The Incremental Predictive Information Associated with Using Theoretical New Keynesian DSGE Models Versus Simple Linear Econometric Models

Oleg Korenok[1] and Norman R. Swanson[2]
[1]Virginia Commonwealth University
[2]Rutgers University

August 2005

Abstract

In this paper we construct output gap and inflation predictions using a variety of DSGE sticky price models. Predictive density accuracy tests related to the test discussed in Corradi and Swanson (2005a) as well as predictive accuracy tests due to Diebold and Mariano (1995) and West (1996) are used to compare the alternative models. A number of simple time series prediction models (such as autoregressive and vector autoregressive (VAR) models) are additionally used as strawman models. Given that DSGE model restrictions are routinely nested within VAR models, the addition of our strawman models allows us to indirectly assess the usefulness of imposing theoretical restrictions implied by DSGE models on unrestricted econometric models. With respect to predictive density evaluation, our results suggest that the standard sticky price model discussed in Calvo (1983) is not outperformed by the same model augmented either with information or indexation, when used to predict the output gap. On the other hand, there are clear gains to using the more recent models when predicting inflation. Results based on mean square forecast error analysis are less clear-cut, although the standard sticky price model fares best at our longest forecast horizon of 3 years, and performs relatively poorly at shorter horizons. When the strawman time series models are added to the picture, we find that the DSGE models still fare very well, often winning our forecast competitions, suggesting that theoretical macroeconomic restrictions yield useful additional information for forming macroeconomic forecasts.

JEL classification: E12, E3, C32
Keywords: sticky price, sticky information, predictive density, model selection.

* Oleg Korenok (okorenok@vcu.edu), VCU School of Business, 1015 Floyd Avenue, Richmond, VA 23284-4000 and Norman R. Swanson (nswanson@econ.rutgers.edu): Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA. This paper has been prepared for the special issue on "Information in Economic Forecasting" co-edited by Michael Clements and David Hendry. The authors owe a great many thanks to both editors for numerous useful comments and guidance on an earlier draft of this paper. The authors also wish to thank Michael Bordo, Roberto Chang, Valentina Corradi, Bruce Mizrach, and Ricardo Reis for useful comments and suggestions. Swanson has benefited from the support of Rutgers University in the form of a Research Council grant.

1 Introduction

In the analysis of stochastic dynamic general equilibrium models one strand of the recent literature focuses on the reconciliation of historical and simulation based empirical evidence.[1] A different strand of the literature focuses on out-of-sample model evaluation and on the fact that all models may well be approximations, and so are misspecified (i.e. no models are "correctly specified"). Evaluation is done in a number of ways, including the construction of Bayesian odds ratios/Kullback-Leibler Information Criterion (KLIC) for comparing RBC models and time series models (see e.g. DeJong, Ingram and Whiteman (2000), Schorfheide (2000), Chang, Gomes and Schorfheide (2002), Fernandez-Villaverde and Rubio-Ramirez (2004), and Del Negro and Schorfheide (2005)). Corradi and Swanson (CS: 2005a,b,c,d) develop an alternative approach applicable in this context, which is based upon the construction of distributional loss measures and test statistics associated with either simulations and /or predictions from DSGE models.[2]

Our intent in this paper is to add to the second strand of research mentioned above. In particular, we apply the methodology developed by CS (2005a,b,c) to the evaluation of the standard sticky price model discussed in Calvo (1983), the sticky price with dynamic indexation model discussed in Christiano, Eichenbaum and Evans (2001), Del Negro and Schorfheide (2005) and Smets and Wouters (2003), and the sticky information model of Mankiw and Reis (2002). Additionally, we consider some simple strawman time series models in our evaluation. In this sense, we attempt to address two issues. First, do recent theoretical advances in the sticky price literature yield models that provide superior predictive densities? This question addresses the issue of whether or not recent theoretical advances in price theory translate into better predictions. Second, how do unrestricted econometric time series models perform, relative to theoretically intricate DSGE models? This question addresses the notion that theoretical restriction imposition may or may not yield superior predictive performance. In particular, as it is well known that the restrictions implied by theoretical DSGE models can be nested in VAR models (see e.g. Bierens and Swanson (2000) and Bierens (2005)), our direct comparison of DSGE model predictions with unrestricted VAR predictions should yield new evidence of the usefulness of theory information in economic forecasting.

[1] See e.g. Watson (1993), Cogley and Nason (1995), Rotemberg and Woodford (1996), Diebold, Ohanian and Berkowitz (1998), Bierens and Swanson (2000), Schmitt-Grohe (2000), and Bierens (2005).

[2] For further details on recent predictive density evaluation techniques, see Clements and Smith (2000,2002) and the references cited therein.

These questions are addressed via examination of marginal and joint predictive densities for the output gap and inflation. In addition, mean square forecast errors (MSFEs) are used to construct predictive accuracy tests using the approach of Diebold and Mariano (1995), West (1996), and Clark and McCracken (2001).[3]

The impetus for this paper derives primarily from the observation that new Keynesian Phillips curves based on standard sticky price assumptions have several shortcomings. For example, Ball (1994) has found that such models yield the controversial result that an announced credible disinflation causes booms rather than recessions. Additionally, Fuhrer and Moore (1995) show that the New Keynesian Phillips curve falls short when used to explain inflation persistence, one of the stylized empirical facts describing US inflation. Finally, Mankiw and Reis (2002) note that such models have trouble explaining why shocks to monetary policy have delayed and gradual effects on inflation.[4] Some of the problems outlined above are addressed in Christiano et al. (2001), Del Negro and Schorfheide (2005), Smets and Wouters (2003), and Mankiw and Reis (2002), via the introduction of sticky prices with dynamic indexation as well as sticky information models. For example, Mankiw and Reis posit that information about macroeconomic conditions spreads slowly because of information acquisition and/or re-optimization costs. Compared to the standard sticky price model, prices in this setup are always readjusted, but decisions about prices are not always based on the latest available information. The model is representative of the wider class of Rational Inattention (RI) models developed by Phelps (1970), Lucas (1973), and more recently by Mankiw and Reis (2002), Sims (2003), and Woodford (2003).

As might be expected, the three models that we consider have very different properties. For example, Ball, Mankiw and Reis (2005) show that implications with regard to optimal monetary policy are quite different for sticky price and sticky information models. In the sticky price model, inflation enters the loss function, which leads to inflation targeting. It is thus optimal to allow inflation drift in this model. On the other hand, in the sticky information model, inflation drift or inflation targeting is a suboptimal policy, as it is optimal to target the price level. These sorts of model implications suggest that the dynamic properties of the alternative models may be quite different, in turn implying that our predictive density comparison of historical and predicted inflation and the output gap measures may uncover interesting new evidence concerning the relative

[3]For further discussion of the use of MSFEs in econometrics, see Clements and Hendry (1993).

[4]See also Bernanke and Gertler (1995) and Christiano and Eichenbaum and Evans (2000).

merits of the models. Put another way, our approach allows us to shed light on the issue of whether theoretical advantages translate into better forecasting performance.

Our findings can be briefly summarized as follows. First, more recent theoretical DSGE models fare better than the standard sticky price model for various forecast horizons and forecast evaluation periods, when recursive estimation and prediction methodologies are implemented. However, there are surprising exceptions to this finding. First, when MSFE loss is the relevant loss measure, the standard model dominates all theoretical models for our longest horizon forecasts (three years). Second, under predictive density loss, the standard model fares surprisingly well for predicting the output gap. These somewhat disparate findings underscore the importance of which loss function is deemed relevant. Second, simple time series models do not dominate DSGE models under either MSFE or CS distributional loss, a finding which is somewhat surprising given the often cited macroeconomic findings supporting the dominance of simple time series models for prediction, and which suggests that economic theory and restrictions implied by economic theory is useful in the context of econometric forecasting.

The rest of the paper is organized as follows. Section 2 outlines our DSGE models, in which the New Keynesian Phillips curve is derived under sticky price, sticky price with indexation, and sticky information assumptions. In Section 3, we describe the data used to construct historical measures of inflation and the output gap; and discuss calibration. Section 4 summarizes the predictive density methodology used to compare the models, and empirical results are gathered in Section 5. Concluding remarks are gathered in Section 6.

2 New Keynesian DSGE Models for Inflation and the Output Gap

In this section we outline the sticky price, sticky information and sticky price with indexation models that will be compared and contrasted in the sequel. Our presentation of the models follows closely along the lines of Gali (2002) and Woodford (2002).

