Working Paper 99-31
Statistics and Econometrics Series 09
April 1999

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax 34 - 91- 624.9849

# EFFECTS OF PARAMETER ESTIMATION ON PREDICTION DENSITIES: A BOOTSTRAP APPROACH

Lorenzo Pascual, Juan Romo and Esther Ruiz*

Abstract

In this paper, we study the impact of parameter estimation on prediction densities using a bootstrap strategy to estimate these densities. We focus on seasonal ARIMA processes with possibly non normal innovations. We compare prediction densities obtained using the Box and Jenkins approach with bootstrap densities which may be constructed taking into account parameter estimation variability (PRR) or using parameter estimates as if they were the true parameters (CB). By means of Monte Carlo experiments, we show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. The effects of parameter estimation are particularly important for small sample sizes and when the error distribution is not Gaussian. We also analyze the effect of the estimation method on the shape of prediction densities comparing prediction densities constructed when the parameters are estimated by OLS and by LAD. We show how, when the error distribution is not Gaussian, the average coverage and length of intervals based on LAD estimates are closer to nominal values than those based on OLS estimates. Finally, the performance of the PRR procedure is illustrated with two empirical examples.

Key Words
Forecasting; Least absolute deviations; non normal distributions; Ordinary least squares.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Spain, e-mail: ortega@est-econ.uc3m.es

# EFFECTS OF PARAMETER ESTIMATION ON PREDICTION DENSITIES: A BOOTSTRAP APPROACH

We use a bootstrap procedure to study the impact of parameter estimation on prediction densities. We focus on seasonal ARIMA processes with possibly non normal innovations. We compare prediction densities obtained using the Box and Jenkins approach with bootstrap densities which may be constructed either taking into account parameter estimation variability or using parameter estimates as if they were known parameters. By means of Monte Carlo experiments, we show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. The effects of parameter estimation are particularly important for small sample sizes and when the error distribution is not Gaussian. We also analyze the effect of the estimation method on the shape of prediction densities comparing prediction densities constructed when the parameters are estimated by OLS and by LAD. We show how, when the error distribution is not Gaussian, the average coverage and length of intervals based on LAD estimates are closer to nominal values than those based on OLS estimates. Finally, the performance of the PRR procedure is illustrated with two empirical examples.

Key Words: Forecasting, Least Absolute Deviations, Non normal distributions, Ordinary Least Squares.

## 1. INTRODUCTION

Our main goal in this paper is to study the impact of parameter estimation on prediction densities and we use the bootstrap as a device to asses its relevance. In the standard approach to construct prediction intervals, based on Box and Jenkins (1976), prediction errors are assumed to be Gaussian and intervals are obtained with center at the point linear predictor and conditioning on parameter estimates. Consequently, Box and Jenkins (BJ) intervals do not take into account the variability due to parameter estimation and may have coverage which is badly different from the nominal one when the errors are not Gaussian. Alternatively, prediction intervals can be built using bootstrap procedures. Bootstrap intervals can incorporate the variability due to parameter estimation without assuming any particular distribution for the errors. We analyze the effect of parameter estimation on the shape of prediction densities using the bootstrap procedure proposed by Pascual, Romo and Ruiz (1998) for ARIMA$(p, d, q)$ models.

First, estimating the parameters by conditional Quasi-Maximum Likelihood (QML), we compare bootstrap intervals (PRR) constructed taking into account parameter variability with intervals obtained by using parameter estimates as if they were the true parameters. The latter approach will be referred as conditional bootstrap (CB). We compare average covering and length of BJ, CB and PRR intervals. The difference between BJ and CB intervals could be assignable to the deviation of the innovation distribution from the Gaussian assumption. The difference between CB and PRR intervals could be due to parameter estimation uncertainty. Consequently, we can distinguish between the two sources which could affect the precision of prediction intervals when the specification of the model is known. As expected, given that the conditional QML estimator is consistent, the variability due to parameter estimation should be taken into account in the construction of prediction intervals when the series sample size is not large enough. In this case, intervals obtained by conditioning on parameter estimates have average coverage under the nominal value. As the sample size increases, the effect of parameter variability is less important. We also study the effect of the estimation method on the shape of prediction densities. In particular, we consider the prediction of future values of ARI$(p, d)$ processes and compare prediction intervals obtained when estimating by Ordinary Least Squares (OLS) and by Least Absolute Deviations (LAD). When the error distribution is not Gaussian, prediction densities based on the LAD estimator have, in general, shapes closer to the corresponding empirical prediction densities. As a second objective of this paper, we show how the bootstrap procedure proposed by Pascual, Romo and Ruiz (1998) can be extended to construct prediction intervals in multiplicative seasonal ARIMA models.

The paper is organized as follows. First, in section 2 we describe the bootstrap procedure proposed by Pascual, Romo and Ruiz (1998) to construct prediction intervals. Then, section 3 contains the Monte Carlo results on the effects of parameter variability on the shape of prediction densities when seasonal ARIMA models are estimated by conditional QML. Also, we carry out experiments to asses the effects of the method used to estimate the model parameters on prediction intervals. In section 4, we apply the bootstrap PRR procedure to obtain prediction densities for two real time series: monthly observations of the Italian Industrial Production Index and levels of a luteinizing hormone measured on a healthy woman. Finally, section 5 contains the conclusions and some suggestions for further research.

## 2. BOOTSTRAP PREDICTION INTERVALS

We now describe the bootstrap procedure proposed in Pascual, Romo and Ruiz (1998) to construct prediction intervals for future values of series generated by ARIMA$(p, d, q)$

processes given by

$$\nabla^d y_t = \phi_0 + \phi_1 \nabla^d y_{t-1} + ... + \phi_p \nabla^d y_{t-p} + a_t + \theta_1 a_{t-1} + ... + \theta_q a_{t-q}, \quad (1)$$

where $a_t$ is a white noise process, $\nabla$ is the difference operator such that $\nabla y_t = y_t - y_{t-1}$ and $\phi_0, \phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ are unknown parameters. From an observed series $\{y_1, y_2, ..., y_T\}$, the parameters can be estimated by a consistent estimator, for example conditional QML. Given $(\widehat{\phi}_0, \widehat{\phi}_1, \ldots, \widehat{\phi}_p, \widehat{\theta}_1, \ldots, \widehat{\theta}_q)$, the residuals are calculated by the following recursion

$$\widehat{a}_t = \nabla^d y_t - \widehat{\phi}_0 - \widehat{\phi}_1 \nabla^d y_{t-1} - ... - \widehat{\phi}_p \nabla^d y_{t-p} - \widehat{\theta}_1 \widehat{a}_{t-1} - ... - \widehat{\theta}_q \widehat{a}_{t-q}, t = p+d+1, ..., T, \quad (2)$$

where the residuals corresponding to periods of time $t = 0, -1, -2, ...$ are set equal to zero. Denote by $\widehat{F}_a$ the empirical distribution function of the centered residuals. Given a set of $p + d$ initial values of the variable $y_t$, say $\{y_1, \ldots, y_{p+d}\}$, a bootstrap replicate of the series $\{y^*_{p+d+1}, \ldots, y^*_T\}$ is constructed by the following recursion

$$\nabla^d y^*_t = \widehat{\phi}_0 + \sum_{j=1}^{p} \widehat{\phi}_j \nabla^d y^*_{t-j} + \sum_{j=1}^{q} \widehat{\theta}_j \widehat{a}^*_{t-j} + \widehat{a}^*_t, \ t = p + d + 1, \ldots, T, \quad (3)$$

where $y^*_t = y_t, t = 1, ..., p + d$ and $\widehat{a}^*_{1+p+d-q}, \ldots, \widehat{a}^*_T$ are random draws from $\widehat{F}_a$. Once the parameters of this bootstrap series are estimated, say $(\widehat{\phi}^*_0, \widehat{\phi}^*_1, \ldots, \widehat{\phi}^*_p, \widehat{\theta}^*_1, \ldots, \widehat{\theta}^*_q)$, the bootstrap forecast $k$ steps ahead is obtained as follows,

$$\nabla^d y^*_{T+k} = \widehat{\phi}^*_0 + \sum_{j=1}^{p} \widehat{\phi}^*_j \nabla^d y^*_{T+k-j} + \sum_{j=1}^{q} \widehat{\theta}^*_j \widehat{a}^*_{T+k-j} + \widehat{a}^*_{T+k}, k = 1, 2, ... \quad (4)$$

where $y^*_{T+k-j} = y_{T+k-j}$ , $j \geqslant k$, and $\widehat{a}^*_{T+k-j} = \widehat{a}_{T+k-j}$, $j \geqslant k$, i.e., the last $p + d$ observations of the series and the last $q$ residuals are fixed in order to obtain the prediction density conditional on the observed data. Finally, in expression (4), $\widehat{a}^*_{T+k-j}, j > k$ are random draws from $\widehat{F}_a$.

