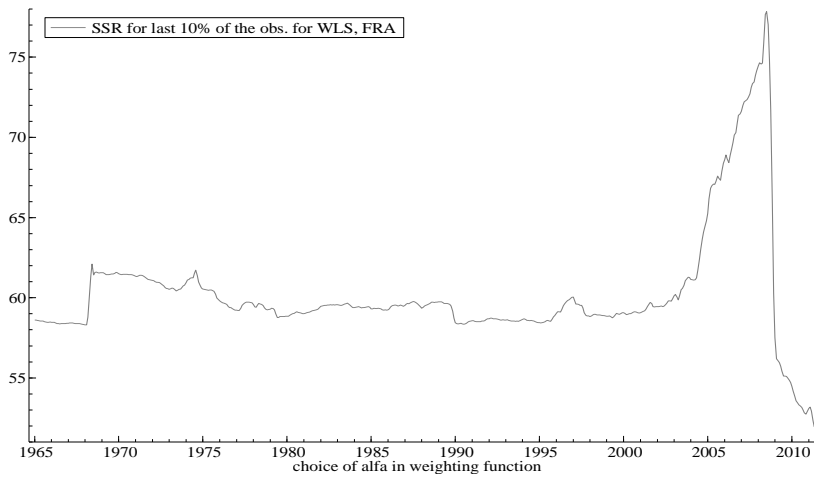
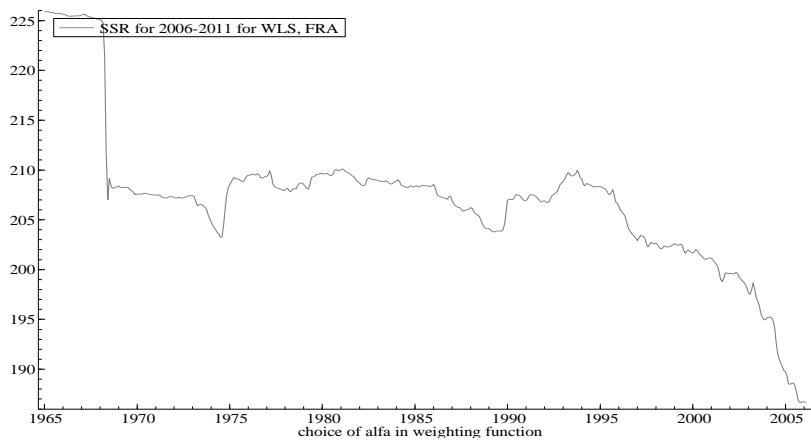
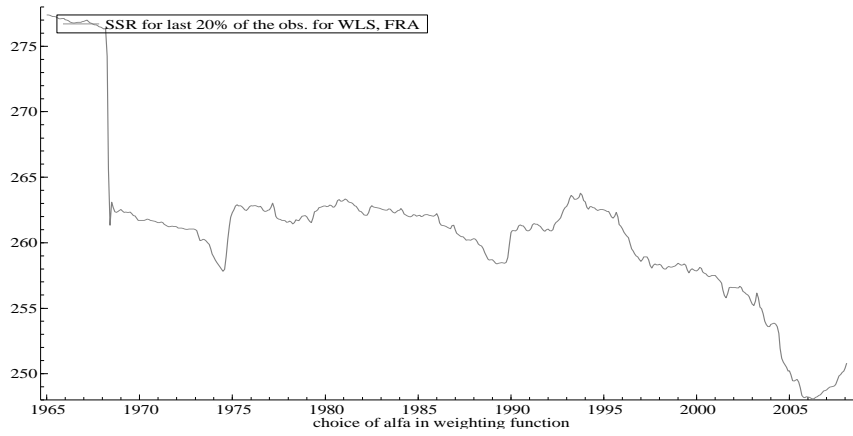


