THE KOREAN ECONOMIC REVIEW
Volume 23, Number 1, Summer 2007

UNCERTAINTY SURROUNDING THE U.S. NAIRU ESTIMATES OF ESTRELLA AND MISHKIN (1999)*

CHARLES HARVIE** · HYEON-SEUNG HUH***

Estrella and Mishkin(1999) propose a promising method to estimate the NAIRU for the U.S. However, their uncertainty measure of the NAIRU estimates is problematic, as the underlying assumption of normality is violated. We apply the block bootstrap techniques to offer a better measure of precision. One implication is that their method underestimates the true uncertainty surrounding the NAIRU estimates.

JEL Classification: C51; E24
Keywords: NAIRU; Uncertainty; Fat tails; Block bootstrap

I. INTRODUCTION

Recently, Estrella and Mishkin(henceforth EM, 1999) proposed a new method to estimate the U.S. NAIRU(Non Accelerating Inflation Rate of Unemployment). Referring to this as the 'short-run NAIRU', they highlighted its usefulness as an empirical basis for predicting future changes in inflation and as a feedback variable in Taylor-type monetary policy rules. Equally important, they argue that the short-run NAIRU can be estimated with more than twice the precision than that of long-run measures, like the natural rate of unemployment, used in many previous

---

Received for publication: Sep. 15, 2006. Revision accepted: Oct. 23, 2006.

* The authors wish to thank two anonymous referees for their constructive comments and suggestions. This research was supported by Yonsei Business and Economics Research Fund.

** Department of Economics, University of Wollongong, Wollongong, NSW, 2522, Australia, Email: charles_harvie@uow.edu.au

*** Corresponding author, School of Economics, Yonsei University, 134 Shinchon-Dong, Seodamoon-Gu, 120-749, Seoul, Korea, Email: hshuh@yonsei.ac.kr

studies. This result should appeal to many, given increasing concern regarding the uncertainty surrounding the NAIRU estimates(Staiger *et al*., 1997; Laubach, 2001). For the construction of their measure of precision, EM first obtained the Newey-West(1987) heteroskedastic and autocorrelation consistent(HAC) standard errors to take account of the resulting serial correlation, and then applied the delta method to the short-run NAIRU measure that includes the ratios of the parameter estimates.

We find, however, several reasons to contradict their uncertainty measure of the NAIRU estimates. First, such ratios of random variables are well known to have fat-tailed distributions in finite samples, while the delta method approximates them by a normal distribution. Staiger *et al*.(1997) find in a Monte Carlo study that this is a main cause for the delta method to underestimate the true uncertainty around the NAIRU estimates. Second, the short-run NAIRU is derived from a multi-horizon prediction regression. Fat tails are not unusual in this type of model as large residuals are frequent.[1] Third, the errors in the Phillips curve are plausibly fat tailed due to truncation errors in the estimation of inflation(Mizon *et al*., 1990; Ball and Mankiw, 1995). Fourth, the presence of fat tails also questions the use of Newey-West HAC standard errors, as their finite-sample properties are dependent on asymptotic normality. Further, there is evidence that such a correction of serial correlation is often insufficient for the multi-horizon prediction regression, and the estimated coefficients tend to have a tail that is too short(Hodrick, 1992; Horowitz, 2001).

This paper aims to offer a better measure of precision for NAIRU estimates. We do so by applying through two block bootstrap methods, the moving blocks bootstrap of Künsch(1989) and the stationary bootstrap of Politis and Romano(1994). The bootstrap technique approximates the exact sampling distribution of the NAIRU estimates without recourse to a specific distribution such as asymptotic normality. The Monte Carlo evidence by Li and Maddala(1999) confirms that this gives better estimates than the delta method in obtaining confidence intervals for ratios of parameter estimates. Unlike the standard bootstrap in the iid case,

[1] Typical examples are tests of market efficiency, such as in exchange rate predictions, the analysis of dividend yields and expected stock returns, and the term structure of interest rates.

however, the resampling is performed on blocks of residuals rather than individual ones. The block sampling technique captures the feature of fat tails more consistently than the standard bootstrap does as Künsch(1989), Lahiri(1995) and Horowitz(2001) highlight. It also deals with the temporal dependency structure, in which the standard bootstrap typically fails. In fact, Lahiri(1992) and Härdle *et al*.(2002) prove that block sampling provides better finite-sample accuracy compared to the asymptotic normal approximation in the presence of dependency.

