WORKING PAPER SERIES

Forecasting with an Adaptive Control Algorithm

Donald S. Allen
Yang Woo Kim
Meenakshi Pasupathy

Working Paper 1996-009A
http://reseach.stlouisfed.org/wp/1996/96-009.pdf

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

FORECASTING WITH AN ADAPTIVE CONTROL ALGORITHM

July 1996

ABSTRACT

We construct a parsimonious model of the U.S. macro economy using a state space representation and recursive estimation. At the core of the estimation procedure is a prediction/correction algorithm based on a recursive least squares estimation with exponential forgetting. The algorithm is a Kalman filter-type update method which minimizes the sum of discounted squared errors. This method reduces the contribution of past errors in the estimate of the current period coefficients and thereby adapts to potential time variation of parameters. The root mean square errors of out-of-sample forecast of the model show improvement over OLS forecasts. One period ahead in-sample forecasts showed better tracking than OLS in-sample forecasts.

JEL CLASSIFICATION: E17, E27

KEYWORDS: Forecasting, Time-varying parameters, Adaptive Control

Donald S. Allen
Economist
Research Department
Federal Reserve Bank of St. Louis
411 Locust Street
St. Louis, MO 63102

Yang Woo Kim
Visiting Scholar-Federal Reserve Bank
The Bank of Korea
Institute for Monetary
and Economic Research

Meenakshi Pasupathy
Visitor-Federal Reserve Bank
Baruch College of the City University
of New York

Acknowledgement: Rob Dittmar provided programming and numerical methods assistance for this article.

Introduction

Implicit in the use of ordinary least squares loss functions in econometric estimation is the assumption of time invariance of system parameters. Intuition suggests that in a macroeconomic environment fundamental relationships among aggregate variables change over time for various reasons such as the Lucas critique or compositional effects of heterogeneity. Dynamic systems with time-varying properties present a fundamental problem in control and signal processing. Sargent (1993) suggests that adaptation is an important part of economic dynamics when considering issues of learning and bounded rationality. Recursive estimation methods play a key role in adaptation and tracking of time-varying dynamics. A good survey of the basic techniques used to derive and analyze algorithms for tracking time-varying systems is provided in Ljung and Gunnarsson (1990).

The purpose of this paper is to use a recursive method to estimate coefficients of a parsimonious model of the U.S. macro economy and compare forecast results to that of a simple OLS model. We use the state space representation approach, where state variables (X) are used to define the state of the economy at each period in time, and the control variable (u) is assumed to be exogenously determined. If we believe that the parameters of the system change slowly over time, the ordinary least squares method can be extended by weighting current information more heavily than past information. In this model we use a recursive least squares with exponential forgetting algorithm to estimate the coefficients of the state space model.

The resulting "transition" matrix is used for a short run forecast of the state variables, assuming time invariance after the end of the sample period. The root mean square of the forecast errors are better for this model than a simple OLS model. The rest of the paper is organized as follows: Section 2 briefly discusses the state space representation and the recursive least squares algorithm. Section 3 discusses the choice of variables and their transformations and the parameters of choice for the algorithm. Section 4 presents the results of the model and Section 5 concludes the paper.

2. State Space Representation and Recursive Estimation

The state space representation of systems[1] is characterized by the choice of n state variables (usually denoted by X) which are assumed to fully describe the system at any point in time, a control variable (denoted by u) which is assumed to be exogenously determined and can be used to move the system from its current state to another state, and a measured or output variable (denoted by y) which is of particular interest. The system can then be described by a triplet A, b, and c as in equations (1) and (2) below.[2]

$$X_t = A X_{t-1} + b u_t \tag{1}$$

$$y_t = c^T X_t \tag{2}$$

[1] For this paper we assume a single-input single-output (SISO) system which means that our output and control variables are scalars.

[2] The system described is fully deterministic. Stochastic errors can be assumed to enter either additively or in the coefficients.

X is an n x 1 vector of state variables which describes the economy,[3] A is an n x n matrix of coefficients, u is a control variable (scalar), b is an n x 1 vector of coefficients, y is the output variable (of interest), c is an n x 1 vector of coefficients. In a stochastic environment we assume that the measurements of the variables are noisy and uncertain and the noise components are independent, identical normally distributed disturbances. For this paper we will focus only on estimating the state transition equation (1).

