BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working papers)

Comparing forecast accuracy: a Monte Carlo investigation

by Fabio Busetti, Juri Marcucci and Giovanni Veronese

Number 723

September 2009

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working papers)

Comparing forecast accuracy: a Monte Carlo investigation

by Fabio Busetti, Juri Marcucci and Giovanni Veronese

Number 723 - September 2009

COMPARING FORECAST ACCURACY:
A MONTE CARLO INVESTIGATION

by Fabio Busetti*, Juri Marcucci* and Giovanni Veronese*

Abstract

The size and power properties of several tests of equal Mean Square Prediction Error (MSPE) and of Forecast Encompassing (FE) are evaluated, using Monte Carlo simulations, in the context of dynamic regressions. For nested models, the F-type test of forecast encompassing proposed by Clark and McCracken (2001) displays overall the best properties. However its power advantage tends to become smaller as the prediction sample increases and for multi-step ahead predictions; in these cases a standard FE test based on Gaussian critical values becomes relatively more attractive. The ranking among the tests remains broadly unaltered for one-step and multi-step ahead predictions, for partially misspecified models and for highly persistent data. A similar setup is then used to analyze the case of non-nested models. Again it is found that FE tests have a significantly better performance than tests of equal MSPE for discriminating between correct and misspecified models. An empirical application evaluates the predictive ability of nested and non-nested models for GDP in Italy and the euro-area.

JEL Classification: C12, C52, C53.
Keywords: forecast encompassing, model evaluation, nested models, non-nested models, equal predictive ability.

Contents

---

* Bank of Italy, Economics, Research and International Relations.

1 Introduction[1]

Evaluating the out-of-sample performance of competing models is an important aspect of economic forecasting and model selection. Diebold and Mariano (1995) have proposed a simple test for the null hypothesis of equal predictive accuracy measured in terms of a general loss function. In most applications, however, little attention is paid to the shape of the loss function and models are generally compared on the basis of their mean square prediction error (MSPE). An alternative approach looks at the out-of-sample correlation between prediction errors, which leads to tests of forecast encompassing (FE) or, in the terminology of Granger and Newbold (1986), of conditional forecast efficiency. A preferred forecast is said to encompass some competing alternative if the latter contains no additional useful information for prediction; see, *inter alia*, Chong and Hendry (1986), Clements and Hendry (1993), Harvey, Leybourne and Newbold (1998).

The recent literature on out-of-sample prediction has highlighted two important issues that may render invalid the standard large sample inference *à la* Diebold and Mariano (1995). First, West (1996) has showed that parameter estimation error may not be asymptotically irrelevant and may therefore affect the limiting distribution of the test statistics. Second, if models are nested, the statistics based on average comparisons of prediction errors have a degenerate limiting variance under the null hypothesis and they are not asymptotically normally distributed. For nested models McCracken (2007) and Clark and McCracken (2001) derive the appropriate non Gaussian limit for tests of, respectively, equal MSPE and FE; the critical values are tabulated across two nuisance parameters (the ratio of the magnitudes of prediction sample to estimation sample and the number of additional regressors in the larger model) and they are, in general, valid only for one-step ahead predictions. The test of forecast encompassing for nested models proposed by Chao, Corradi and Swanson (2001) does not suffer from this degeneracy: its limiting distribution is a chi-square under the null hypothesis. A different approach is taken by Giacomini and White (2006) that focus on comparing

[1]We wish to thank Joerg Breitung, Graham Elliott, Raffaella Giacomini, Lutz Kilian, Michael McCracken and Ken West for useful suggestions on an earlier draft. We also thank seminar participants at the Federal Reserve Bank of Boston, the Bank of Italy workshop on "Factor models, high-frequency data and short-term forecasting" the I IFO-INSEE-ISAE Macroeconomic Forecasting Conference and the XVII Annual Symposium of the Society for Nonlinear Dynamics and Econometrics for their comments. The views expressed are those of the authors and do not necessarily reflect those of the Bank of Italy. Corresponding author: Juri Marcucci. Emails: fabio.busetti@bancaditalia.it (Fabio Busetti), juri@sssup.it (Juri Marcucci), giovanni.veronese@bancaditalia.it (Giovanni Veronese).

forecasting methods as opposed to forecasting models: their test statistic of equal conditional predictive ability has a chi-square null distribution as the prediction sample size tends to infinite for a finite length of the estimation sample. A comprehensive survey on asymptotic inference for predictive ability for nested and non-nested models is West (2006).