Figure 17: France

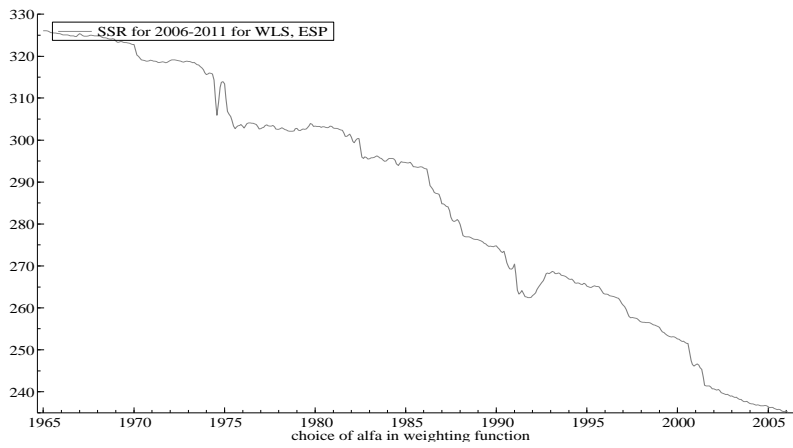
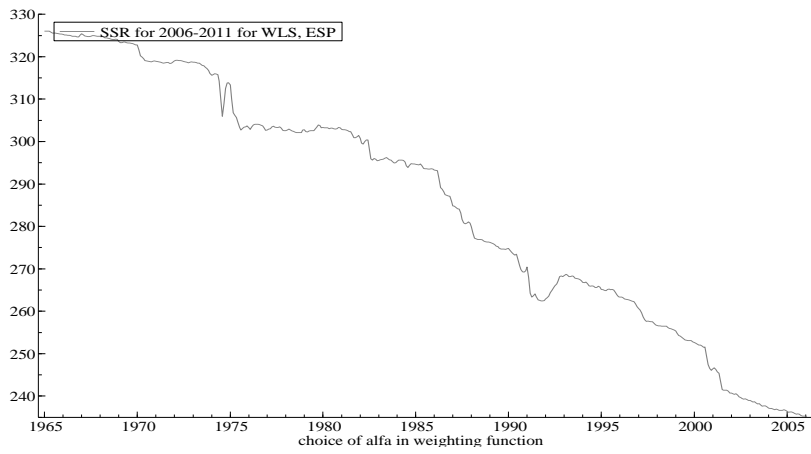
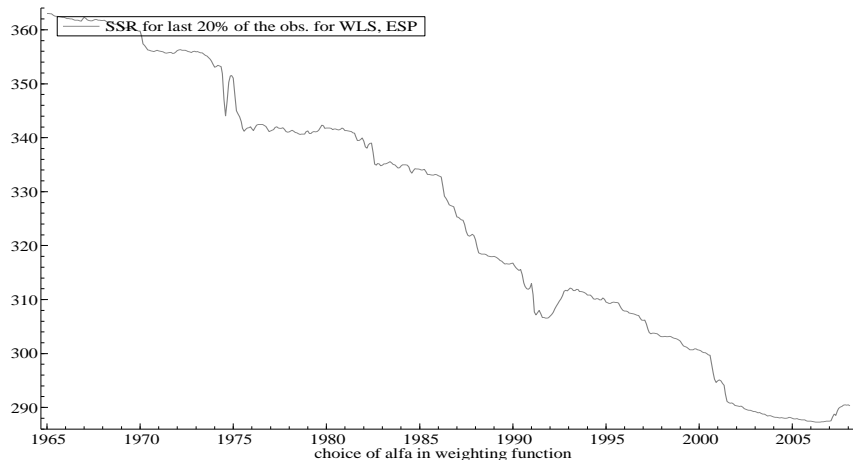