As an illustration, we consider an ARIMA$(1, 1, 1)$ model without constant term

$$\nabla y_t = \phi \nabla y_{t-1} + a_t + \theta a_{t-1}. \quad (5)$$

Once the parameters of model (5) have been estimated and the bootstrap draws $\widehat{a}_2^*, \ldots, \widehat{a}_T^*$ are available, a bootstrap replicate of the series is constructed by

$$y_t^* = (1+\widehat{\phi})\ y_{t-1}^* - \widehat{\phi} y_{t-2}^* + \widehat{a}_t^* + \widehat{\theta}\widehat{a}_{t-1}^*, \qquad t = 3, \ldots, T, \tag{6}$$

where $y_1^* = y_1$ and $y_2^* = y_2$. Then, bootstrap estimates $\widehat{\phi}^*$ and $\widehat{\theta}^*$ are obtained for the bootstrap series and bootstrap replicates of future values of the series are generated by

$$\begin{aligned} y_{T+1}^* &= (1+\widehat{\phi}^*)y_T - \widehat{\phi}^* y_{T-1} + \widehat{a}_{T+1}^* + \widehat{\theta}^* \widehat{a}_T, \\ y_{T+2}^* &= (1+\widehat{\phi}^*)y_{T+1}^* - \widehat{\phi}^* y_T + \widehat{a}_{T+2}^* + \widehat{\theta}^* \widehat{a}_{T+1}^*, \\ y_{T+3}^* &= (1+\widehat{\phi}^*)y_{T+2}^* - \widehat{\phi}^* y_{T+1}^* + \widehat{a}_{T+3}^* + \widehat{\theta}^* \widehat{a}_{T+2}^*, \end{aligned} \tag{7}$$

and so on.

This procedure is repeated $B$ times to produce a set of $B$ bootstrap replicates $\left\{y_{T+k}^{*(1)}, \ldots, y_{T+k}^{*(B)}\right\}$. Finally, the prediction limits are defined as the quantiles of the bootstrap distribution function of $y_{T+k}^*$, i.e., if $G^*(h) = \Pr(y_{T+k}^* \leq h)$ is the distribution function of $y_{T+k}^*$ and its Monte Carlo estimate is $G_B^*(h) = \#(y_{T+k}^{*(b)} \leq h)/B$, a $100\alpha\%$ prediction interval for $Y_{T+k}^*$ is given by

$$[L_B^*, U_B^*] = \left[Q_B^*\left(\frac{1-\alpha}{2}\right), Q_B^*\left(\frac{1+\alpha}{2}\right)\right], \tag{8}$$

where $Q_B^* = G_B^{*-1}$.

Notice that in the procedure just described, the last $p+d$ observations of the series and the final $q$ residuals are fixed in all bootstrap replicates of future values so we can obtain the prediction density conditional on the observed sample. However, we do not fix any observation when generating bootstrap replicates of the series used to obtain bootstrap estimates of the parameters of the model.

In the bootstrap procedure proposed by Thombs and Schucany's (1990) for AR($p$) processes, they fix the last $p$ observed values of $y_t$ to obtain bootstrap replicates of the series used to estimate the parameters and, consequently, they need to used the backward representation of the AR($p$) process. Therefore, the main advantage of the method just described over the technique in Thombs and Schucany (1990) is that the computational burden associated with resampling through the backward representation is avoided. Consequently, the PRR bootstrap procedure can be easily applied to models with moving average components while the procedure proposed by Thombs and Schucany (1990) can only be directly applied to autoregressive models. In Pascual, Romo and Ruiz (1998) can be seen a proof of the asymptotic validity of this bootstrap resampling and a Monte Carlo comparison between both proposals.

Alternatively, the bootstrap procedure just described could be also applied to construct prediction intervals conditional on the parameter estimates (CB). In this case, it is not necessary to generate bootstrap replicates of the series as in (3) and the bootstrap forecast $k$ steps ahead depends only on the resampled residuals and is given by

$$\nabla^d y^*_{T+k} = \widehat{\phi}_0 + \sum_{j=1}^{p} \widehat{\phi}_j \nabla^d y^*_{T+k-j} + \sum_{j=1}^{q} \widehat{\theta}_j \widehat{a}^*_{T+k-j} + \widehat{a}^*_{T+k}, k = 1, 2, ..., \qquad (9)$$

where $y^*_{T+k-j}$ and $\widehat{a}^*_{T+k-j}$ are defined as in (4). The parameter estimates are kept fixed in all bootstrap replicates of future values so the CB prediction intervals do not incorporate the uncertainty due to parameter estimation. In the case of ARI($p$) processes, the conditional bootstrap was proposed by Cao *et al.* (1997).

## 3. EFFECTS OF ESTIMATION ON PREDICTION DENSITIES

In this section, several Monte Carlo experiments are carried out to study the effect of parameter estimation variability on the shape of estimated prediction densities. The focus is on prediction of values of multiplicative seasonal ARIMA$(p,d,q)x(P,D,Q)_s$ processes, where $s$ is the seasonal period. For example, for monthly data, we consider the model

$$\phi_p(L)\,\Phi_P\left(L^{12}\right)\nabla^d\nabla_{12}^D Y_t = \theta_q(L)\,\Theta_Q\left(L^{12}\right)a_t, \tag{10}$$

where $L$ is the backshift operator, $\nabla^d = (1-L)^d$, $\nabla_{12}^D = \left(1-L^{12}\right)^D$, and the autoregressive and moving average polynomials $\phi_p(L) = (1-\phi_1 L - \cdots - \phi_p L^p)$, $\Phi_P(L^{12}) = \left(1-\Phi_1 L^{12} - \cdots - \Phi_P L^{12P}\right)$, $\theta_q(L) = (1+\theta_1 L + \cdots + \theta_q L^q)$, $\Theta_Q(L^{12}) = (1+\Theta_1 L^{12} + \cdots + \Theta_Q L^{12Q})$ have all their roots out of the unit circle to ensure stationarity and invertibility. Artificial series are generated by model (10) for several choices of parameter values and error distributions. In particular, we consider Gaussian and exponential errors. First, all the models considered are estimated by conditional QML that coincides with OLS when the model lacks of a moving average component. Prediction densities are constructed by the bootstrap procedure described in the previous section, either conditioning on parameter estimates (CB) or introducing the variability due to parameter estimation (PRR).

The features of the estimated prediction intervals depend on the properties of the estimation method used. Thus, in this section we will also compare intervals for ARI$(p,d)$ processes constructed when the parameters of the model are estimated either by OLS or by LAD.

To illustrate the effect of parameter variability on estimated prediction densities, we generated 1000 time series with the ARMA$(1,1)$ process

$$y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}, \tag{11}$$

where $a_t$ is Gaussian. For each series, we compute the empirical prediction density by generating future values of $y_{T+k}$ conditional on $\{y_1, y_2, ..., y_T\}$. We also calculate the bootstrap prediction densities obtained conditioning on the parameter estimates (CB) and by using the technique described in the previous section. Finally, we construct prediction intervals based on the Box and Jenkins procedure (BJ). Notice that the difference between BJ and CB intervals could be associated with departures of the error distribution from Gaussianity. On the other hand, the differences between CB and PRR intervals are assignable to the uncertainty in the estimation of the parameters. The average coverage, the average coverage for each tail and the average length of intervals constructed with a 95% nominal coverage are reported in Table 1 for predictions one and three steps ahead and T=25,50 and 100. For the Gaussian error distribution in this table, CB intervals have lower average coverage than PRR intervals, the latter having average coverage closer to the nominal value. Note that the average length of CB intervals is also shorter than the empirical length. This effect is more evident for small sample sizes. Consequently, it seems that for relatively small sample sizes, it is important to include the uncertainty due to parameter estimation in prediction intervals in order to obtain coverages closer to the nominal values. As expected, since the conditional QML estimator is consistent, CB and PRR intervals get closer in terms of coverage and length as the sample size increases. The results are similar for predictions made one and three steps ahead.