With respect to households, assume that the representative consumer's preferences are represented by the following utility function:

$$U(C_t, N_t(i)) = \frac{C_t^{(1-\sigma)}}{1-\sigma} - \int_0^1 \frac{N_t(i)^{(1+\varphi)}}{1+\varphi} di, \tag{1}$$

where $N_t(i)$ denotes the quantity of labor supplied by a consumer of type "i", and C_t is an index of the different goods consumed. We assume a factor specific labor market, so that production of good i requires labor of type i to be used. The parameter σ is the inverse of the intertemporal elasticity of substitution, and the parameter φ is the inverse of the elasticity of labor supply. Assume further that C_t is a constant-elasticity-of-substitution index, namely, $C_t = \left(\int_0^1 C_t(i)^{(\frac{\varepsilon-1}{\varepsilon})}\, di\right)^{(\frac{\varepsilon}{\varepsilon-1})}$, where $\varepsilon < 0$. The corresponding price index, P_t, is given by $P_t = \left(\int_0^1 P_t(i)^{(1-\varepsilon)}\, di\right)^{(\frac{1}{1-\varepsilon})}$, where $P_t(i)$ denotes the price of good $i \in [0,1]$. Subject to a standard sequence of budget constraints and a solvency condition, the solution to the consumer's optimization problem can be summarized in log-linear ($x_t = \ln X_t$) form by two static conditions:

$$c_t(i) = -\varepsilon\left(p_t(i) - p_t\right) + c_t \tag{2}$$

$$w_t(i) - p_t = \sigma\, c_t + \varphi\, n_t(i), \tag{3}$$

where $w_t(i)$ is the log nominal wage paid for labor type i; and by the intertemporal Euler equation:

$$c_t = -\frac{1}{\sigma}\left(r_t - E_t\, \pi_{t+1} - \rho\right) + E_t\, c_{t+1}, \tag{4}$$

where r_t is the yield on a nominal riskless one period bond (i.e. the nominal interest rate), π_{t+1} is the rate of inflation between t and $t+1$, $\rho = -\ln\beta$ represents the time discount rate (as well as the steady state real interest rate, given the absence of secular growth), and β is the subjective discount factor. Finally, following Gali (2002) we postulate (without derivation) a standard money demand equation. Namely, $m_t - p_t = y_t - \eta\, r_t$, which has unit income elasticity.

With respect to firms, we assume that there exists a continuum of firms, each producing a differentiated good, $Y_t(i) = A_t\, N_t(i)^{\alpha}$, where log of productivity evolves according to the following process: $\triangle a_t = \rho_a \triangle a_{t-1} + \varepsilon_{a,t}$, which is an exogenous, difference-stationary stochastic process. Assume further that the producer is a wage taker, so that the real marginal cost of supplying good i is equal to:

$$MC_t\,(i) = \frac{1}{\alpha}\frac{W_t\,(i)}{P_t A_t}\left(\frac{Y_t\,(i)}{A_t}\right)^{\frac{1}{\alpha}-1}. \tag{5}$$

Total demand for good i is thus given by $Y_t(i) = C_t(i)$. Now, let $Y_t = \left(\int_0^1 Y_t(i)^{(\frac{\varepsilon-1}{\varepsilon})}\, di\right)^{(\frac{\varepsilon}{\varepsilon-1})}$ denote aggregate output. Then equilibrium in the goods market implies that $Y_t = C_t$. Combining the real marginal cost equation together with a market clearing condition and the static first order condition from the consumer optimization problem, and taking a log transformation yields the equilibrium real marginal cost of the individual firm in terms of output produced by the individual firm, aggregate output and productivity. Namely:

$$mc_t(i) = \sigma\, y_t + \omega y_t(i) - (1+\omega)\, a_t - \ln(\alpha), \tag{6}$$

where $\omega = \frac{\psi}{\alpha} + \frac{1}{\alpha} - 1$. We also can combine the Euler equation with the market clearing condition to get another equilibrium condition, as follows:

$$y_t = -\frac{1}{\sigma}\left(r_t - E_t\, \pi_{t+1} - \rho\right) + E_t\, y_{t+1}. \tag{7}$$

In deriving equilibrium behavior it remains to discuss how firms set prices. In this section we describe four alternative models of price setting behavior, the final three of which will be examined in the sequel.

I. Flexible Prices: First, suppose that all firms choose the price of good i each period, independent of prices that were charged in the past, and with full information about current demand and cost. Due to the fact that real marginal costs are increasing in $y_t(i)$, the same quantity of each good is supplied, and it is equal to Y_t. This implies that all firms will choose a common constant markup given by $\mu = \frac{\epsilon}{(\epsilon-1)}$. The flexible price equilibrium process for output, consumption, and the expected real rate is given by:

$$y_t^n = \gamma + \psi_a\, a_t, \tag{8}$$

$$c_t^n = \gamma + \psi_a\, a_t, \tag{9}$$

$$r_t^n = \rho + \sigma\, \phi_a\, \rho_a \triangle a_{t-1}, \tag{10}$$

where $\psi_a = \frac{1+\omega}{\sigma+\omega}$ and $\gamma = \frac{\ln\alpha - \mu}{\sigma+\omega}$. We will refer to the above equilibrium conditions as a natural levels of the corresponding variables.

II. The Sticky Price Model: Following Calvo (1983), assume that in every period, a fraction, $(1-\theta_1)$, of firms can set a new price, independent of the past history of price changes. This set-up implies that the expected time between price changes is $\frac{1}{1-\theta_1}$. Also assume that firms that cannot set their prices optimally have to keep last periods' price (i.e. $P_t(i) = P_{t-1}(i)$).

III. The Sticky Price Model with Indexation: Modifications of the standard sticky price model have been shown by numerous authors to perform better in empirical applications. For example, we follow Christiano et al. (2001), Smets and Wouters (2003), and Del Negro and Schorfheide (2005), who use dynamic inflation indexation. In this model, as in Calvo (1983), only a proportion of firms, $(1-\theta_2)$, can reset their prices during the current period; but other firms, unable to set prices optimally, set their price equal to: $P_t(i) = \pi_t P_{t-1}(i)$.

IV. The Sticky Information Model: Following Mankiw and Reis (2002), assume that all firms reset prices each period. A fraction of firms, $(1-\theta_3)$, use current information in pricing decisions, so that the probability that a firm acts upon the newest information available in a given quarter is $1-\theta_3$, independent of the past history of price changes. The remaining fraction of firms use past or outdated information when they set prices. The sticky information model can be interpreted as a model where firms, which are unable to set prices optimally, use even more complex updating schemes than in the case of the sticky price model with indexation. Instead of using past inflation for indexation, when they have opportunity to use current information, firms in the sticky information model solve not only for the optimal current price, but also for the infinite path of future prices. Later, when firms do not have the opportunity to update information, they set price equal to the appropriate value in their solution set; a set which was calculated based on the old information set.

In these models, the fact that a fraction of firms is not able to adjust prices optimally implies a difference between the actual and the potential (natural) level of output. We denote this difference by $y_t^g = y_t - y_t^n$, and refer to it as the output gap. Now, solving the associated optimization problems and using a log-linear transformation, we can write expressions for the Phillips curve for each model.[5] In particular, the dynamics of inflation in the sticky price economy is characterized by New Keynesian Phillips Curve:

$$\pi_t = \beta \, E_t \, \pi_{t+1} + \lambda_1 \, y_t^g, \tag{11}$$

where $\lambda_1 = \frac{(1-\theta_1)(1-\beta\theta_1)\xi}{\theta_1}$ and $\xi = \frac{\omega+\sigma}{1+\varepsilon\,\omega}$. In the sticky price model with indexation the above equation has a hybrid New Keynesian Phillips Curve analog:

$$\pi_t = \frac{1}{1+\beta}\pi_{t-1} + \frac{\beta}{1+\beta} \, E_t \, \pi_{t+1} + \frac{\lambda_2}{1+\beta} \, y_t^g, \tag{12}$$

where $\lambda_2 = \frac{(1-\theta_2)(1-\beta\theta_2)\xi}{\theta_2}$. Finally, in the sticky information model, dynamics of inflation are governed by a sticky information Phillips Curve:

$$\pi_t = \frac{(1-\theta_3)\,\xi}{\theta_3} \, y_t^g + (1-\theta_3) \sum_{k=0}^{\infty} E_{t-k-1} \theta_3^k (\pi_t + \xi \, \triangle y_t^g). \tag{13}$$

[5] For a detailed derivation for the sticky price and the sticky price with indexation models, see Woodford (2003). For derivation using the sticky information model, see Khan and Zhu (2002).

Notice that the Euler equation above can be written in terms of the output gap. Namely $y_t^g = -\frac{1}{\sigma}\left(r_t - E_t\,\pi_{t+1} - r_t^n\right) + E_t\,y_{t+1}^g$.

To close our models, specify a monetary policy rule by assuming that an exogenous path for the growth rate of the money supply is given by the following stationary process, $\triangle m_t = \rho_m \triangle m_{t-1} + \varepsilon_{m,t}$, where $\rho_m \in [0, 1]$. This yields the desired outcome that: (i) the money demand equation, (ii) the equilibrium Euler equation, (iii) one of three of the Phillips curve equations: (11), (12) or (13), (iv) the specification of an exogenous process for technology, and (v) the exogenous process for the money supply fully describe the equilibrium dynamics of the economy, and in particular, the dynamics of the (endogenous) output gap and inflation variables in the models. The system is solved out using standard solution techniques; and the solution for all three models will have the following form: $x_t = \Theta_{t-1} + \Psi z_t$, where Θ and Ψ are solution matrices, which are functions of the structural parameters of the models, x_t is a vector of variables in economy, and z_t is a vector of exogenous disturbances. For example, for the sticky price model, $x_t = (\pi_t, y_t^g, r_t, \triangle m_t, \triangle a_t)'$ and $z_t = (\varepsilon_{m,t}, \varepsilon_{a,t})'$.