In this paper we evaluate the properties of several tests of equal MSPE and tests of FE, with the goal to provide practical guidance to forecasters needing to choose among a set of predictions from (a small number of) competing models[2]. We use Monte Carlo simulation methods to compute empirical size and empirical power functions in the context of dynamic regression models. One-step and multi-step ahead predictions are considered for correctly specified and misspecified regressions; we also investigate the properties of the tests across different values of the ratio between prediction and estimation sample sizes and for various degrees of persistence of the data generating process.

The tests under scrutiny are the followings: (i) the standard Diebold-Mariano test of equal MSPE; (ii) the $MSE-t$ and (iii) the $MSE-F$ modifications of McCracken (2007) for nested models; (iv) the forecast encompassing test of Harvey, Leybourne and Newbold (1998); (v) the $ENC-t$ and (vi) $ENC-F$ modifications of Clark and McCracken (2001) for nested models; (vii) the forecast encompassing test of Chao, Corradi and Swanson (2001) for nested models.[3]

Our results extend previous analyses (mostly concerned with the size properties of the tests) by providing empirical power functions in a variety of settings, including misspecification of the regression models and high persistence in the data generating process. For nested models we generally confirm the findings of Clark and McCracken (2001, 2005a) that the ENC-F test has, overall, the best properties, noticing however that its power advantage tends to become smaller when the dimension of the prediction sample is larger and for multi-step ahead prediction. We find that, the relative ranking among the different tests changes according to whether the number of out-of-sample observations is "small" or "large". In many situations, a standard FE test, based on a null Gaussian distribution, becomes relatively attractive, as for

[2]For issues arising on comparing a large number of models, see White (2000), Hansen (2005), Hubrich and West (2007).

[3]In the comparison we do not include the method of Giacomini and White (2006) because it relates to a different null hypothesis from the other tests. In their framework, under the null hypothesis the true model is the larger one, implying a bias-variance tradeoff in the forecasts. In our experiments for the nested case, the null hypothesis is the smaller, or restricted, model and there is no bias. See the discussion of Giacomini and White (2006, p.1559-1561)

example it does not require bootstrapping the critical values for multi-step ahead forecasts. A similar simulation set-up is then used to analyze non-nested models. Again we find a significant advantage for the tests of FE over those of equal MSPE in discriminating between a correct and a mis-specified model.

In summary, the paper proceeds as follows. Section 2 briefly reviews the test statistics under scrutiny. Section 3 and 4 contain the simulation results for nested and non-nested models respectively. In section 5 we provide an application that evaluates the prediction ability of nested and non-nested models for GDP in Italy and the Euro area. Section 6 concludes.

2 The setup and the tests under scrutiny

We consider a sample of T observations on a target series y_t and two k_i-dimensional vectors of (non mutually exclusive) predictors $X_{it},\ i = 1, 2$. The sample is divided into R in-sample and P out-of-sample observations, with $T = R + P$.

We want to compare two sets of h -step ahead forecasts, $h \geq 1$, generated by the linear models

$$\widehat{y}_{it} = X'_{i,t-h}\widehat{\beta}_{i,t-h}, \qquad t = R + h, R + h + 1, ..., T \tag{1}$$

where $\widehat{\beta}_{i,t-h}$ is the least square estimate for model i constructed using observations up to time $t - h$ and the predictors $X_{i,t-h}$ may include lags of the dependent variable y_{t-j} for $j \geq h$. The models are estimated under the recursive or the rolling scheme: the recursive least square estimates are constructed using observations indexed from 1 to $t - h$, while the rolling coefficients are estimated using the R observations indexed from $t - R - h + 1$ to $t - h$.

The forecasting performance of the models is evaluated using the two sets of h-step ahead forecast errors $e_{it} = y_t - \widehat{y}_{it}$, $i = 1, 2$, for $t = R + h, R + h + 1, ..., R + P$; for simplicity we have suppressed the dependency on h in the notation. The tests under scrutiny are briefly detailed below.