*Insert Table* 1

Table 2 reports the Monte Carlo results for the same model but with innovations generated by an exponential distribution centered to have zero mean. It can be seen that differences between CB and PRR intervals are even larger than for Gaussian errors. Therefore, when the innovations are not normal and estimation is carried out by conditional QML, it seems important to include the variability due to parameter estimation in prediction intervals. For the shake of comparison, Table 2 also includes the results of BJ intervals. As pointed out by Pascual, Romo and Ruiz (1998), BJ intervals are clearly distorted when the error distribution is not Gaussian. As an illustration, Figure 1 represents the empirical, BJ, CB and PRR densities obtained for one step ahead predictions of one of the series generated by model (11) with exponential errors and T=100. This figure shows that the density constructed by taking into account the variability due to parameter estimation is much closer to the empirical density than when parameter estimates are considered as fixed. Furthermore, it can be seen that BJ density is clearly distorted.

*Insert Table* 2

To analyze how the presence of unit roots may affect the previous conclusions, we generated 1000 series from the ARI(2, 1) process

$$\nabla^2 y_t = 0.5\nabla^2 y_{t-1} + a_t \tag{12}$$

with exponential errors. The results for 95% prediction intervals are reported in Table 3 where it can be observed that the one-step ahead intervals have similar behavior to the previously commented, i.e., BJ intervals are not able to capture the asymmetry present in the data and the average coverage of the intervals built conditional on parameter estimates is generally under nominal coverage. Also, notice that CB intervals are not able to correctly capture the asymmetry of the error distribution. Finally, PRR intervals have average coverage close to the nominal value and they capture properly the error prediction asymmetry. Notice that when predictions are made three steps ahead, the average coverage of BJ intervals is over the nominal value, i.e. standard intervals are too wide, implying more uncertainty about the future than they should. On the other hand, Table 4 reports the results obtained when the nominal coverage is 80%. In this case, the behavior of BJ intervals is even worse than before. For example, the average coverage of BJ intervals for three steps ahead predictions constructed with 100 observations is 93.54%, i.e. 13.54 % bigger than nominal. Of course, BJ intervals are not able to capture the asymmetry in the error distribution. The behavior of CB and PRR intervals is similar to that for 95% intervals.

*Insert Table* 3

*Insert Table* 4

Next, we consider a model with both stochastic trend and stochastic seasonal components. In particular, we generate 1000 series with the following ARIMA$(0,1,1)x(0,1,1)_{12}$ model, usually known as *airline* model:

$$\nabla\nabla^{12}y_t = (1-0.33L)(1-0.82L^{12})a_t, \tag{13}$$

where $a_t$ is a Gaussian error. Table 5 reports the results obtained for sample sizes 120 and 240. For Gaussian errors and sample sizes rather large, the properties of the prediction densities constructed by the three methods considered in this paper are rather similar. The results for different innovation distributions and sample sizes are similar to the ones previously commented for models (11) and (12) and are available from the authors. The goal of this Monte Carlo experiment is to show how the PRR procedure can be extended to seasonal models with good results. As an illustration, the empirical, BJ, CB and PRR densities of one-step ahead predictions for a particular series of size 240 generated by model (13) appear in Figure 2 where it can be seen that all densities are very similar.

*Insert Table* 5

Finally, Table 6 reports the results of the Monte Carlo experiments designed to study the effect of the parameter estimation method on the shape of prediction densities. The results on Table 6 are based on 1000 series generated by model (12) with exponential innovations and with the model parameters estimated by LAD. The results for the same model estimated by OLS were reported in Table 3. Comparing tables 3 and 6, we observe that when parameters are estimated by LAD, the average coverage of CB and PRR intervals is closer to the nominal value of 95%. For example, with a sample size of 100 and intervals constructed for one-step ahead predictions, the average coverage when OLS is used is 93.45 and when the parameters are estimated by LAD, the average coverage is 94.79. The same holds for three-steps ahead predictions. Notice that the BJ intervals behavior is quite similar when estimating either by OLS or by LAD. It seems that, in this case, the estimation method does not have any effect on prediction intervals.

*Insert Table* 6

## 4. EMPIRICAL APPLICATION

In this section, we study the implementation of the PRR bootstrap procedure to the prediction of future values of two real time series, monthly observations of the Italian Industrial Production Index (IPI) and observations of the levels of a luteinizing hormone taken from Diggle (1990) that where previously analyzed by Efron and Tibshirani (1993). The Italian IPI observed monthly from January 1983 to September 1998 can be seen in Figure 3 and it presents a strong seasonal component and a stochastic trend. The first 165 observations of the series, corresponding to the period up to September 1996, are used to estimate the ARIMA$(p,d,q)x(P,D,Q)_{12}$ model which describes the dynamic behavior of the Italian IPI. The last 24 observations are used to asses the predictive performance of the Box-Jenkins and bootstrap prediction intervals.

Before estimating the model, the effects of several outliers have been removed from the original series using the program TRAMO; see Gómez and Maravall (1996). The model estimated by conditional QML from the series without outliers is given by

$$\nabla\nabla_{12}y_t = (1-\underset{(0.07)}{0.59}\,L)(1-\underset{(0.07)}{0.57}\,L^{12})a_t. \qquad (14)$$

The standard deviations in parenthesis have been calculated using the habitual approximation to the asymptotic distribution. Residuals from model (14) have skewness -0.099 and excess kurtosis of -0.16, so the Gaussianity hypothesis is not rejected at any usual level. Since the distribution of the residuals is not far from normality and the sample size is big enough, the intervals constructed using the standard and PRR approaches are very similar. In Figure 4, where 95% prediction intervals for $\widehat{y}_{T+k}$, $k = 1, ..., 24$, are plotted together with the actual observations and the linear point predictions, it can be seen that BJ and PRR intervals essentially coincide.

Next, we analyze the levels of the luteinizing hormone measured in a healthy woman every 10 minutes during 8 hours. The data set is studied by Efron and Tibshirani (1993) and has been plotted in Figure 5. The first 40 observations have been used to estimate the model to obtain

$$y_t = \underset{(0.36)}{1.19} + \underset{(0.16)}{0.48}\, y_{t-1} + a_t.$$

The histogram of the residuals together with the normal density appears in Figure 6. The empirical distribution of the residuals has a long tail to the right. The skewness coefficient is 0.83 and the excess kurtosis is 0.20, both significantly different from the values under normality. We implement our procedure to construct the prediction density of the luteinizing hormone $k$ steps ahead for $k$ = 1, ..., 8. The estimated densities for $k$ =1 and 3 appear in Figure 7, with the asymmetry observed in the residuals distribution; see Figure 6. Finally, from these densities we construct prediction intervals. Figure 8 provides the point linear prediction, the observed levels of hormone and 80% and 95% prediction intervals constructed using Box-Jenkins and bootstrap procedures. It is clear the improvement in constructing prediction intervals using the PRR procedure over standard intervals. The 80% Box-Jenkins prediction intervals contains 3 out of 8 observations while the PRR intervals are able to cope with the asymmetry in the error distribution and include 5 observations without increasing the length of the intervals. Even when looking at the 95% prediction intervals, BJ intervals leave out 2 observations while PRR intervals do not leave out any observation. We have also computed bootstrap prediction intervals conditional on parameter estimates (CB). However, although the sample size is small, CB intervals are hardly distinguishable from PRR intervals and, consequently, we have not plotted them in Figure 8. Therefore, it seems that for the values of the luteinizing hormone analyzed in this paper, the difference between BJ and PRR intervals is due to non-normality of the errors and not to parameter estimation. Efron and Tibshirani (1993) give the bootstrap distribution of the OLS autoregressive parameter estimates with observations centered at the sample mean and using all 48 observations; the bootstrap standard error for $\widehat{\phi}$ based on 200 bootstrap replicates is 0.12. The standard deviation of $\widehat{\phi}$ is rather small with respect to the standard deviation of the errors (0.43) and this could explain why the parameter variability does not affect the shape of prediction intervals. This example shows how for small sample sizes and non-normal error distributions it could be worth considering bootstrap prediction intervals in order to improve the prediction performance of ARIMA models.

Since the error distribution of the luteinizing hormone is not Gaussian, we have also estimated the parameters of the AR(1) model by LAD with the following results:

$$y_t = \underset{(0.37)}{0.73} + \underset{(0.17)}{0.68}\, y_{t-1} + a_t.$$

The standard deviations in parenthesis have been calculated using the suggestion by Bassett and Koenker (1978). The point linear predictions of the luteinizing hormone provided by the OLS and LAD estimators have been plotted in Figure 9. It can be seen that LAD predictions are systematically larger than OLS predictions and usually closer to the observed values. The MSE of the OLS predictions is 0.51 while for the LAD estimator the prediction MSE is reduced to 0.38. Figure 9 also represents the PRR 80% and 95% bootstrap intervals constructed using the OLS and LAD estimators. The LAD intervals adapt better to the asymmetry of the error distribution than the OLS intervals. Remember that the 80% PRR intervals constructed with OLS estimates leave out 3 observations when they were supposed to leave approximately one out. When PRR intervals are constructed using LAD estimates, they leave out only one observation. Consequently, it seems, as expected, that using parameter estimators more appropriate to the innovations distribution improves the performance of PRR prediction intervals.