The rest of this paper is organized as follows. Section 2 details the EM procedure and presents its empirical application to the U.S. Section 3 discusses the uncertainty measure of this NAIRU estimates and constructs new confidence intervals using the moving blocks bootstrap and stationary bootstrap techniques. Section 4 offers a Monte Carlo experiment to illustrate the finite-sample performance of these moving blocks bootstrap and delta methods for the application at hand. Section 5 concludes the paper.

II. THE NAIRU MEASURE OF EM AND ITS UNCERTAINTY

EM define the short-run NAIRU as the unemployment rate which would correspond to a *forecast* of no inflation change over the policy horizon. Suppose that the policy horizon for inflation is from c to $c+k$ months ahead. They construct the following equation, which forecasts the difference between current annual inflation and inflation over the policy horizon:

$$\pi_t^{(c,k)} - \overline{\pi}_t = \alpha + \sum_{i=0}^{p} \beta_i u_{t-i} + \sum_{j=0}^{q} \gamma_j \Delta\pi_{t-j} + \varepsilon_t \,, \tag{1}$$

where $\pi_t^{(c,k)} = (1200/\mathrm{k})ln(p_{t+c+k}/p_{t+\mathrm{c}})$ is the k-period inflation in the price level p_t, reported at an annual rate; $\overline{\pi}_t = 100\,ln(p_t/p_{t-12})$ is the annual rate of inflation; $\Delta\pi_t = \pi_t - \pi_{t-1}$; $\pi_t = 1200\,ln(p_t/p_{t-1})$ is monthly inflation at an annual rate; u_t is the unemployment rate; and ε_t is a disturbance term. Eq. (1) can always be rewritten as:

$$\pi_t^{(c,k)} - \overline{\pi}_t = \beta_0(u_t - u_t^N) + \varepsilon_t , \tag{2}$$

with the short-run NAIRU

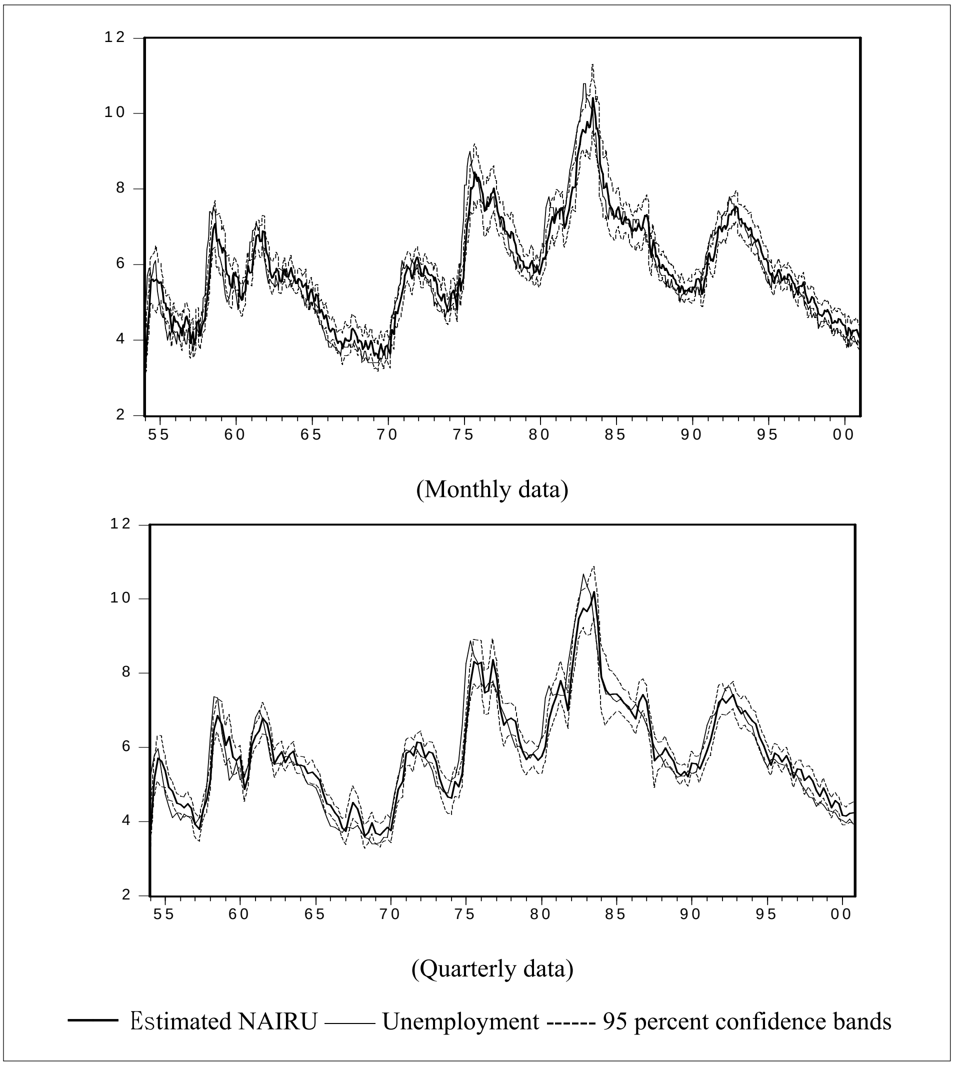
$$u_t^N = -[\alpha + \sum_{i=1}^{p} \beta_i u_{t-1} + \sum_{j=0}^{q} \gamma_i \Delta\pi_{t-j}] / \beta_0 . \tag{3}$$