2.1 Tests of equal MSPE

The test of equal mean square prediction error of Diebold and Mariano (1995) is based on the following t-type statistic

$$DM = \widehat{P}^{\frac{1}{2}}\overline{d} \ / \ \widehat{\sigma}_{DM}(m), \tag{2}$$

where $\overline{d} = \widehat{P}^{-1} \sum_{t=R+h}^{T} d_t$, $d_t = e_{1t}^2 - e_{2t}^2$, $\widehat{P} = P - h + 1$, and $\widehat{\sigma}_{DM}^2\,(m)$ is the non-parametric estimator of the long run variance of d_t

$$\widehat{\sigma}_{DM}^2(m) = \widehat{P}^{-1} \sum_{t=R+h}^{T} \left(d_t - \overline{d}\right)^2 + 2\widehat{P}^{-1} \sum_{j=1}^{m} w(j,m) \sum_{t=j+R+h}^{T} \left(d_t - \overline{d}\right)\left(d_{t-j} - \overline{d}\right), \tag{3}$$

where $w(j,m)$ is a weight function truncated at $m << T$; e.g. $w(j,m) = 1 - j/(m+1)$ as in Newey and West (1987); note that, in large samples, P can replace $\widehat{P}$ in the definition of (2). The DM statistic tests the null hypothesis of equal forecast accuracy $H_0 : E\, d_t^* = 0$, where d_t^* is the population version of d_t, net of parameter estimation error. If the models are *non nested*, the limiting null distribution of (2) is a standard Gaussian. By contrast, if the models are nested the denominator converges to zero under the null and the limiting distribution of the DM statistic is non-Gaussian.[4]

McCracken (2007) obtains the correct null limiting distribution of the DM statistic for the case of one-step ahead forecasts between *nested* models: the test, based on McCracken critical values, will be called $MSE - t$. The following F-type statistic is also proposed

$$MSE - F = \widehat{P}\overline{d}\,/\,\widehat{\sigma}_2^2 \tag{4}$$

where $\widehat{\sigma}_2^2 = \widehat{P}^{-1} \sum_{t=R+h}^{T} e_{2t}^2$ is the estimate of the second moment of the forecast errors of the nesting model. The distributions of $MSE - t$ and $MSE - F$ depend on the ratio P/R and on the number $k_2 - k_1$ of excess parameters in the nesting model; critical values are tabulated for recursive and rolling one-step ahead forecasts. The limiting distributions change for the case of multi-step ahead predictions, but critical values can be obtained by bootstrap; see Clark and McCracken (2005a).

For the case of *nested* models the standard DM test turns out to be heavily undersized and with low power. Although the correct limiting distribution is non-Gaussian, Clark and West (2006, 2007) argue that most of the bias can be corrected by a simple adjustment in the statistic: this leads to a test with Gaussian critical values that has size close to, but a little less than, the nominal one. Specifically, the Clark-West adjusted statistic is

$$DM_{AD} = \widehat{P}^{\frac{1}{2}}\overline{d}_{AD}\,/\,\widehat{\sigma}_{AD}(m), \tag{5}$$

where $\overline{d}_{AD} = \widehat{P}^{-1} \sum_{t=R+h}^{T} d_{AD,\,t}$, $d_{AD,\,t} = e_{1t}^2 - e_{2t}^2 + (\widehat{y}_{1t} - \widehat{y}_{2t})^2$ and $\widehat{\sigma}_{AD}^2\,(m)$ is the non-parametric estimator of the long run variance of $d_{AD,\,t}$, that par-

[4] It is however argued that the Gaussian critical values would still approximately hold if P/R is small (e.g. less than 0.1, see West, 2006).

allels the definition in (3). Since $(\widehat{y}_{1t}-\widehat{y}_{2t})^2 = (e_{2t}-e_{1t})^2$ one can write $d_{AD,\ t} = 2e_{1t}\,(e_{1t}-e_{2t})$. Thus, as noted in West (2006), the DM_{AD} statistic is based on the covariance between e_{1t} and $e_{1t}-e_{2t}$, and it corresponds to the test of forecast encompassing given in (6) below.