## 5. CONCLUSIONS

This paper focuses on the effects of parameter estimation on the shape of prediction densities for seasonal ARIMA models. Box-Jenkins prediction intervals are constructed assuming Gaussianity of the innovation distribution and considering the estimated parameters as true parameters. Alternatively, prediction intervals can be obtained using bootstrap procedures which do not assume any distribution for the errors and can incorporate the variability due to parameter estimation. In particular, we consider the bootstrap technique proposed by Pascual, Romo and Ruiz (1998) for ARIMA models extending it to models with seasonal components. By means of Monte Carlo experiments, we have first studied how coverage and length of prediction intervals are affected by not taking into account the variability due to parameter estimation. We show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. As expected, since we are considering consistent estimators, the effects of parameter estimation are particularly important for small sample sizes. Furthermore, these effects are more important when the error distribution is not Gaussian. We also analyze the effect of the estimation method on the shape of prediction densities. In particular, we compare prediction densities constructed when the parameters of ARI$(p, d)$ models are estimated by OLS and by LAD. We show how, when the error distribution is not Gaussian, the average coverage and length of intervals based on LAD estimates are closer to nominal values than those based on OLS estimates. Since both estimates are consistent, this effect is less evident as the sample size grows. It is remarkable how bootstrap prediction intervals adapt to the asymmetry of the problem providing asymmetric prediction intervals improving on the necessarily symmetric BJ prediction intervals (see, e.g., Figure 8).

Finally, the performance of the PRR technique is illustrated with two empirical examples. First, we estimate prediction densities for a monthly series of the Italian IPI. Since the sample size is rather big (165 observations) and the innovation distribution is not far from normality, the prediction intervals obtained by BJ and PRR procedures are very similar. However, BJ and PRR prediction intervals constructed for the levels of a luteinizing hormone differ significantly.

Several questions remain open for further research using resampling techniques; for example, the effect of the uncertainty on the specification of the model over prediction densities. This question has been addressed for autoregressions and using a different bootstrap strategy by Masarotto (1990) and Grigoletto (1998). However, they center the prediction intervals at a linear combination of past observations and this strategy may not be adequate when the distribution of the innovations is not Gaussian.

## ACKNOWLEDGMENTS

The authors are very grateful to Eva Senra for providing the Italian IPI data and to Regina Kaiser to help us to clean this series of outliers. Financial support was provided by the European Union project ERBCHRXCT 940514 and by projects CICYT PB95-0299 and DGICYT PB96-0111 by the Spanish Government.

## REFERENCES


[1] BASSETT, G. AND KOENKER, R., "Asymptotic Theory of Least Absolute Error Regression", Journal of the American Statistical Association, 73 (1976), 618-622.

[2] BOX, G.E.P., AND JENKINS, G.M., Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, 1976.

[3] CAO, R., FEBRERO-BANDE, M., GONZÁLEZ-MANTEIGA, W., PRADA-SÁNCHEZ, J.M. AND GARCÍA-JURADO, I., "Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes", Communications in Statistics, Simulation and Computation, 26 (1997), 961-978.

[4] DIGGLE, Time Series, a Biostatistical Introduction, Oxford: Clarendon Press, 1990.

[5] EFRON AND TIBSHIRANI, An Introduction to the Bootstrap, New York: Chapman & Hall, 1993.

[6] GÓMEZ, V. AND MARAVALL, A., Programs TRAMO and SEATS: Instructions for the User. Documento de Trabajo no. 9628, Banco de España (Spain), 1996.

[7] GRIGOLETTO, M., "Bootstrap Prediction Intervals for Autoregressions: Some Alternatives", International Journal of Forecasting, 14 (1998) 447-456.

[8] MASAROTTO, G., "Bootstrap Prediction Intervals for Autoregressions", International Journal of Forecasting, 6 (1990), 229-239.

[9] PASCUAL, L. ROMO, J. AND RUIZ, E. (1998) "Bootstrap Predictive Inference for ARIMA Processes", Working Paper 98-86, Universidad Carlos III de Madrid.

[10] THOMBS, L.A. AND SCHUCANY, W.R., "Bootstrap Prediction Intervals for Autoregression", Journal of the American Statistical Association, 85 (1990), 486-92.

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **95%** | **2.5%/2.5%** | **3.92** |
| | 25 | BJ | 92.74(.05) | 3.8/3.4 | 3.90(.57) |
| | | CB | 90.68(.07) | 4.87/4.44 | 3.84(.70) |
| | | PRR | 92.70(.05) | 3.7/3.6 | 3.99(.70) |
| | 50 | BJ | 94.04(.03) | 3.1/2.8 | 3.92(.39) |
| | | CB | 92.12(.04) | 4.04/3.84 | 3.79(.52) |
| | | PRR | 93.46(.03) | 3.3/3.2 | 3.93(.51) |
| | 100 | BJ | 94.49(.02) | 2.7/2.7 | 3.92(.29) |
| | | CB | 93.66(.03) | 3.06/3.28 | 3.87(.39) |
| | | PRR | 94.04(.02) | 2.91/3.05 | 3.91(.37) |
| 3 | n | **Empirical** | **95%** | **2.5%/2.5%** | **4.35** |
| | 25 | BJ | 93.59(.05) | 3.4/3.0 | 4.48(.79) |
| | | CB | 90.79(.07) | 4.81/4.40 | 4.18(.74) |
| | | PRR | 93.14(.04) | 3.4/3.4 | 4.38(.77) |
| | 50 | BJ | 94.39(.03) | 2.9/2.8 | 4.44(.54) |
| | | CB | 92.97(.04) | 3.67/3.36 | 4.26(.55) |
| | | PRR | 93.84(.03) | 3.1/3.1 | 4.36(.55) |
| | 100 | BJ | 94.78(.02) | 2.6/2.6 | 4.42(.39) |
| | | CB | 93.95(.25) | 2.93/3.12 | 4.30(.40) |
| | | PRR | 94.27(.02) | 2.8/2.9 | 4.34(.41) |

Table 1. Monte Carlo results for model $y_t = .7y_{t-1} + a_t - .3a_{t-1}$ with Gaussian errors

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **95%** | **2.5%/2.5%** | **3.64** |
| | 25 | BJ | 92.97(.06) | .58/6.44 | 3.83(1.02) |
| | | CB | 89.52(.12) | 6.34/4.14 | 3.86(1.30) |
| | | PRR | 93.28(.08) | 2.8/3.9 | 4.05(1.39) |
| | 50 | BJ | 94.09(.03) | .01/5.81 | 3.88(.79) |
| | | CB | 90.92(.09) | 5.22/3.86 | 3.64(.96) |
| | | PRR | 94.27(.06) | 2.2/3.5 | 3.86(1.04) |
| | 100 | BJ | 94.44(.02) | 0.0/5.56 | 3.89(.55) |
| | | CB | 93.14(.06) | 3.64/3.21 | 3.65(.64) |
| | | PRR | 94.91(.05) | 1.97/3.12 | 3.74(.66) |
| 3 | n | **Empirical** | **95%** | **2.5%/2.5%** | **4.20** |
| | 25 | BJ | 93.64(.04) | .39/5.96 | 4.40(1.25) |
| | | CB | 89.40(.10) | 6.0/4.59 | 4.14(1.29) |
| | | PRR | 93.25(.06) | 2.7/4.07 | 4.39(1.39) |
| | 50 | BJ | 94.28(.03) | .02/5.53 | 4.38(.91) |
| | | CB | 91.32(.07) | 5.0/3.68 | 4.18(1.00) |
| | | PRR | 93.48(.05) | 3.08/3.4 | 4.30(1.01) |
| | 100 | BJ | 94.83(.02) | .003/5.14 | 4.38(.64) |
| | | CB | 93.06(.05) | 3.81/3.12 | 4.17(.67) |
| | | PRR | 93.94(.04) | 2.98/3.08 | 4.22(.69) |

Table 2. Monte Carlo results for model $y_t = .7y_{t-1} + a_t - .3a_{t-1}$ with Exponential errors

