Helsinki Center of Economic Research

Discussion Papers

Has U.S. Inflation Really Become Harder to Forecast?

Markku Lanne
University of Helsinki and HECER

and

Jani Luoto
University of Helsinki and HECER

Discussion Paper No. 307
October 2010

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No.307

Has U.S. Inflation Really Become Harder to Forecast?*

Abstract

Recently Stock and Watson (2007) showed that since the mid-1980s it has been hard for backward-looking Phillips curve models to improve on simple univariate models in forecasting U.S. inflation. While this indeed is the case when the benchmark is a causal autoregression, little change in forecast accuracy is detected when a noncausal autoregression is taken as the benchmark. In this note, we argue that a noncausal autoregression indeed provides a better characterization of U.S. inflation dynamics than the conventional causal autoregression and it is, therefore, the appropriate univariate benchmark model.

JEL Classification: C22, C53, E31

Keywords: Inflation forecast, Noncausal time series, Phillips curve.

Markku Lanne

Department of Political and Economic Studies
University of Helsinki
P.O. Box 17 (Arkadiankatu 7)
FI-00014 University of Helsinki
FINLAND

e-mail: markku.lanne@helsinki.fi

Jani Luoto

Department of Political and Economic Studies
University of Helsinki
P.O. Box 17 (Arkadiankatu 7),
FI-00014 University of Helsinki
FINLAND

e-mail: jani.luoto@helsinki.fi

* Financial support from the Academy of Finland and the OP-Pohjola Group Research Foundation is gratefully acknowledged.

1 Introduction

In their recent, widely cited article, Stock and Watson (2007, SW henceforth) argued that while U.S. inflation in general has become easier to forecast after 1983, it has also become more difficult to improve upon univariate models by means of backward-looking Phillips curve (PC) forecasts. Specifically, they claim that before 1983, PC models were superior to the univariate autoregressive (AR) model, but after 1984, the situation has reversed.We argue that SW's benchmark model is not the appropriate univariate model, especially in the 1970–1983 period, but, in fact, inflation dynamics are better captured by a noncausal, instead of a conventional causal AR model. This claim is backed up by the findings of Lanne and Saikkonen (2008) and Lanne et al. (2009) for the CPI inflation and Lanne et al. (2010) for the GDP price inflation. Also, in contrast to SW, we do not force a unit root in the inflation process.

Our results show that once the noncausal AR benchmark is adopted, the changes in the forecastability of U.S. GDP inflation are minor, and mainly confined to the two-year forecast horizon. As to the other inflation measures (personal consumption expenditure deflator for core items (PCE-core) and all items (PCE-all), and the consumer price index (CPI-U)) considered by SW, the PC forecasts very rarely beat the noncausal AR forecast in either forecast period.

Causal and noncausal AR models are both univariate and not discernible under Gaussian errors. However, as shown by Lanne and Saikkonen (2008), and Lanne et al. (2009, 2010), the errors of AR models estimated for U.S. inflation series are not well characterized as Gaussian, but exhibit excess kurtosis. In those papers, Student's t-distribution is assumed for the errors and it turns out to provide a good fit. Under this distributional assumption, noncausality can be checked, and a noncausal AR model indeed proves superior for U.S. inflation series based on the GDP deflator and the consumer price index. Moreover, it is the purely noncausal AR model without lagged inflation that, in general, yields the most accurate forecasts, and, therefore, we employ that model as well.

The plan of the paper is as follows. In Section 2, we present the noncausal AR model, and discuss estimation and forecasting. Section 3 presents the forecasting results and comparisons to SW's findings. Finally, Section 4 concludes.

2 Noncausal AR Model

Let us consider the following noncausal AR model for inflation π_t $(t = 0, \pm 1, \pm 2, ...)$,

$$\varphi\left(B^{-1}\right)\phi\left(B\right)\pi_t = \epsilon_t, \tag{1}$$

where $\phi\left(B\right) = 1 - \phi_1 B - \cdots - \phi_r B^r$, $\varphi\left(B^{-1}\right) = 1 - \varphi_1 B^{-1} - \cdots - \varphi_s B^{-s}$, and ϵ_t is a sequence of independent, identically distributed (continuous) random variables with mean zero and variance σ^2 or, briefly, $\epsilon_t \sim i.i.d.\left(0, \sigma^2\right)$. Moreover, B is the usual backward shift operator, that is, $B^k \pi_t = \pi_{t-k}$ $(k = 0, \pm 1, ...)$, and the polynomials $\phi\left(z\right)$ and $\varphi\left(z\right)$ have their zeros outside the unit circle so that

$$\phi\left(z\right) \neq 0 \quad \text{for} \quad |z| \leq 1 \quad \text{and} \quad \varphi\left(z\right) \neq 0 \quad \text{for} \quad |z| \leq 1. \tag{2}$$

