Time-varying coefficient methods to measure inflation persistence

Zsolt Darvas and Balázs Varga•

August 7, 2009

Abstract

This paper studies inflation persistence with time-varying-coefficient models. To this end, using Monte Carlo methods we compare the statistical properties of the well known maximum likelihood estimation using the Kalman-filter to the less known Flexible Least Squares estimator. We also suggest a procedure for selecting the weight for FLS based on an iterative Monte Carlo simulation technique calibrated to the time series in question. We apply the methods for the study of inflation persistence of the US, the euro-area and the new members of the EU.

Keywords: flexible least squares, inflation persistence, Kalman-filter, time-varying coefficient models

JEL classifications: C22, E31

• Darvas is Assistant Professor, Department of Mathematical Economics and Economic Analysis, Corvinus University of Budapest, e-mail: zsolt.darvas@uni-corvinus.hu. Varga is Ph.D. Candidate, Corvinus University of Budapest, e-mail: balazs.varga@ uni-corvinus.hu. A current version of the paper, the GAUSS codes used, and the data set are available at http://www.uni-corvinus.hu/darvas. Financial support from OTKA grant No. K76868 is also acknowledged.

1 Introduction

It is widely accepted that inflation is often exposed to numerous macroeconomic shocks which pull it away from its mean, which is generally identified by the central bank's inflation target. Shocks can be persistent or could have persistent effects on inflation due to, for example, nominal rigidities, leading to persistent deviations of inflation from its target. Therefore, knowing the persistence of these shocks and inflation deviations from target plays an essential role for the central bank whose primal aim is to achieve price stability. The adjustment of inflation towards its long run level after a shock can be characterized by the speed with which it converges back to its mean. The higher this speed is, the less complicated may be the central bank's price stability maintaining task. Inflation persistence (IP) is a measure of this convergence speed, based on different kinds of properties of the impulse response function within the model built to describe inflation.

Although the analysis of IP in the euro area and the US has received much attention[1], there has been almost no research regarding the new members of the EU. Understanding inflation persistence (IP) in the new EU members is crucial for the central banks of these countries on the one hand, and has implications for their future euro area entry on the other.

Inflation persistence has been studied by various models, ranging from simple autoregressions to well-structured dynamic general equilibrium models. In studying univariate autoregressive time series models, many authors found very high persistence or even could not reject the hypothesis of a unit root for a 50 year long sample stretching from the post world war II era, both in the US and in the euro-area. Latter studies found that inflation series have several structural breaks[2] and most of these could be explained by corresponding historical events, for example, the oil crises of the 70's. When studying the properties of the estimated autoregressive models for sub-periods identified by the break points, persistence turned out to be significantly smaller, particularly in the more recent periods. Hence, inflation persistence could be changing in time. Naturally, a change in inflation persistence could be the result of

[1] Eurosystem central banks even set up an Inflation Persistence Network (IPN) and directed substantial resources for the study of various aspects of IP; see Altissimo, Ehrmann and Smets (2006) for a summary of IPN. A debate whether IP has declined in the US has also received much attention in the academic literature; see, for example Pivetta and Reis (2007) for a recent work and a summary of the debate.

[2] Pivetta and Reis (2007) challenge this view and claims that IP was reasonably stable in the post WW-II US.

(a) change in the type of underlying shocks, (b) change on the persistence of the underlying shocks, (c) change in the monetary policy reaction function, (d) change in the way the economy responds to shocks or monetary policy actions, or (e) the fact that a linear approximation of an otherwise non-linear underlying structure is poor. A univariate autoregressive model estimated on different samples can not discriminate among these alternatives. Obviously, a time-varying coefficient autoregression also can not discriminate among these alternatives, but allows us to investigate changes in persistence more accurately and particularly, to highlight the dating and amplitude of breaks. Time-varying coefficient models were used for either or both the euro area and the US, for example, in Cogley and Sargent (2001, 2005), Dossche and Evaraert (2005) and Pivetta and Reis (2007); all of these papers adopted a Bayesian estimation framework.

Time-varying coefficient analysis of IP in the new member states of the EU seems inevitable. Ten of the twelve of these countries went through substantial structural changes when transformed their economies and institutions from a socialist to a market one. The transformation process was a gradual one and the economies of these countries probably still changing in a faster pace than mature economies. These arguments imply that it is rather difficult to set a date from which constancy of the parameters could be assumed on safe grounds.

This paper has two main goals. First, we compare the statistical properties of two time-varying coefficient (TVC) methods: the maximum likelihood estimation of the state-space representation of TVC autoregression using the Kalman-filter, and the less known, distribution-free estimator of the same model via Flexible Least Squares (FLS), which was suggested by Kalaba and Tesfatsion (1988, 1989, 1990). We also propose a procedure to set the weight parameter needed for the FLS: to our knowledge, such a procedure is not yet available. The proposed procedure is based on an iterative Monte Carlo simulation technique calibrated to the time series in question. Second, we study inflation persistence in four new members of the EU (Czech Republic, Hungary, Poland and Slovakia) in comparison with the euro-area and the US. Comparison to the euro area also has an implication for the optimum currency area literature.

The rest of the paper is organized as follows. Section 2 describes Kalman-filtering and the less known FLS and presents are our iterative procedure to determine the weighting parameter of FLS. In Section 3 we perform a Monte Carlo study of the properties of Kalman-filtering and FLS for various data generating processes. Section 4 presents the empirical results for the

US, euro-area, and new members of the EU. Finally, some concluding remarks are presented in Section 5.

2 Methodology

Time-variation in parameters of an economic model could be due to various reasons, for example, change in the behavior or agents, or could be the consequence that the functional form of the estimated model differs from the functional form of the underlying data generating process.[3] A standard approach to estimate time-varying coefficient models is the maximum likelihood estimation of a state-space representation of unobserved components models using the Kalman-filter to evaluate the likelihood function. However, this approach is built on the assumption of a certain distribution of the innovations, which is usually set to be the Gaussian distribution, which could not be valid.

Therefore, in this paper we also aim to compare Kalman-filtering to a less frequently used method, the so called Flexible Least Squares (FLS) introduced by Kalaba and Tesfatsion (1988, 1989, 1990). Since this estimation approach is less frequently used in economic applications we briefly describe it below.

2.1 Flexible Least Squares

The FLS algorithm solves the time-varying linear regression problem with a minimal set of assumptions. Suppose y_t is the time t realization of a time series for which a time-varying coefficient model is to be fitted,

$$(1) \qquad y_t = x_t'\beta_t + u_t, \qquad\qquad t = 1, \ldots, T,$$

where $x_t = \left(x_{0,t}, \ldots, x_{K-1,t}\right)$ denotes a $K \times 1$ vector of known exogenous regressors (which can also contain the lagged values of y_t), $\beta_t = \left(\beta_{0,t}, \ldots, \beta_{K-1,t}\right)$ denotes the $K \times 1$ vector of unknown coefficients to be estimated, which can change in time, and u_t is the approximation error.

Two main assumptions are needed for the formulation of a cost function to be minimized. First, the prior measurement specification states that the residual errors of the regression are small, that is,

---

[3] For example, most estimated economic models are linear, while the true data generating process could be non-linear.

(2) $y_t - x_t'\beta_t \approx 0$, $t = 1,\ldots,T$.

Second, the prior dynamic specification declares that the vector of coefficients evolves slowly over time:

(3) $\beta_{t+1} - \beta_t \approx 0$, $t = 1,\ldots,T\text{-}1$.

A basic problem for the investigator is to find a coefficient sequence estimate, $\left(\hat{\beta}_1,\ldots,\hat{\beta}_T\right)$, which satisfies both of these prior assumptions in an acceptable manner. The idea of the FLS method is to assign two types of residual error to each possible coefficient sequence estimate. One of these consists of the sum of squared residual measurement errors:

(4) $$r_M^2(\beta,T) = \sum_{t=1}^{T}\left(y_t - x_t'\beta_t\right)^2,$$

matching the above mentioned measurement prior. The other – following the smoothness prior – is the sum of squared residual dynamic error, formally

(5) $$r_D^2(\beta,T) = \sum_{t=1}^{T-1}\left(\beta_{t+1} - \beta_t\right)'\left(\beta_{t+1} - \beta_t\right).$$

Taking the collection *P(T)* of all pairs of these two kinds of sum of squared residual errors as β ranges over the space of possible coefficient sequences defines the $\left(r_D^2(\beta,T), r_M^2(\beta,T)\right)$ residual possibility set on the positive quarter of the two-dimensional plane. The vector minimums of this set, analogously with the idea of a Pareto-efficient frontier, yield the $P_F(T)$ residual efficiency frontier. Each point of this frontier represents a corresponding sequence of estimates which can be found by minimizing the weighted sum of the two types of measurement errors. Thus, with a given μ weighting parameter, Kalaba and Tesfatsion (1988) define the incompatibility cost assigned to any β coefficient sequence as

(6) $$C(\beta,\mu,T) = \mu \cdot r_D^2(\beta,T) + r_M^2(\beta,T).$$

Minimization of the incompatibility cost for β, given any $\mu>0$, leads to a unique estimate for β as $\hat{\beta}^{FLS}(\mu,T) = \left(\hat{\beta}_1^{FLS}(\mu,T),\ldots,\hat{\beta}_T^{FLS}(\mu,T)\right)$, which is the flexible least squares solution, conditional on μ and *N*. This conditional minimization is performed with a dynamic programming algorithm.

Consequently, there are a continuum number of FLS solutions for a given set of observations, depending on the weight parameter μ. The solutions lie between two extremes. First, if μ

approaches zero, the incompatibility cost function places absolutely no weight on the smoothness prior. This means that while r_D^2 stays relatively large, r_M^2 will be brought down close to zero, resulting in a rather erratic sequence of estimates. Second, as μ becomes arbitrarily large, the cost function assigns all importance to the dynamic specification. This case yields the ordinary least squares (OLS) solution, i.e. r_M^2 is minimized subject to $r_D^2 = 0$.[4] Therefore, the selection of the weighing parameter is a highly critical part of the FLS procedure, as the appropriate coefficient sequence lies somewhere between the most variable and the fully stable – OLS – solution.

Kalaba and Tesfatsion's estimation mechanism uses the fact that the incompatibility cost at a certain point of time can be written recursively as

$$(7) \qquad C(\beta_{t+1},\mu,t) = \min_{\beta_t}\left\{(y_t - x_t'\beta_t)^2 + \mu(\beta_{t+1}-\beta_t)'(\beta_{t+1}-\beta_t) + C(\beta_t,\mu,t-1)\right\}.$$

Besides the recursion, we only need to assume that the incompatibility cost is a quadratic function of the unknown parameter β_t:

$$(8) \qquad C(\beta_t,\mu,t-1) = \beta_t' S_{t-1}\beta_t - 2\beta_t' s_{t-1} + r_{t-1}$$

Now if we plug this into the recursive updating equation and differentiate with respect to β_t, we arrive at the updating equation of the time-varying parameter, conditioned on β_{t+1}.This allows us – using all the data from time 1 until time T – first to compute our estimate of β_T in a given way (see Kalaba – Tesfatsion, 1988), and then roll back the recursion in order to get the members of the unknown sequence one after another.