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **95%** | **2.5%/2.5%** | **3.65** |
| | 25 | BJ | 93.33(.04) | .14/6.5 | 3.77(1.04) |
| | | CB | 88.01(.12) | 7.6/4.4 | 3.62(1.23) |
| | | PRR | 91.02(.09) | 4.7/4.3 | 3.72(1.26) |
| | 50 | BJ | 94.03(.03) | .11/5.9 | 3.84(.72) |
| | | CB | 90.15(.09) | 6.07/3.77 | 3.60(.94) |
| | | PRR | 92.65(.07) | 3.8/3.5 | 3.70(.89) |
| | 100 | BJ | 94.44(.02) | .00/5.56 | 3.87(.53) |
| | | CB | 91.85(.07) | 5.04/3.11 | 3.65(.66) |
| | | PRR | 93.45(.06) | 3.5/3.05 | 3.72(.68) |
| 3 | n | **Empirical** | **95%** | **2.5%/2.5%** | **19.05** |
| | 25 | BJ | 96.48(.03) | .18/3.34 | 26.09(7.82) |
| | | CB | 88.07(.10) | 6.61/5.31 | 17.34(5.39) |
| | | PRR | 90.46(.08) | 4.4/5.1 | 18.01(5.72) |
| | 50 | BJ | 97.41(.02) | .01/2.6 | 26.96(5.37) |
| | | CB | 91.60(.07) | 4.79/3.61 | 18.62(3.99) |
| | | PRR | 92.75(.06) | 3.7/3.6 | 18.85(3.90) |
| | 100 | BJ | 97.75(.01) | .00/2.25 | 27.35(4.02) |
| | | CB | 92.96(.05) | 3.91/3.13 | 18.75(2.90) |
| | | PRR | 93.55(.04) | 3.3/3.13 | 18.89(2.96) |

Table 3. Monte Carlo results for model $(1 - B)^2(1 - 0.5B)y_t = a_t$ with Exponential errors

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **80%** | **10%/10%** | **2.19** |
| | 25 | BJ | 84.60(.10) | 3.7/11.7 | 2.46(.68) |
| | | CB | 73.13(.15) | 14.81/12.05 | 2.11(.60) |
| | | PRR | 76.06(.14) | 12.72/11.21 | 2.23(.61) |
| | 50 | BJ | 87.44(.07) | 1.6/10.98 | 2.51(.47) |
| | | CB | 76.68(.12) | 12.30/11.02 | 2.16(.43) |
| | | PRR | 77.96(.11) | 11.25/10.97 | 2.20(.42) |
| | 100 | BJ | 88.86(.03) | .52/10.62 | 2.53(.35) |
| | | CB | 77.96(.10) | 11.56/10.48 | 2.18(.31) |
| | | PRR | 78.62(.09) | 10.99/10.4 | 2.19(.31) |
| 3 | n | **Empirical** | **80%** | **10%/10%** | **11.82** |
| | 25 | BJ | 90.14(.08) | 1.97/7.9 | 17.03(5.10) |
| | | CB | 73.74(.13) | 14.12/12.14 | 11.28(3.07) |
| | | PRR | 75.12(.13) | 12.92/11.94 | 11.47(3.16) |
| | 50 | BJ | 92.51(.04) | .69/6.8 | 17.60(3.51) |
| | | CB | 76.80(.10) | 12.08/11.11 | 11.56(2.15) |
| | | PRR | 77.37(.090) | 11.55/11.07 | 11.63(2.15) |
| | 100 | BJ | 93.54(.02) | .17/6.3 | 17.86(2.63) |
| | | CB | 78.08(.07) | 11.34/10.58 | 11.67(1.55) |
| | | PRR | 78.48(.07) | 10.99/10.5 | 11.73(1.58) |

Table 4. Monte Carlo results for model $(1 - B)^2(1 - 0.5B)y_t = a_t$ with Exponential errors

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **95%** | **2.5%/2.5%** | **3.92** |
| | 120 | BJ | 95.95(.02) | 2.06/1.99 | 4.25(.31) |
| | | CB | 95.14(.03) | 2.44/2.42 | 4.20(.42) |
| | | PRR | 95.26(.03) | 2.39/2.34 | 4.23(.41) |
| | 240 | BJ | 95.66(.01) | 2.20/2.15 | 4.10(.20) |
| | | CB | 95.15(.02) | 2.38/2.47 | 4.07(.30) |
| | | PRR | 95.19(.02) | 2.34/2.46 | 4.08(.28) |
| 3 | n | **Empirical** | **95%** | **2.5%/2.5%** | **5.40** |
| | 120 | BJ | 95.89(.02) | 2.09/2.02 | 5.87(.58) |
| | | CB | 95.18(.03) | 2.42/2.40 | 5.77(.60) |
| | | PRR | 95.41(.03) | 2.29/2.31 | 5.87(.61) |
| | 240 | BJ | 95.61(.02) | 2.19/2.20 | 5.65(.38) |
| | | CB | 95.20(.02) | 2.35/2.45 | 5.60(.42) |
| | | PRR | 95.27(.02) | 2.29/2.43 | 5.65(.43) |
| 12 | n | **Empirical** | **95%** | **2.5%/2.5%** | **9.54** |
| | 120 | BJ | 95.53(.03) | 2.26/2.20 | 10.38(1.45) |
| | | CB | 94.06(.04) | 2.96/2.97 | 10.22(1.44) |
| | | PRR | 94.55(.04) | 2.70/2.71 | 10.43(1.45) |
| | 240 | BJ | 95.42(.02) | 2.26/2.32 | 9.99(.95) |
| | | CB | 94.59(.03) | 2.63/2.77 | 9.90(1.01) |
| | | PRR | 94.82(.03) | 2.49/2.68 | 10.01(.99) |
| 24 | n | **Empirical** | **95%** | **2.5%/2.5%** | **14.42** |
| | 120 | BJ | 96.35(.03) | 1.85/1.80 | 16.72(2.61) |
| | | CB | 94.59(.05) | 2.71/2.70 | 16.50(2.64) |
| | | PRR | 95.77(.04) | 2.11/2.12 | 17.50(2.64) |
| | 240 | BJ | 95.97(.02) | 1.99/2.04 | 15.64(1.67) |
| | | CB | 94.92(.03) | 2.47/2.61 | 15.52(1.73) |
| | | PRR | 95.66(.02) | 2.09/2.25 | 16.12(1.73) |

Table 5. Monte Carlo results for model $(1-B)(1-B^{12})y_t = (1-.33B)(1-.82B^{12})a_t$ with Gaussian errors

| Lead time | Sample Size | Method | Average Coverage(se) | Coverage below/above | Average Length |
|---|---|---|---|---|---|
| 1 | n | **Empirical** | **95%** | **2.5%/2.5%** | **3.65** |
| | 25 | BJ | 93.11(.05) | .38/6.51 | 3.83(1.04) |
| | | CB | 89.91(.11) | 5.72/4.37 | 3.75(1.23) |
| | | PRR | 93.17(.08) | 2.62/4.21 | 3.89(1.26) |
| | 50 | BJ | 93.90(.04) | .24/5.87 | 3.88(.72) |
| | | CB | 91.13(.10) | 5.10/3.77 | 3.68(.94) |
| | | PRR | 94.01(.07) | 2.51/3.48 | 3.81(.88) |
| | 100 | BJ | 94.41(.02) | .03/5.56 | 3.89(.53) |
| | | CB | 92.74(.07) | 4.17/3.10 | 3.70(.64) |
| | | PRR | 94.79(.05) | 2.21/3.00 | 3.80(.68) |
| 3 | n | **Empirical** | **95%** | **2.5%/2.5%** | **19.05** |
| | 25 | BJ | 96.55(.04) | .30/3.15 | 27.11(7.91) |
| | | CB | 89.25(.10) | 5.66/5.09 | 18.25(5.61) |
| | | PRR | 92.42(.08) | 2.98/4.60 | 19.77(6.23) |
| | 50 | BJ | 97.51(.02) | .03/2.45 | 27.64(5.37) |
| | | CB | 92.01(.08) | 4.50/3.49 | 19.21(3.99) |
| | | PRR | 93.96(.06) | 2.72/3.32 | 19.99(4.10) |
| | 100 | BJ | 97.83(.01) | .00/2.16 | 27.76(3.85) |
| | | CB | 93.47(.05) | 3.48/3.04 | 19.09(2.80) |
| | | PRR | 94.70(.04) | 2.36/2.95 | 19.64(2.99) |

Table 6. Monte Carlo results for model $(1 - B)^2(1 - 0.5B)y_t = a_t$ with Exponential errors and LAD estimation