This formulation was recently suggested by Lanne and Saikkonen (2008). We use the abbreviation AR(r, s) for the model defined by (1). If $\varphi_1 = \cdots = \varphi_s = 0$, model (1) reduces to the conventional causal AR(r) model.

The conditions in (2) imply that π_t has the two-sided moving average representation

$$\pi_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}, \tag{3}$$

where ψ_j is the coefficient of z^j in the Laurent series expansion of $\phi(z)^{-1} \varphi(z^{-1})^{-1} \overset{def}{=} \psi(z)$. Note that this implies that past observations can be used to predict future errors. From (1) one also obtains the representation

$$\pi_t = \phi_1 \pi_{t-1} + \cdots + \phi_r \pi_{t-r} + v_t, \tag{4}$$

where $v_t = \varphi(B^{-1})^{-1} \epsilon_t = \sum_{j=0}^{\infty} \beta_j \epsilon_{t+j}$ with β_j the coefficient of z^j in the power series expansion of $\varphi(B^{-1})^{-1}$. This representation can be used to obtain forecasts. Taking conditional expectations conditional on past and present inflation of (4) yields

$$\pi_t = \phi_1 \pi_{t-1} + \cdots + \phi_r \pi_{t-r} + E_t \left(\sum_{j=0}^{\infty} \beta_j \epsilon_{t+j} \right),$$

which shows that in a noncausal AR model, future errors are predictable by past values of inflation.

A well-known feature of noncausal autoregressions is that a non-Gaussian error term is required to achieve identification. Therefore, following Lanne and Saikkonen (2008), we specify Student's t-distribution for ϵ_t. In addition to these authors, also Lanne et al. (2009, 2010) have shown this distribution to fit U.S. inflation series well. Under this assumption, the noncausal AR model can be estimated by maximizing the approximate likelihood function proposed by Lanne and Saikkonen (2008).

To compute forecasts based on representation (4), simulation methods are called for. Let $E_T(\cdot)$ signify the conditional expectation operator given the observed data vector $\boldsymbol{\pi} = (\pi_1, ..., \pi_T)$. From (4) it is seen that the optimal predictor of π_{T+h} $(h > 0)$ based on $\boldsymbol{\pi}$ satisfies

$$E_T(\pi_{T+h}) = \phi_1 E_T(\pi_{T+h-1}) + \cdots + \phi_r E_T(\pi_{T+h-r}) + E_T(v_{T+h}).$$

Thus, if we are able to forecast the variable v_{T+h}, we can compute forecasts of inflation recursively. In the purely noncausal case of particular interest in this paper, the optimal forecast of π_{T+h} reduces to $E_T(v_{T+h})$. To calculate v_{T+h} in practice we use the approximation $v_{T+h} \approx \sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}$,where the integer M is supposed to be so large that the approximation error is negligible for all forecast horizons h of interest. To a close approximation we then have

$$E_T(\pi_{T+h}) \approx \phi_1 E_T(\pi_{T+h-1}) + \cdots + \phi_r E_T(\pi_{T+h-r}) + E_T\left(\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}\right). \quad (5)$$

Lanne et al. (2010) show how to generate by simulation the conditional density of future errors needed in the computation of the conditional expectation of $\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}$. Following their recommendations based on simulation experiments, we set $M = 50$, and the number of replications, N, in the simulation procedure equals 100 000.

3 Forecast Results

We focus on quarterly GDP price index inflation, but also present results for a number of other inflation measures. All data are downloaded from Mark Watson's web page. In addition to the univariate causal and noncausal AR models, random walk forecasts of Atkeson and Ohanian (2001) and moving average forecasts are considered. The PC forecasts are calculated using autoregressive distributed lag models with various activity variables and potentially gap variables based on them as additional regressors (SW's equation (3)). The specifications PC-Δu, PC-Δy, PC-ΔCapUtil and PC-ΔPermits omit gap variables. For detailed variable definitions, see SW.