3 Data and Calibration

Our empirical investigation is based upon the use of quarterly U.S. data for the period 1964:1 - 2004:4. For our measure of inflation, we use the consumer price index (CPI).[6] We use the output gap measure constructed by the OECD.[7] For the measures of nominal interest rate and money supply we used the federal funds rate and M2. The federal funds rate and money supply data were taken from the OECD Main Economic Indicators database (Vol 2005 release 06 Database Edition (ISSN 1608-1234)). Output gap estimates and the CPI were obtained from OECD Economic Outlook database (Vol 2005 release 01 Database Edition (ISSN 1608-1153)).

With regard to calibration, we follow the approach of Gali (2002). Namely, assume log utility of consumption, so that $\sigma = 1$. Also, set the labor wage elasticity as $\psi = 1$, and set the value of the elasticity of money demand with respect to the interest rate as $\eta = 1$, which is consistent with the interest rate elasticity found in empirical work and used in other calibration studies (see e.g.

[6] The GDP deflator was also used in order to check for the robustness of our results, which are qualitatively similar, regardless of which index is used. Complete results are available upon request.

[7] We also constructed the output gap using the Hodrick-Prescott (H-P) filter and the one-sided optimal bandpass filter. Empirical results were qualitatively the same as those reported here, and are available upon request.

Chari, Kehoe, and McGrattan (1996)). The Dixit-Stiglitz elasticity of substitution is set to $\epsilon = 11$, which implies a 10% markup of price over marginal cost; and the consumer discount factor is set to $\beta = 0.99$, which implies an average annual interest rate 4%. We set the labor share parameter to $\alpha = 2/3$.

The degree of information and price stickiness, θ, was chosen to be common across all models and is set to $\theta = 0.5$.[8] This implies twice yearly price or information updating. The motivation for this level of stickiness comes from Bils and Klenow (2004), who study price stickiness by examining 350 categories of goods and services, constituting about 70% of consumer spending, and find evidence of more frequent price changes than hitherto suspected. In addition in Korenok and Swanson (2005), we find that simulated inflation and output gap data from theoretical models is much "closer" (we evaluate "closeness" using a simulation based model selection framework) to historical levels when adjustment occurs twice a year instead of the more standard annual adjustment, where $\theta = 0.75$.[9]

Finally, the exogenous processes are calibrated in the following way. For the technology growth rate, we set the value of the autoregression coefficient, ρ_a, equal to zero, and the standard deviation equal to $\sigma_a = 0.007$. The low value of ρ_a accounts for the low autocorrelation of output growth. Of further note is that the usual standard deviation for the technology growth rate is at or below 1% (see e.g. Gali (2002) or Gali et al. (2003)). The autoregression coefficient of growth in the money supply is set equal to $\rho_m = 0.5$, and the standard deviation is set equal to $\sigma_m = 0.007$; a value which is close to the estimated parameters for autoregressive processes describing M0, M1 or M2 growth rates in the United States.[10] Finally we make a standard assumption for exogenous shocks, $\varepsilon_{a,t} \sim i.i.d.N(0, \sigma_a)$ and $\varepsilon_{m,t} \sim i.i.d.N(0, \sigma_m)$.

[8] Our motivation for a common value for information and price stickiness comes from the fact that empirical estimates of information and price stickiness are quite close.

[9] See e.g. Blinder, Canetti, Lebow and Rudd (1998), Gali and Gertler (1999), Khan and Zhu (2002), Gali (2002), Sbordone (2002), Gali, Lopez-Salido and Valles (2003), Smets and Wouters (2003), and Woodford (2003), and Korenok (2004).

[10] See Mankiw and Reis (2002), Cooley and Hansen (1989), Walsh (1998), and Yun (1996) for further justification of this calibration.

4 Empirical Methodology

We begin by briefly summarizing an out-of-sample version of the distributional accuracy test discussed in CS (2005a). In particular, the test discussed below is the same as that in CS, except that the "simulation period" used in the formation of the simulated distributions in the statistic in CS is replaced with one constructed using predictions. For discussion of asymptotic distribution theory closest to the current context, the reader is referred to Bhardwaj, Corradi and Swanson (2005) and Corradi and Swanson (2005b,c,d). Assume that the objective is to compare the joint (predictive) distribution of the historical data with the joint distribution of the predicted series. Consider m DSGE models, and set model 1 as the benchmark model. Let $\Delta \log X_t$, $t = 1, ..., T$ denote the actual historical (output) series, and let $\Delta \log X_{j,n}$, $j = 1, ..., m$ and $n = 1, ..., P$, denote the output series predicted under model j, where P denotes the length of the prediction period. Along these lines, denote $\Delta \log X_{j,n}(\widehat{\theta}_{j,T})$, $n = 1, ..., P$, $j = 1, ..., m$ to be a sample of length P drawn (simulated or constructed using standard prediction approaches) from model j and evaluated at the parameters estimated, under model j, using the R available historical observations, where $T = R + P$.[11] For simplicity, assume that the forecast horizon is $\tau = 1$, although empirical results are reported for a number of forecast horizons. The reason why we use differences is that stationarity is assumed in our subsequent analysis (and is borne out in our data via application of standard unit root tests). For ease of exposition, and in keeping with our focus on current and lagged values of the variable of interest when we evaluate marginal distributions, let $Y_t = (\Delta \log X_t, \Delta \log X_{t-1})$, $Y_{j,n}(\widehat{\theta}_{j,T}) = (\Delta \log X_{j,n}(\widehat{\theta}_{j,T}), \Delta \log X_{j,n-1}(\widehat{\theta}_{j,T}))$. Also, let $F_0(u; \theta_0)$ denote the distribution of Y_t evaluated at u and $F_j(u; \theta_j^\dagger)$ denote the distribution of $Y_{j,n}(\theta_j^\dagger)$, where $\theta_j^\dagger$ is the probability limit of $\widehat{\theta}_{j,T}$, taken as $T \to \infty$, and where $u \in U \subset \Re^2$, possibly unbounded. Accuracy is measured in terms of forecast square error. In particular, the squared (approximation) error associated with model j, $j = 1, ..., m$, is measured in terms of the (weighted) average over U of $E\left(\left(F_j(u; \theta_j^\dagger) - F_0(u; \theta_0)\right)^2\right)$, where $u \in U$, and U is a possibly unbounded set on $\Re^2$. Thus, the rule is to choose Model 1 over Model 2 if

$$\Xi = \int_U E\left(\left(F_1(u; \theta_1^\dagger) - F_0(u; \theta_0)\right)^2\right) \phi(u) du - \int_U E\left(\left(F_2(u; \theta_2^\dagger) - F_0(u; \theta_0)\right)^2\right) \phi(u) du < 0,$$

[11]The T in $\Delta \log X_{j,n}(\widehat{\theta}_{j,T})$ is used for notational convenience, and should be replaced with R (for non-recursive estimation), and with a t when estimation and prediction is carried our recursively. See Section 5 for further details.

where $\int_U \phi(u)du = 1$ and $\phi(u) \geq 0$ for all $u \in U \subset \Re^2$. For any evaluation point, this measure defines a norm and it implies a usual goodness of fit measure. The hypotheses of interest are: $H_0 : \max_{j=2,\ldots,m} \Xi \leq 0$ versus $H_A : \Xi > 0$. Thus, under H_0, no model can provide a better approximation (in square error sense) to the distribution of Y_t than the approximation provided by model 1. In order to test H_0 versus H_A, the relevant test statistic is $\sqrt{T}Z_{T,S}$, where $Z_{T,S} = \max_{j=2,\ldots,m} \int_U Z_{j,T,S}(u)\phi(u)du$, and

$$Z_{j,P,S}(u) = \frac{1}{P}\sum_{t=R+1}^{T}\left\{\left(1\{Y_t \leq u\} - \frac{1}{P}\sum_{n=1}^{P} 1\{Y_{1,n}(\widehat{\theta}_{1,T}) \leq u\}\right)^2 - \left(1\{Y_t \leq u\} - \frac{1}{P}\sum_{n=1}^{P} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \leq u\}\right)^2\right\},$$

where $\widehat{\theta}_{j,T}$ is an estimator of $\theta_j^\dagger$, which is estimated either recursively, or using a single in-sample period. In this way, all predictions are truly *ex ante.* For the above test, we report the simple variety of critical values constructed as in CS (2005a), although these are not valid for our recursively estimated models (see CS (2005b,e) for a complete discussion of bootstrap techniques in recursive prediction contexts), and are only shown in the non-recursive case to be valid when used in-sample (see CS (2005a)). In light of this, CS type critical values reported here are meant only as "rough guides" for inference. For these reasons, we also report the CS distributional loss measures used in the construction of $Z_{j,P,S}(u)$, defined as:

$$CS = \int_U \frac{1}{P}\sum_{t=1}^{P}\left(1\{Y_t \leq u\} - \frac{1}{P}\sum_{n=1}^{P} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \leq u\}\right)^2 \phi(u)du,$$

where j denotes a particular model. This measure is useful for direct pointwise comparison of alternative models.