The resulting NAIRU gap, ($u_t - u_t^N$), subsumes all the predictive power of the equation and hence, it can be a good predictor of inflation over the policy horizon. When $u_t = u_t^N$, inflation is forecast to neither accelerate nor decelerate over the policy horizon.

EM apply the OLS technique to Eq. (1) with a 12-month-ahead, 12-month horizon ($c = k = 12$) and a lag length of $p = q = 12$ for the estimation of the U.S. NAIRU. The standard errors are estimated using the Newey-West technique with a 24-lag window, as the errors in regressions will exhibit serial correlation. To offer a measure of precision they employ the delta method, which involves making a first-order Taylor series approximation to the NAIRU measure in Eq. (3) and then using the formula to calculate the asymptotic variance of this linearized function. Suppose that $\theta = (\alpha, \beta_0, \cdots, \beta_p, \gamma_0, \cdots, \gamma_q)'$ is a vector of regression coefficients, $\hat{\theta}$ is its estimate, and $g(\hat{\theta})$ is the function of interest corresponding to Eq. (3). The delta method approximates the distribution of $g(\hat{\theta})$ by a normal distribution with mean $g(\hat{\theta})$ and variance $(\partial g / \partial \theta)' \hat{V} (\partial g / \partial \theta)$, where $\hat{V}$ is the estimated Newey-West HAC variance-covariance matrix of $\hat{\theta}$ and $(\partial g / \partial \theta)$ is the first derivative of g, evaluated at $\hat{\theta}$.

EM initially estimated the U.S. NAIRU for the period 1954:M1 to 1997:M11, but the dataset is updated to include 2000:M12. We further construct the corresponding NAIRU measures from quarterly data (1954:Q1-2000:Q4) for two reasons. First, to guard against the concern that the imprecision in the NAIRU estimates may be unduly large as a consequence of using noisy monthly data. The estimates are expected to be more precise when temporally aggregated data are used. Second, for the sake of compatibility with previous studies in the field, as they were typically undertaken using quarterly data.[2] Figure 1 shows both estimates of short-run NAIRU and the actual unemployment rate. Also depicted are 95 percent(two standard error) confidence intervals generated using the delta method. Note that the standard error of $\hat{u}_t^N$ is a time-varying function of the values of the variables in Eq. (3).

[Figure 1] Estimates of the NAIRU and 95% delta method confidence intervals



[2] For the quarterly model, $\pi_t^{(c,k)} = (400/\mathrm{k})ln(p_{t+c+k}/p_{t+c})$ with $c = k = 4$, $\overline{\pi}_t = 100\,ln(p_t/p_{t-4})$, and a lag length of $p = q = 4$. The series on prices and the unemployment rate are averaged to quarterly values. An 8-lag window is used for the Newey-West corrections of standard errors.