The noncausal AR models are estimated recursively, with data from 1960:I–1969:IV used for initial parameter estimation. Following SW, forecast results are presented separately for the periods 1970:I–1983:IV and 1984:I–2004:IV. Unlike SW, we only consider iterated multistep forecasts that SW found quantitatively quite similar to their direct forecasts. Lanne and Saikkonen (2008) propose a model selection procedure that was employed in forecasting by Lanne et al. (2010). However, in this paper all noncausal forecasts are based on the recursively estimated fixed AR(0,4) model that should be adequate for quarterly data. SW mainly rely on the Akaike information criterion (AIC) in model selection, i.e., they recursively select the order of the AR model (denoted AR(AIC) below). However, they also show that the fixed AR(4,0) model produces similar results.

Table 1 reproduces the root mean squared forecast errors (RMSFE) of the AR(AIC) forecast and the relative mean squred forecast errors (MSFE) of a number of alternative models in relation to that model from SW's Tables 1 and 4. Compared to the benchmark AR(AIC) model, the predictive performance of virtually all PC models is inferior in the latter compared to the former subsample period at all horizons. This is even more clearly seen in the left panel of Table 3 that presents the percentage changes of the relative MSFEs. There are only two negative entries, both of which are small in absolute value compared to the positive percentage changes. Moreover, while in the 1970–1983 period, the relative MSFEs in Table 1 are, in general, less than unity, indicating the superiority of the PC models, the situation is reversed in the 1984–2004 period. This evidence warrants SW's claim that since the mid-1980s, it has been difficult for inflation forecasts to improve on univariate models.

If the purely noncausal AR(0,4) model is used as the benchmark, the results are drastically different. As the figures in Table 2 show, the PC forecasts are, in general inferior to this univariate benchmark model. In contrast to Table 1, this is the case also in the 1970–1983 period, while the performance of the AR(AIC) and AR(0,4) models is similar in the 1984–2004 period. As a result, the changes in predictive accuracy of the PC models are, in general, much smaller than SW's results in Table 1 lead one to believe. Moreover, the right panel of Table 3 shows that in many cases, the predictive performance of these models has improved, especially at horizons of four quarters or less, and in case of relative deterioration, it is much lesser than suggested by SW. Particularly noteworthy is the result that the model with the change in building permits as the predictor (PC-ΔPermits) is the only model that beats the AR(0,4) benchmark at all horizons in the latter subsample period, and shows great improvement in predictive accuracy over the 1970–1983 period.

In addition to the PC forecasts, SW also considered the random walk forecasts (AO) of Atkeson and Ohanian (2001) and first-order integrated moving average (IMA(1,1)) forecasts proposed in the previous literature. As to the former, our conclusion is similar to SW's: these forecasts improve upon the AR(0,4) and PC forecasts at the one- and two-year horizons in the 1984–2004 period but not at other horizons or in the first subsample period. SW interpret this finding as evidence in favor of inflation following an IMA(1,1) process, and this model indeed is superior to the AR(AIC) and PC forecasts. However, with the exception of the two-year horizon in the latter subsample period, the noncausal AR(0,4) model either beats the IMA(1,1) forecasts or produces very similar accuracy. Hence, the simple AR(0,4) model seems to provide, in general, a better description of the inflation process in both subsample periods at least in terms of predictive accuracy.

According to SW, inflation dynamics are well characterized by an IMA(1,1) process with time-varying parameters. They present forecast results with the moving average (MA) parameter θ equal to 0.25 and 0.65, in the first and second subsample periods, respectively. For the AR(AIC) model, these results are reproduced in Table 1, whereas Table 2 shows the corresponding figures related to the AR(0,4) model. While the IMA(1,1) model with $\theta = 0.25$ produces more accurate forecasts in the 1970–1983 period compared to the AR(AIC) model, it is clearly inferior to the AR(0,4) model at the one- and two-year horizons. In the 1984–2004 period, the IMA(1,1,) model with $\theta = 0.65$ is clearly superior to both of these models. These findings undermine the alleged superiority of the time-varying IMA model as a description of U.S. inflation dynamics.

Following SW, we also considered other inflation measures (PCE-core, PCE-all and CPI-U) besides the GDP price inflation. The results (not shown) are similar to those discussed above in that the differences in accuracy of the PC forecasts between the two subsample periods are minor vis-à-vis the AR(0,4) benchmark. With the exception of the one-quarter horizon, the noncausal AR(0,4) model, in general, turned out to be more accurate. Interestingly, also for both PCE inflation measures, the PC-ΔPermits model is the only model to show substantial improvement in predictive accuracy in the latter compared to the first subsample period.[1]

[1] The detailed results are available upon request.