Figure 18: Spain

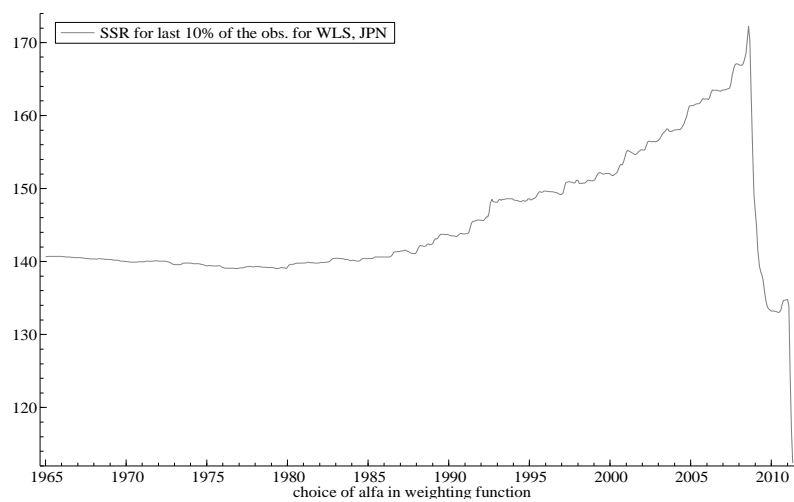
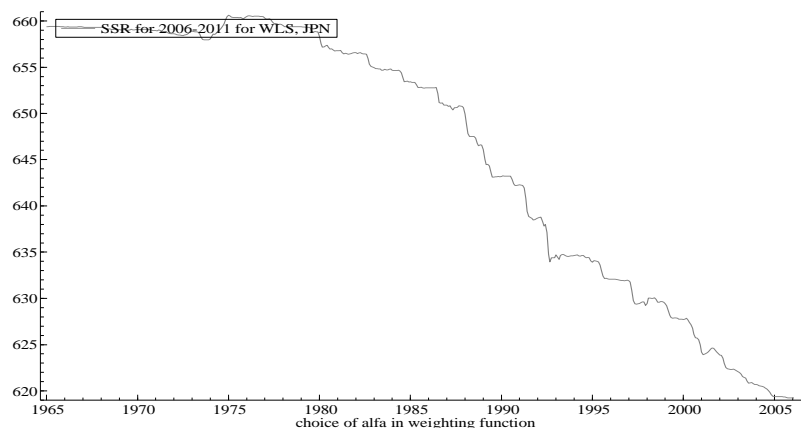
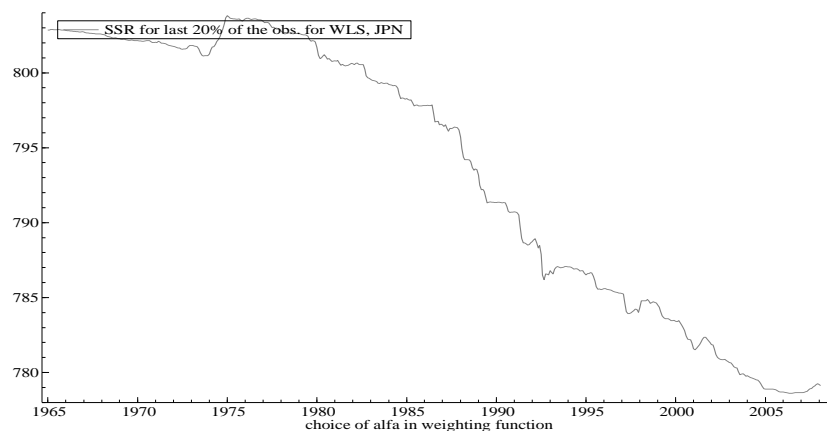


Figure 19: Japan

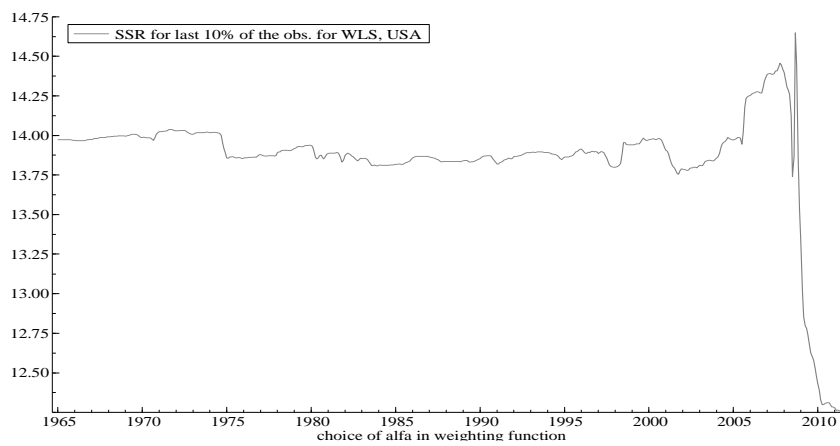
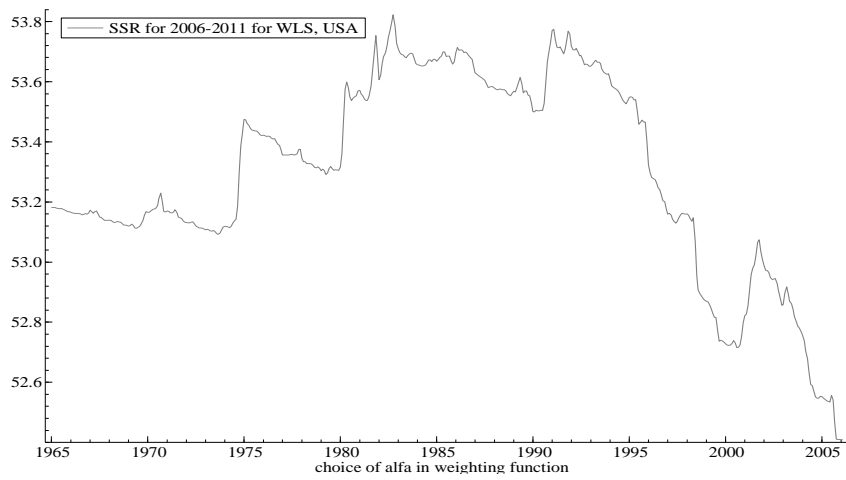
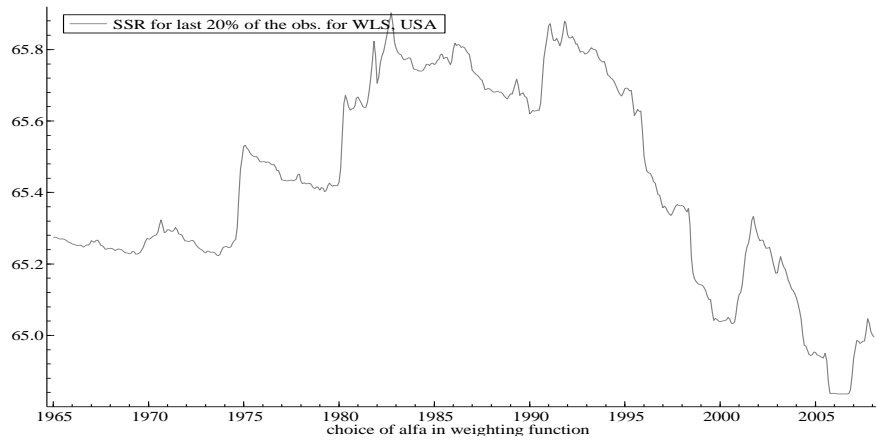