2.2 Relation to Kalman filtering

Kalaba and Tesfatsion (1990) clarify the relationship of FLS with Kalman filtering. In particular, they emphasize that the two methods address conceptually distinct problems, but also prove that the Kalman filter recurrence relations could be derived by means of simple intuitive cost considerations (similarly to FLS), without reliance on probabilistic arguments. In a recent paper Montana et al. (2007) shed new light on the relation of FLS to Kalman filtering by adding mild probabilistic assumptions to FLS and weakening the assumptions behind the Kalman filter. Specifically, they assume that the errors of both observed and latent variables have finite first and second moments. Formally speaking, the dynamic and

[4] Our estimations will also illustrate the convergence of the FLS estimate to the OLS estimate as μ increases.

measurement priors are expressed in the state and measurement equations of the model, respectively, as follows:

(9) $\beta_{t+1} = \beta_t + \omega_t$, $\qquad t = 1,...,T\text{-}1,$

(10) $y_t = x_t{}' \beta_t + \varepsilon_t$, $\qquad t = 1,...,T.$

In essence, the requirement that innovations ω_t and ε_t are mutually and individually uncorrelated and have finite expected values and covariance matrices is close in spirit to the assumptions of FLS. The key difference is the randomness of the unknown parameter vector: recall that the smoothness prior of FLS doesn't require β_t to be random – only the smooth change in time is postulated.

Montana et al. (2007) first proved that the Kalman filter recursions work perfectly well under this distribution free circumstance – in fact, the derivation is even simpler and doesn't require any matrix inversion which makes it easy to implement even in higher dimension spaces with long streams of observations. The authors also show that the recursive updating equations of the Kalman filter are equivalent to those of FLS under the new assumptions and that maximizing the likelihood function of the Kalman filter is the same as minimizing the quantity

(11) $$\sum_{t=1}^{T} \left(y_t - x_t{}' \beta_t \right)^2 + V_\omega^{-1} \sum_{t=1}^{T-1} (\beta_{t+1} - \beta_t)' (\beta_{t+1} - \beta_t),$$

where V_ω is the covariance matrix of the ω_t errors of the parameter vector. The proof thus sheds light on the role of the μ smoothing parameter of FLS: comparing (11) to the definition (6) of incompatibility cost we get

(12) $V_\omega = \mu^{-1} I_K$,

where I_K is the $K \times K$ identity matrix. This means that by choosing a certain value for μ we also set the variance of the innovations of the estimated parameter vector: as μ approaches infinity, we make the variance of β_t diminish and hence get close to the OLS solution, and vice versa: the $\mu \to 0$ extreme makes V_ω larger and larger, thus making the parameter sequence rather erratic.

The algorithm of Montana et al. (2007) thus could be a useful extension of the FLS since it does not require matrix inversion which reduces calculation time on the one hand[5], and allows the calculation of confidence bands for FLS estimates on the other, which could not be made with the classic probability-free algorithm of Kabala and Tesfatsion.

In the current version of the paper we still use the original FLS algorithm which does not set the weighting parameter, although it would be advisable to get a unique solution instead of the parameter-based optimization process. Therefore we detail a draft of a procedure for determining an 'optimal' weighting parameter of the FLS in the Appendix.

3 The ability of FLS and Kalman-filtering to capture time-varying parameters: a simulation study

In this section we compare the properties of Kalman-filtering and the FLS for the estimation of linear autoregressive models with time-varying coefficients by Monte Carlo simulation. To this end, we set up and calibrate various data generating processes (DGP), stochastically simulate them, and apply the FLS estimation technique to simulated time series with the aim of comparing the estimation results to the known properties of the DGP. Calibration of the DGPs is based on inflation time series of the euro-area and the US. Our study includes four types of time-varying coefficient DGPs:

– Discrete shifts in parameters;

– Linear deterministic change in parameters;

– Sinusoidal deterministic change in parameters,

– Unit root process of the parameters (stochastic).

The simulation study helps us to determine the accuracy of the Kalman-filtering and FLS method.

In the current version of the paper we study simple AR(1) processes, that is,

(13) $y_t = \beta_{0,t} + \beta_{1,t} \cdot y_{t-1} + u_t, \qquad t = 1, \ldots, T,$

[5] The empirical question studied by Montana et al. (2007) is the real time application of an FLS-based algorithmic trading system for the S&P500 Futures Index.

where y_t is the time series under study (i.e., simulated series in this section and the inflation rate in the next empirical section); $\beta_{0,t}$ and $\beta_{1,t}$ are the time-varying coefficients, for which we assume the four alternative processes described above for the simulation study; and u_t is the error term, which will be set to a series of Gaussian random numbers. The initial condition, y_0, is always set to be equal to the initial mean of the process, $y_0 \equiv \beta_{0,0}/(1-\beta_{1,0})$. All simulations were performed using a sample size T = 200, which is typically used in the literature for studying inflation persistence.

As a starting point we apply OLS, FLS (with different μ-*s*) and Kalman-filtering/smoothing to a model with fixed parameters. The upper block of Table 1 shows the RMSE (Root Mean Squared Error) of the various estimation techniques when the parameters of the data generating process are constant. That is, assuming that $\beta_{0,t} = \beta_0 = 0.2$, $\beta_{1,t} = \beta_1 = 0.9$, and $\text{var}(u_t) = 0.25^2$ in equation (13), we simulated equation (13) and applied the various estimation techniques. Repeating the simulation and estimation exercise N=1000 times, we arrive at 1000 estimated parameter sequences for each estimator (the OLS estimator, obviously, yields a parameter, not a time-varying parameter sequence). We computed the quadratic difference between the estimated and true sequence, which – in the form of RMSE – are averaged across all repetitions and reported in the table. Among the various estimators the OLS captures the most adequately the constant nature of the parameters. All time-varying coefficient methods attribute some of the error variance to parameter changes. This result highlights the importance of applying tests for structural breaks before estimating any time-varying coefficient methods.

The rest of Table 1 shows RMSE of the estimators when the parameters change in time; the exact specifications will be detailed below. The plot of the parameter sequences are shown in the top panel of Figures 1 to 4. In Figure 4, which shows the case with unit root in parameters, obviously only a certain realization is shown for the parameters. The top panels of these figures also show a certain realization for y_t. The middle panels of the figures show, in addition to the true autoregressive parameter, $\beta_{1,t}$, the OLS estimate based on a fixed coefficient model, and four different FLS estimates corresponding to given weights: μ = 10, 10^2, 10^3, and 10^4, and the Kalman-filter and Kalman-smoother estimates.[6] Because in this

[6] For Kalman-filtering we assumed random walk processes for the parameters in the state-space representation in all cases. Probably Kalman-filter estimates could be improved if we assumed the true DGP for the time-varying

experiment we know both the underlying parameter sequences and the realizations of the process, we could also find the μ weight value which minimizes the sum of squared error between the true sequence and its FLS-estimate. As the best possible FLS estimator, this is also presented in the table, along with the found μ values, which is called as *"optimal μ"*. The bottom panels show the true intercept, $\beta_{0,t}$ with its corresponding OLS, FLS and Kalman-filter estimates.

Figures 1-4 help us to get a visual impact of the nature of the FLS-estimator and Kalman-filter and smoother.

3.1 Discrete shifts in parameters

In Figure 1 we study the properties of our TVC model estimators when there are single discrete shifts in parameters of the data generating process. Specifically, the intercept of our simulated process shifts at observation $t = 60$ from 0.15 to 0.45, and the autoregressive parameter changes at observation $t = 140$ from 0.9 to 0.7. Notice that these parameters imply the same mean for the series until $t = 60$ and from $t = 140$. The simulated series, shown on the top panel of the figure, resembles somewhat to US inflation (see Figure 5).

The OLS does an awful job. The estimated parameters indicate a random walk without drift, as the estimated autoregressive parameter is almost 1 and the intercept is close to zero. This is not surprising and the biases of the OLS estimates in the face additive outlier have already been reported in the literature.

The FLS, although better than the OLS, but can not capture well the sudden changes in parameters. Depending on the value of μ, the estimated time-varying parameters are either rather smooth, or change at both shifts. For example, when $\mu = 10$, the autoregressive parameter rises from about 0.4 to 0.7 at observation $t = 60$, when only the intercept changed, and falls from 0.75 to 0.4 at $t = 140$. The estimated intercept also changes at both dates, although more smoothly, and has values much larger than implied by the data generating process. The Kalman-filter and smoother also can not capture well sudden discrete changes. The middle and bottom panels of Figure 1 indicates that for the certain draw shown both the

parameters, but we motivate our choice for the following two reasons. First, the random walk assumption is the counterpart of the FLS assumptions. Second, when applying the Kalman-filter to a real-world series, the DGP of time variation is also unknown.

Kalman-filtered and smoothed parameter estimate move along quite close to the FLS μ = 10^2 sequence.

The results of the stochastic simulation shown in the second panel of Table 1 indicates that, the FLS estimator with μ = 10^2 and μ = 10^3 are the best and hence better than the Kalman-Filter and smoother.

3.2 Continuous deterministic change in parameters

In the above example we ignored the assumption of smooth changes in the underlying parameter sequences, which is one of the priors of FLS. Our next DGP thus assumes a perfectly smooth change. $\beta_{0,t}$ rises from 0.1 to 0.9 and $\beta_{1,t}$ the decreases from 0.9 to 0.1 linearly in time. These assumptions leave the mean of the process unchanged through time.

Figure 2 shows the results. The OLS crosses the true parameter values at about one third of the sample. The FLS does quite well now, especially with μ = 10^2. In this case, after some initial misdirection up to about where the OLS line crosses the true sequence, it follows closely the parameter paths of the data generating process. In the stochastic simulation the Kalman-filter and smoother perform only a little worse than the FLS estimator with μ = 10^2, which can be seen on the third panel of Table 1.

3.3 Sinusoid change in $\beta_{1,t}$

The next data generating process assumes that the intercept remains constant, $\beta_{0,t} = 0.5$, but the autoregressive parameter follows a sinusoid process, namely, $\beta_{1,t} = 0.6 + 0.3 \cdot \sin(\pi / 100 \cdot t)$. Thus, $0.3 \le \beta_{1,t} \le 0.9$.Sinusoid change is a good example of smooth nonlinear variations, as we will see, the capabilities of the FLS estimator can be illustrated well through this case.

Results of a certain Monte-Carlo draw can be seen in Figure 3. The OLS estimates indicate a random walk process for y_t with the autoregressive parameter very close to one and even higher than the highest value of the whole data generating parameter sequence (0.9). The FLS estimator, on the other hand, does again very well when μ = 10^2. It can nicely capture both the sinusoid path of $\beta_{1,t}$ and the constant value of $\beta_{0,t}$, although with a little remaining sinusoid variation in the latter. The Kalman-filter and smoother's estimates are similar to those of the

FLS at $\mu = 10^2$ but again, the Monte Carlo simulation shows an advantage for the former in the fourth panel of Table 1.