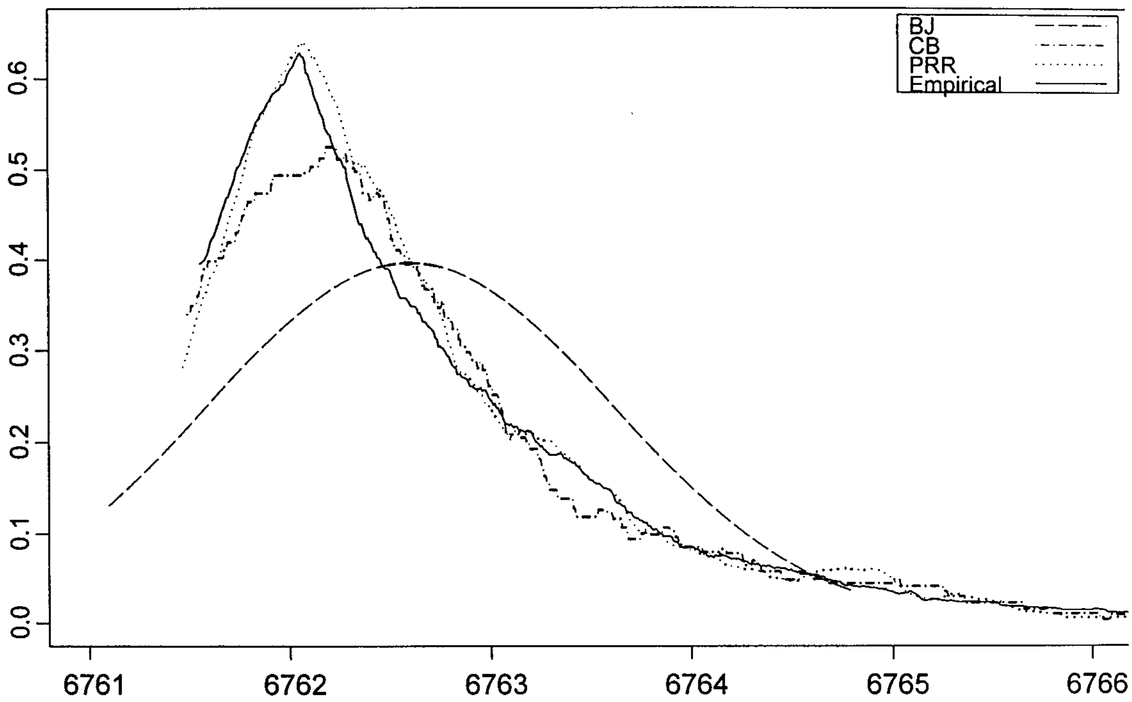


Figure 1: Densities of one-step ahead predictions of one series of size 100 generated by model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ with exponential innovations.

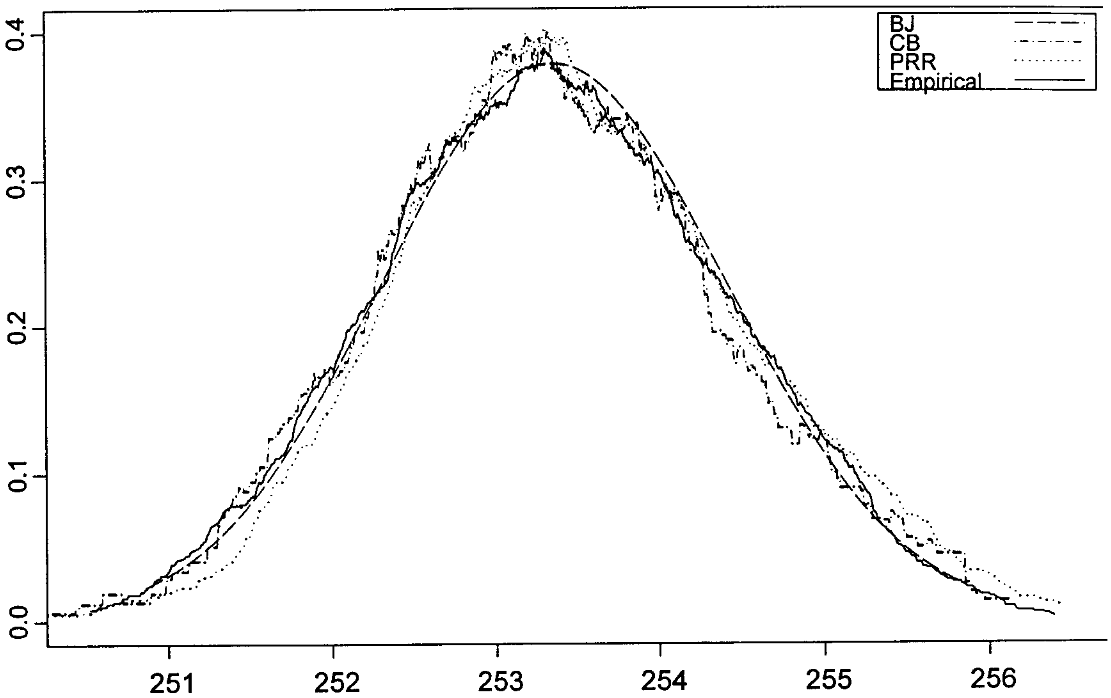


Figure 2: One-step ahead prediction densities of one series of size 240 generated by model (13) with Gaussian innovations.

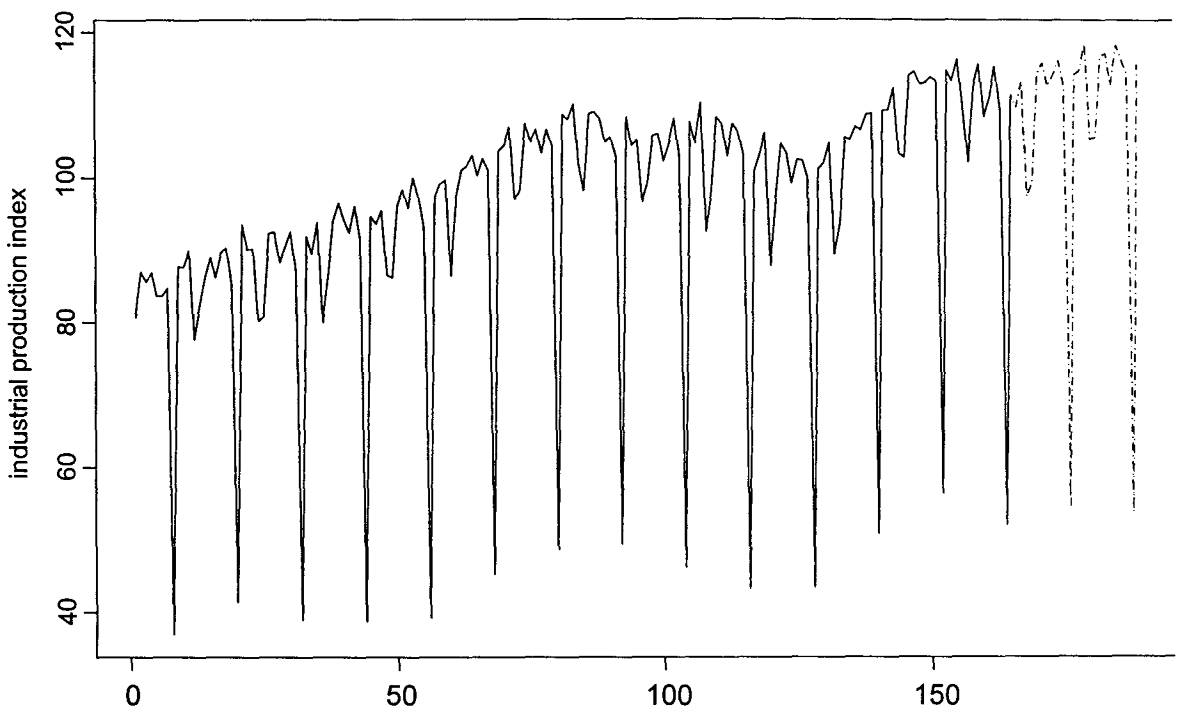


Figure 3: Italian Industrial Production Index observed monthly from January 1983 to September 1998. Continuous line corresponds to estimation period and dashed line to prediction period.