Figure 20: United States

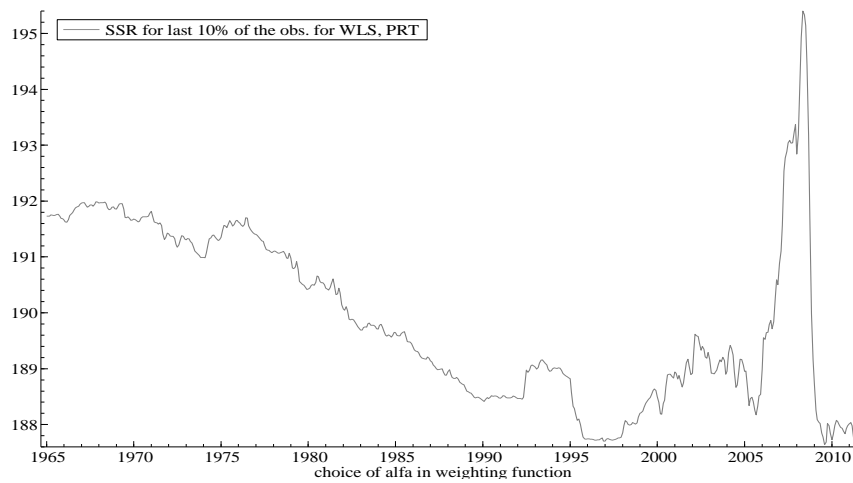
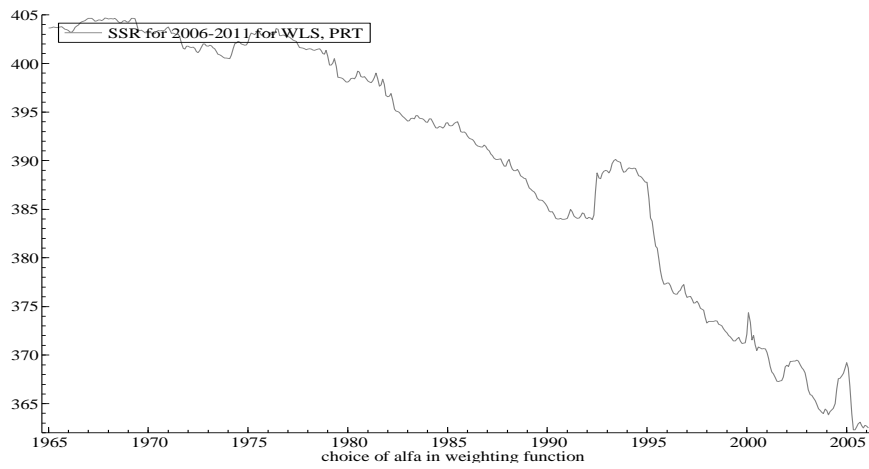
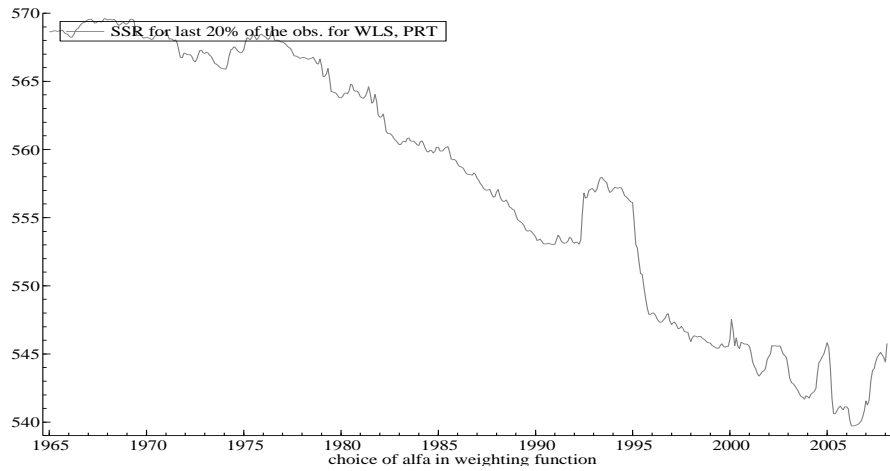