3.4 Random walk parameters

Figure 4 shows estimation results of a certain Monte-Carlo draw when both parameters follow independent random walks without drift. Standard error of the innovations was set to 0.05 and the initial value of both parameters was set to 0.5. The OLS estimator indicates a value of 0.7 for the autoregressive parameter, which is even larger than the maximum of $\beta_{1,t}$ in this certain outcome shown on the figure. The FLS estimator, on the other hand, does quite well again when $\mu = 10^2$, and captures well the time-variation of $\beta_{1,t}$, and a little bit better again than the Kalman-smoother. For $\beta_{0,t}$, on the other hand, the Kalman-smoother slightly outperforms the FLS, as can be seen on the bottom panel of Table 1.

3.5 Summing up

Let us summarize the main findings of the simulation study.

1. The OLS is the best when parameters are constant, but is generally upward biased even compared to the mean (or time-average) of the autoregressive parameter of the DGP when there is time variation.
2. Neither FLS, nor the Kalman-filter and Kalman-smoother can capture sudden changes in parameters
3. The FLS with a given μ of 10^2 can reasonably well capture more gradual changes in parameters, like continuous deterministic change, deterministic sinusoid path and random walk behavior, especially for the autoregressive parameter, which determines persistence and thus carries more importance for us.
4. Looking at the results of Table 1 it can be obviously seen that being able to determine the weight parameter of the FLS could imply a much more appropriate estimation for the time-varying parameter sequences. With ‘optimal μ’, which was found in the range [$10^{1.53}$, $10^{2.55}$] in the four time-varying DGPs indicated in Table 1, the FLS strongly outperforms the Kalman-filter and smoother.
5. The Kalman-filter and smoother is similarly good for capturing gradual changes in parameters, although they are slightly worse than the FLS with given weights, and

reasonably worse than the FLS with optimal weights. The Kalman-smoother dominates the Kalman-filter in all cases we studied

4 Estimates of inflation persistence for the US, Euro-area, and newly accessed countries of the European Union

In this section we study IP in four new members of the EU (Czech Republic, Hungary, Poland, Slovakia) in comparison with the euro area and the US. The case of euro-area candidate countries raises the issue of convergence. These countries must fulfill, among others, the Maastricht criterion related to the level of inflation. However, similar persistence to that of the euro-area will be crucial for the optimality of the common monetary policy. Consequently, our paper also has an implication for the optimum currency area literature.

Time-varying coefficient analysis is especially inevitable in the case of the euro-area and the Central and Eastern European Countries (CEEC's). The euro-area did not exist before 1999 and its data were constructed by aggregating country time series. It is rather likely therefore that euro-area data include structural breaks. The CEEC's had a socialist economic system till 1989, when transition to a market economy had speeded up, accompanied by a substantial economic downturn and pick up in inflation (partly due to price liberalization) in the early nineties.[7] Thus, it would be nonsense to model the inflation of these states with a fixed coefficient model incorporating both the planned and the market systems.[8] However, with time-varying coefficient models we can study changes in the parameters of the inflation process on a longer time span.

4.1 Data

In the spirit of time varying coefficient modeling we used the longest available samples for estimation. We study inflation time series of the US and the euro area as benchmarks and for comparison with the literature and four new members of the EU: the Czech Republic, Hungary, Poland and Slovakia. Price level time series for two of these four new EU countries are available at least since the eighties. The sample period covers 1957Q1 – 2006Q4 for the

[7] Moreover, some countries separated in the early nineties and introduced their own new currencies (Czech Republic and Slovakia in our sample).

[8] In applied research the starting year is usually set to 1995, which renders the number of observations short.

US, 1970Q1 – 2006Q4 for the euro area, 1976Q1 – 2006Q4 for Hungary, 1980Q1-2006Q4 for Poland and 1993Q1 – 2006Q4 for both the Czech Republic and Slovakia. Euro area data are taken from the AWM database; data for the other five countries are from the IMF: International Financial Statistics. We define inflation as *Δln(seasonally adjusted consumer price level)x100* and we measure persistence as the sum of estimated autoregressive coefficients. This definition is mostly used by the literature and practically in our case means the size of the AR(1) coefficient.[9]

4.2 Estimation results

Estimation results are shown on Figures 5-10. We show two results for the FLS: with μ equal to $10^{1.5}$ and $10^{2.5}$, because almost all the values which proved to be the best in the simulation study were between these two values. For the Kalman-filter, we show both the filtered and smoothed estimate, however, as also learned from our simulation study, we put most of our trust in the latter one.

It is evident for all countries except Slovakia that OLS estimates are likely upward biased[10]: OLS parameter estimates are almost the highest considering all time-varying coefficient estimations and significantly larger than the time-average of them.[11] Recall from the previous section that the OLS estimator proved to be upward biased compared to the time average of the parameters, when there were changes in the parameters of the data generating process.

Inflation persistence tends to be higher in times of high inflation in the cases of all countries. For the Kalman-filter we also calculated the error bands (not shown on the figures) which indicated that changes in inflation persistence are significant for most of the countries.

Considering the magnitudes, in the US, the high inflation persistence episode following the oil crises showed an autoregressive parameter value of around 0.7-0.8, which declined to around 0.05-0.25 in recent years. In the euro-area, the autoregressive parameter declined from 0.4-0.6 to zero. In fact, three of the four estimates showed in the lower panel of Figure 6 even

[9] NOTE: we currently work with higher order AR models.

[10] Note that it is widely documented in the literature that the OLS estimate of the autoregressive coefficient (or the dominant autoregressive root) is *downward* biased when parameters are fixed.

[11] The significance was tested for the OLS AR(1) estimate and the average values of time-varying AR(1) estimates using a standard *t*-statistic. We could reject the null of equality at the 5% level for almost all cases except Slovakia. For Poland's FLS estimates the rejection could be made at the 10% level.

indicates a negative parameter, while our preferred FLS estimate with $\mu = 10^{2.5}$ indicates a value of 0.4. Hence, considering our preferred estimator inflation persistence in the US and the euro-area are similar.

For Hungary, the FLS estimator with $\mu = 10^{2.5}$ indicates a parameter value of around 0.1-0.2 in the socialist era. Persistence then jumps to a very high level – near 0.8 – at the time of the transition, starts sinking in 1996 and falls to a range between 0.2 and 0.4 around 2001. Hence, the historically different inflation peak causes the autoregressive parameter sequence to behave differently in the case of Hungary, but persistence in the 2000s is not much larger in magnitude than that of the US and euro-area.

The results for Poland are slightly different from that of Hungary probably due to the following two reasons. First, preceding the inflation peak of the transition in 1990, there has been another one in 1982; and second, these highs have been extremely large (much larger than in Hungary). For these reasons, our estimators show erratic movements, for example, the persistence parameter estimates hover in the range of 1.5-2 in the periods of high and hyperinflations, implying an explosive inflationary dynamics. From 1992 we can observe a downward trend of the FLS estimates which shows that the steady downturn of the inflation in the last decade was accompanied also by a decline of persistence.

Unfortunately our samples for the Czech Republic and Slovakia are short because these countries broke up the former Czechoslovakia. In both countries we can observe a high inflation period (around 1998-1999, following their exchange rate crises), which was probably too short to produce persistence estimates close to unity similarly to other countries in our sample. Apart from the inflation highs, our FLS results show a permanent decline of persistence throughout the whole period, from 0.4-0.6 to 0-0.2, which is in line with the hypothesis of positive correlation between high inflation and high persistence. Note that especially in Slovakia, the OLS also produces a notably low persistence value, which is probably due to the small sample lacking any remarkable high-inflation periods.

We have considered only an AR(1) process by now. However, autocorrelation tests for the innovation of the Kalman-filter indicate the US and euro-area innovations are highly autocorrelated, clearly indicating the simple AR(1) model is not adequate (Hungarian innovations, on the other hand, are not autocorrelated). Thus we will continue with the study of higher order models.

5 Summary

This paper studied inflation persistence with time-varying-coefficient autoregressions for the US, euro-area, and four Central-European countries waiting for euro area accession. We first compared the statistical properties of the well known classical maximum likelihood estimation using the Kalman-filter to the less known Flexible Least Squares estimator by Monte Carlo simulation. We showed that the FLS estimator does capture many types of smooth parameter changes and works definitely not worse than the more conventional Kalman-filter and smoother, and even somewhat better. We have also suggested a procedure to set the weighting parameter of the FLS based on an iterative Monte Carlo simulation technique calibrated to the time series in question.

The FLS estimator with $\mu = 10^{2.5}$ seems to yield plausible results for empirical inflation series: parameters of the autoregressive model of inflation in the cases of the US, the Euro-area, and the four CEEC's are time-varying, and the parameter sequences are interpretable in economic terms. Inflation persistence tends to be higher in times of high inflation in the cases of all countries. In the US and euro-area inflation persistence declined to historically low levels, while a similar process went on in the once communist countries during the first two decades of the post-transition era. We argued that similar persistence is an important structural similarity in a currency union and progress on this front of the new EU members could contribute to the economic arguments in favor of their entry to the euro area.

6 Further work

In the coming weeks we plan to extend our paper with the following tasks:

- Include a Bayesian autoregression as a third method
- Study the relation between the Bayesian autoregression and the FLS
- Allow higher order autoregressive models
- Include formal tests for constant persistence
- Finalize the procedure for setting the weighting parameter of FLS (see the draft in the Appendix of this version of our paper)

- Compare our results for the euro area and the US to the literature

7 References

Altissimo, Filippo, Michael Ehrmann and Frank Smets (2006): Inflation Persistence and Price Setting Behaviour in the Euro Area - a Summary of the IPN Evidence, ECB Occasional Paper No. 46 (Jun 2006), ECB, Frankfurt am Main.

Cogley, Timothy and Thomas J. Sargent (2001): Evolving Post-World War II U.S. Inflation Dynamics, NBER Macroeconomics Annual 2001, 331-372.

Cogley, Timothy and Thomas J. Sargent (2005): Drifts and Volatilities: Monetary Policies and Outcomes in the Post World War U.S., Review of Economic Dynamics 8(2), 262-302.

Dossche, Maarten and Gerdie Everaert (2005): Measuring Inflation Persistence – A Structural Time Series Approach. ECB Working Paper No. 495 (Jun 2005), ECB, Frankfurt am Main.

Gadzinski, Gregory and Fabrice Orlandi (2004): Inflation persistence in the European Union, the euro area, and the United States. ECB Working Paper No. 414 (Nov 2004), ECB, Frankfurt am Main.

Hansen, Bruce E. (2001): The New Econometrics of Structural Change: Dating Changes in U.S. Labor Productivity, Journal of Economic Perspectives 15, 117-128.