2.2 Tests of forecast encompassing

It is said that the forecast $\widehat{y}_{1t}$ encompasses $\widehat{y}_{2t}$ if there is no gain from combining them into a composite forecast $\widehat{y}_{ct} = (1-\lambda)\,\widehat{y}_{1t} + \lambda\widehat{y}_{2t}$, for some weight $\lambda > 0$; see *inter alia* Chong and Hendry (1986), Granger and Newbold (1986), Clements and Hendry (1993) and the early empirical work of Nelson (1972). As the combined forecast error e_{ct} satisfies the relation $e_{1t} = \lambda\,(e_{1t}-e_{2t}) + e_{ct}$, Ericsson (1992) tests the null hypothesis of forecast encompassing, $H_0 : \lambda = 0$, by a t-test on λ in the regression of e_{1t} on $e_{1t}-e_{2t}$. In a similar way, Harvey, Leybourne and Newbold (1998) write the null hypothesis of forecast encompassing as $H_0 : Ef_t^* = 0$, where f_t^* is the population version of $f_t = e_{1t}\,(e_{1t}-e_{2t})$, and they construct a t-test on $\overline{f} = \widehat{P}^{-1}\sum_{t=R+h}^{T} f_t$; precisely their statistic is

$$HLN = \widehat{P}^{\frac{1}{2}}\overline{f}\ /\ \widehat{\sigma}_{HLN}(m), \tag{6}$$

where $\widehat{\sigma}^2_{HLN}(m)$ is a non-parametric estimator of the long run variance of f_t, that parallels the definition in (3). If the models are *non nested*, the limiting null distribution of the HLN statistic is a standard Gaussian.

Clark and McCracken (2001) show that when applied to nested models the HLN statistic is no longer asymptotically Gaussian and they obtain the correct null limiting distribution for one-step ahead forecasts: the test that uses their critical value will be called $ENC-t$. They also propose the F-type statistic

$$ENC-F = \widehat{P}\overline{f}\ /\widehat{\sigma}_2^2, \tag{7}$$

where $\widehat{\sigma}_2^2$ is the mean squared forecast error of the nesting model as in (4). The distributions of $ENC-t$ and $ENC-F$ depend on the ratio P/R and on the number k_2-k_1 of excess parameters in the nesting model; critical values are tabulated for recursive and rolling one-step ahead forecasts. The extension to multi-step ahead forecasts is given in Clark and McCracken (2005a).

A different test of forecast encompassing for nested models has been proposed by Chao, Corradi and Swanson (2001): the null hypothesis is $H_0 : Ec_t^* = 0,$ where c_t^* is the population version of $c_t = e_{1t}\left(Z_{2t}-\overline{Z}_2\right)$

and Z_{2t} are the additional $k_2 - k_1$ predictors in X_{2t} not included in X_{1t}. [5] This is again a Wald-type test with statistic given by

$$CCS = \widehat{P}\overline{c}' \left(\widehat{\Sigma}_{CCS}(m)\right)^{-1} \overline{c}, \tag{8}$$

where $\overline{c} = \widehat{P}^{-1} \sum_{t=R+h}^{T} c_t$ and $\widehat{\Sigma}_{CCS}(m)$ is a non-parametric estimator of the long run variance-covariance matrix of c_t, that parallels the definition in (3). Under the null hypothesis of forecast encompassing CCS is asymptotically distributed as chi square with $k_2 - k_1$ degrees of freedom.[6]

3 Monte Carlo evaluation for nested models

To evaluate the properties of the tests for nested models we start by considering the following VAR(1) data generating process (for $t = 1, 2, \ldots, T$)

$$y_t = \mu_y + \phi_y y_{t-1} + c x_{t-1} + \varepsilon_t, \tag{9}$$

$$x_t = \mu_x + \phi_x x_{t-1} + u_t, \tag{10}$$

with Gaussian i.i.d. innovations

$$\begin{pmatrix} \varepsilon_t \\ u_t \end{pmatrix} \sim NIID\left(0, \begin{pmatrix} 1 & \rho_{\varepsilon u} q \\ \rho_{\varepsilon u} q & q^2 \end{pmatrix}\right). \tag{11}$$

Note that, if $c \neq 0$, y_t can be represented as a Gaussian ARMA(2,1) process with degree of persistence, as measured by the sum of the autoregressive roots, equal to $\phi_x + \phi_y - \phi_x \phi_y$. If $c = 0$ then y_t is not Granger-caused by x_t.