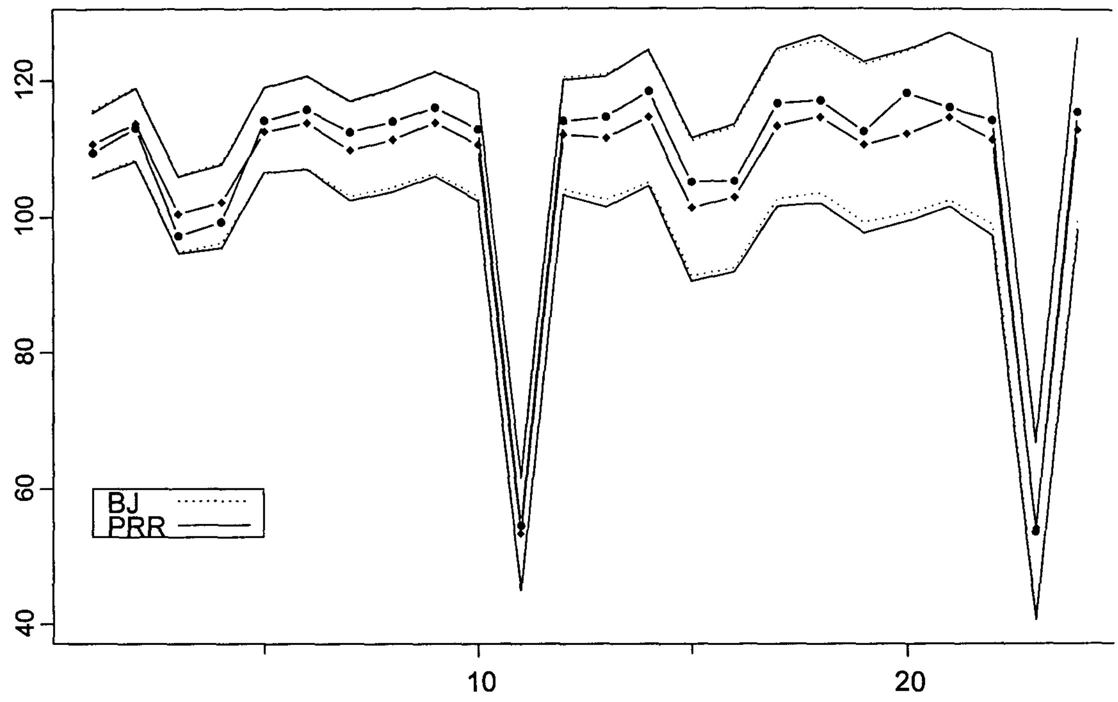


Figure 4: Real observations of IPI (•) together with point linear predictions (♦). 95% prediction intervals constructed by Box-Jenkins and bootstrap procedures.

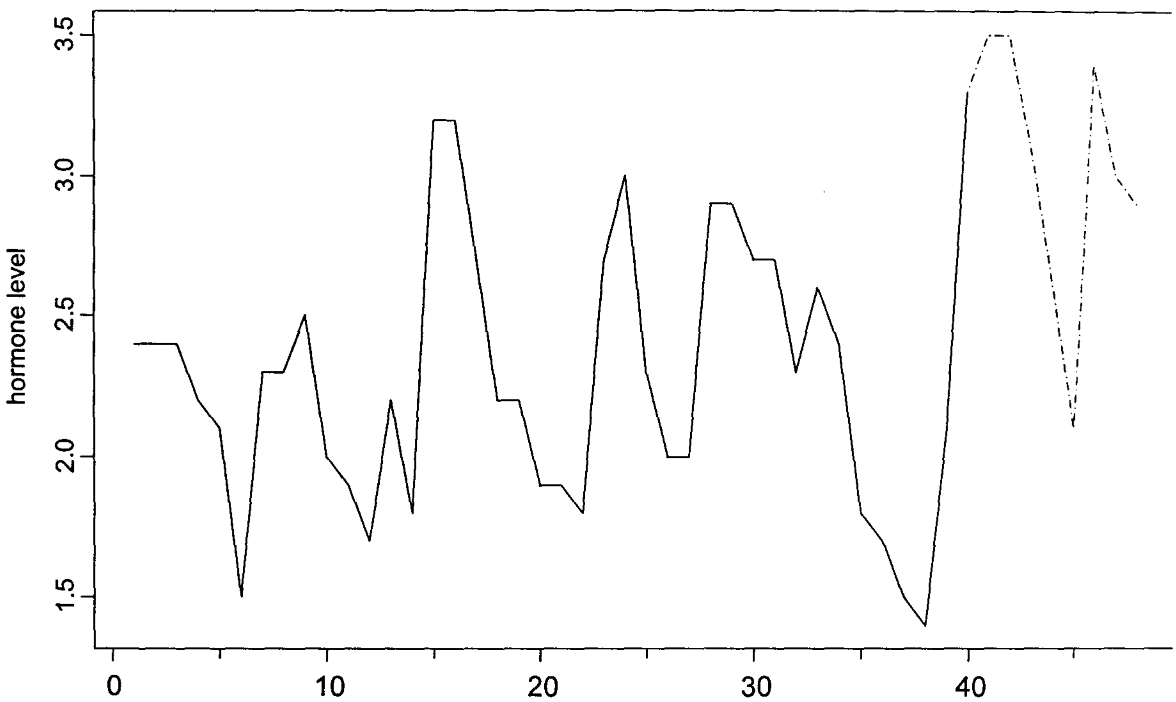


Figure 5: Observations of the luteinizing hormone measured in a healthy woman every minute during 8 hours. Continuous line corresponds to estimation period and dashed line to prediction period.

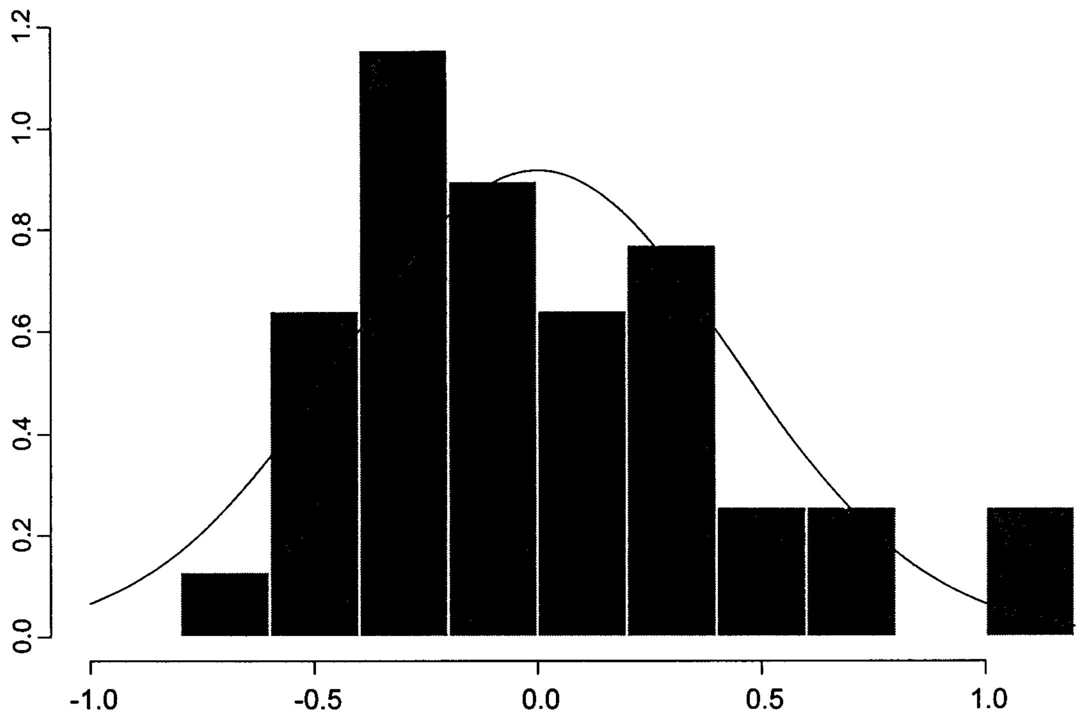


Figure 6: Histogram of residuals from AR(1) model for the luteinizing hormone and normal density.