5 Empirical Results

The predictive densities are constructed for the three prediction periods: 1970:1-2004:4, 1982:4-2004:4 and 1990:1-2004:4. Thus, $R = \{24, 75, 104\}$, For each forecast period consider $\tau = \{4, 8, 12\}$, where τ corresponds to forecast horizons in quarters. All predictions (and prediction models) are formed (and estimated) both recursively and using models that were estimated only once, at the beginning of each prediction period. Fixed estimation window results are qualitatively the same, and are not reported, for the sake of brevity (complete results are available upon request). The three theoretical models discussed above as well as two time series models (a univariate autoregressive model (AR) model and a vector autoregressive (VAR) model are evaluated. In addition forecasts are constructed using a naive "no change" model.[12]

Predictive densities for theoretical models are constructed based on the DSGE solution outlined above. The structure of the DSGE models requires that we first remove the long run average of the series in order to make the data directly comparable to analogous data simulated using the DSGE models. Since forecasts are in real time, to compute the mean we use only observations from 1 to $R-(\tau-1)$. Second, we use historical observations for the output gap, inflation, nominal interest rate and money supply, and a steady state value for the unobservable processes as a starting point (i.e. for the sticky price model $x_t = (y_t^g, \pi_t, r_t, \triangle m_t, \triangle a^{ss})'$ is the starting point, where the steady state of technology growth, $\triangle a^{ss}$ is equal to zero). Third, we generate a technology shock, $\varepsilon_{a,t+1}$, and money supply shock, $\varepsilon_{m,t+1}$ with parameters specified in our calibration discussion above, where $z_{t+1} = (\varepsilon_{m,t+1}, \varepsilon_{a,t+1})'$.[13] Fourth, having shocks for the period $t+1$ and observations for period t we can compute forecast for the period $t+1$. Namely:

$$x_{t+1|t} = \Theta x_t + \Psi z_{t+1}, \quad R-(\tau-1) \leq t \leq T-\tau,$$

using the theoretical model solution. The τ-period ahead forecast of the theoretical models is thus obtained by iterating:

$$z_{t+k|t} = \Theta z_{t+k-1|t} + \Psi z_{t+k}, \quad R-(\tau-1) \leq t \leq T-\tau,$$

for $k = 2, 3, ..., \tau$. Finally we add back the mean to make the series directly comparable to our historical observations, where only that part of the historical distribution corresponding to the *ex ante* prediction period is used in subsequent evaluation of the models. We repeat the above procedure for each forecast period and each forecast horizon.

[12] For a comprehensive discussion of the uses and misuses of econometrics models, as well as a discussion of nonlinear models, structural breaks, forecast pooling, and a whole host of relevant econometric issues that are not discussed in this paper, the reader is referred to Clements and Hendry (2002,2003,2004), and the references cited therein. The issues raised in this series of papers are potentially of consequence, but are left to future research.

[13] When comparing mean square forecast errors, DSGE predictions are formed assuming that the shocks are equal to zero, corresponding to the standard practice in time series models of comparing conditional mean predictions, and thus making our time series and theoretical predictions directly comparable. Shocks, however, are added to the theoretical models when forming predictive densities. Results for the case where predictive densities are constructed for DSGE models without shocks have been calculated, and yield inferior results. Complete tabulated findings are available upon request.

Predictive densities for the AR(p) models are based on the following equation:

$$x_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + \epsilon_t, \tag{14}$$

which is estimated using maximum likelihood, with lags selected via the Schwarz information criterion (SIC). In the current context, we set $x_t = y_t^g$ for the output gap, and $x_t = \pi_t$ for inflation. Our VAR(p) model is:

$$x_t = c + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + ... + \Phi_p x_{t-p} + \epsilon_t, \tag{15}$$

where $x_t = (y_t^g, \pi_t, r_t, \triangle m_t)'$, Φ_j is (4×4) matrix of autoregressive coefficients for $i = 1, 2, ..., p$, c is (4×1) vector of constants and ϵ_t is a (4×1) vector of errors. Lags are again selected via the (matrix version of the) SIC, and estimation is via maximum likelihood. An alternative analysis was carried out with lags selected via the Akaike information criterion. However, predictions were inferior to those formed using the SIC.

Conditional τ-step ahead forecasts for the AR(p) model are obtained by iterating:

$$x_{t+k|t} = \mu + \hat{\phi}_1(x_{t+k-1|t} - \mu) + \hat{\phi}_2(x_{t+k-2|t} - \mu) + ... + \hat{\phi}_p(x_{t+k-p|t} - \mu), \quad R - (\tau - 1) \leq t \leq T - \tau,$$

for $k = 1, 2, ..., \tau$, where $\mu = c/(1 - \hat{\phi}_1 - \hat{\phi}_2 - ... - \hat{\phi}_p)$ and $x_{s|t} = x_s$ for $s \leq t$. Recursive estimation and prediction is carried out by re-estimating the model before each successive forecast in the *ex ante* period is formed. Conditional τ-step ahead forecasts for VAR(p) model are formed analogously.

Our empirical finding are gather in Tables 1-4 and Figure 1. A number of conclusions emerge upon inspection of the results.

Predictive Density Analysis

Note that Figure 1 plots predictive densities for the standard sticky price model. It is apparent that the dispersion of 1-year ahead predictions of the output gap is lower than that for the longer horizons. The same is not the case for inflation. In the density plots, note also that densities based upon our smaller out-of-sample periods are apparently further from their historical counterparts than when longer out-of-sample periods are used. This may be indicative of the need to use many predictions when forming densities in the manner done in this paper (i.e. we form densities across a long out-of-sample period rather than for individual observations, allowing for the application of the CS tests discussed later). As expected, predictive densities appear to "drift" further from the truth as the forecast horizon is increased from 1 to 3 years. Interestingly, output gap densities appear, in general, to be much closer to their historical counterparts than inflation densities. This may in part be due to the fact that inflation is more volatile from period to period than the output gap, as can be noted via examination of the last two rows of plots in Figure 1.

We now turn to a comparison of the three alternative sticky price models. Table 1 contains CS distributional loss measures (see above discussion), and Table 2 contains associated test statistics where the standard sticky price model is set equal to the benchmark, against which all other models (including the time series models) are compared. Notice in Table 1 that when y_t^g and π_t jointly evaluated (Panel A), the sticky price model yields lower loss in only 2 of 9 forecast horizon/predictive period cases, and none of these wins occur for the more recent predictive periods. Thus, we have some evidence that the alternative sticky price models are useful from a predictive standpoint, confirming earlier theoretical evidence in their support. However, it should also be noted that there appears to be little to choose between the indexation and information models. Interestingly, when attention is instead focused on distributions of y_t^g or on π_t (see Panels B and C of the table), then a different story emerges. Namely, the standard sticky price model dominates the other two theoretical models in 5 of 9 cases for the output gap. In opposition to this finding, the standard model "wins" only 1 of 9 times for inflation. Thus, the newer models do not appear to universally dominate the standard model. In light of the above results, it appears that target variable of interest is crucial to assessment of alternative models. If the models are to be used only to form predictive densities for the output gap, then using the standard model may be quite reasonable; the same cannot be said if interest focuses on inflation. This result is not surprising, and is somewhat analogous to the oft cited finding that the choice of loss function is crucial to model assessment.

When the time series models are added to the mix, a very interesting finding emerges. Namely, the simple time series models do not dominate the DSGE models. This is opposed to the usual finding that VAR models often outpredict DSGE models, for example. Furthermore, given that simple time series models are difficult to beat in head-on predictive horserace with more complicated nonlinear models, such as regime switching and threshold models, we have some evidence that there is little to choose between time series and simple theoretical models, in our context. For example, in Panel A of Table 1, notice that the theoretical models "win" in 7 of 9 cases ("wins" are denoted in bold). This sort of surprising result holds also in Panels B and C, where the output gap and inflation are individually examined.[14]

[14]In order to illustrate the varieties of models specified and estimated when forming our predictions based upon

Thus far, we have discussed only point loss measure estimates. Table 2 contains results analogous to those reported in Table 1, but presents instead the CS statistics and associated critical values discussed above. It is immediate upon inspection of this table that there is little to choose between the different models, with very few rejections of the null hypothesis that no model beats the standard sticky price model. Of course, it should be stressed that the use of larger samples and/or prediction periods may yield additional evidence that could be useful for understanding the whether there is any statistically significant difference between our models when used to form predictive densities.

Mean Square Forecast Error Analysis

The picture that emerges when MSFEs are examined (see Table 4) is qualitatively the same as when predictive densities are examined. In the table, MSFEs and Diebold and Mariano (DM: 1995) predictive accuracy test statistics are reported. Starred entries denote rejection of the null hypothesis that the DM test finds nothing to choose between the standard sticky price model and each of the alternative models listed across the first row of the table (see Clark and McCracken (2001) and McCracken (2004) for a detailed discussion of appropriate critical values for the DM test). Panel A reports results for the output gap, and Panel B contains results for inflation. Amongst the theoretical models, it again appears that the standard model is sub-optimal. Indeed, all rejections of the DM test favor the newer theoretical models. There is a surprising exception to this rule, however. Namely, for inflation, the standard model yields statistically significant advantage over the other theoretical models whenever the forecast horizon is 3 years, regardless of which *ex ante* period is being used. Additionally, the same finding holds for the output gap when the longest *ex ante* period is considered. This finding does not manifest itself when predictive densities are evaluated, and underscores the obvious differences between examining MSFEs versus examining entire predictive densities. In short, if interest focuses on conditional mean prediction, and square error loss is deemed the relevant loss measure, then the standard model is clearly preferable at our longest horizon of three years ahead. It remains to be seen whether this finding will hold up when

---

time series models, Table 3 gathers estimation results for typical AR and VAR models (the estimation period in all cases is the first estimation period used for a particular *ex ante* period). Of note is that the models generally fit quite well, even though issues of structural breaks are ignored.