Figure 21: Portugal

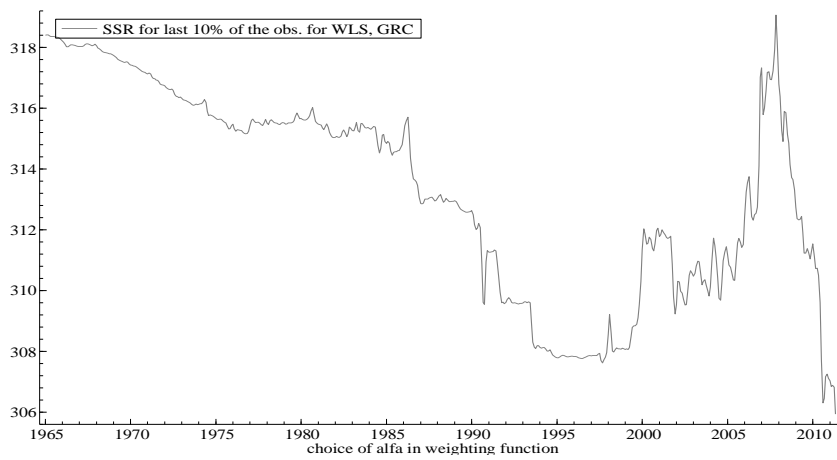
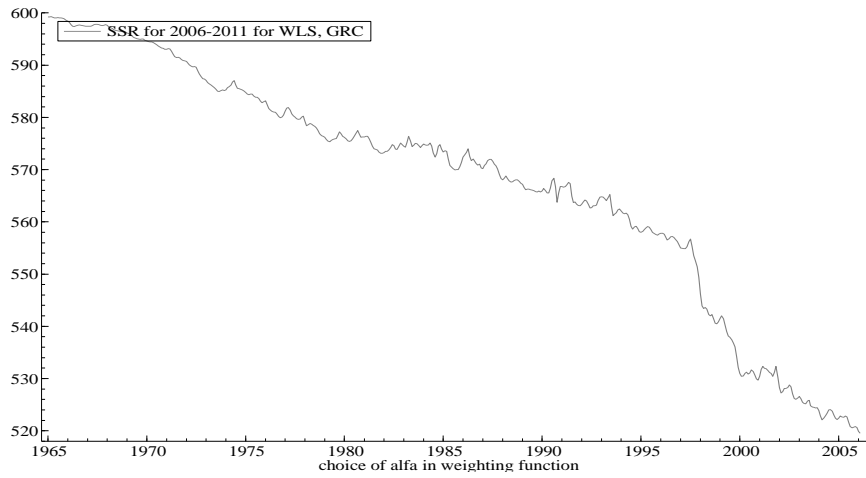
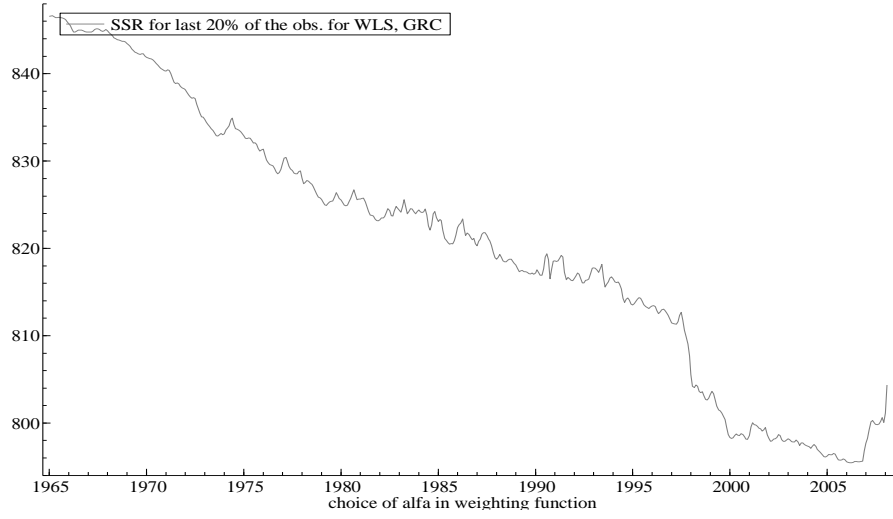