The object is to forecast y_t by a dynamic univariate regression. We compare two sets of out-of-sample forecasts: the first one is obtained by an autoregression of order 1 (the restricted model), the other by including additional predictors (the unrestricted or nesting model). The case $c = 0$ measures the size of the tests of equal MSPE and FE, while $c \neq 0$ provides the power. All tests are one-sided, in the sense that the alternative hypothesis is that the nesting model yields better forecasts.[7] Given that the null

[5] While in the original formulation of Chao et al. (2001) the regressors Z_{2t} are not demeaned in the expression for c_t, we find better properties of the test after demeaning.

[6] Chao et al. (2001) also propose a version of the test that takes into account estimation uncertainty, with $\widehat{\Sigma}_{CCS}(m)$ replaced by a more complicated expression which depends on the sampling scheme. However they also argue that the modified test does not provide a clear advantage in terms of size and it turns out to be less powerful.

[7] For the HLN test, it can be shown that if x_{t-1} has predictive power for y_t then the covariance between e_{1t} and $e_{1t} - e_{2t}$ is positive. Thus the test is one-sided in the right tail; see Clark and McCracken (2005a, p.376).

hypothesis is $c = 0$, the tests can also be interpreted as out-of-sample tests of Granger causality.

We consider sample sizes of $T = R+P$ where $R = (100, 200)$ and $P = \pi R$ with $\pi = (.1, .25, .5, 1)$. The properties of the tests clearly depend on the number of out-of-sample observations P, with power expected to increase with π (for given R). Since a constant term will always be included in the set of predictors, without loss of generality we set $\mu_y = \mu_x = 0$ in the data generating process (9)-(11).

In the first subsection below we evaluate the properties of the tests under the case of one-step ahead forecasts and correct specification, in the sense that the estimated unrestricted model is the same as the true data generating process. The second subsection investigates the impact on the properties of the tests of different degrees of persistence in the data. The third subsection studies the effect of mis-specification and overparameterization, while the fourth one considers the case of multi-step ahead predictions.

3.1 The nesting model is correctly specified

The restricted model is the regression of y_t on $X_{1t} = (1, y_{t-1})'$; in the unrestricted model the predictors are given by $X_{2t} = (1, y_{t-1}, x_{t-1})'$. Since there is no additional temporal dependence to be taken into account, we calculate the statistics (2), (5), (6), and (8) for $m = 0$, i.e. with scaling provided by the sample variance instead of the long-run variance.

Table 1 provides the empirical sizes of the tests ($c = 0$) run at 5% and 10% level of significance for $R = (100, 200)$ for the case of recursive forecasts (the results for rolling regressions are nearly identical, and therefore are not presented). We present figures where the values of the parameters in the data generating process are set to $\phi_y = \phi_x = 0.8$, $q = 1$, $\rho_{\varepsilon u} = 0$. The size of the tests does not change in any significant way if different values of these parameters are considered.

Consider first the case $R = 100$ with tests run at 10% level of significance. For $\pi = 0.1$ (10 out-of-sample observations) all tests except DM are oversized, in particular MSE-t (0.17), ENC-t (0.16) and CCS (0.15). As π increases size improves for all tests except DM and HLN; however while DM is deeply undersized for $\pi \geq 0.5$, the rejection frequencies for HLN do not fall below 7% consistently with the arguments of Clark and West (2006, 2007) for the (equivalent) adjusted DM statistic. Doubling the sample ($R = 200$) yields more reliable sizes for all tests except, to some extent, DM and HLN. Qualitatively similar arguments apply for the tests run at 5%.

Figure 1 shows the empirical power functions (with respect to the parameter c governing the distance from the null hypothesis) of tests[8] run at 10% significance level for $R = 200$; results for $R = 100$ and for tests run at 5% significance are qualitatively similar and therefore will be not discussed. Power is affected by the parameter q that controls the variance of x_t (for given c, the higher q the more powerful the tests) and, to a lesser extent, by the value of the correlation $\rho_{\varepsilon u}$; however, as the relative ranking among the tests turns out to be unaffected by the values of q and $\rho_{\varepsilon u}$, to save space we only present results for $q = 1$ and $\rho_{\varepsilon u} = 0$. The four panels of figure 1 refer to different magnitudes of the prediction sample, $\pi = (0.1, 0.25, 0.5, 1)$; clearly, for fixed R, the larger π the more power.