4 Conclusion

In this note, we have shown that compared to the noncausal AR benchmark model, U.S. inflation has hardly become more difficult to forecast by means of backward-looking PC models since the mid-1980s, contrary to the claims of SW, who used the causal autoregression as the benchmark. At least the differences are much smaller than SW found, and in one case (the change in building permits as the predictor in the PC model), inflation forecasts even seem to have improved. Based on the findings in the previous literature, U.S. inflation dynamics are better described by a noncausal than a causal AR model, and hence, the noncausal model should be taken as the benchmark model against which the PC forecasts are judged. These findings are reinforced by the fact that the noncausal AR(0,4) model also consistently produces more accurate forecasts than the causal AR(4,0) or AR(AIC) models in the 1970–1983 period and has comparable accuracy in the 1984–2004 sample. Our results show that compared to this univariate benchmark, the PC models provide poor forecasts both before and after the mid-1980s. We also find some evidence against U.S. inflation being well characterized as a time-varying IMA(1,1) process, as these forecasts can be clearly inferior to the AR(0,4) forecasts in the 1970–1983 period.

The question why the noncausal AR model seems to forecast U.S. inflation better than causal AR or PC models, remains unanswered in this note. One potential explanation that we are working on, is related to the predictability of the errors of the noncausal AR model pointed out in Section 2. We conjecture that these errors are able to approximate information that is missing in the simple autoregressive model, and because they are predictable, part of this information is made use of in forecasting.

References

Atkeson, Andrew, and Lee E. Ohanian (2001). "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 25:1, 2–11.

Lanne, M., and P. Saikkonen (2008). "Modeling Expectations with Noncausal Autoregressions." HECER Discussion Paper 212.

Lanne, M., A. Luoma, and J. Luoto (2009). "Bayesian Model Selection and Forecasting in Noncausal Autoregressive Models." *Journal of Applied Econometrics* (forthcoming).

Lanne, M., J. Luoto, and P. Saikkonen (2010). "Optimal Forecasting of Noncausal Autoregressive Time Series." HECER Discussion Paper 286.

Stock, James H., and Mark W. Watson (2007) "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, Supplement to Vol. 39, 3–33.

Table 1: Pseudo out-of-sample forecasting results for GDP inflation with the AR(AIC) model as the benchmark.

	1970:I–1983:IV				1984:I–2004:IV			
	$h=1$	$h=2$	$h=4$	$h=8$	$h=1$	$h=2$	$h=4$	$h=8$
AR(AIC) RMSFE	1.72	1.75	1.89	2.38	0.78	0.68	0.62	0.73
Relative MSFEs								
AR(AIC)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AO	1.95	1.57	1.06	1.00	1.22	1.10	0.89	0.84
PC-u	0.85	0.92	0.88	0.61	0.95	1.11	1.48	1.78
PC-Δu	0.87	0.87	0.86	0.64	1.06	1.27	1.83	2.21
PC-$ugap^{1\text{-}sided}$	0.88	0.99	0.98	0.87	1.06	1.29	1.84	2.39
PC-Δy	0.99	1.06	0.93	0.58	1.05	1.06	1.23	1.53
PC-$ygap^{1\text{-}sided}$	0.94	0.97	0.99	0.78	0.97	0.97	1.25	1.55
PC-CapUtil	0.85	0.88	0.79	0.55	0.95	1.01	1.35	1.52
PC-ΔCapUtil	1.02	1.00	0.87	0.64	1.03	1.10	1.30	1.51
PC-Permits	0.93	1.02	0.98	0.78	1.08	1.23	1.31	1.52
PC-ΔPermits	1.02	1.04	0.99	0.86	1.00	1.00	1.00	1.02
AR(4,0)	0.95	1.08	1.05	0.93	0.93	0.96	0.99	0.94
IMA(1,1)	0.82	0.83	0.87	0.89	1.01	1.03	0.98	0.89
IMA(1,1), $\theta = 0.25$	0.79	0.80	0.82	0.87	1.05	1.11	1.05	0.93
IMA(1,1), $\theta = 0.65$	0.97	0.94	0.96	0.90	0.90	0.87	0.89	0.82
AR(0,4)	0.87	0.81	0.75	0.66	1.03	1.02	1.01	1.07

The table reproduces results in Stock and Watson's (2007) Tables 1 (top panel) and 4 (middle panel) and presents our results based on the noncausal AR(0,4) model (bottom panel). The first row reports the root mean squared forecast errors of the causal AR(AIC) benchmark forecast. The rest of the entries are the relative mean squared forecast errors relative to the AR(AIC) benchmark.