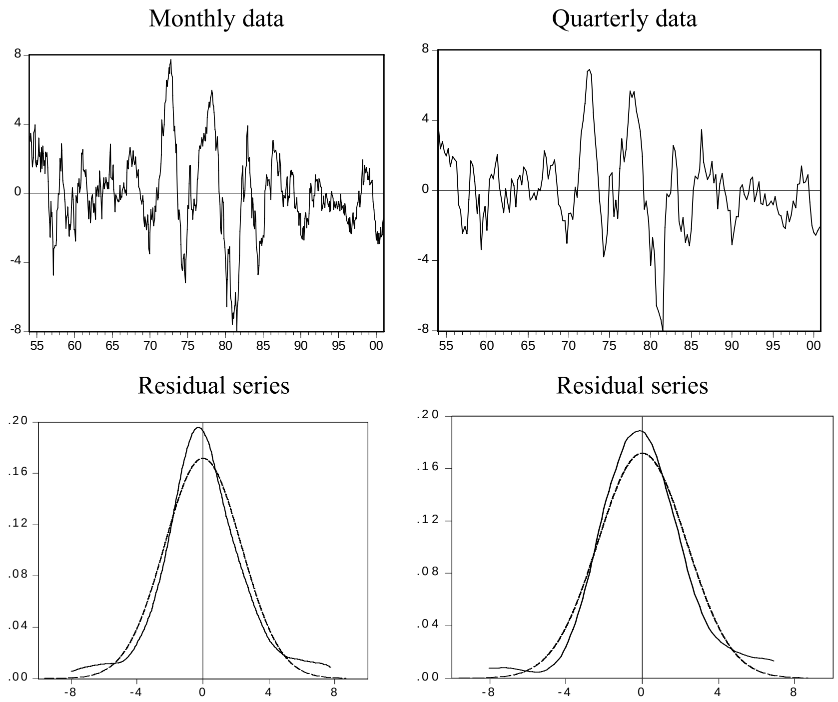
Both monthly and quarterly figures demonstrate the high variability of the short-run NAIRU. This is in contrast with long-run measures designed to estimate a natural rate. For example, Staiger *et al.*(1997) estimate a constant NAIRU of 6.2 percent using monthly and quarterly data samples for 1955 to 1994. Laubach(2001) applies a time-varying NAIRU model in which all the estimates remain in the 5 to 7 percent range using quarterly data over period of 1970:Q1 to 1998:Q4. Nevertheless, short-run NAIRU measures are estimated more precisely. The standard errors range from 0.11(0.10) to 0.45(0.46) with a mean of 0.2(0.18) over the sample period of monthly (quarterly) data. Whereas Staiger *et al*. report standard errors of 0.52 and 0.46 using the delta method for monthly and quarterly data, respectively. Staiger *et al*. also considered the Fieller(1954) method which resulted in a much larger standard error. In Laubach, the average standard errors are in the range of 0.54 and 1.7, depending on the specifications for the NAIRU.

III. NEW CONFIDENCE INTERVALS FOR THE NAIRU ESTIMATES

The upper panel in Figure 2 shows the raw residual series for monthly and quarterly data. Both series exhibit large values frequent, pointing to the presence of fat tails as discussed in the introduction. The lower panel shows the empirical density functions of the residuals estimated nonparametrically using an Epanechnikov kernel. The bandwidth of the kernel is set at one quarter of the interquartile range(that is, between the 75 percentile and the 25 percentile) of the residual series. Fat tails are indeed visible in the series by comparison to the normal distribution that has the same size of mean and variance.[3] We report the results of several tests in Table 1 to see the statistical significance of this departure from normality.

[3] In addition, there appears to be a smaller extent of finite-sample bias in the monthly model. But, the t-test statistic is close to zero, comfortably accepting the null hypothesis that the mean of the residual series is zero.