Figure 22: Greece

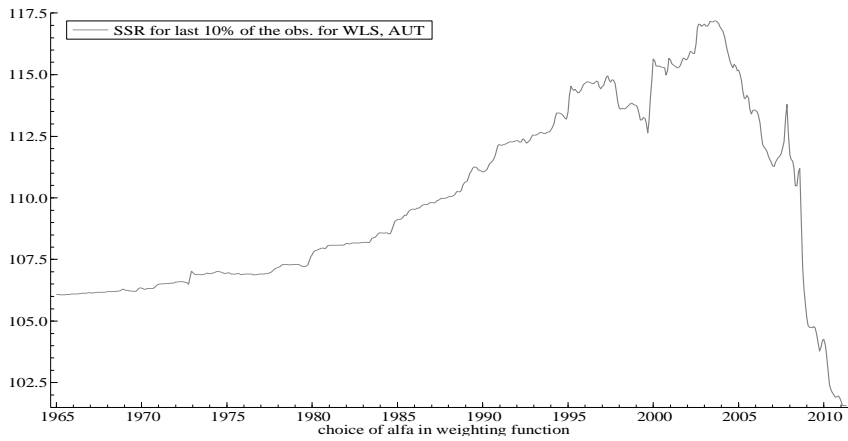
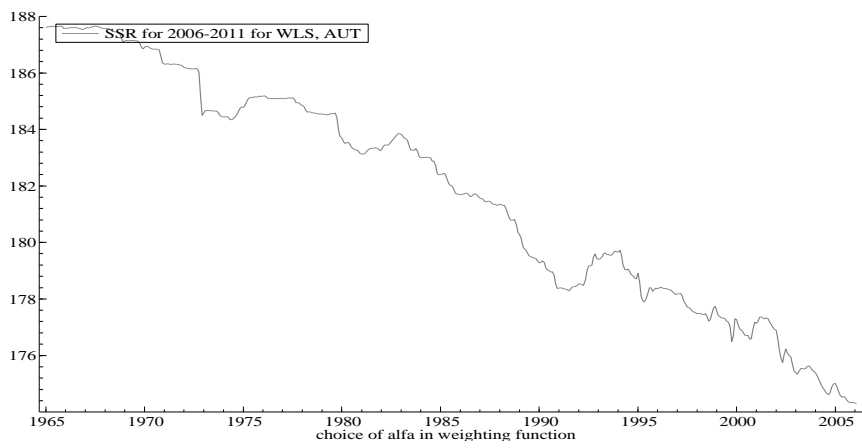
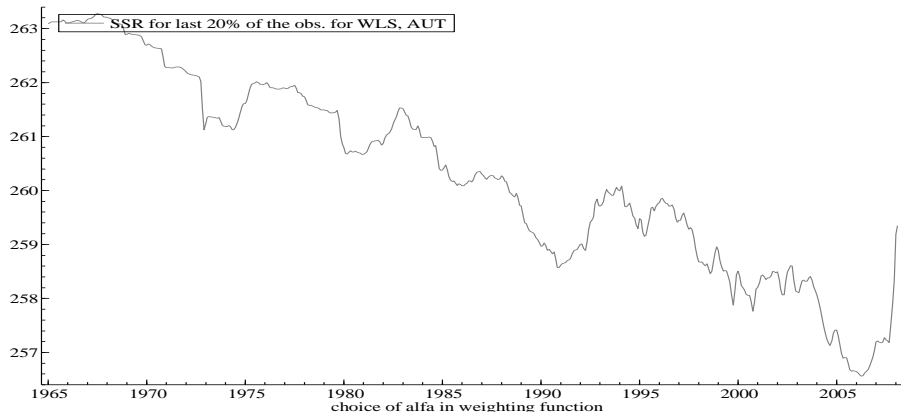