Recursive Least Squares with exponential forgetting:

Recursive Least Squares (RLS) estimation is a special case of the Kalman filter which can be used to avoid the numerical difficulties of matrix inversion present in ordinary least squares (OLS) estimation. OLS is applied to the first k observations of the data sample to determine a starting point for parameter estimates. Each additional observation is used to update coefficient estimates recursively, thus avoiding the need for further matrix inversion. With proper choice of the initial conditions, the final estimator at the end of the sample period is equal to the OLS estimator. Harvey (1993) summarizes the method. The RLS method with exponential forgetting used here modifies the basic RLS updating algorithm to weigh new information more heavily. The method is appealing for cases where time-varying parameters are suspected.

For an equation of the form

$$z(t) = \varphi^T(t)\ \Theta \qquad \textbf{(3)}$$

[3] The accepted format is a first order difference equation. If additional lags of particular variables are desired then the list of variables is expanded appropriately by defining lagged values of these variables as X's. It can be shown that an ARMA representation can be modeled by this first order difference equation model.

where Θ is a vector of model parameters and $\varphi(t)$ is a set of explanatory variables, the usual quadratic loss function is replaced by a discounted loss function of the form

$$V(\Theta,t)=1/2\sum_{i=1}^{t} \lambda^{t-i}(z(i)-\varphi^T(i)\Theta)^2 \tag{4}$$

where λ is a number less than or equal to one and is referred to as the forgetting factor. The recursive algorithm is given by

$$\begin{aligned} \hat{\Theta}(t) &= \hat{\Theta}(t-1) + K(t)\ (z(t) - \varphi^T(t)\ \hat{\Theta}(t-1)) \\ K(t) = P(t)\ \varphi(t) &= P(t-1)\ \varphi(t)\ (\lambda I+\varphi^T(t)P(t-1)\ \varphi(t))^{-1} \\ P(t) &= (I - K(t)\varphi^T(t))\ P(t-1)/\lambda \end{aligned} \tag{5}$$

The essential feature of the algorithm is that t-1 estimates of Θ are adjusted with new information by a transformation of the error in predicting z using Θ_{t-1} and current φ's. The adjustment to the error, K(t), is called the Kalman gain and is a function of the rate of change in the errors and is weighted by the discount factor λ. P(t) is the covariance matrix at time t. Both K(t) and the moment matrix P(t) are updated recursively. In the state space model of equations (1) and (2), z(t) is X_t, $\varphi(t)$ is X_{t-1} and $\Theta(t)$ is A(t), or $\varphi(t)$ can be $[X_t, u_t]$ and $\Theta(t)$ would correspond to [A(t), b(t)].

3. Data and Transformations Used

The choice of state variables was based on a variety of factors including availability on a monthly frequency, explanatory capacity, timeliness of data release and available sample size. From an initial list of 18 variables we decided that the U.S. macro economy can be parsimoniously described by twelve variables (transformed appropriately). We define the state variable vector X as the following: consumption (CBM), investment(CPB), industrial production (IP), changes in manufacturing inventory (MIM), changes in retail inventory (TRIT), manufacturing inventory/sales ratio (MR1), retail inventory/sales ratio (TSRR), urban CPI (PCU), total employment (LE), 3-month Treasury Bill interest rate (FTB3M), and M2 monetary aggregate (FM2). We choose the control variable u as the Fed Funds rate (FFED). New construction is used as a proxy for investment. The data are of monthly frequency, and the sample period considered is January, 1981 through December, 1995.

The variable names are listed in Table 1 and following transformations were made: log levels of CBM, CPB, IP, LE, FM2 and PCU; log differences of MIM, and TRIT; MR1, TSRR, FTB3M, and, FFED were untransformed.

First-differencing and deflation of data The objective of the forecasting model is to track variables that may be changing over time. Despite evidence of unit roots in all variables except MR1, TSRR and 3-month T-Bill rate (FTB3M), detrending of data via first differencing or filtering was not deemed necessary for tracking purposes.[4] Inventory data was differenced for two reasons: first, because change in inventory is a component of Gross Domestic Product (GDP) and second, because including first-differences in inventories (instead of levels) appeared to improve the model's ability to capture turning points in the business cycle. Nominal values were used for all variables along with the consumer price index as one of the state variables to observe the effect of changes in fed funds rate on inflation.