[Figure 2] The estimated residual series and their empirical density functions



Empirical density function of the residuals (thick lines) and normal distribution (dotted lines)

[Table 1] The results of normality tests for the residual series

	Skewness		Kurtosis		JB		KS
Monthly data	0.06	(0.54)	1.22	(0.00)	35.29	(0.00)	0.056*
Quarterly data	0.05	(0.77)	1.80	(0.00)	25.43	(0.00)	0.084*

Notes: Entries are the test statistics and their marginal significance levels(*p*-values) are reported in parentheses, if applicable. The first and second columns test for skewness and kurtosis, respectively, both of which are distributed as $\chi^2(1)$. The third column (JB) reports the results of the Jarque-Bera test for normality, which is distributed as $\chi^2(2)$. Reported in the final column(KS) are the Kolmogorov-Smirnov test statistics for normality. The critical values are tabulated in Lilliefors(1967). They are 0.037 and 0.043 at the 5% and 1% significance levels, respectively, for the monthly model (number of effective observations =564) while the corresponding critical values are 0.065 and 0.075 for the quarterly model(number of effective observations=188). An asterisk (*) indicates significance at the 1% level.

While there is no evidence of skewness, the test of kurtosis strongly rejects the null hypothesis, suggesting that the residual series have fatter tails than a normal distribution. The strong presence of fail tails is confirmed by the Jarque-Bera statistic, which tests for normality based on the skewness and kurtosis measures combined. Its marginal significance level is virtually zero, rejecting the null hypothesis of normality. The table also reports the results of the nonparametric Kolmogorov-Smirnov test, which is often more powerful than chi-squared tests like the Jarque-Bera for any sample size. This test is performed by comparing the cumulated frequencies of the observed distribution to those expected under the null hypothesis of normality. The mean and variance of the residual series are used for construction of the expected normal distribution. The test statistics rejects the null hypothesis comfortably at all significance levels, concluding that departures from normality are statistically significant in both models.

We now turn to the moving blocks bootstrap(MB) and stationary bootstrap(SB) techniques. The MB divides the data of n observations into overlapping blocks of b consecutive observations, and resamples $k = n / b$ blocks randomly with replacement from the set of $n - b + 1$ blocks.[4] By resampling the blocks of length b, the correlation present in observations less than b, units apart is retained. All observations of the k-sampled blocks are then pasted together in succession to form a bootstrap sample of the same length as the original data. The basic steps for the SB are the same as those of the MB with one exception. The SB resamples overlapping blocks of random length, where the length of each block has a geometric distribution with the probability parameter p. The average length of a block is $1/p$ and this corresponds to the fixed block length of b in the MB.[5]