Figure 23: Austria

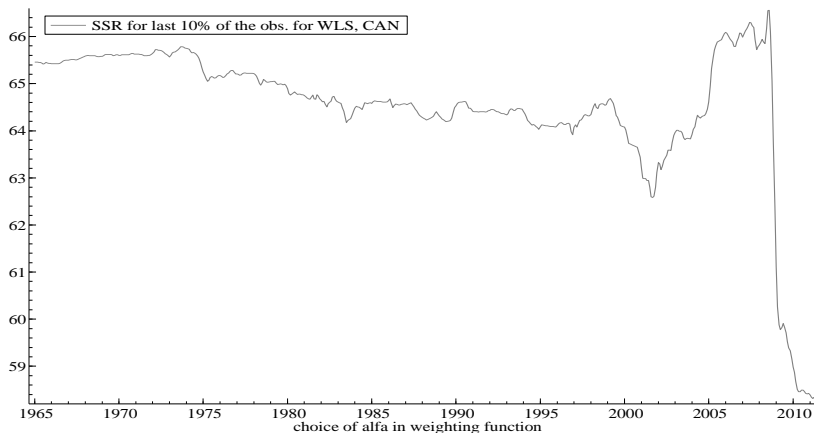
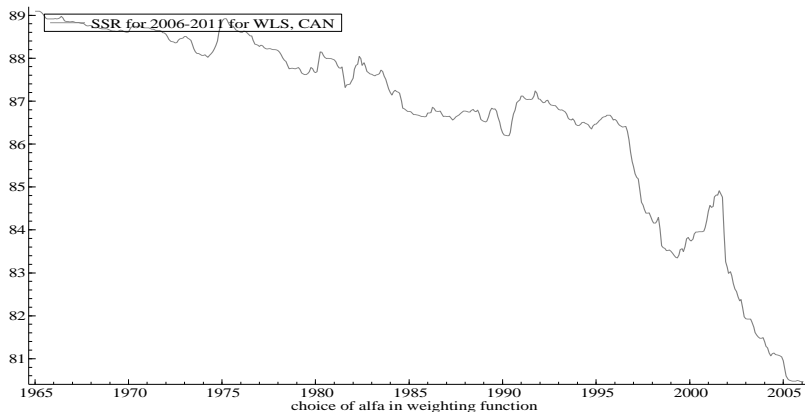
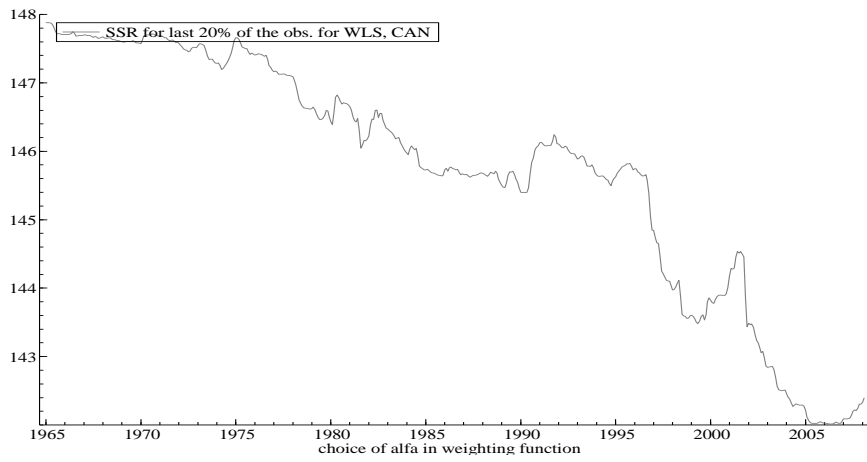