With no doubt, for all values of π the $ENC - F$ test of Clark and McCracken (2001) turns out to be the most powerful. The second ranked test depends on the value of π, i.e. on the length of the prediction sample relative to the estimation sample. If the prediction sample is short ($\pi = 0.1$) then the $MSE - F$ test is preferable, otherwise $ENC - t$ is better. For large π the HLN test, that uses Gaussian critical values, behaves not so differently from $MSE - F$, and it is quite more powerful than $MSE - t$. The DM test has by far the lowest power, while the CCS test (that uses χ^2 critical values) has relatively good power only for large π, but it is always dominated by HLN.

Larger differences in the behavior of the tests occur when the number of out-of-sample observations is small. In particular, when $\pi = 0.1$ the better sized tests are DM, HLN and $MSE - F$, but only the latter has high rejection rates under the alternative hypothesis (being second only to $ENC - F$). For higher π tests tend to behave more similarly: while $ENC - F$ clearly dominates, the HLN test may become attractive being based on Gaussian critical values. Detailed simulation results are available upon request.

To sum up, this sub-section extends the findings of Clark and McCracken (2001) by providing empirical power functions of the tests (while they only reported two specific values of the alternative) and by including HLN and CCS in the comparison. We confirm that, overall, the ENC-F test has better properties, noticing however that its power advantage tends to become smaller as the dimension of the prediction sample increases. In addition, we find that the relative ranking among the different tests changes according to whether the number of out-of-sample observations is "small" or "large".

Overall we would rank the test according to the following order: ENC-$F \succ MSE$-F, ENC-$t \succ HLN \succ MSE$-t, $CCS \succ DM$.

---

[8] The empirical power functions refer again to the recursive case. The results for rolling regressions as very similar and therefore they are not reported.

3.2 The impact of different degrees of persistence

While the limiting power of the tests should not be affected by the degree of autocorrelation of x_t (as long as it remains weakly dependent), one can expect non-negligible finite sample effects as the series gets closer to the region of non-stationarity. In terms of our simulated data generating process, increasing the persistence parameter ϕ_x yields higher variance of the regressor x_t and thus higher rejection rates for all tests, for given q and $c > 0$. However, the finite sample effects of varying ϕ_x can be studied simply by holding constant the variance of x_t, $\sigma_x^2 = q^2/(1 - \phi_x^2)$, by changing the parameter q correspondingly.

In Figure 2 we compare the results reported in the previous section (where $\phi_x = .8$) with cases of lower and higher persistence, namely $\phi'_x = .5, .95, .99$, where q is selected such that $1/(1 - .8^2) = q^2/(1 - \phi'_x)$. For brevity we only report results for the statistics $MSE - F$, $ENC - F$, HLN, CCS (one for each quadrant), with $\pi = .5$ and $R = 200$. While it is interesting to notice the absence of size distorsions, in all cases power gets reduced as ϕ_x tends to 1; e.g. when $\phi'_x = .99$ for $ENC - F$ the rejection rates at $c = .10$ are about 20% lower than in the baseline case of $\phi = .8$. Overall, the CCS test appears to be mostly affected, losing most of its power when there is a near-unit root. Except for CCS, the relative ranking among the tests remains broadly the same as in the previous section.[9]

Finally, unreported experiments show that increasing only the coefficient, ϕ_y, attached to the lagged dependent variable has no effect on the properties of the tests (even for $\phi_y = 1$), which remain nearly identical to those described in the previous section.

3.3 Mis-specification of the nesting model

We consider three cases where the nesting model is somehow different from the true data generating process, so to understand to what extent the tests of equal MSPE and FE are still effective and whether some of them are more robust to misspecification.

(1) *Error-in-variables.* In the unrestricted model we take as predictors $(1, y_{t-1}, w_{t-1})'$ instead of $(1, y_{t-1}, x_{t-1})'$, where

$$w_t = x_t + u_{w,t}, \qquad u_{w,t} \sim NIID\left(0,\ q_w^2 \sigma_x^2\right), \tag{12}$$

so that w_t and x_t are positively correlated with coefficient $\rho_{xw} = 1/\left(1 + q_w^2\right)$. Figure 3 reports the empirical power functions for a correlation parameter

[9] Clearly, for a given distance c from the null hypothesis, the power loss gets diminished as R gets larger for each test; detailed results are available upon request.