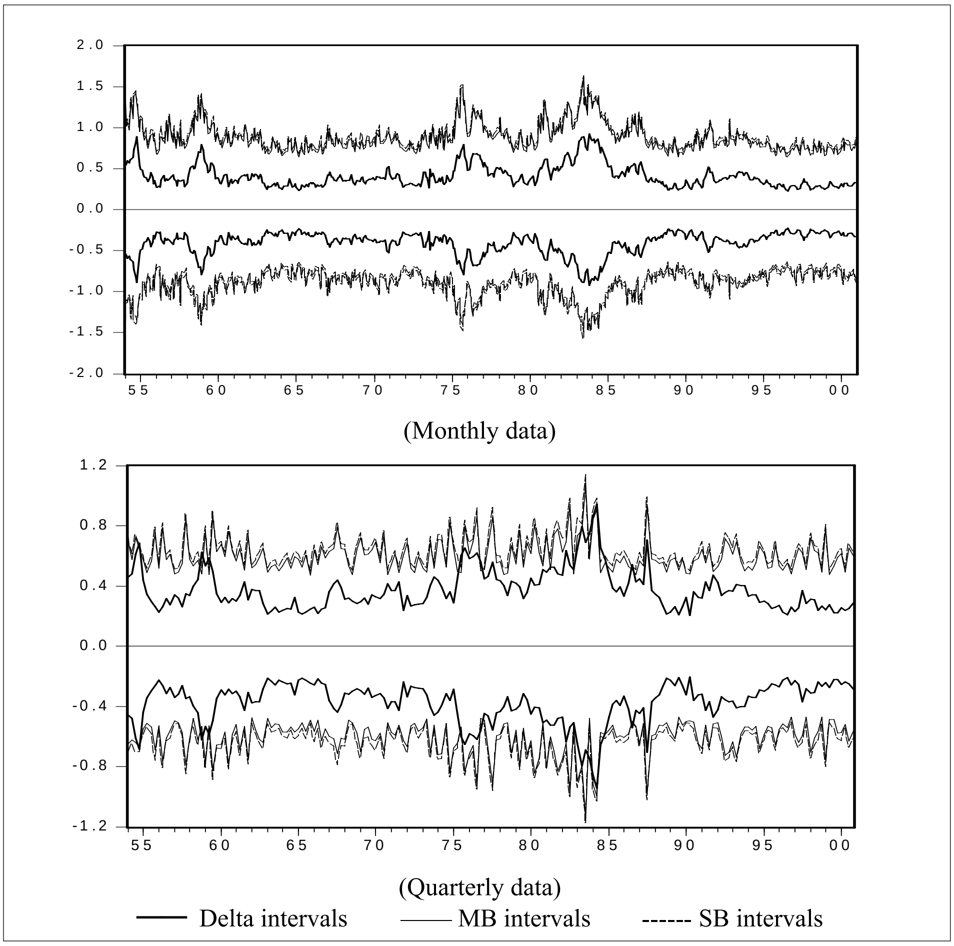
The choice of the block length can be quite important, and has received some attention in recent literature(see the reviews by Berkowitz and Kilian(2000) and Horowitz(2001)). However, the rules are mostly suggestive in finite samples, and, in this light, Härdle *et al*.(2002) comment that satisfactory data-based methods for selecting block lengths

[4] When the sample size n is not simply a multiple of the block size b, one can choose k as the smallest integer for which $bk \geq n$, generate a bootstrap series as above, and discard the last $bk-n$ bootstrap observations.

[5] A consequence is that the bootstrap series generated by the SB is stationary, whereas they are not with the MB even if the original series is stationary. Lahiri(1999) shows, on the other hand, that the errors made by the SB are larger than those of the MB having non-stochastic lengths.

are not yet available. Rather than appeal to an asymptotically correct block length we follow Li and Maddala(1996) who infer that Künsch's 1989 suggestion to use "subjective judgement based on sample correlations" is an acceptable way to proceed. That is, the block lengths b and $1/p$ are set at 24 for monthly data and at 8 for quarterly data to take account of serial correlation in the errors. This block length for monthly data, of course, matches the 24-lag window used by EM for the Newey-West HAC standard errors.

[Figure 3] The 95% confidence intervals from the delta, MB and SB methods



We repeated the MB and SB procedures outlined above 500 times. The 95 percent confidence interval for $\hat{u}_t^N$ are then constructed, which is the interval between the 2.5 and 97.5 percentiles of the bootstrap distribution of the estimated short-run NAIRU. Figure 3 shows the results. For a better visual comparison, depicted are the upper and lower bounds minus the estimated short-run NAIRU in Figure 1. The 95 percent confidence intervals from the delta method are also reproduced. The key finding is that the two block bootstrap methods have wider confidence intervals than the delta method. Using quarterly data, for example, the average spreads of the MB and SB intervals are 1.22 and 1.29 percentage points, respectively, while the average spread of the delta method interval is 0.76 percentage points. The differences get somewhat bigger with monthly data, which may not be surprising in light of a stronger presence of fat tails. On average, the MB and SB intervals are over two times wider at 1.75 and 1.86 percentage points, respectively, compared to the spread of the delta method interval at 0.8 percentage points. An implication is that asymptotic normal approximation of the delta method in the presence of fat tails may lead to an unwarranted reduction in the uncertainty around the NAIRU estimates, which will be investigated further in the following section. Apparently, the two block bootstrap methods generate very similar confidence intervals.