Of further note is that the specification of more complex time series models that allow for structural breaks, for example, is left to future research.

even longer forecast horizons are examined.

Of final note is that the naive model rarely dominates (notice that the lowest MSFE is achieved by the naive model in only three cases, across both panels of Table 4).

In summary, the newer theoretical models appear to dominate the standard model in many cases. Notable exceptions include the case where output gap predictive densities are of interest, and where 3-year ahead MSFE predictive performance is the yardstick being used to evaluate models. Furthermore, the simple time series models examined here clearly do not dominate the DSGE models, even though the time series models are estimated recursively, and use the parsimonious SIC for lag selection.

6 Conclusion

In this paper we carry out a predictive analysis of various sticky price models using predictive density and mean square forecast error (MSFE) loss. We find that more recent information and indexation type sticky price models fare well and often dominate the standard sticky price model. An important exception to this finding is when MSFE loss is specified, in which case the standard model dominates the newer models at our longest forecast horizon of 3 years ahead. We also find that theoretical macroeconomic model based predictions are not dominated by simple econometric model based predictions, suggesting a role for economic theory in econometric forecasting. It remains to see whether these findings will hold up when longer forecast horizons are considered. It also remains to ascertain whether the finding in this paper that simple linear time series models offer little predictive advantage over DSGE models when predicting the output gap and inflation holds up under further scrutiny, where, for example, more complex nonlinear econometric models are specified, such as those discussed in Dahl and Hylleberg (2003).

7 References

Ball, L., (1994), "Credible Disinflation with Staggered Price Setting," *American Economic Review*, LXXXIV, 282-289.

Ball, L., Mankiw, N.,G., and R. Reis, (2005), "Monetary Policy for Inattentive Economies," *Journal of Monetary Economics* 52(4), 703-725.

Bernanke, B.S., and M. Gertler, (1995), "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *Journal of Economic Perspectives*, vol. 9(4), 27-48.

Bhardwaj, G., V. Corradi and N.R. Swanson, (2005), "A Simulation Based Specification Test for Diffusion Processes", Working Paper, Rutgers University.

Bierens, H.J., (2005), Econometric Analysis of Singular Dynamic Stochastic General Equilibrium Models with an Application to the King-Plosser-Rebelo Stochastic Growth Model, *Journal of Econometrics*, forthcoming.

Bierens, H.J., and N.R. Swanson, (2000), The Econometric Consequences of the Ceteris Paribus Condition in Theoretical Economics, *Journal of Econometrics*, 95, 223-253.

Bils, M., and P.J. Klenow, (2004),"Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy* 112, 947-985.

Blinder, A.,S., Canetti E.,R., Lebow D., and J.,B.,Rudd, (1998), Asking About Prices: A New Approach to Understand Price Stickiness, New York: Russell Sage Foundation.

Calvo, G. A., (1983), "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, XII, 383-398.

Chang, Y.S., J.F. Gomes, and F. Schorfheide, (2002), Learning-by-Doing as a Propagation Mechanism, *American Economic Review*, 92, 1498-1520.

Chari, V.V., Kehoe P.,J., and E.R. McGrattan, (1996), "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?," *Econometrica*, 68(5), 1151-1180.

Christiano, L., Eichenbaum, M., and Evans, C. (2000), "Monetary Policy Shocks: what have we Learned and to what End?", in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Amsterdam, The Netherlands: Elsevier.

Christiano, L.J., Eichenbaum, M. and C.L. Evans (2001), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *NBER Working Paper* No. 8403.

Clark, T.E., and M.W., McCracken, (2001), "Tests of Equal Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics* 105, 85-110.

Clements, M.P. and D.F. Hendry, (1993), "On the Limitations of Comparing Mean Squared Forecast Errors, *Journal of Forecasting* 12, 617–637.

Clements, M.P. and D.F. Hendry, (2002), "Modelling Methodology and Forecast Failure, *The Econometrics Journal* 5, 319-344.

Clements, M.P. and D.F. Hendry, (2003), "Economic Forecasting: Some Lessons from Recent Research," *Economic Modeling* 20, 301-329.

Clements, M.P. and J. Smith, (2000), "Evaluating the Forecast Densities of Linear and Non-linear Models: Applications to Output Growth and Unemployment," *Journal of Forecasting* 19, 255-276.

Clements, M.P. and J. Smith, (2002), Evaluating multivariate forecast densities: A comparison of two approaches, *International Journal of Forecasting* 18, 397-407.

Cogley, T., and J.M. Nason, (1995), Output Dynamics for Real Business Cycles Models, *American Economic Review,* 85, 492-511.

Cooley, Thomas F. and Gary D. Hansen (1989), "The Inflation Tax in a Real Business Cycle Model," *American Economic Review*, 79, 733-48.

Corradi, V. and N.R. Swanson, (2005a), "Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data," *Journal of Econometrics*, forthcoming.

Corradi, V. and N.R. Swanson, (2005b), "Predictive Density Evaluation", in: Handbook of Economic Forecasting, eds. Clive W.J. Granger, Graham Elliot and Allan Timmerman, Elsevier, Amsterdam, forthcoming.

Corradi, V. and N.R. Swanson, (2005c), "Predictive Density and Conditional Confidence Interval Accuracy Tests," *Journal of Econometrics*, forthcoming.

Corradi, V. and N.R. Swanson, (2005d), "Simulation Based Predictive Density Estimation and Testing for Diffusion Processes," Working Paper, Rutgers University.

Corradi, V. and N.R. Swanson, (2005e), "Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes," Working Paper, Rutgers University.

Dahl, C. and S. Hylleberg, (2003), "Flexible Regression Models and Relative Forecast Performance," Working Paper, Purdue University.

DeJong, D., Ingram, B., and D. Whiteman, (2000). "A Bayesian Approach to Dy- namic Macroeconomics," *Journal of Econometrics*, 98(2), 203-223.

Del Negro, M. and F. Schorfheide, (2005), "Monetary Policy Analysis with Potentially Misspecified Models," ECB Working Paper No. 475.

Diebold, F.X., and R.S. Mariano, (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X., L.E. Ohanian, and J. Berkowitz, (1998), Dynamic Equilibrium Economies: A Framework for Comparing Models and Data, *Review of Economic Studies,* 65, 433-451.

Fernandez-Villaverde, J., and J.F. Rubio-Ramirez, (2004), "Comparing Dynamic Equilibrium Models to Data: a Bayesian Approach," *Journal of Econometrics*, 123, 153 – 187.

Fuhrer, J., and G. Moore, (1995), "Inflation Persistence," *Quarterly Journal of Economics*, CX, 127-160.

Gali., J. and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 44, 195-222.

Gali, J., (2002), "New Perspectives on Monetary Policy, Inflation, and the Business Cycle," *NBER Working Paper* No. 8767.

Gali, J., Lopez-Salido, J.,D., and J. Valles, (2003), "Technology shocks and monetary policy: assessing the Fed's performance," *Journal of Monetary Economics*, 50, 723-743.

Hendry, D.F. and M.P. Clements, (2004), "Pooling of Forecasts," *Econometrics Journal* 7, 1-31.

Khan, H. and Z. Zhu., (2002), "Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom," Bank of Canada Working Paper No. 2002-19.

Korenok, O., (2004), "Empirical Comparison of Sticky Price and Sticky Information Models," Working Paper, Rutgers University.

Korenok, O. and N.R. Swanson (2005), "How Sticky Is Sticky Enough? A Distributional and Impulse Response Analysis of New Keynesian DSGE Models," Working Paper, Rutgers University.

Lucas, Robert E., Jr., (1973) "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review* 63, 326-334.

Mankiw, N.,G., and R. Reis, (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, Vol. CXVII(IV), 1295-1328.

McCracken, M.W., (2004), Asymptotics for Out of Sample Tests of Causality, Working Paper, University of Missouri-Columbia.

Phelps, E. S., (1970), Introduction: The New Microeconomics in Employment and Inflation Theory, in E.S. Phelps et al., Microeconomic Foundations of Employment and Inflation Theory, New York: Norton.

Rotemberg, J.J., and M. Woodford (1996), "Real Business Cycle Models and the Forecastable Movements in Output, Hours and Consumption," *American Economic Review* 86, 71-89.

Sbordone, A., (2002), "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics*, vol. 49 (2), 265-292.

Schmitt-Grohe, S., (2000), Endogenous Business Cycles and the Dynamics of Output, Hours and Consumption, *American Economic Review,* 90, 1136-1159.

Schorfheide, F., (2000), Loss Function Based Evaluation of DSGE Models, *Journal of Applied Econometrics,* 15, 645-670.

Sims, C. A., (2003), "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), April 2003.

Smets, F. and R. Wouters, (2003), "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123-1175.

Walsh, C.,E., (1998), Monetary Theory and Policy, MIT Press, chapters 2-4.

Watson, M.W., (1993), Measure of Fit for Calibrated Models, *Journal of Political Economy,* 101, 1011-1041.

West, K., (1996), Asymptotic Inference About Predictive Ability, *Econometrica*, 64, 1067-1084.

Woodford, M., (2002), Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press, Princeton University.

Woodford, M., (2003), Imperfect Common Knowledge and the Effects of Monetary Policy, Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, Edited by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, *Princeton University Press*, 25-59.