Kalaba, Robert and Leigh Tesfatsion (1988): The Flexible Least Squares Approach to Time-varying Linear Regression. Journal of Economic Dynamics and Control, Vol. 12, 43-48.

Kalaba, Robert and Leigh Tesfatsion (1989): Time-Varying Linear Regression Via Flexible Least Squares, Computers and Mathematics with Applications, Vol. 17, 1215-1245.

Kalaba, Robert and Leigh Tesfatsion (1990): Flexible Least Squares for Approximately Linear Systems, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, 978-989.

Kalaba, Robert and Leigh Tesfatsion (1996): A Multicriteria Approach to Model Specification and Estimation, Computational Statistics and Data Analysis, Vol. 21, 193-214.

Levin, Andrew T. and Jeremy M. Piger (2004): Is Inflation Persistence Intrinsic in Industrial Economies? ECB Working Paper No. 334 (Apr 2004), ECB, Frankfurt am Main

Marques, Carlos Robalo (2004): Inflation Persistence: Facts Or Artefacts? ECB Working Paper No. 371 (Jun 2004), ECB, Frankfurt am Main.

Montana, Giovanni, Kostas Triantafyllopoulos, Theodoros Tsagaris (2007): Flexible Least Squares for Temporal Data Mining and Statistical Arbitrage, http://arxiv.org/pdf/0709.3884

O'Reilly, Gerard and Karl Whelan (2004): Has Euro-area Inflation Persistence Changed over Time? ECB Working Paper No. 335 (Apr 2004), ECB, Frankfurt am Main

Pivetta, Frederic and Ricardo Reis (2007): The persistence of inflation in the United States, Journal of Economic Dynamics and Control, forthcoming.

Tesfatsion, Leigh and John Veitch (1990): U.S. Money Demand Instability: A Flexible Least Squares Approach, Journal of Economic Dynamics and Control, Vol. 14, 151-173.

8 Appendix: A first draft of a new procedure to set the weighting parameter of FLS

Kalaba and Tesfatsion (1988, 1989, 1990) suggested to study the properties of the FLS estimates along the residual efficiency frontier and to draw conclusions based on some general features of the estimates. Their approach is motivated by the fact that model estimation for economic processes is, intrinsically, a *multicriteria* optimization problem. The assumptions generally made for estimation typically fall into three categories: measurement, dynamic, and probabilistic.[12] However, an actual economic process will behave in a manner that is incompatible to some degree with each of these assumptions. Associated with any set of estimates there will be a set of discrepancy terms reflecting the incompatibility of the assumptions with the data and an econometrician undertaking the estimation would presumably want each type of discrepancy term to be small in some sense. Kalaba and Tesfatsion therefore suggest dropping probabilistic assumptions and studying the incompatibility of the data with assumptions by giving different weights to measurement and dynamic errors. Consequently, Kalaba and Tesfatsion do not suggest any particular way to set a given value for the weighting parameter, and we are not aware of such a suggestion by other authors as well.

Econometricians, however, are usually interested in a single estimate having certain properties, for example, unbiasedness, 'small' variance, and so on. To this end, both classical and Bayesian econometricians generally postulate certain probabilistic assumptions in order to derive a point estimate and a confidence band reflecting the uncertainty of the point estimate. The probabilistic assumptions, though, could be dubious.

We aim to bridge the probability-free FLS approach yielding a continuum number of solutions with the standard probability-based techniques leading to a unique estimate by suggesting a procedure, which maintains the probability-free assumption of FLS but leads to a unique estimate. This estimate can be regarded, to a certain extent, as the 'most likely' estimate of the time-varying parameter sequence for the particular model and time series under study. In practice, our procedure sets the weighting parameter of the FLS and regards the FLS estimate associated with this weighting parameter to be the 'most likely' estimate.

[12] Assumptions on the cross-sectional dimension of the model add a further type.

Our suggested procedure is based on the following iterative algorithm:

STEP 1. As a starting point, estimate an initial time-varying coefficient model (equation (1)) for the times series studied, that is, set an initial arbitrary μ and calculate the FLS solution.

STEP 2. Perform a stochastic simulation of the estimated model with N trials. This step can be viewed as a simulation of a model that fits to the times series under study (taking into account time variation in parameters). Note that while the FLS estimator is distribution free, we have to assume a certain distribution for the stochastic simulation.

For each of the simulated series select μ which minimizes, say, the mean quadratic difference between the FLS estimate and the parameter sequence of the data generating process used. That is, for each of the simulated series perform many FLS estimates for various μ values with $\mu \in [\mu^L, \mu^U]$ and select the one that attains the minimum. The mean of all μ-estimates can be regarded as a first estimate of μ, which we denote as $\mu^{(1)}$.

STEP 3. Estimate the time-varying parameters of the model with FLS for the time series studied, using the chosen $\mu^{(1)}$ in Step 2.

STEP 4: Return to Step 2: perform a stochastic simulation of the estimated model of Step 3; then select μ for each of the simulated series and calculate their mean to arrive a new μ-estimate to be denoted as $\mu^{(2)}$; and to Step 3: estimate the model for the data using $\mu^{(2)}$ in the FLS algorithm.

Continue this iterative procedure till $\mu^{(i)}$ converges, that is, till $\mu^{(i)}$- $\mu^{(i-1)}$<ε, where ε is a 'small' number. ε

The following selections should be made in order to make this procedure operative and to study its robustness.[13]

The initial μ in Step 1. We study the sensitivity of our procedure for a wide selection of initial μ.

[13] Some technical parameters are set to the following values: the number of Monte-Carlo draws, N, in Step 2 is set to 1000; the possible range of μ is determined between $\mu^L = 0$ and $\mu^U = 10^6$; and the criteria for convergence, ε, is set to 10^{-5}.

The probability distribution of the random numbers to be used for the stochastic simulation in Step 2. As a benchmark we use Gaussian random numbers but will study the sensitivity of the results to the choice of the distribution.

The loss function to be used in the selection of μ for each simulated series. In Step 2 above we indicated the mean quadratic difference, but we also study the sensitivity of the results other loss functions, such as the mean absolute difference.

NOTE: we have not yet completed all of these tasks and therefore do not report our preliminary results in this version of the paper. In both the simulation exercise of Section 3 and in the empirical estimation of Section 4 we report FLS results for various μ but not for our 'optimal' μ.

Constant Parameters

	OLS	FLS ($\mu = 10^1$)	FLS ($\mu = 10^2$)	FLS ($\mu = 10^3$)	FLS ($\mu = 10^4$)	FLS optimal ($\mu = 10^{3.91}$)	Kalman-Filter	Kalman-Smoother
Intercept	*0.061*	*0.887*	*0.380*	*0.151*	*0.079*	*0.075*	*0.410*	*0.335*
AR(1)	*0.030*	*0.461*	*0.201*	*0.080*	*0.040*	*0.040*	*0.202*	*0.165*

Discrete Shift in Parameters

	OLS	FLS ($\mu = 10^1$)	FLS ($\mu = 10^2$)	FLS ($\mu = 10^3$)	FLS ($\mu = 10^4$)	FLS optimal ($\mu = 10^{2.35}$)	Kalman-Filter	Kalman-Smoother
Intercept	*0.338*	*0.653*	*0.174*	*0.202*	*0.312*	*0.106*	*0.461*	*0.416*
AR(1)	*0.168*	*0.332*	*0.104*	*0.096*	*0.151*	*0.064*	*0.199*	*0.172*

Linear Deterministic Change in Parameters

	OLS	FLS ($\mu = 10^1$)	FLS ($\mu = 10^2$)	FLS ($\mu = 10^3$)	FLS ($\mu = 10^4$)	FLS optimal ($\mu = 10^{2.55}$)	Kalman-Filter	Kalman-Smoother
Intercept	*0.440*	*0.545*	*0.182*	*0.117*	*0.363*	*0.062*	*0.214*	*0.084*
AR(1)	*0.448*	*0.307*	*0.111*	*0.147*	*0.377*	*0.061*	*0.102*	*0.076*

Sinusoid Deterministic Change in Autoregressive Parameter

	OLS	FLS ($\mu = 10^1$)	FLS ($\mu = 10^2$)	FLS ($\mu = 10^3$)	FLS ($\mu = 10^4$)	FLS optimal ($\mu = 10^{2.17}$)	Kalman-Filter	Kalman-Smoother
Intercept	*0.457*	*0.538*	*0.111*	*0.252*	*0.411*	*0.060*	*0.180*	*0.118*
AR(1)	*0.432*	*0.286*	*0.074*	*0.249*	*0.390*	*0.064*	*0.137*	*0.093*

Random Walk Change in Parameters

	OLS	FLS ($\mu = 10^1$)	FLS ($\mu = 10^2$)	FLS ($\mu = 10^3$)	FLS ($\mu = 10^4$)	FLS optimal ($\mu = 10^{1.53}$)	Kalman-Filter	Kalman-Smoother
Intercept	*0.416*	*0.206*	*0.197*	*0.329*	*0.398*	*0.131*	*0.217*	*0.188*
AR(1)	*0.346*	*0.195*	*0.150*	*0.265*	*0.329*	*0.111*	*0.182*	*0.156*

Table 1: Average RMSE of estimated parameter sequences in the Monte Carlo simulation

Notes: Number of Monte-Carlo draws is 1000 for each specification. The last column showing results for FLS is based on results with the average of 'optimal µ'-s.