Table 2: Pseudo out-of-sample forecasting results for GDP inflation with the AR(0,4) model as the benchmark.

	1970:I–1983:IV				1984:I–2004:IV			
	$h = 1$	$h = 2$	$h = 4$	$h = 8$	$h = 1$	$h = 2$	$h = 4$	$h = 8$
AR(0,4) RMSFE	1.60	1.58	1.64	1.93	0.79	0.69	0.62	0.76
Relative MSFEs								
AR(0,4)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AO	2.25	1.93	1.41	1.51	1.19	1.08	0.88	0.78
PC-u	0.98	1.14	1.18	0.93	0.92	1.09	1.46	1.66
PC-Δu	1.00	1.07	1.15	0.96	1.03	1.25	1.82	2.06
PC-$ugap^{1\text{-}sided}$	1.01	1.21	1.31	1.32	1.04	1.27	1.82	2.23
PC-Δy	1.14	1.30	1.25	0.88	1.02	1.04	1.22	1.43
PC-$ygap^{1\text{-}sided}$	1.09	1.20	1.33	1.18	0.94	0.96	1.24	1.44
PC-CapUtil	0.98	1.08	1.06	0.84	0.93	0.99	1.34	1.41
PC-ΔCapUtil	1.18	1.23	1.16	0.97	1.00	1.08	1.29	1.41
PC-Permits	1.08	1.25	1.31	1.19	1.05	1.21	1.29	1.42
PC-ΔPermits	1.18	1.28	1.32	1.30	0.97	0.98	0.99	0.95
AR(4,0)	1.10	1.33	1.34	1.51	0.97	0.98	0.99	0.93
IMA(1,1)	0.95	1.02	1.17	1.35	0.98	1.01	0.97	0.83
IMA(1,1), $\theta = 0.25$	0.91	0.98	1.10	1.32	1.02	1.09	1.04	0.87
IMA(1,1), $\theta = 0.65$	1.12	1.16	1.29	1.36	0.88	0.86	0.88	0.76

The first row reports the root mean squared forecast errors of the AR(0,4) benchmark forecast. The rest of the entries are the relative mean squared forecast errors relative to the AR(0,4) benchmark.

Table 3: Percentage changes in the relative MSFE in relation to the AR(AIC) model (left panel) and the AR(0,4) model (right panel) between the 1970–1983 and 1984–2004 periods.

	Benchmark Model							
	AR(AIC)				AR(0,4)			
	$h=1$	$h=2$	$h=4$	$h=8$	$h=1$	$h=2$	$h=4$	$h=8$
AO	–46.3%	–35.7%	–17.6%	–17.0%	–63.5%	–57.8%	–47.8%	–65.4%
PC-u	11.3%	18.1%	51.6%	106.7%	–5.8%	–4.1%	21.3%	58.3%
PC-Δu	20.2%	38.3%	76.3%	124.2%	3.1%	16.2%	46.1%	75.8%
PC-$ugap^{1\text{-}sided}$	19.4%	26.6%	63.0%	100.6%	2.2%	4.4%	32.8%	52.2%
PC-Δy	6.1%	0.3%	27.6%	96.4%	–11.0%	–21.9%	–2.6%	48.0%
PC-$ygap^{1\text{-}sided}$	2.7%	0.0%	23.8%	68.1%	–14.5%	–22.2%	–6.4%	19.7%
PC-CapUtil	11.9%	13.4%	53.2%	100.8%	–5.3%	–8.8%	23.0%	52.4%
PC-ΔCapUtil	0.0%	9.5%	40.2%	86.3%	–17.1%	–12.7%	10.0%	37.9%
PC-Permits	14.6%	18.5%	28.6%	66.3%	–2.5%	–3.7%	–1.6%	17.9%
PC-ΔPermits	–2.4%	–4.3%	1.5%	16.9%	–19.5%	–26.5%	–28.7%	–31.5%