[4] The existence of a trend does not adversely impact the performance of tracking algorithms. Results of unit root tests are not reported here but are available.

Estimation Procedure

The algorithm shown in equation 5 was used to estimate the coefficients for each variable individually. The right hand side variables are the one period lags of all the state variables plus the current period value of the control variable. The final period estimate of coefficients was assembled into the equivalent A matrix and b vector for forecasting.

Choice of λ, Θ_0, and P_0 The P matrix was initialized as the identity matrix and the Θ's were initialized as an AR(1) process with coefficient 0.8. The forgetting factor, λ, was chosen to be 0.96411. This value of the forgetting factor reduces the weight of the error after 63 months, (a period equivalent to the average length of postwar business cycles), to 10 percent. Lower values of λ, which represent higher rates of forgetting, led to improvement in tracking but resulted in a higher variance of the estimated Θ's over time. The model converges relatively quickly and initial values affect only the early estimates in the sample. Using VAR coefficient estimates as starting values for the parameters did no better than using unit roots as initial starts. Starting coefficients of 0.8 on an AR(1) model were chosen to avoid any biases toward a unit root. High variance on the initial P matrix, which is equivalent to a diffuse prior, resulted in higher variance of Θ and exaggerated the "turning points" in the forecast. Estimates of the A matrix were nonsingular and had stable eigenvalues with the "typical" assumptions for λ, Θ_0, P_0.

4. Results

The model is distinguished by the recursive technique which updates past estimates of coefficients using the error in the one period ahead forecast. As a first test of the model's performance, the one period ahead forecast of the model for each state variable was recorded for each period in the sample and compared to the actual values. Since the initial values, Θ_0, of the parameters were chosen arbitrarily, and the data used to adjust errors are limited initially,[5] the estimates took a few periods to converge. Once convergence was achieved, the estimates tracked the actual values very closely. The Root Mean Square Error (RMSE) of the in-sample forecasts of consumption, industrial production, employment, and CPI for the most recent 60 month period (AKP) were monitored. The model was also estimated using OLS for the period February, 1981 through December, 1995. In-sample forecasts from OLS and the one-period ahead forecasts obtained using the recursive least squares procedure were considered for the last sixty sample periods (91:01 - 95:12) and the percentage RMSE were computed for the four variables of interest.[6] Table 2 shows the comparison of the RMSE to the OLS in-sample forecast for the same period. For all four variables the RLS with forgetting had lower RMSE than the OLS.

Out-of-sample forecasts for 5-months (January, 1996 to May, 1996), were made using the recursive least squares and OLS estimation procedures. Two different RLS out-of-sample forecasts were made for comparison purposes. One assumed no new information was available for updating the coefficients, the other assumed that coefficients were updated each period. The 5-month forecast using the constant parameter estimates (A matrix) at the end of

---

[5] The recursive technique by definition uses only information from the past. Hence early estimates use limited data.

[6] The RMSE were computed after transforming the data back to their original form by exponentiating the forecasted values. % RMSE was computed using the formula

$$\sqrt{1/n\sum_{i=1}^{n}\left(\frac{y_i-\hat{y}_i}{y_i}\times 100\right)^2}.$$

95:12 (without update) is called AKP5. The second set of out-of-sample one period ahead forecasts using the updated coefficients is called AKP1. The percentage RMSE from the out-of sample OLS forecasts and the recursive least squares procedure forecasts are compared in Table 3. For all variables except industrial production using the AKP1 method, the recursive least squares model had a lower % RMSE than OLS. These results suggest that the recursive least squares procedure provides better tracking and forecasting (both in-sample and out-of-sample) than OLS for the model developed in this paper. Figures 1 - 4 represent the one period ahead in-sample forecast for the last 24 months of the sample, the 5 period forecast (AKP5) and the one period ahead forecast (AKP1) compared to actual values (for the last 29 periods) for the four variables of interest. As the figures show, the tracking is quite close in all four cases.