IV. A MONTE CARLO ILLUSTRATION

This section presents the results of Monte Carlo experiments that illustrate the finite-sample performance of the MB, SB and delta methods. We follow the Monte Carlo procedure of Staiger *et al.*(1997), which evaluate the quality of asymptotic-based methods in measuring the uncertainty surrounding the estimates of NAIRU. The design is empirically based and is intended to capture key features of the empirical models considered here. There arises, however, an extra complication due to the fact that the short-run NAIRU in Eq. (3) is a time-varying function. To get around this, all the explanatory variables in Eq. (1) are dropped with the exception of the term u_t, rendering the NAIRU measure a constant, $u_t^N = -\alpha / \beta_0$. The biannual 1954:Y1-2000:Y2 sample is then used to mitigate the effects of serial correlation that may occur as a consequence of removing the lagged terms. This simplified setup is not innocuous, but can have a capacity to address the issues such as departures from normality. The estimation results for the biannual 1954:Y1-2000:Y2 sample are:

$$u_t = 0.678 + 0.888u_{t-1} + v_{1t} \tag{4a}$$

$$\pi_t^{(2,2)} - \overline{\pi}_t = \alpha + \beta u_t + v_{2t} \tag{4b}$$

where $(\hat{\alpha}, \hat{\beta})$=(2.542, -0.419) and $\hat{u}_t^N = 6.07$.[6]

Two methods were used to generate the pseudorandom errors. In the first, the bivariate errors from the 1954-2000 regression were randomly sampled with replacement, and used to generate the artificial draws. In the second, $\{v_t\}$ was drawn from an iid bivariate normal with the covariance matrix set to the sample covariance matrix of the residuals. The values of (α, β) for which the performance of the procedures is investigated, are (2.542, -0.419), (2.169, -0.606), (4.074, -0.707), (1.009, -0.131), and (4.281, -0.495). The first set contains the point estimates for the biannual 1954-2000 sample, which corresponds to an estimate of the NAIRU of 6.07. The next four are values which lie on the boundary of the

[Table 2] Monte Carlo coverage rates of the delta, MB and SB confidence intervals

	Delta intervals		MB intervals		SB intervals	
NAIRU estimates	90%	95%	90%	95%	90%	95%
	Errors drawn from the empirical distribution					
3.58	0.96	0.99	0.89	0.96	0.91	0.96
5.76	0.97	0.99	0.90	0.95	0.89	0.94
6.07	0.98	0.99	0.90	0.95	0.89	0.94
7.70	0.97	0.99	0.90	0.94	0.89	0.95
8.65	0.98	0.99	0.89	0.94	0.88	0.94
	Errors drawn from a normal distribution					
3.58	0.93	0.97	0.89	0.95	0.90	0.96
5.76	0.94	0.97	0.90	0.95	0.89	0.95
6.07	0.96	0.99	0.90	0.95	0.90	0.95
7.70	0.95	0.96	0.90	0.95	0.90	0.94
8.65	0.94	0.98	0.90	0.95	0.89	0.95

6 $\pi_t^{(2,2)} = 100\,ln(p_{t+4} / p_{t+2})$ and $\overline{\pi}_t = 100\,ln(p_t / p_{t-2})$.

usual 80 percent confidence ellipse for (α, β) estimated from that regression.[7] The corresponding NAIRU measures are 3.58, 5.76, 7.70, and 8.65, respectively. The total number of Monte Carlo samples is 5,000. The delta method employs the Newey-West HAC standard errors with a 4-lag window. For the MB and SB methods the block lengths are set at 4, and the number of bootstrap samples is 500 as before.

The results of the experiments are shown in Table 2, which gives the empirical coverage probability of the nominal 90 and 95 percent confidence intervals. The MB and SB intervals are shown to have better finite-sample coverage rates than the delta method interval. In fact, the Monte Carlo coverage rates of the MS and SB intervals are generally close to their normal confidence levels. By contrast, the delta method intervals overcover consistently. The overcoverage is more evident when the errors are from the empirical distribution. The coverage rates are all 99 percent for the 95 percent confidence intervals while those of the 90 percent intervals range from 96 to 98 percent, depending on α and β. Evidently, the presence of fat tails in the errors creates further complications to the delta method that relies on asymptotic normality. The Monte Carlo simulations confirm that the delta method is biased towards producing tighter confidence intervals. This consolidates our earlier finding that the method used by EM results in intervals that underestimate the true extent of the imprecision attached to the NAIRU estimates.