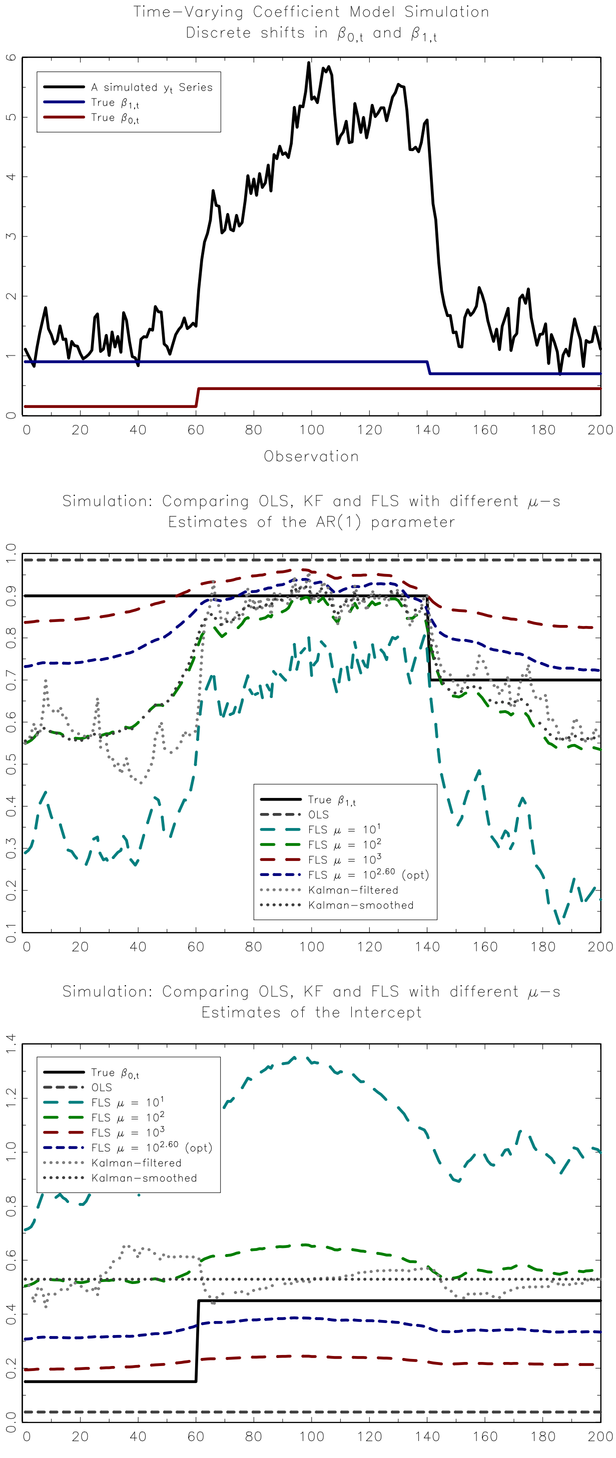


Figure 1: The ability of FLS and the Kalman-filter to capture discrete shift in parameters

Top panel: true processes for $\beta_{0,t}$ and $\beta_{1,t}$ and a certain realization for y_t.

Middle panel: the true process for $\beta_{1,t}$, its OLS estimate, and four different FLS estimates corresponding to weights: μ = 10, 10^2, 10^3 and the optimal (SSE minimizing) value, and the Kalman-filter and Kalman-smoother estimate.

Bottom panel: the true process for the constant, $\beta_{0,t}$, and its estimates.

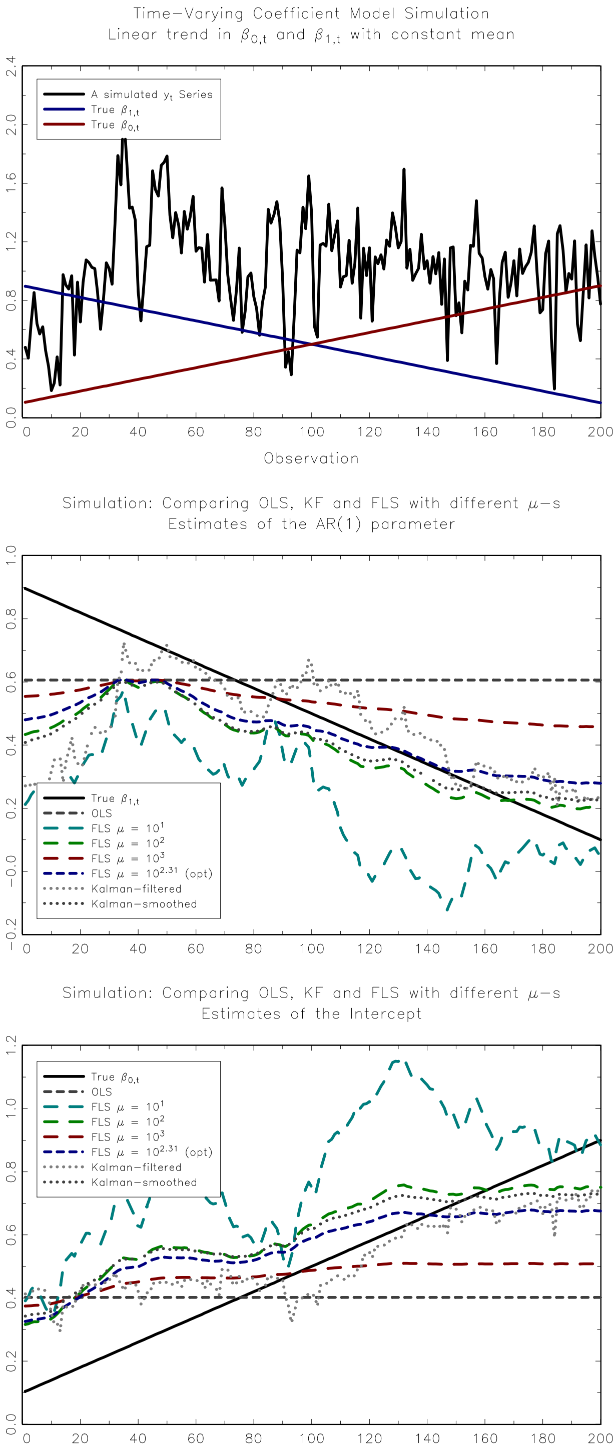


Figure 2: The ability of FLS and the Kalman-filter to capture continuous deterministic change in parameters

Top panel: true processes for $\beta_{0,t}$ and $\beta_{1,t}$ and a certain realization for y_t.

Middle panel: the true process for $\beta_{1,t}$, its OLS estimate, and four different FLS estimates corresponding to weights: μ = 10, 10^2, 10^3 and the optimal (SSE minimizing) value, and the Kalman-filter and Kalman-smoother estimate.

Bottom panel: the true process for the constant, $\beta_{0,t}$,and its corresponding estimates.

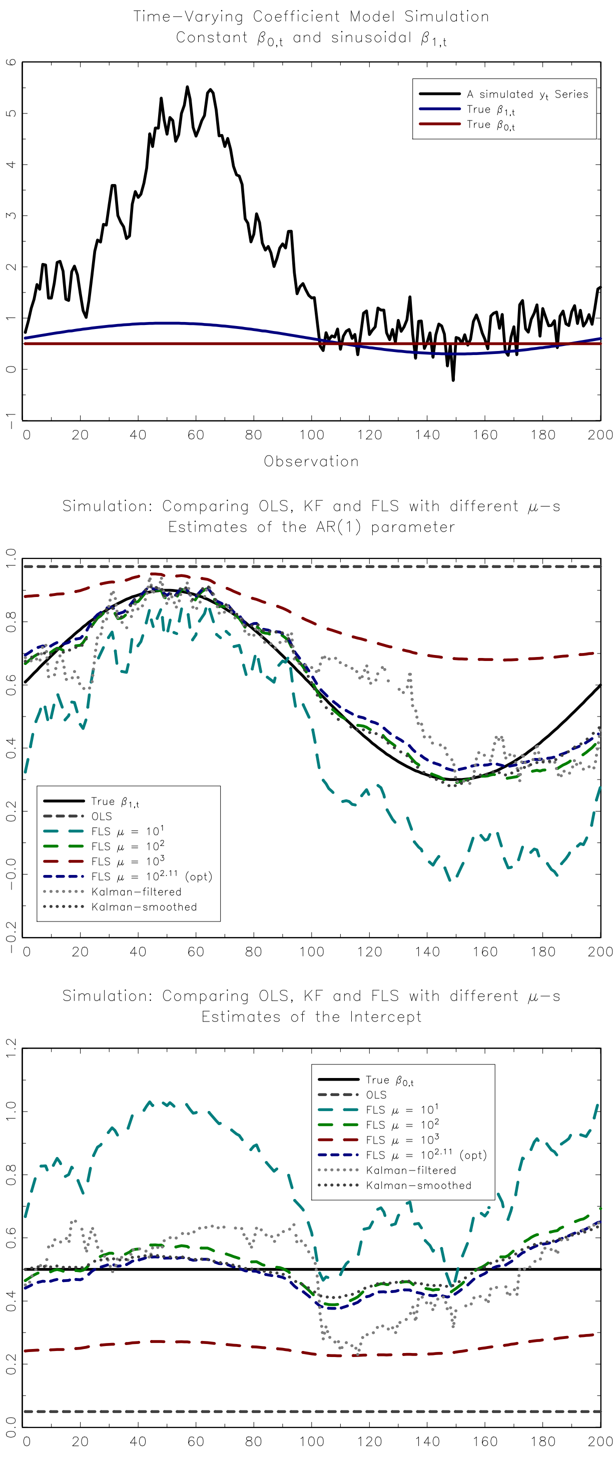


Figure 3: The ability of FLS and the Kalman-filter to capture sinusoid change in $\beta_{1,t}$

<u>*Top panel:*</u> true processes for $\beta_{0,t}$ and $\beta_{1,t}$ and a certain realization for y_t .

<u>*Middle panel:*</u> the true process for $\beta_{1,t}$, its OLS estimate, four different FLS estimates corresponding to weights: μ = 10, 10^2, 10^3 and the optimal (SSE minimizing) value, and the Kalman-filter and Kalman-smoother estimate.

<u>*Bottom panel:*</u> the true process for the constant, $\beta_{0,t}$, with its different estimates.

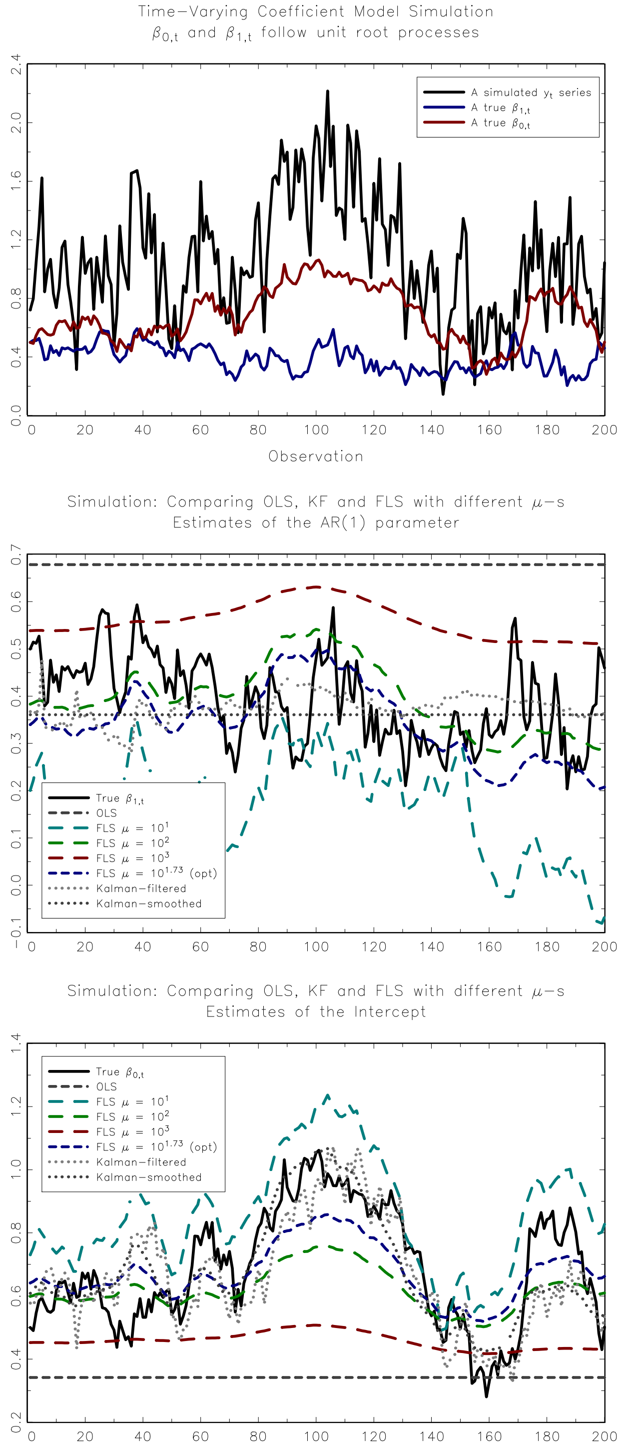


Figure 4: The ability of FLS and the Kalman-filter to capture random walk change in parameters

<u>*Top panel:*</u> true processes for $\beta_{0,t}$ and $\beta_{1,t}$ and a certain realization for y_t.