Yun, T., (1996), "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles," *Journal of Monetary Economics* 37, 345–370.

Table 1: Corradi and Swanson Distributional Loss Measures for Various Models

Panel A: Joint Distribution of the Output Gap and Inflation							
Fcst.Period, τ	*sp*	*spi*	*si*	*AR,R*	*AR,NR*	*VAR,R*	*VAR,NR*
1970:1,4	1.0736	1.0989	**1.0680**	1.1063	1.9858	1.0940	1.7342
1970:1,8	**1.1123**	1.1336	1.1649	1.2098	2.3933	1.1781	2.5163
1970:1,12	**1.1450**	1.2027	1.1675	1.2240	2.4500	1.2504	2.8099
1982:4,4	0.7134	0.6880	0.7212	**0.6847**	0.7848	0.7395	0.8941
1982:4,8	0.8471	**0.8147**	0.9138	0.8786	1.2249	0.9103	1.1737
1982:4,12	0.9275	0.8923	**0.8836**	1.0036	1.4299	1.0242	1.2834
1990:1,4	0.5441	0.5163	0.5123	**0.5023**	0.5552	0.6001	0.8367
1990:1,8	0.7087	**0.6351**	0.6655	0.6678	0.7617	0.8925	1.2643
1990:1,12	0.7364	**0.6992**	0.7121	0.7885	0.9596	1.0102	1.3239
Panel B: Distribution of the Output Gap							
Fcst.Period, τ	*sp*	*spi*	*si*	*AR,R*	*AR,NR*	*VAR,R*	*VAR,NR*
1970:1,4	1.0784	1.0746	1.0950	**1.0745**	2.6712	1.0937	2.1516
1970:1,8	**1.1357**	1.1979	1.2398	1.1762	3.2760	1.1395	3.7948
1970:1,12	1.2067	1.2898	1.2009	**1.1404**	3.3715	1.1753	4.2559
1982:4,4	0.6469	0.6865	0.6530	**0.6411**	0.6510	0.7656	0.9337
1982:4,8	**0.6759**	0.7126	0.7882	0.7216	1.0478	0.8518	0.8097
1982:4,12	**0.7387**	0.7427	0.7415	0.7977	1.2537	0.8162	1.0992
1990:1,4	0.4274	**0.4177**	0.4433	0.4409	0.4593	0.6545	0.9824
1990:1,8	**0.4723**	0.4749	0.5207	0.5397	0.5578	0.9487	1.3717
1990:1,12	0.5918	0.5590	**0.5307**	0.6260	0.6616	0.9577	1.2753
Panel C: Distribution of Inflation							
Fcst.Period, τ	*sp*	*spi*	*si*	*AR,R*	*AR,NR*	*VAR,R*	*VAR,NR*
1970:1,4	1.1068	1.1196	1.0890	1.1817	1.1849	**1.0779**	1.2834
1970:1,8	1.1076	1.1022	**1.0959**	1.2849	1.5253	1.1522	1.1644
1970:1,12	**1.1026**	1.1094	1.1252	1.3313	1.8553	1.2818	1.3472
1982:4,4	0.6164	0.4952	0.6299	0.5379	0.7656	**0.4562**	0.5126
1982:4,8	0.8536	**0.6305**	0.8354	0.8817	1.2680	0.6553	1.1116
1982:4,12	0.9433	0.8275	**0.8171**	1.0840	1.5173	0.9981	1.1575
1990:1,4	0.5631	0.5234	0.4675	0.4351	0.5294	**0.3170**	0.3183
1990:1,8	0.8625	0.6143	0.6869	0.6753	0.8782	**0.5200**	0.8399
1990:1,12	0.6943	**0.6480**	0.7882	0.8447	1.2152	0.8511	1.2739

Notes: Entries in the table are Corradi and Swanson (CS: 2005a) distributional loss measures (i.e. an estimate of $E\left(\left(F_j(u;\theta_j^\dagger) - F_0(u;\theta_0)\right)^2\right)$. Namely, we report $CS = \int_U \frac{1}{P}\sum_{t=1}^{P}\left(1\{Y_t \leq u\} - \frac{1}{P}\sum_{n=1}^{P} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \leq u\}\right)^2 \phi(u)du$, where j denotes a particular model (see Section 4 for complete details). The CS distributional loss measure is calculated for the sticky price (sp), sticky price with indexation (spi) and sticky information (si) theoretical models, as well as for AR and VAR models. For the latter two models, predictions and predictive densities are constructed using two different estimation and prediction construction methods; namely recursive (R) and non-recursive (NR). Hence the notation AR,R; AR,NR; VAR,R; and VAR,NR. Boldface entries indicate the lowest CS measure for a particular model.

Table 2: Distributional Accuracy Tests Based on the Comparison of Actual and Predicted π_t and y_t^g

Panel A: Tests Based on the Joint Distribution of the Output Gap and Inflation									
	1970:1			*1982:4*			*1990:1*		
		Critical Value			*Critical Value.*			*Critical Value.*	
τ, l	Z	5%	10%	Z	5%	10%	Z	5%	10%
4,l1	0.0056	0.2295	0.2087	0.0287	0.1586	0.1396	0.0418	0.2760	0.2383
4,l2	0.0056	0.3042	0.2622	0.0287	0.1786	0.1532	0.0418	0.2571	0.2223
4,l3	0.0056	0.3681	0.2669	0.0287	0.1720	0.1624	0.0418	0.3079	0.2509
4,l4	0.0056	0.3709	0.3220	0.0287	0.1809	0.1575	0.0418	0.3297	0.2678
4,l5	0.0056	0.3045	0.2829	0.0287	0.1968	0.1656	0.0418	0.3026	0.2556
8,l1	-0.0213	0.3501	0.2806	0.0323	0.2223	0.1882	0.0736	0.1916	0.1723
8,l2	-0.0213	0.4654	0.3638	0.0323	0.2615	0.2015	0.0736	0.1933	0.1583
8,l3	-0.0213	0.3726	0.3152	0.0323	0.2333	0.2100	0.0736	0.1815	0.1576
8,l4	-0.0213	0.3895	0.3675	0.0323	0.2445	0.2104	0.0736	0.1733	0.1613
8,l5	-0.0213	0.3595	0.2997	0.0323	0.2408	0.1937	0.0736	0.1670	0.1079
12,l1	-0.0224	0.3665	0.3339	0.0439	0.2179	0.1904	0.0371	0.2118	0.1473
12,l2	-0.0224	0.4223	0.4058	0.0439	0.2616	0.2022	0.0371	0.2494	0.1709
12,l3	-0.0224	0.4250	0.3606	0.0439	0.2764	0.2458	0.0371	0.2325	0.1846
12,l4	-0.0224	0.5739	0.5039	0.0439	0.3010	0.2342	0.0371	0.2749	0.2495
12,l5	-0.0224	0.5088	0.4475	0.0439	0.2441	0.2292	0.0371	0.2553	0.1919
Panel B: Tests Based on the Distribution of the Output Gap									
	1970:1			*1982:4*			*1990:1*		
		Critical Value.			*Critical Value.*			*Critical Value.*	
τ, l	Z	5%	10%	Z	5%	10%	Z	5%	10%
4,l1	0.0039	0.3891	0.3258	0.0058	0.1813	0.1697	0.0096	0.2494	0.2321
4,l2	0.0039	0.4768	0.3687	0.0058	0.2036	0.1679	0.0096	0.2562	0.2167
4,l3	0.0039	0.4870	0.3793	0.0058	0.2324	0.2000	0.0096	0.2845	0.2346
4,l4	0.0039	0.5226	0.4511	0.0058	0.2312	0.1843	0.0096	0.3957	0.2928
4,l5	0.0039	0.3989	0.3500	0.0058	0.2024	0.1791	0.0096	0.3336	0.2535
8,l1	-0.0037	0.6340	0.4304	-0.0368	0.2530	0.2155	-0.0026	0.3151	0.2742
8,l2	-0.0037	0.6947	0.5334	-0.0368	0.2800	0.2082	-0.0026	0.3854	0.3099
8,l3	-0.0037	0.6250	0.4839	-0.0368	0.2669	0.2471	-0.0026	0.3675	0.2934
8,l4	-0.0037	0.6045	0.5472	-0.0368	0.2953	0.2427	-0.0026	0.3478	0.3056
8,l5	-0.0037	0.4803	0.4549	-0.0368	0.2935	0.2403	-0.0026	0.2756	0.2221
12,l1	0.0663	0.5392	0.4821	-0.0028	0.2513	0.2358	0.0611	0.1616	0.1390
12,l2	0.0663	0.6096	0.5376	-0.0028	0.3330	0.2888	0.0611	0.2539	0.1956
12,l3	0.0663	0.5814	0.5048	-0.0028	0.3555	0.3253	0.0611	0.2222	0.1696
12,l4	0.0663	0.6905	0.5877	-0.0028	0.3670	0.3286	0.0611	0.3138	0.2372
12,l5	0.0663	0.5791	0.5126	-0.0028	0.3741	0.3068	0.0611	0.2492	0.1894
Panel C: Tests Based on the Distribution of Inflation									
	1970:1			*1982:4*			*1990:1*		
		Critical Value.			*Critical Value.*			*Critical Value.*	
τ, l	Z	5%	10%	Z	5%	10%	Z	5%	10%
4,l1	0.0289	0.1987	0.1731	0.1602	0.3544	0.2744	0.2461	0.3581	0.3241
4,l2	0.0289	0.2589	0.2371	0.1602	0.3333	0.2971	0.2461	0.3787	0.3043
4,l3	0.0289	0.3160	0.2523	0.1602	0.3434	0.3053	0.2461	0.4213	0.3714
4,l4	0.0289	0.4085	0.3112	0.1602	0.3524	0.3208	0.2461	0.3805	0.3125
4,l5	0.0289	0.4123	0.3156	0.1602	0.3779	0.3089	0.2461	0.3723	0.3474
8,l1	0.0118	0.3320	0.2916	0.2231*	0.2615	0.2216	0.3425**	0.1710	0.1333
8,l2	0.0118	0.4041	0.2948	0.2231	0.3046	0.2634	0.3425**	0.1781	0.1208
8,l3	0.0118	0.4680	0.3147	0.2231	0.3267	0.2890	0.3425**	0.1784	0.1190
8,l4	0.0118	0.5453	0.4450	0.2231	0.3989	0.3032	0.3425**	0.1005	0.0724
8,l5	0.0118	0.5630	0.4591	0.2231	0.4294	0.2899	0.3425**	0.1355	0.1218
12,l1	-0.0068	0.3769	0.3147	0.1262	0.2272	0.1905	0.0463	0.4400	0.3369
12,l2	-0.0068	0.4858	0.4239	0.1262	0.2761	0.2212	0.0463	0.4114	0.3570
12,l3	-0.0068	0.5517	0.4425	0.1262	0.3241	0.2589	0.0463	0.4174	0.3528
12,l4	-0.0068	0.7061	0.5930	0.1262	0.2496	0.2246	0.0463	0.3903	0.3491
12,l5	-0.0068	0.6667	0.5797	0.1262	0.2407	0.2281	0.0463	0.4188	0.3612