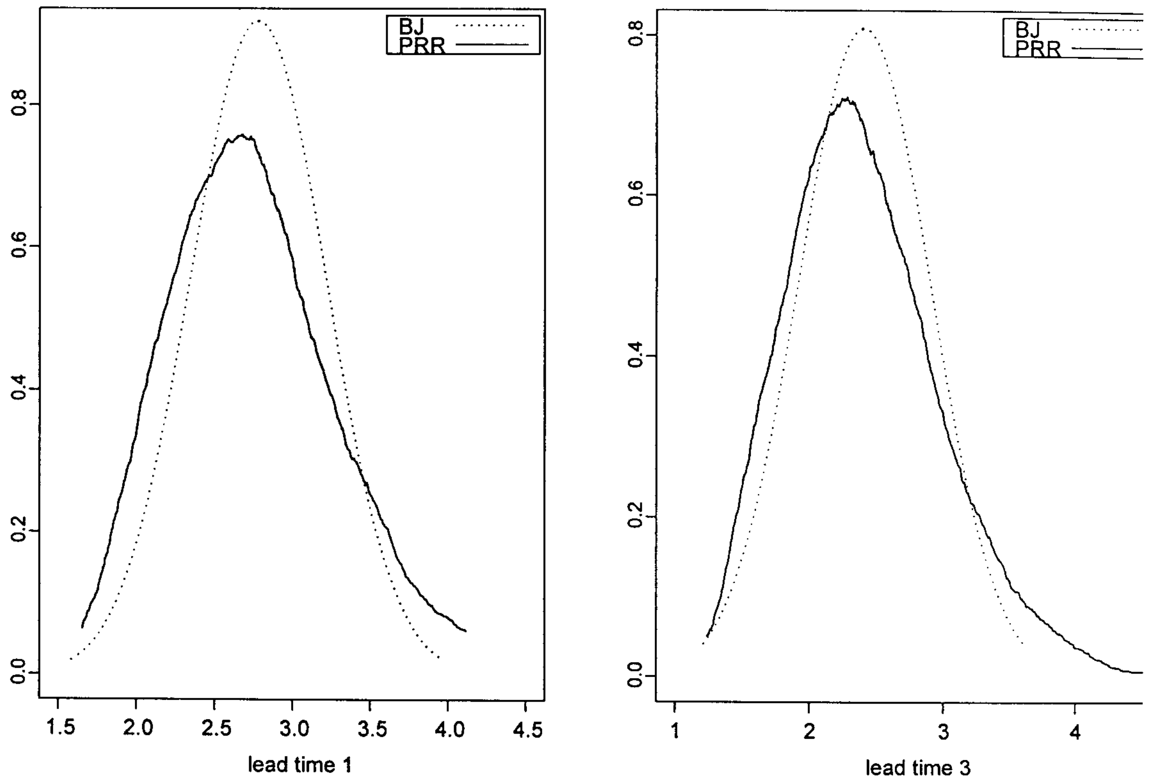


Figure 7: Densities of one and three steps ahead predictions of the luteinizing hormone constructed by BJ and PRR procedures.

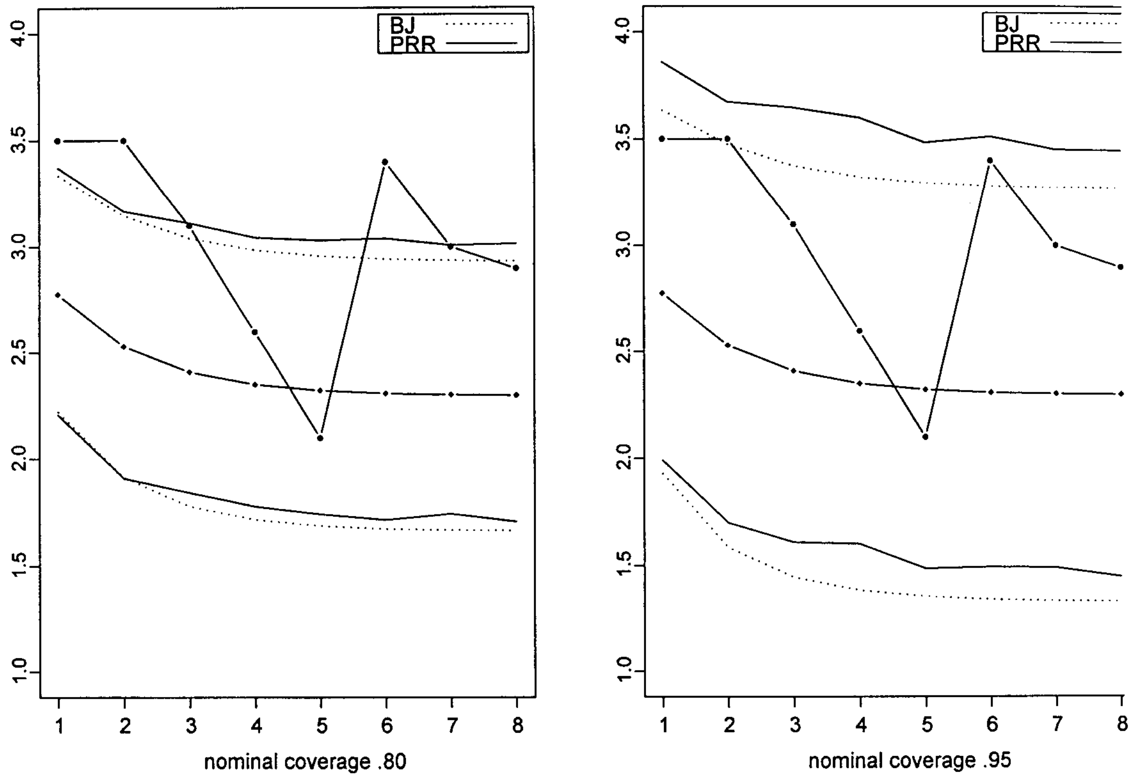


Figure 8: Observations of luteinizing hormone (•) and point linear predictions (♦). 80% and 95% intervals constructed by BJ and PRR procedures.

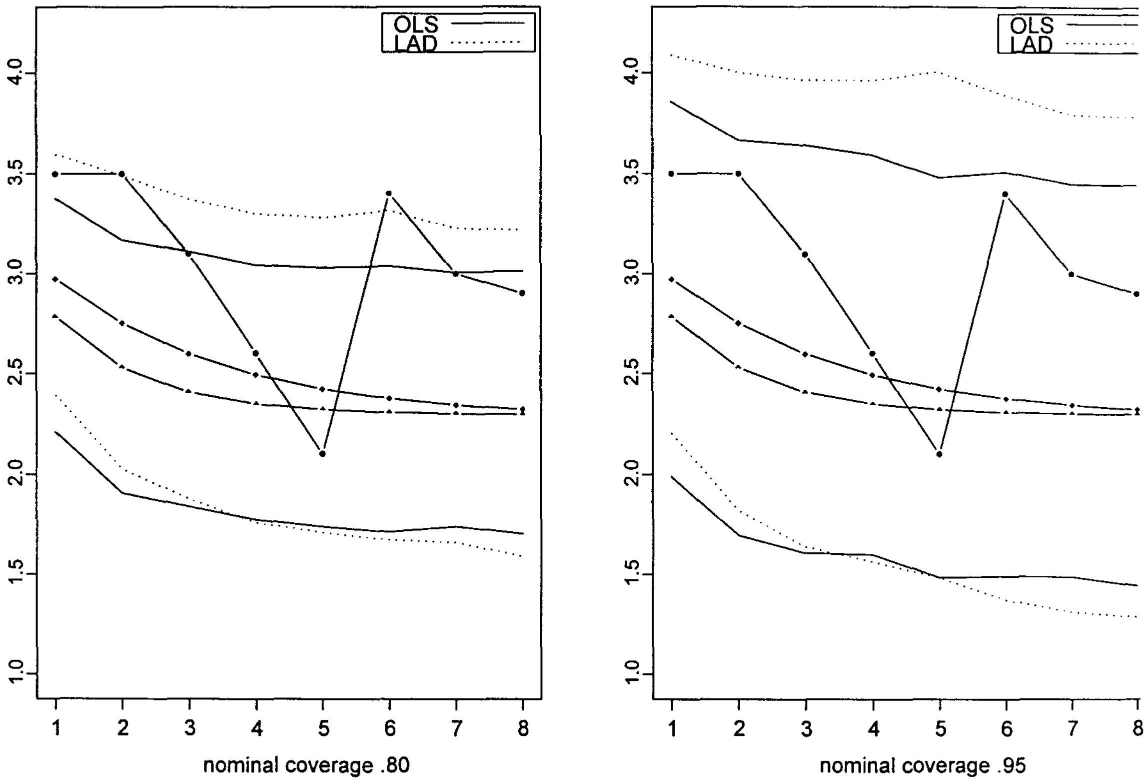


Figure 9: Observations of luteinizing hormone (•) and point linear predictions obtained using OLS (▲) and LAD (♦). 80% and 95% PRR intervals constructed using OLS and LAD.