<u>*Middle panel:*</u> the true process for $\beta_{1,t}$, its OLS estimate, and four different FLS estimates corresponding to weights: μ = 10, 10^2, 10^3 and the optimal (SSE minimizing) value, and the Kalman-filter and Kalman-smoother estimate.

<u>*Bottom panel:*</u> the true process for the constant, $\beta_{0,t}$ and its estimates.

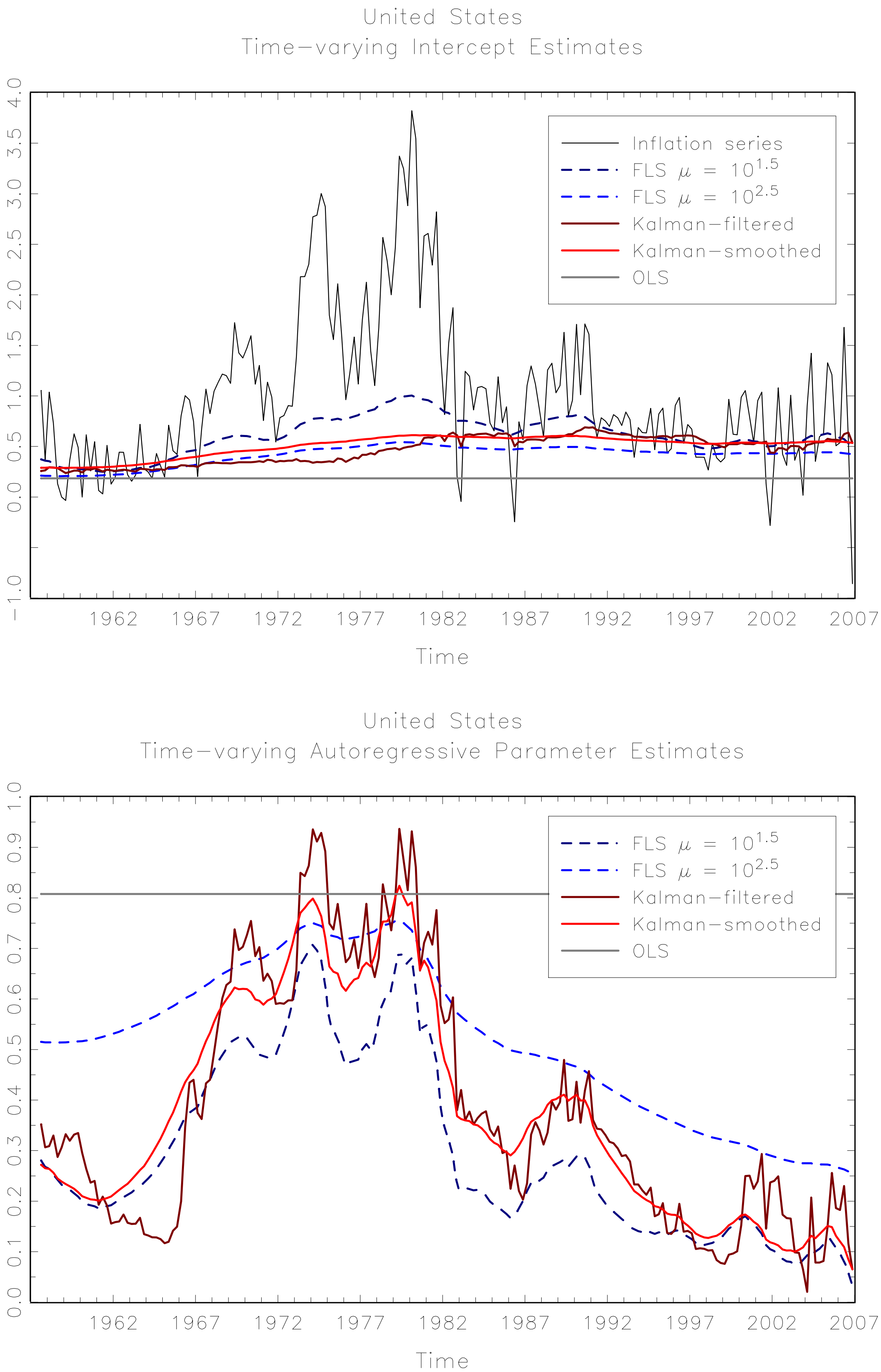
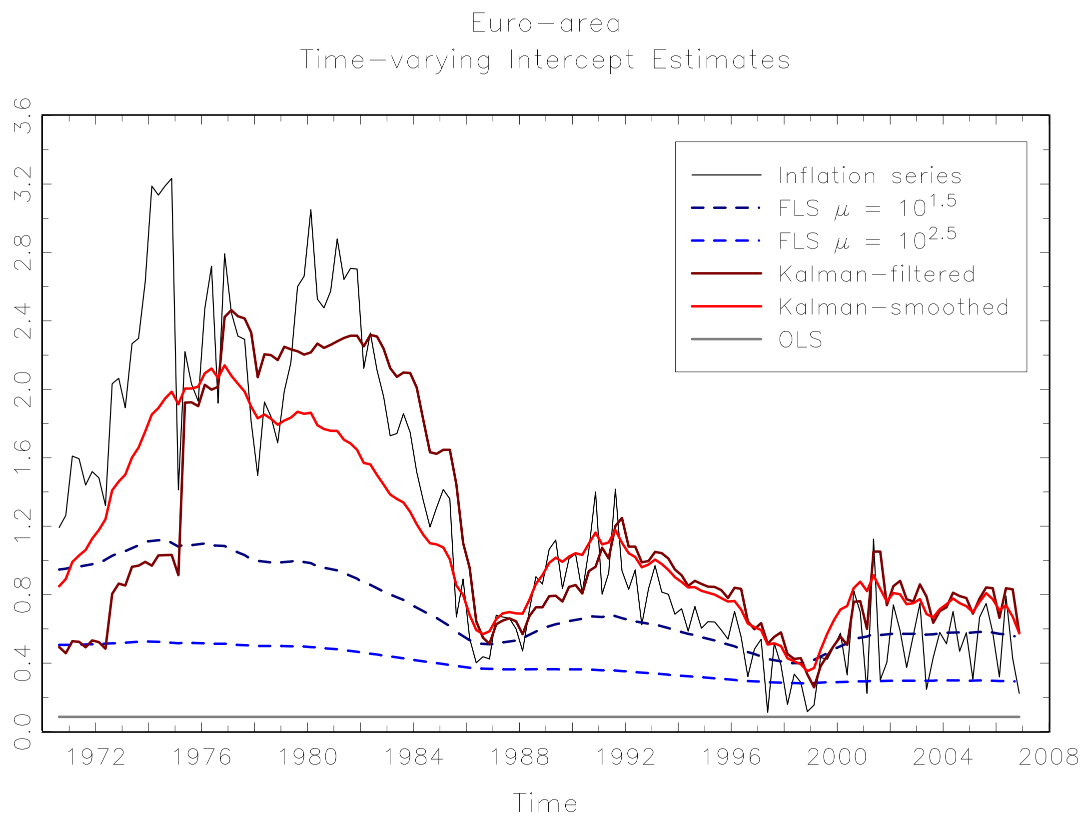