Sensitivity to λ: The estimation and forecasting results discussed above are obtained by setting the forgetting factor equal to 0.96411. Sensitivity of the results to changes in the rate at which past errors are discounted was studied by setting λ equal to 0.8 and 1.0. The percentage RMSE for the in-sample and out-of-sample forecasts of industrial production for the three different choices of λ are given in Table 4. RMSE is the lowest for the choice λ=0.96411 in almost all cases. Figures 5 and 6 compare the OLS forecasts with the recursive least squares procedure (AKP5) forecasts for two different values of λ. The graphs show that the forecasts obtained using λ=0.96 perform much better than OLS forecasts, while the forecasts corresponding to λ=1.0 are worse than OLS forecasts. Figures 8 - 10 compare the two sensitivity cases with the actual and the λ=0.96 case. As expected, forecasts are more volatile with lower λ. This is because lower values of λ weigh recent errors more heavily and the correction tends to be sharper in forecasts. For similar reasons we expect forecasts with higher values of λ to be smoother, and the figures show that the forecasts are indeed smoother when λ=1.0. Although our choice of λ was not an attempt to optimize, it does better than the alternatives used in the sensitivity tests.

5. Summary and Conclusions

The purpose of this exercise was to develop a relatively parsimonious macroeconomic model which might be useful for short term forecasts and would recognize the potential for time varying parameters. Using a measure of per cent RMSE of in-sample and out of sample forecasts, the model does better than a simple OLS model. Because there was little attempt to steep the model in theoretical microfoundations, we would not recommend it at this time for use in long-term forecasting. However, it is parsimonious enough to be used in the decision making process.

Table 1 - List of Variables Used

Name	Variable
IP	Industrial Production
CPB	Total New Construction
MIM	Manufacturers Inventories
TRIT	Retailers Inventories
MR1	Manufacturers Inventory/Sales Ratio
TSRR	Retailers Inventory/Sales Ratio
FM2	M2 Money Stock
FTB3M	3-Month Treasury Bill (Auction Average)
PCU	Urban Consumers Price Index (All Items)
LE	Number of Civilians over 16 employed
CBM	Personal Consumption Expenditure
FFED	Fed Funds Rate

Table 2 - % RMSE for (60 Period) In - Sample Forecasts

	IP	CONS	LE	PCU
AKP	0.418	0.485	0.230	0.132
OLS	0.738	0.828	0.248	0.374

Table 3 - % RMSE for (5 Period) Out of Sample Forecasts

	IP	CONS	LE	PCU
AKP1	0.652	0.468	0.214	0.137
AKP5	0.608	0.389	0.429	0.301
OLS	0.626	0.808	0.582	0.393

Table 4 - % RMSE for Industrial Production For Different Values of Lambda

	AKP5			AKP1		
	λ=0.8	λ=0.96	λ= 1.0	λ=0.8	λ=0.96	λ= 1.0
RMSE 60	0.699984	0.418171	0.521724	0.699984	0.418171	0.521724
RMSE 5	0.649919	0.608086	1.481615	0.413427	0.651874	0.749956