Notes: The benchmark is the sticky price model, which is tested against all other models. For each forecast period, there are three columns of entries: the 1st column reports the numerical values of the test statistic; the next two columns report 5% and 10% bootstrap critical values based on bootstrap statistics constructed allowing for parameter estimation error. * indicates rejection of null hypothesis at a 10% and ** at a 5% significance level. We compare three τ step ahead forecasts, where we choose $\tau = 4$, 8,or 12. The block length used in the bootstrap is set as follows: $l1 = 5$, $l2 = 8$, $l3 = 10$, $l4 = 16$, and $l5 = 20$. All statistics are constructed using grids of 20x20 values for u, distributed uniformly across the historical data ranges of π_t and y_t^g. Bootstrap empirical distributions are constructed using 100 bootstrap replications. See Sections 4 and 5 for complete details.

Table 3: Estimation Summary Statistics for Selected AR and VAR(p) Models

Panel A: AR(p) models

	$x_t = y_t^g$			$x_t = \pi_t$		
	1964:1-1969:1	1964:1-1981:4	1964:1-1989:1	1964:1-1969:1	1964:1-1981:4	1964:1-1989:1
c	0.0104**	0.0001	-0.0001	0.0037*	0.0025**	0.0021**
	(-2.2923)	(0.0516)	(-0.1187)	(2.1449)	(2.2413)	(2.3274)
x_{t-1}	1.0002**	1.1339**	1.2227**	0.6093**	0.8082**	0.8180**
	(3.9946)	(9.1465)	(12.539)	(3.0481)	(6.3919)	(8.1346)
x_{t-2}	−0.3635*	−0.2502**	−0.3149**		-0.134882	−0.2169*
	(-1.7818)	(-2.0093)	(-3.2239)		(-0.8665)	(-1.8118)
x_{t-3}					0.5421**	0.5299**
					(3.3887)	(4.4245)
x_{t-4}					−0.3625**	−0.2749**
					(-2.7921)	(-2.7639)
R^2	0.6114	0.8074	0.8778	0.3989	0.7738	0.7418
Adj. R^2	0.5516	0.8014	0.8752	0.3560	0.7593	0.7305
Log Likl.	59.763	211.74	313.15	73.909	275.84	391.72
DW	1.7674	2.0114	2.1168	1.9119	1.8722	2.0040
AIC	-7.0954	-6.2312	-6.4615	-8.9887	-8.0848	-8.0567
SIC	-6.9506	-6.1325	-6.3813	-8.8921	-7.9202	-7.9232

Table 3 (cont): Estimation Summary Statistics for Selected AR and VAR(p) Models

Panel B: VAR(p) models												
	1964:1-1969:1				1964:1-1981:4				1964:1-1989:1			
	y_t^g	π_t	r_t	$\triangle m_t$	y_t^g	π_t	r_t	$\triangle m_t$	y_t^g	π_t	r_t	$\triangle m_t$
c	0.0099	-0.0102^{**}	0.0012	0.0066	0.0039	-0.0031	-0.0020	0.0163^{**}	0.0096^{**}	-0.0017	-0.0031	0.0148^{**}
	(0.5115)	(-2.4149)	(0.5323)	(0.6132)	(0.6896)	(-1.3048)	(-0.7233)	(4.6632)	(2.2014)	(-0.8905)	(-1.4776)	(3.8478)
y_{t-1}^g	0.7384^{**}	0.0888^{**}	0.0020	-0.1315	0.9304^{**}	0.0650^{**}	0.0658^{**}	-0.1605^{**}	0.9547^{**}	-0.0039	0.1290^{**}	-0.0732
	(-3.9127)	(2.1537)	(0.0899)	(-1.2535)	(15.318)	(2.5615)	(2.2149)	(-4.2753)	(9.6603)	(-0.0869)	(2.7128)	(-0.8435)
y_{t-2}^g									-0.0943	0.0500	-0.0633	-0.0380
									(-0.9694)	(1.1452)	(-1.3511)	(-0.4444)
π_{t-1}	0.3414	0.4433^{**}	0.2594^{**}	-0.6303	-0.3319	0.6583^{**}	0.1355	0.0343	-0.0930	0.6744^{**}	-0.3337^{**}	-0.1135
	(0.4028)	(2.3950)	(2.5750)	(-1.3377)	(-1.4574)	(6.9167)	(1.2166)	(0.2440)	(-0.4337)	(7.0104)	(-3.2340)	(-0.6026)
π_{t-2}									0.0093	0.1255	-0.4332^{**}	0.3222^{*}
									(0.0447)	(1.3535)	(4.3550)	(1.7737)
r_{t-1}	-0.2461	0.4561	0.6734^{**}	0.9938	-0.1778	0.2742^{**}	0.8131^{**}	-0.1347	0.2732	0.6317^{**}	0.6838^{**}	-0.4168^{**}
	(-0.1426)	(1.2105)	(3.2829)	(1.0360)	(-0.8168)	(3.0135)	(7.6346)	(-1.0011)	(1.2951)	(6.6764)	(6.7377)	(-2.2492)
r_{t-2}									-0.8271^{**}	-0.5659^{**}	0.2420^{**}	0.1907
									(-3.5275)	(-5.3812)	(2.1457)	(0.9261)
$\triangle m_{t-1}$	-0.1168	0.4061^{**}	0.0586	0.4114	0.2188	0.1588^{*}	0.1555^{*}	0.3599^{**}	-0.0335	0.2078^{**}	0.0047	0.2842^{*}
	(-0.2610)	(4.1555)	(1.1009)	(1.6535)	(1.2315)	(2.1384)	(1.7894)	(3.2777)	(-0.2629)	(3.6407)	(0.0769)	(2.5416)
$\triangle m_{t-2}$									0.1432	-0.0430	0.1643^{**}	0.0517
									(1.0872)	(-0.7273)	(2.5945)	(0.4473)
R^2	0.6846	0.8034	0.8631	0.3838	0.8405	0.7783	0.7527	0.5158	0.9158	0.8161	0.8003	0.3994
Adj. R^2	0.5944	0.7472	0.8240	0.2078	0.8302	0.7640	0.7367	0.4846	0.9080	0.7992	0.7820	0.3442
DW	1.4558	2.0933	1.7403	1.7283	2.0185	1.9680	2.3498	1.9872	1.9850	2.1200	2.0453	2.0080
Log Likl.	346.02				1010.0				1486.3			
AIC	-34.318				-29.552				-30.214			
SIC	-33.324				-28.894				-29.252			