V. CONCLUDING REMARKS

The new NAIRU measure by Estrella and Mishkin(1999) can be a valuable tool for predicting future changes in inflation, and thus for guiding policy decisions. However, their uncertainty measure of the NAIRU estimates is misleading, as the resulting presence of fat tails violates the assumptions of normality underlying the delta method. Empirical results show that their procedure underestimates the true uncertainty around the NAIRU estimates. Here, we offer a better measure of precision by applying the moving blocks bootstrap and the stationary

[7] See Judge *et al*. (1988, Chapter 6) for the construction of joint confidence intervals.

bootstrap. These techniques are distribution free and capture the feature of fat tails consistently. Monte Carlo simulations confirm that they produce more accurate estimates than the delta method in obtaining confidence intervals for the application at hand.

References

Ball, L. and G. Mankiw (1995), "Relative Price Changes as Aggregate Supply Shocks," *Quarterly Journal of Economics*, Vol. 110, 161-193.

Berkowitz, J. and L. Kilian (2000), "Recent Developments in Bootstrapping Time Series," *Econometric Reviews*, Vol. 19, 1-54.

Estrella, A. and F. Mishkin (1999), "Rethinking the Role of NAIRU in Monetary Policy: Implications of Model Formation and Uncertainty," in J. Taylor, ed., *Monetary Policy Rules*, University of Chicago Press, 405-430.

Fieller, E. (1954), "Some Problems in Interval Estimation," *Journal of the Royal Statistical Society Series B*, Vol. 16, 175-185.

Härdle, W., J. Horowitz and J.-P., Kreiss (2002), "Bootstrap Methods in Time Series," manuscript, downloadable from http://faculty.econ.northwestern.edu/faculty/horowitz.

Hodrick, R. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, Vol. 5, 357-386.

Horowitz, J. (2001), "The Bootstrap," in J. Heckman and E. Leamer, eds, *Handbook of Econometrics IV*, North-Holland, 3159-3228.

Judge, G., C. Hill, W. Griffiths and H. Lütkepohl (1988), *Introduction to the Theory and Practice of Econometrics*, John Wiley and Sons.

Künsch, H. (1989), "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, Vol. 17, 1217-1241.

Lahiri, S.N. (1992), "Edgeworth Correction by 'Moving Block' Bootstrap for Stationary and Nonstationary Data," in R. LePage and L. Billard, eds, *Exploring the Limits of Bootstrap,* John Wiley and Sons, 183-214.

Lahiri, S.N. (1995), "On the Asymptotic Behavior of the Moving Block Bootstrap for Normalized Sums of Heavy-tail Random Variables," *Annals of Statistics*, Vol. 23, 1331-1349.

Lahiri, S.N. (1999), "Theoretical Comparisons of Block Bootstrap Methods," *Annals of Statistics*, Vol. 27, 386-404.

Laubach, T. (2001), "Measuring the NAIRU: Evidence from Seven Countries," *Review of Economics and Statistics*, Vol. 83, 218-231.

Li, H. and G.S. Maddala (1996), "Bootstrapping Time Series Models," *Econometric Reviews*, Vol. 15, 115-158.

Li, H. and G.S. Maddala (1999), "Bootstrap Variance Estimation of Nonlinear Functions of Parameters: An Implication to Long-run Elasticities of Energy Demand," *Review of Economics and Statistics*, Vol. 81, 728-733.

Lilliefors, H. (1967), “On the Kolmoorov-Smirnov Test for Normality with Mean and Variance Known,” *Journal of the American Statistical Association*, Vol. 62, 399-402.

Mizon, G., C. Safford and S. Thomas (1990), “The Distribution of Consumer Price Changes in the United Kingdom, *Economica*, Vol. 57, 249-262.

Newey, W. and K. West (1987), “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, Vol. 55, 703-708.

Politis, D. and J. Romano (1994), “The Stationary Bootstrap,” *Journal of the American Statistical Association*, Vol. 89, 1303-1313.

Staiger, D., J. Stock, and M. Watson (1997), “How Precise are Estimates of the Natural Rate of Unemployment?” in C. Romer and D. Romer, eds, *Reducing Inflation: Motivation and Strategy*, University of Chicago Press, 195-242.