Figure 1: Actual vs. Forecast of Consumption

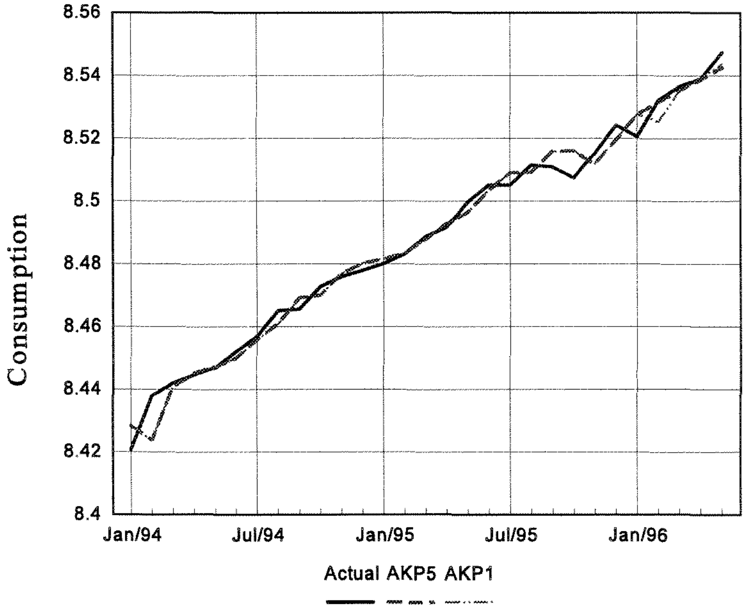


Figure 2: Actual vs.Forecast of Industrial Production

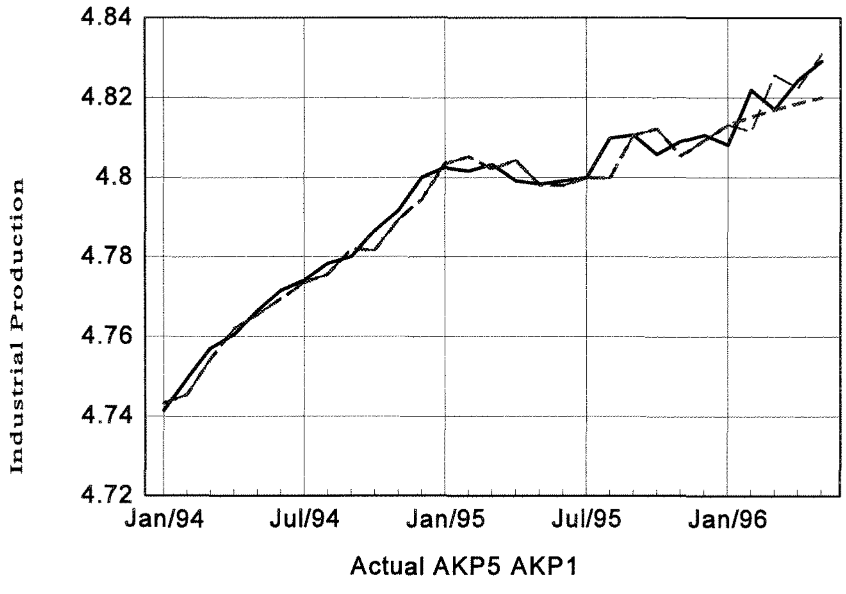


Figure 3: Actual vs. Forecast Employment

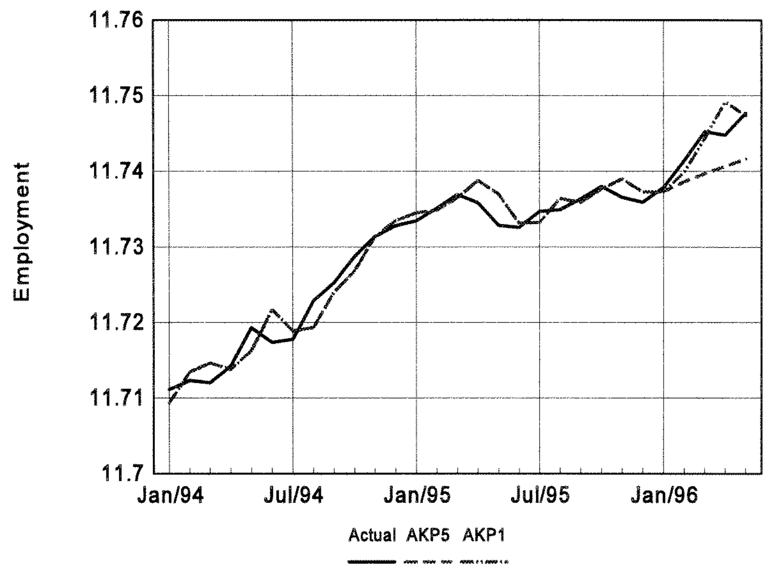


Figure 4: Actual vs. Forecast CPI

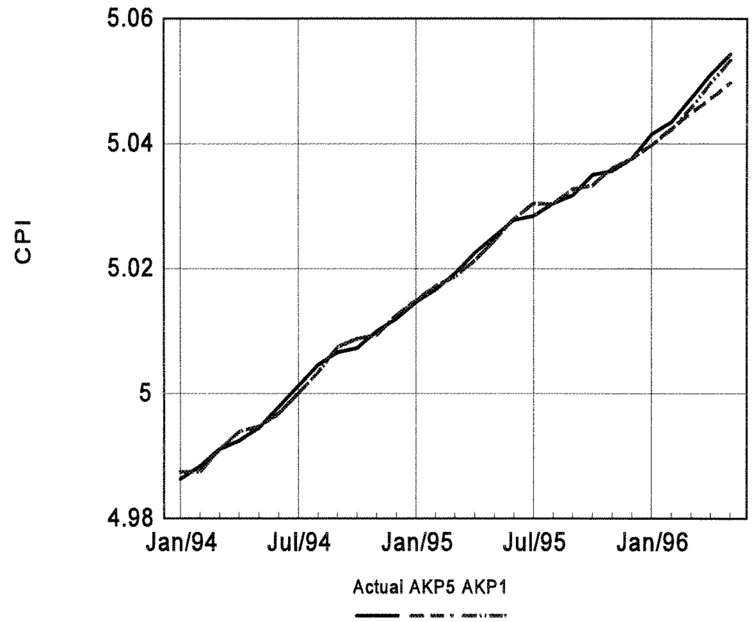


Figure 5: AKP5 (lambda=0.96) and OLS Forecasts for IP

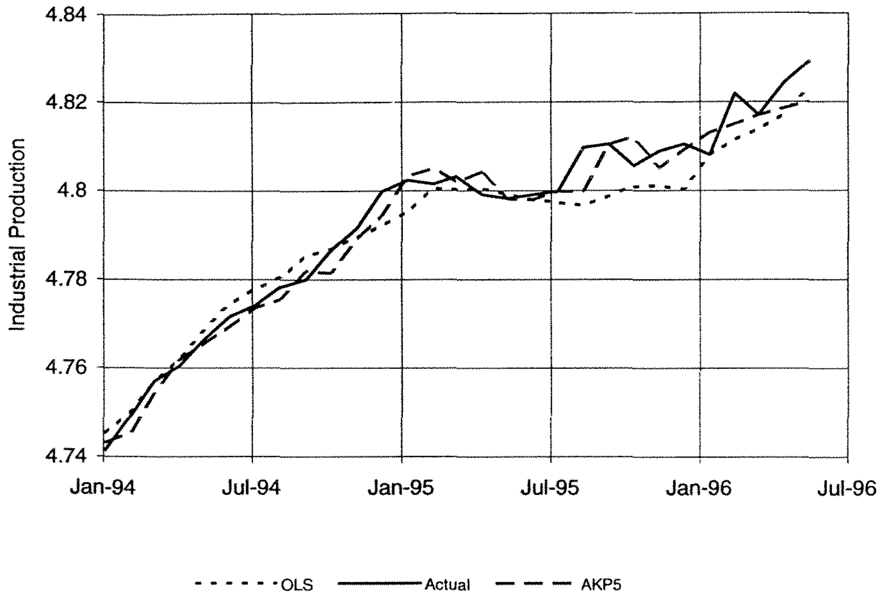