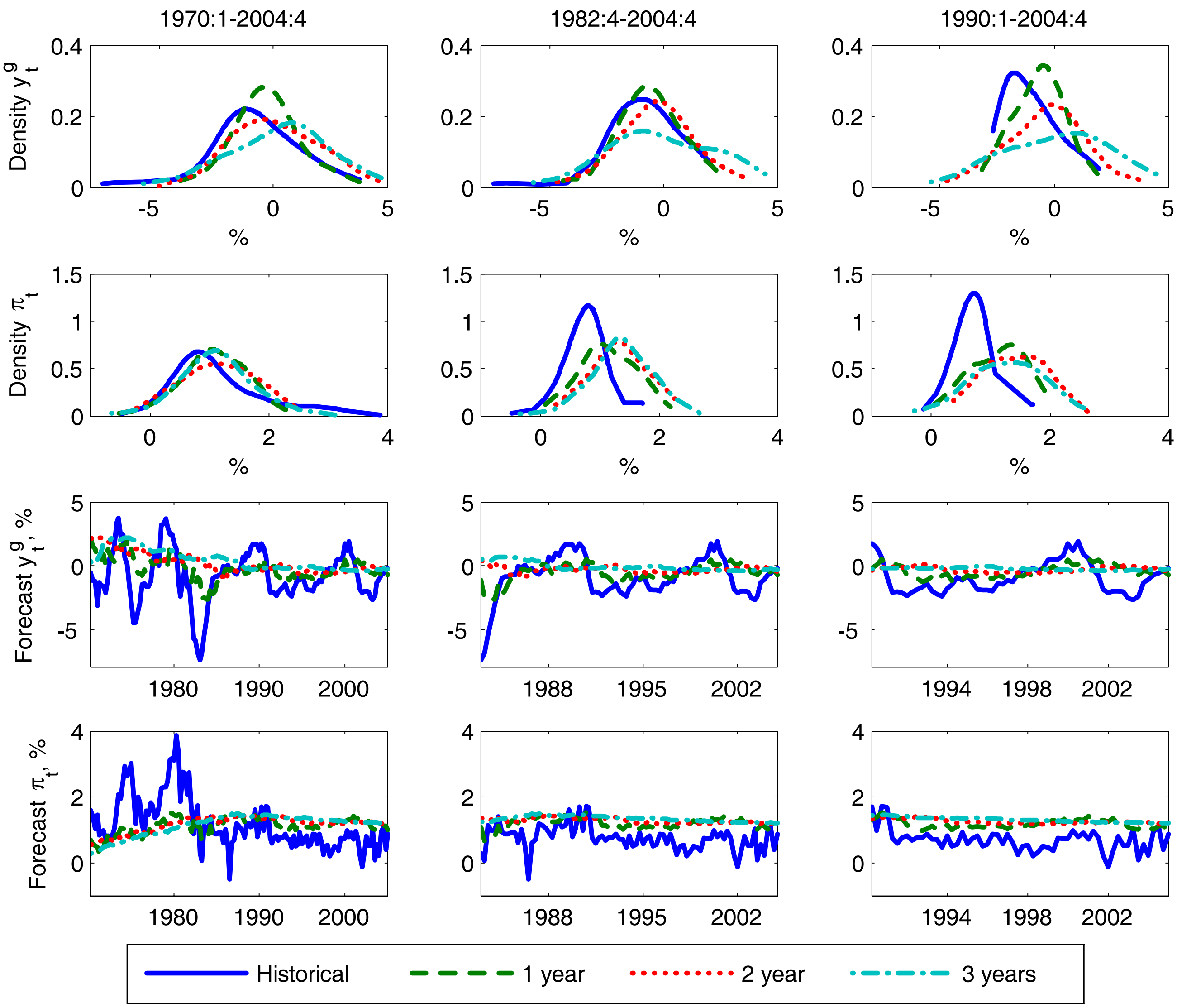
Notes: The first row specifies in each panel of the table states the estimation sample size for each regression, and the second row of specifies the endogenous variable in the regression. First column of the table specifies the exogenous variables used in the regressions, as well as regression diagnostics where we use following notation: R^2 for R-squared; Adj. R^2 to denote adjusted R-squared; Log Likl. for log likelihood; DW for the Durbin-Watson statistic; AIC for the Akaike information criterion; and SIC for the Schwarz information criterion. Columns from two to six contain estimates of regression coefficients with t-statistic in parentheses, * indicates rejection of null hypothesis associated with a standard t-test at 10% and ** and 5% significance levels.

Table 4: Mean Square Forecast Errors for Various Models

Fcst. Period	τ	*sp*	*spi*	*si*	*AR,R*	*AR,NR*	*VAR,R*	*VAR,NR*	*Naive*
				Panel *A: Output Gap*					
1970:1	4	0.0357	0.0322	0.0424**	0.0422**	0.1218**	**0.0317**	0.1166**	0.0473**
DM			-0.8083	4.4794	2.4926	10.9531	-0.7596	7.7678	2.5308
1970:1	8	0.0617	.0534**	0.0636**	0.0729**	0.1599**	0.0563	0.3822**	0.1027**
DM			-3.8321	2.0709	3.2168	12.7517	-0.5117	8.4801	4.4708
1970:1	12	**0.0609**	0.0660**	0.0617*	0.0974**	0.1953**	0.1065*	0.2994**	0.126**
DM			3.8714	1.9042	4.2189	13.7508	1.7093	14.0641	5.7827
1982:4	4	0.0200	0.0230	0.0208	0.0221	**0.0179**	0.0227	0.0322**	0.0254
DM			1.0506	1.1867	0.5059	-0.7557	0.7513	3.1006	0.9182
1982:4	8	0.0367	0.0371	0.0395**	0.0396	0.0428**	**0.0359**	0.0746**	0.0629**
DM			0.2487	2.2959	1.1747	2.9171	-0.0826	2.9531	2.6272
1982:4	12	0.0397	0.0426**	**0.0393**	0.0439**	0.0544**	0.0439	0.0695**	0.0843**
DM			2.6717	-1.4498	2.0405	4.9313	0.4887	2.2628	3.8381
1990:1	4	0.0140	**0.0128**	0.0133*	0.0129	0.0130	0.0251**	0.0550**	0.0179
DM			-0.4503	-1.6457	-0.6359	-0.5726	4.7504	6.9762	1.2321
1990:1	8	0.0210	.0185*	0.0207	0.0215	0.0206	0.0538**	0.1065**	0.0367**
DM			-1.9215	-0.5335	0.2727	-0.2801	5.0880	7.3993	3.0334
1990:1	12	0.0202	0.0223**	**0.0201**	0.0234**	0.0212	0.0583**	0.0842**	0.0479**
DM			2.3149	-0.4305	2.3315	0.9324	5.0360	7.3419	5.2353
				Panel *B: Inflation*					
1970:1	4	0.0651	0.0595*	0.0517**	0.0437**	0.0452**	0.0510	0.1116**	.0431**
DM			-1.8729	-4.3657	-3.6129	-5.4637	-1.4436	2.3029	-2.7485
1970:1	8	0.0734	0.0789*	.0710**	0.0767	0.0820	0.1223	0.0712	0.0795
DM			1.6423	-3.3187	0.4212	1.3495	1.3694	-0.3944	0.5498
1970:1	12	0.0893	0.0953**	0.0894	0.0854	0.1024*	0.4037	.0750**	0.0914
DM			4.4845	0.0217	-0.4389	1.7208	1.0972	-3.1820	0.1477
1982:4	4	0.0335	0.0283**	0.0322	.0261**	0.0367	0.0263**	0.0292	0.0266
DM			-2.0838	-0.7349	-2.1117	1.0240	-1.9635	-1.1660	-1.2187
1982:4	8	0.0400	.0245**	0.0420**	0.0435	0.0654**	0.0485	0.0898**	0.0437
DM			-7.6745	4.1755	0.7679	5.8424	0.8970	4.2343	0.3526
1982:4	12	**0.0459**	0.0494**	0.0481**	0.0627	0.0874**	0.0610	0.0644**	0.0677
DM			2.9649	5.2596	1.4847	5.8120	1.0055	3.3275	1.1183
1990:1	4	0.0345	0.0251**	0.0306**	0.0204**	0.0246**	0.0173**	.0161**	0.0195**
DM			-6.4908	-2.7655	-3.6930	-3.0479	-4.9228	-5.3171	-2.8851
1990:1	8	0.0411	0.0267**	0.0414	0.0295**	0.0395	0.0239**	0.0386	.0202**
DM			-9.5679	1.1816	-5.3215	-1.1083	-6.6900	-1.0043	-4.3158
1990:1	12	0.0476	0.0501**	0.0490**	0.0385**	0.0653**	0.0386**	0.0776**	.0193**
DM			2.6164	4.4862	-5.8662	10.7507	-3.5972	8.6139	-4.9834

Notes: Forecast accuracy results are presented in this table for models of the output gap (Panel A) and inflation (Panel B). Numerical entries are mean square forecast errors (x100 in Panel A; x1000 in Panel B) and Diebold and Mariano (DM: 1995) statistics for the sticky price (sp), sticky price with indexation (spi) and sticky information (si) theoretical models; as well as for AR and VAR, and naive no change models based on recursive (R) and non-recursive (NR) estimation and forecast construction schemes. Boldface entries indicate the minimum mean square error among alternative models. The first column of entries reports the start date of the forecast period, and the second column gives the forecast horizon. The null hypothesis of the DM test is that the sticky price model and each alternative model perform equally well based on their mean square forecast errors. * indicates rejection of null hypothesis at 10% and ** denotes rejection at 5% significance level. See Section 5 for complete details.

Figure 1: Predictive Summary Plots for Standard Sticky Price Model



Notes: Predictive densities, rows one and two, and forecasts, rows three and four, are computed using the theoretical sticky price model for three forecast periods: 1970:1, 1982:4 and 1990:1 in the first, second and third columns, respectively. X-axes on each density graph are growth rates in percent, Y-axes are empirical probability density functions, f(x). Y-axes on each forecast graph are growth rates in percent. Each graph reports actual observations - blue solid line; one year ahead forecasts - green, dashed line; two year ahead forecasts - red, dotted line; and 3 year ahead forecasts - light blue, dash-dotted line.