Figure 5: US – Estimation results using OLS, FLS, and the Kalman-filter



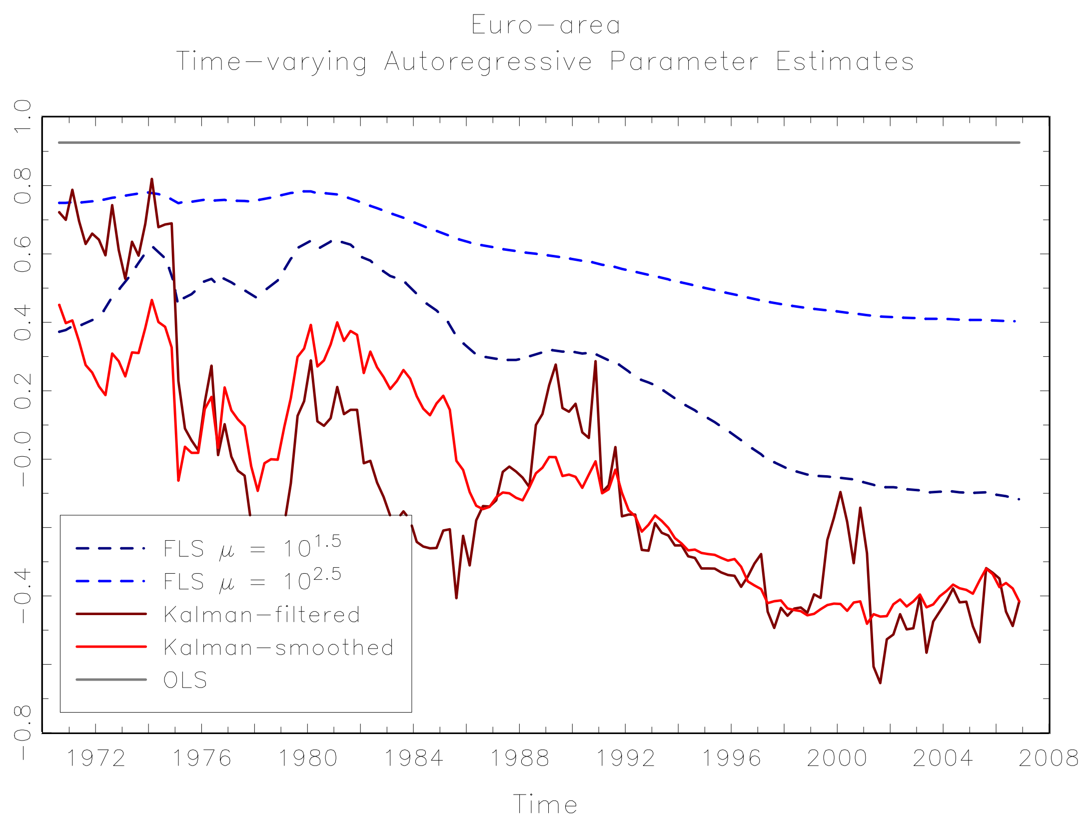
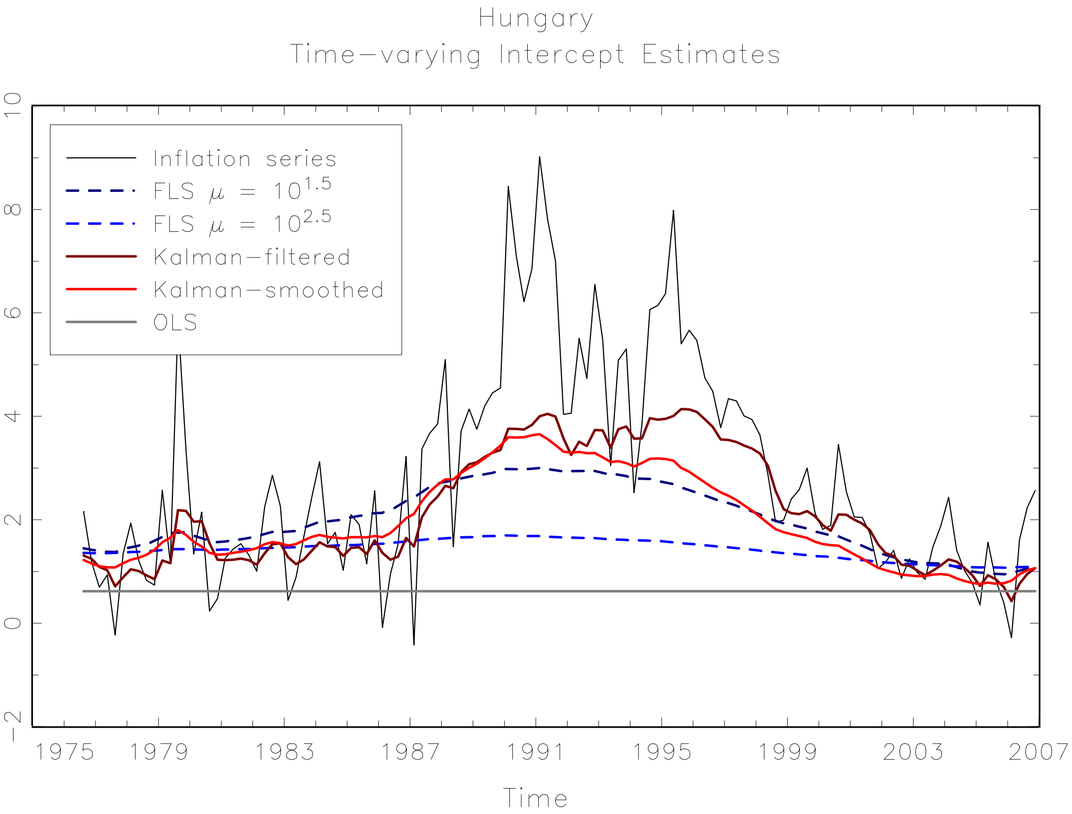


Figure 6: Euro-area – Estimation results using OLS, FLS, and the Kalman-filter



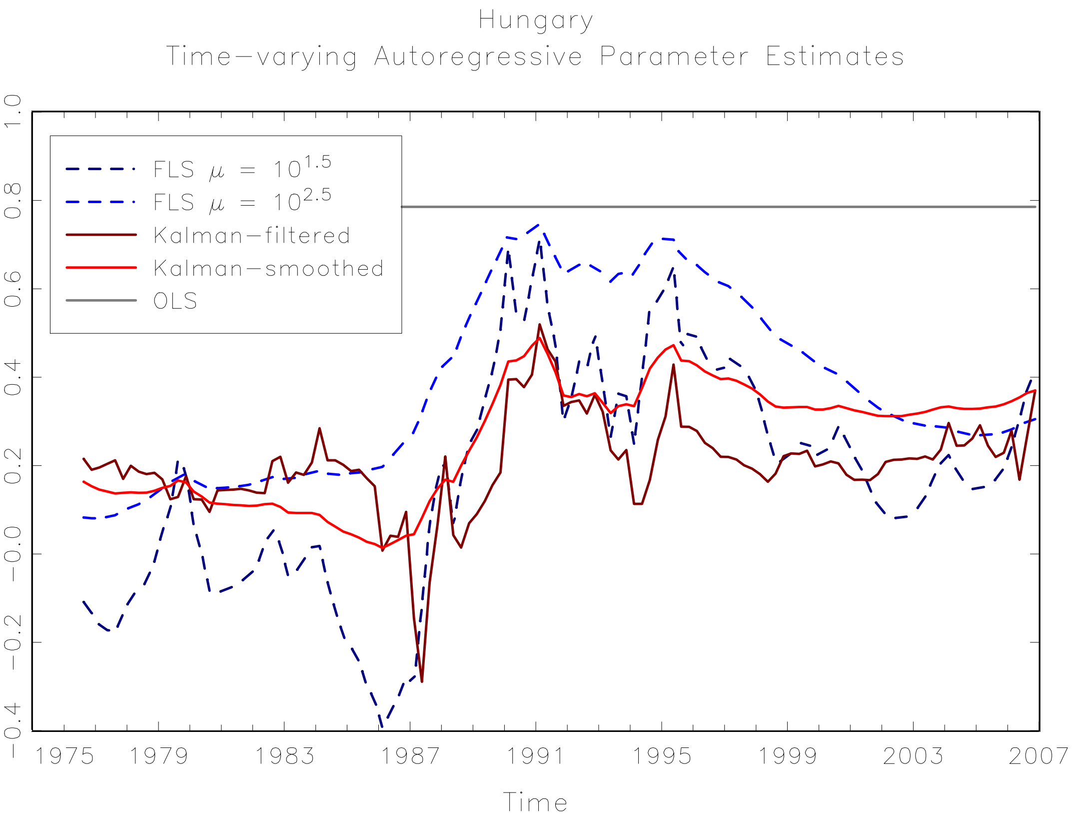
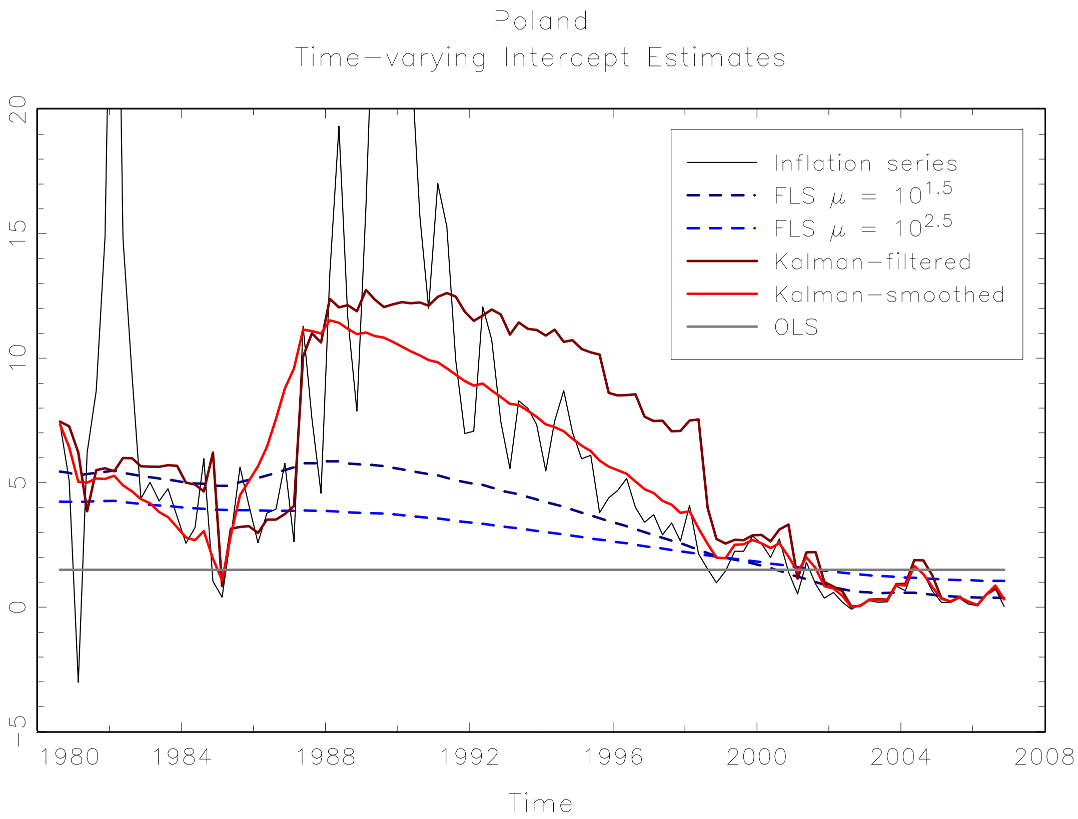