Figure 6: AKP5 (lambda=1.0) and OLS Forecasts for IP

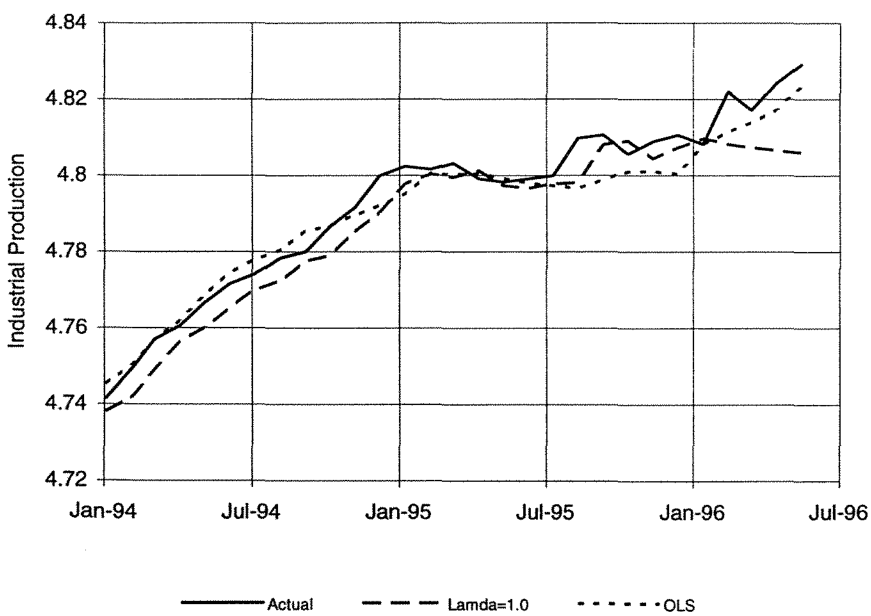


Figure 7: Forecasts (AKP5) of IP for Different Values of Lambda

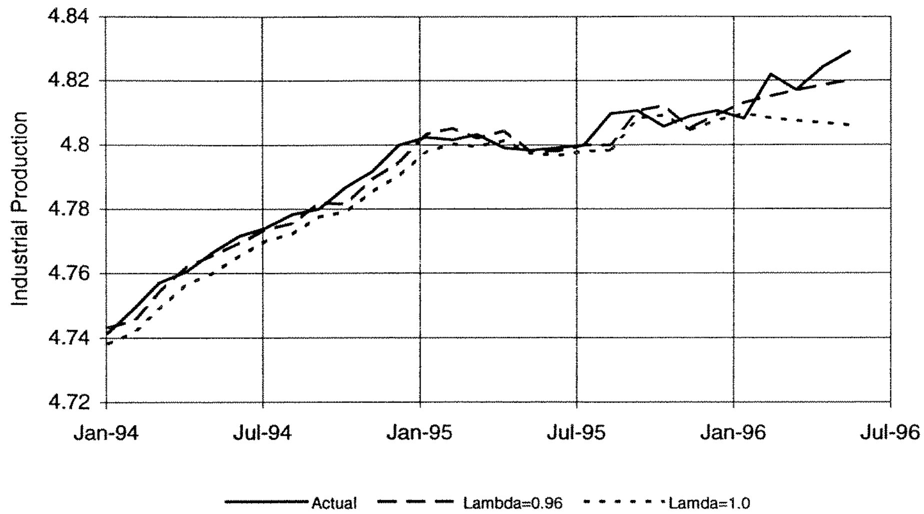


Figure 8: Forecasts (AKP5) of IP for Different Values of Lambda

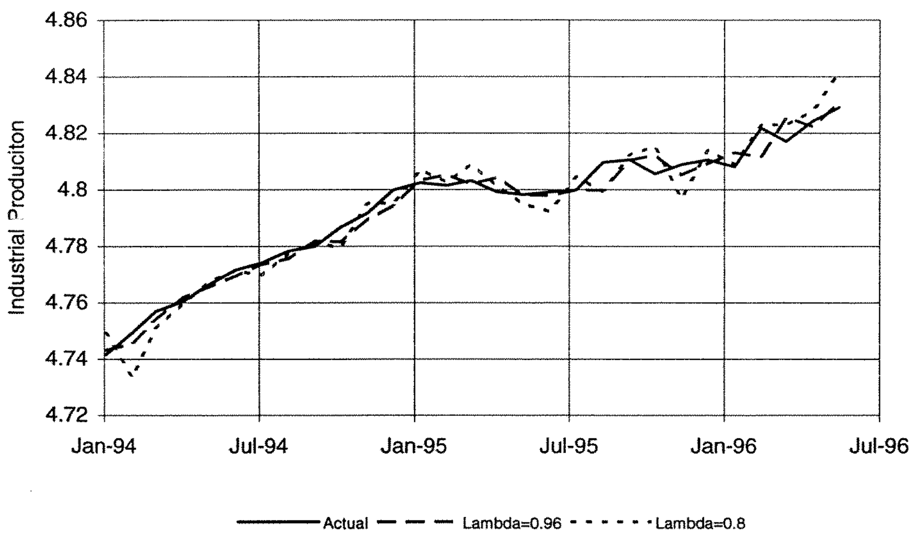


Figure 9: Forecasts (AKP1) of IP for Different Values of Lambda

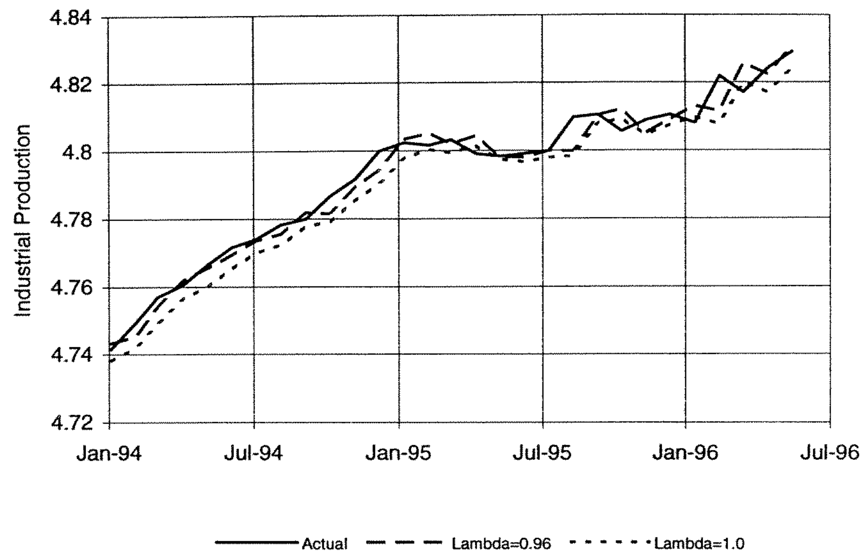
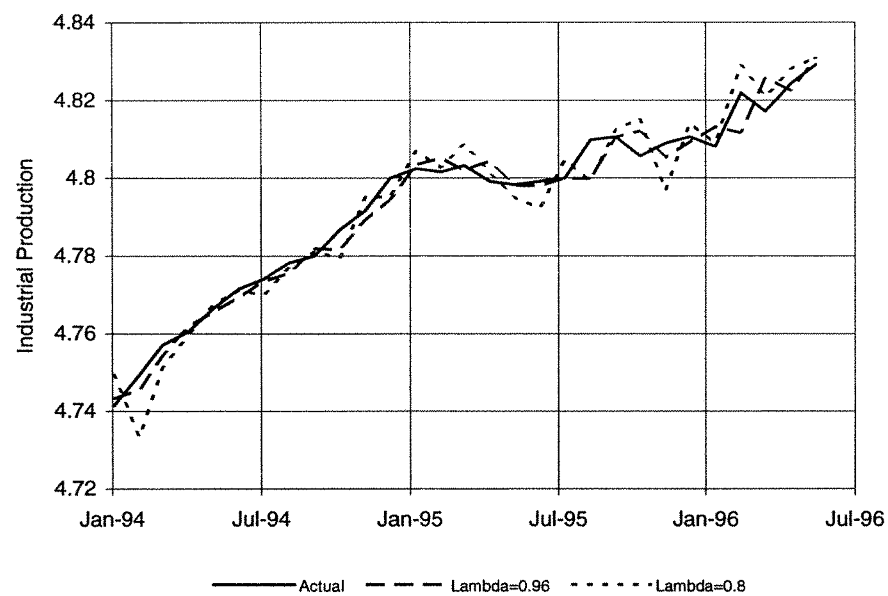


Figure 10: Forecasts (AKP1) of IP for Different Values of Lambda



References

Äström, Karl Johan, and Björn Wittenmark, *Adaptive Control*, Addison-Wesley Publishing Company, New York, 1989.

Bryson, Arthur E., and Yu-Chi Ho, *Applied Optimal Control: Optimization, Estimation, and Control* John Wiley & Sons, New York, 1975.

Goldberger, A. S., "Unbiased Prediction by Recursive Least Square", *Econometrica*, Vol. 38, No. 2, p.367, March, 1970.

Gunnarsson, Svante, and Lennart Ljung, "Adaptation and Tracking in System Identification - A Survey," *Automatica*, Vol. 26, No. 1, pp. 7-21, 1990.

Harvey, Andrew C., *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, 1990.

___________, *Time Series Models*, Second Edition, The MIT Press, Cambridge, Massachusetts, 1993.

Sargent, Thomas J., *Bounded Rationality in Macroeconomics: The Arne Ryde Memorial Lectures*, Clarendon Press, Oxford, 1993.