Figure 24: Canada

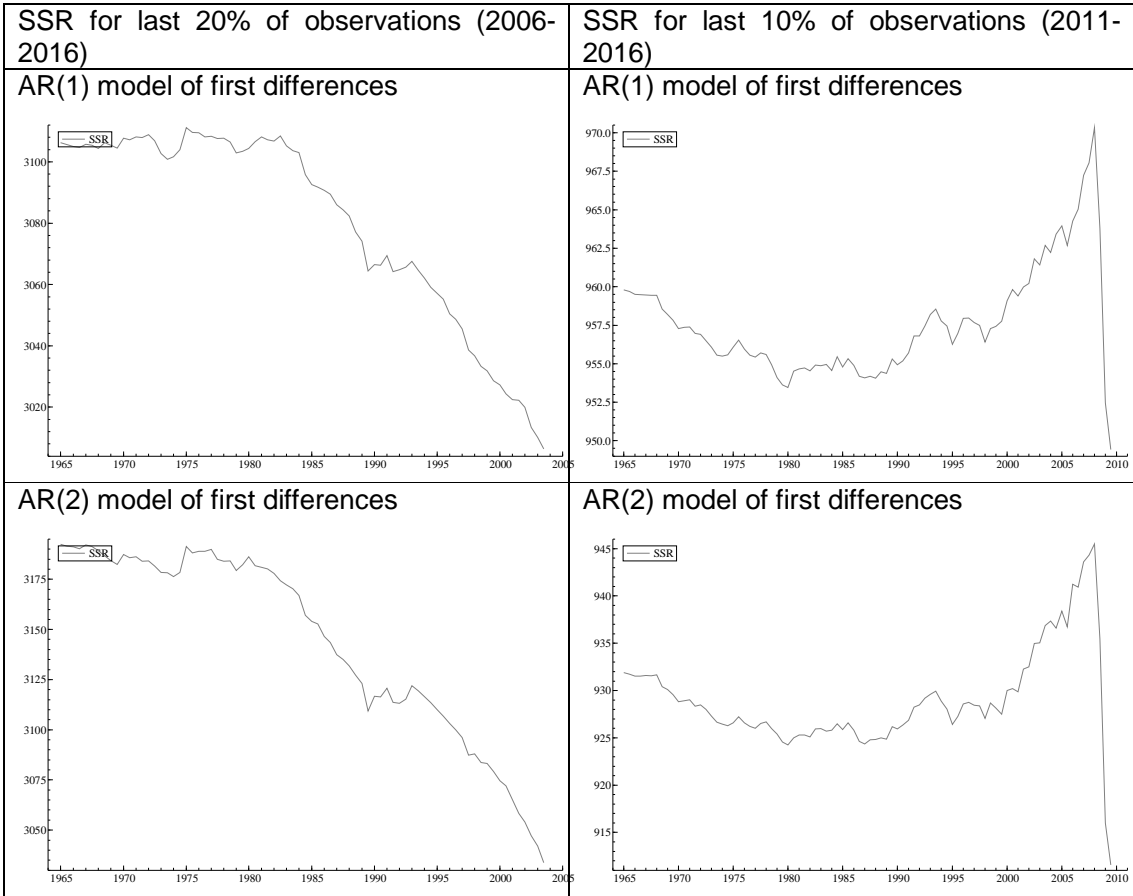


Figure 25: SUR results

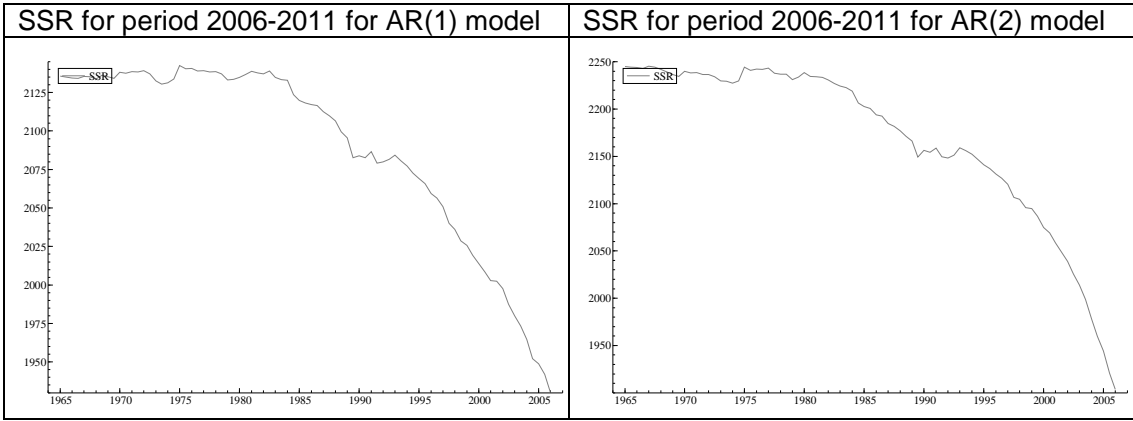


Figure 26: SUR results

References

Giraitis, L., G. Kapetanios, and S. Price (2013), Adaptive forecasting in the presence of recent and ongoing structural change, *Journal of Econometrics*, 177, 153-170.

Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*, Oxford: Oxford University Press.

Hamilton, J.D. (1989), A new approach to the economic analysis of nonstationary time series subject to changes in regime, *Econometrica*, 57, 357-384.

Inoue, A. L. Jin, and B. Rossi (2017), Rolling window selection for out-of-sample forecasting with time-varying parameters, *Journal of Econometrics*, 196, 55-67.

Judd, J. P., and Rudebusch, G. D. (1998), Taylor's Rule and the Fed: 1970-1997, *Economic Review-Federal Reserve Bank of San Francisco*, 3.

Refenes, A.N., Y. Bentz, D.W. Bunn, A.N. Burgess and A.D. Zapranis (1997), Financial time series modelling with discounted least squares backpropagation, *Neurocomputing*, 14, 123-138.

Van Dijk, D.J.C., T. Teräsvirta, and P.H. Franses (2002), Smooth transition autoregressive models – survey of recent developments, *Econometric Reviews*, 21, 1-47.