Figure 7: Hungary – Estimation results using OLS, FLS, and the Kalman-filter



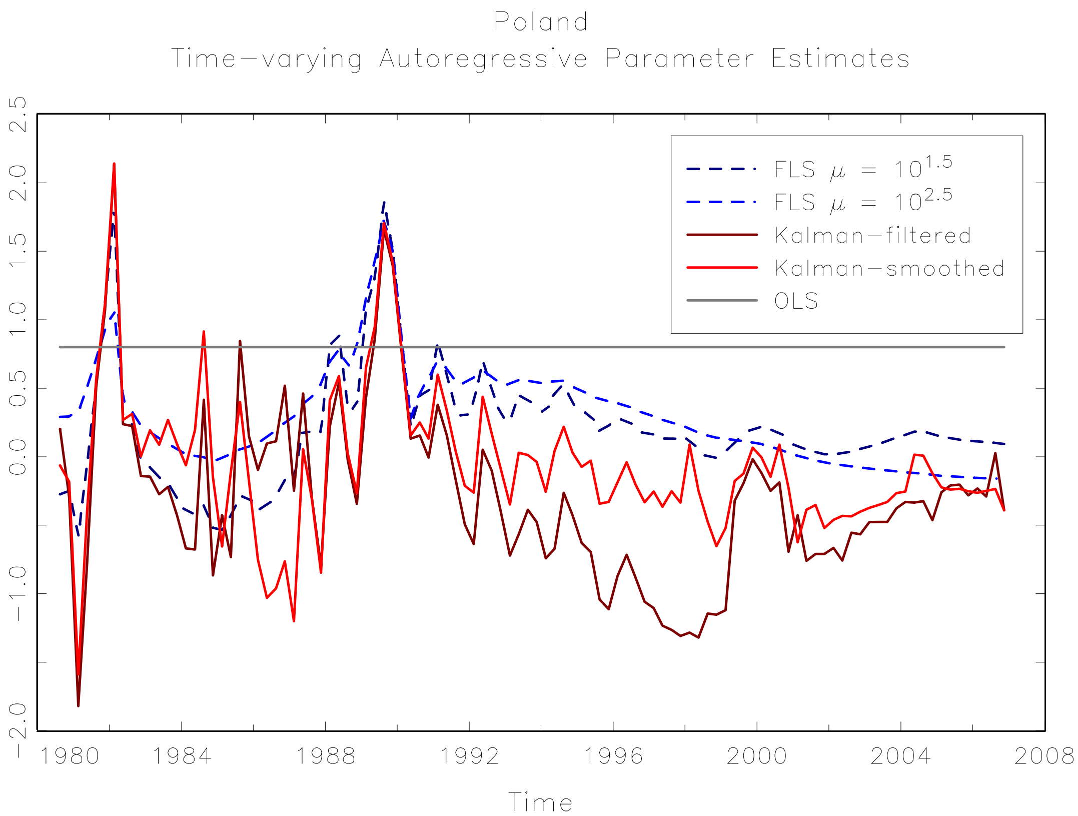


Figure 8: Poland – Estimation results using OLS, FLS, and the Kalman-filter

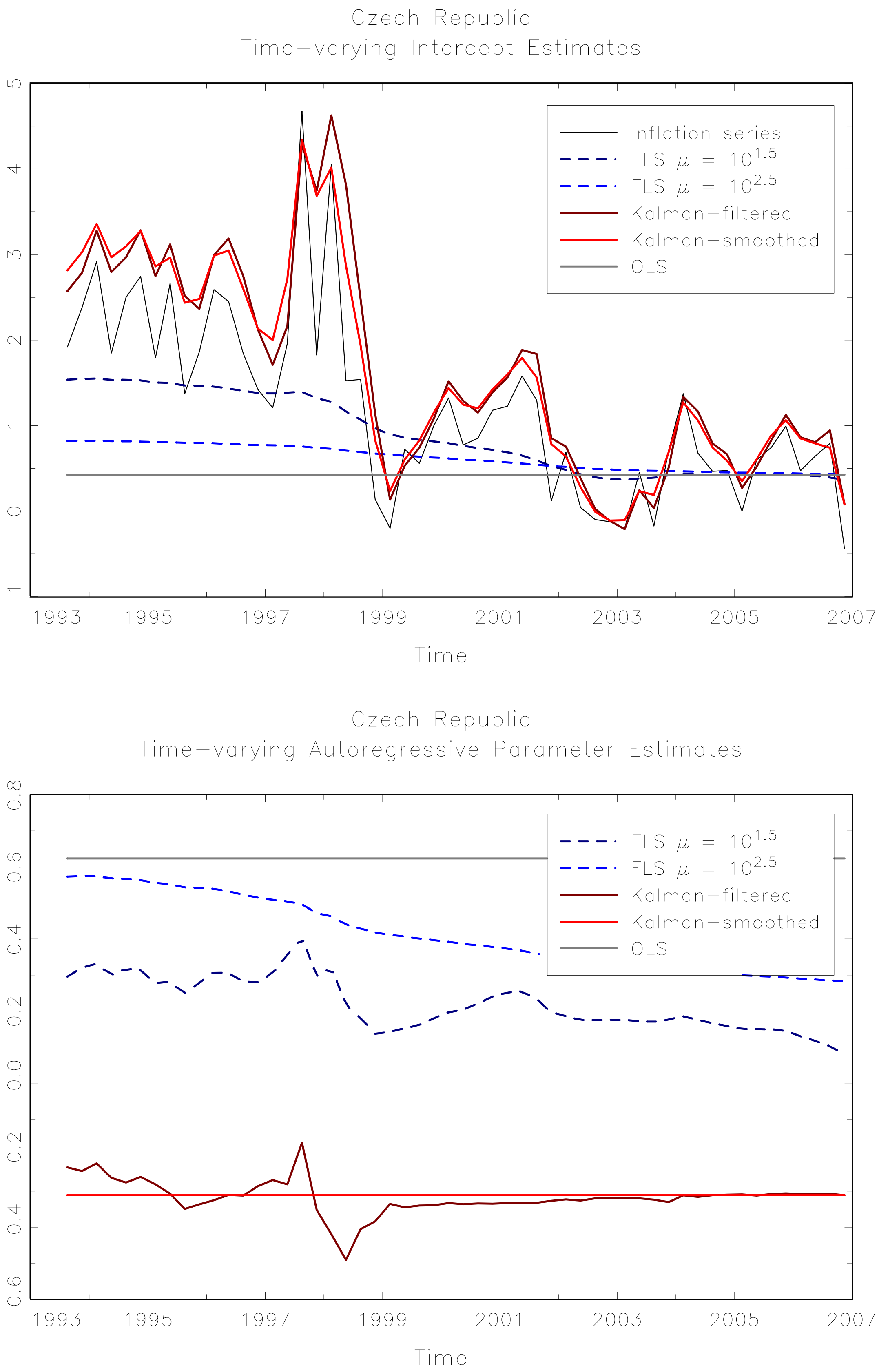
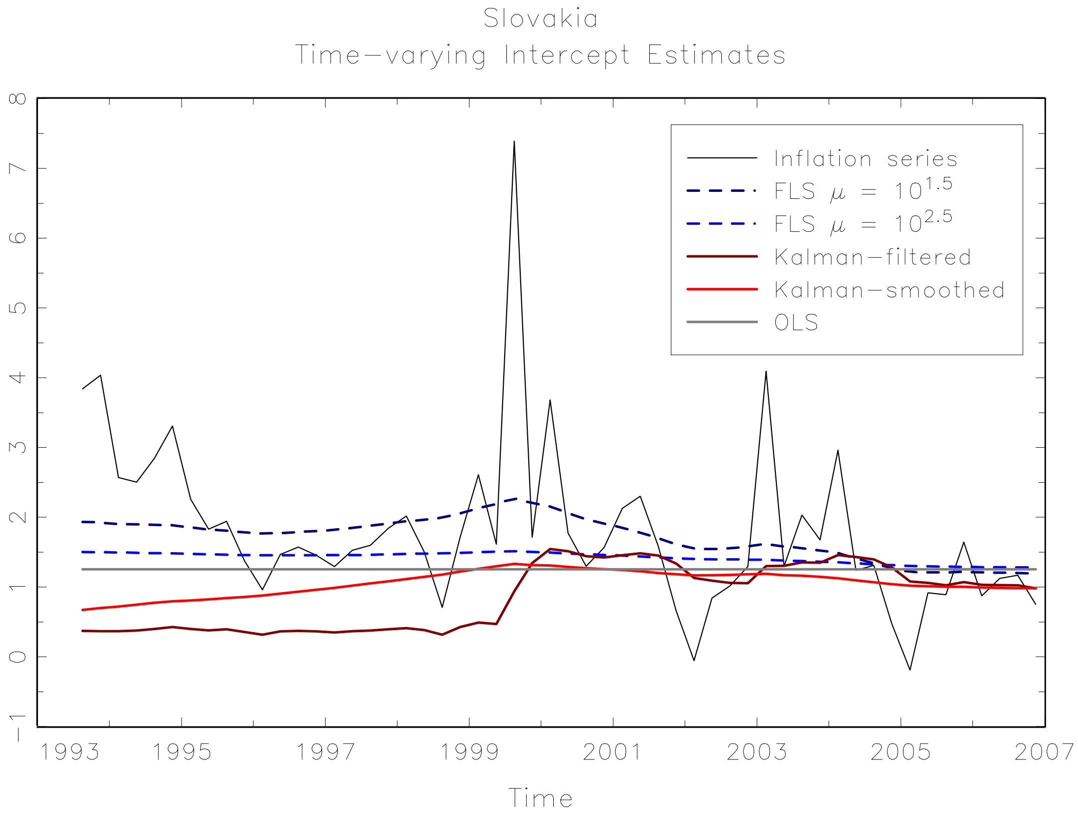


Figure 9: Czech Republic – Estimation results using OLS, FLS, and the Kalman-filter



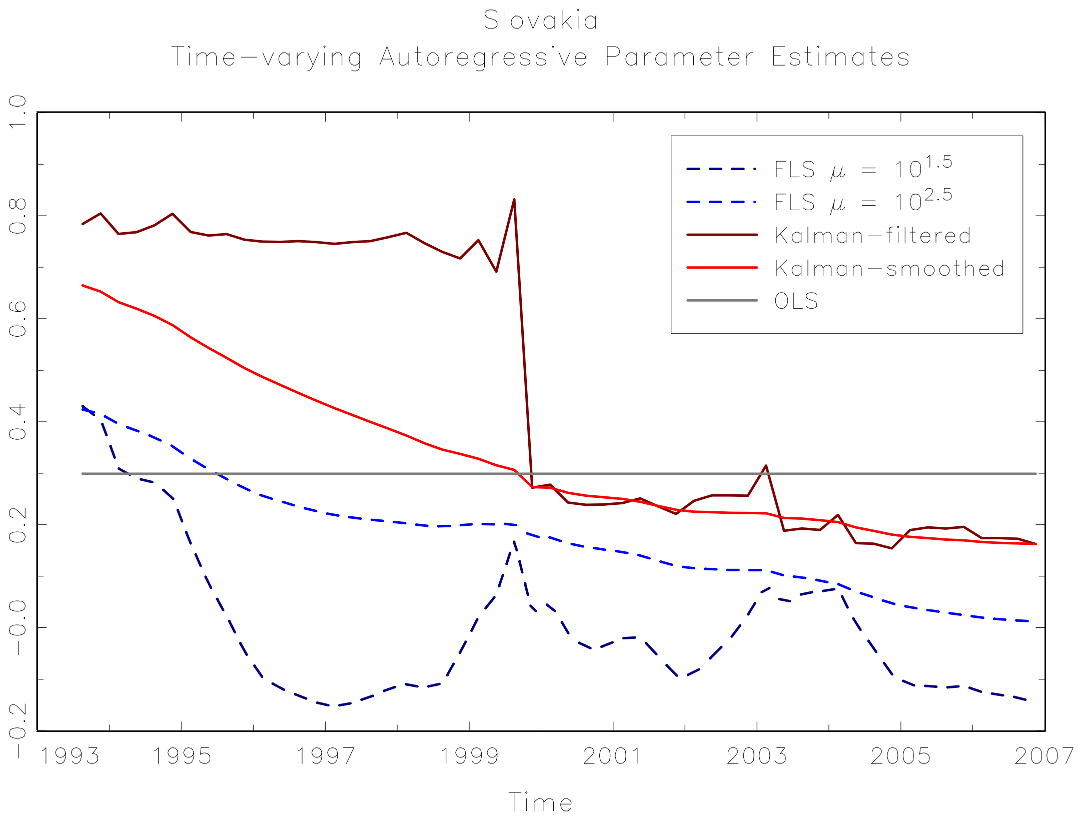


Figure 10: Slovakia – Estimation results using OLS, FLS, and the Kalman-filter