$\rho_{xw} = 0.5$ (i.e. $q_w = 1$) and $R = 200$. All tests undergo some reduction of power with respect to the case of correct specification[10], but interestingly the relative ranking among them remains the same. Clearly, high values of ρ_{xw} would generate small losses of power from misspecification.

(2) *Autoregression.* We take as unrestricted model an autoregression of order p, where $2 \leq p \leq 8$ is chosen according to the BIC method. As the true data generating process is an $ARMA(2,1)$, it is plausible that an autoregressive model provides a reasonable approximation. Here, the power loss from misspecification turns out to be very relevant; for example, if $c = 0.50$ and $\pi = 0.25$, the ENC-F and CCS tests reject, respectively, 56% and 22% of the times, against 100% and 98% for the case of correct specification. The empirical power functions of DM, MSE-F, ENC-F, HLN, CCS are depicted in Figure 4, for the case of $R = 200$. The power loss is most extreme for the CCS test, probably in connection with the fact that now several nuisance parameters are embedded in the statistic (now distributed as χ^2_{p-1} under the null, instead of χ^2_1). Except for CCS (which no longer outperforms DM), the relative power among the other tests is broadly unaltered.

(3) *Over-parameterization.* In the unrestricted model we take as predictors $(1, y_{t-1}, x_{t-1}, w_{t-1})'$ instead of $(1, y_{t-1}, x_{t-1})'$, where w_t is given by (12). Again the relative ranking is mostly unaffected. However, it turns out that, while ENC-F is still the most powerful test, for HLN the power loss from over-parameterization is rather small; in fact now HLN becomes more attractive than MSE-F for $\pi \geq 0.50$. Detailed results for this case are available upon request.

3.4 Multi-step ahead forecasts

Clark and McCracken (2005a) argue that for multi-step ahead predictions the critical values of the $ENC-t$, $MSE-t$, $ENC-F$ and $MSE-F$ tests should be obtained by bootstrap or simulation methods, as the limiting approximation generally depends on several nuisance parameters, which makes it infeasible to tabulate. However, in the case of a single additional regressor in the unrestricted model (as in the simulation experiment of this section), the asymptotic critical values for $ENC-t$ and $MSE-t$ coincide with those tabulated for the case of one-step ahead forecasts.

Here we consider multi-step ahead predictions for correctly specified models as in section 3.1. For the $ENC-F$ and $MSE-F$ tests we provide results using bootstrap critical values; for $ENC-t$, $MSE-t$ and CCS we consider

[10] When $\rho_{xw} = .5$, $\pi = .25$ and $R = 200$ the power loss for all tests is roughly about 30-35% with respect to the case of correct specification.

both asymptotic and bootstrap critical values; the DM and HLN tests are computed as usual. The statistics have been calculated setting $m = 1.5h$ in the long run variance estimator (3). The bootstrap algorithm is that in Kilian (1998), as also implemented by Clark and McCracken (2005a). We denote the bootstrap version of the tests by adding a * to the original name, e.g. $MSE - F^*$.

Table 2 provides the empirical size of the tests, run at 10% significance level, for the cases of $h = 2$ and 4 step ahead predictions, $R = (100, 200)$, and recursive regressions. For $\pi = .10$ the tests not based on bootstrap critical values are grossly oversized, with huge distorsions affecting CCS, MSE-t and ENC-t; the bootstrap allows to control size in all cases.[11] When $R = 200$ and $\pi \geq .25$ the HLN test has reasonably good size properties both for 2-step and for 4-step ahead projections, while DM displays a strong tendency toward under-rejection.

Figure 5 provides the empirical power functions of the DM, MSE-t^*, $MSE-F^*, HLN, CCS^*, ENC - t^*, ENC - F^*$ tests for $\pi = (.25, .50)$, $h = (2, 4)$, $R = 200$. The bootstrap version of the F-type test by Clark and McCracken (2001), $ENC - F^*$, displays the highest power, as for one-step ahead predictions. However the HLN test now seems to perform better: it is broadly equivalent to $MSE - F$ and $ENC - t$ for $h = 2$ and it does even better for $h = 4$.

Overall we would rank the test according to the following order: ENC-$F \succ HLN$, MSE-F, ENC-$t \succ MSE$-t, $CCS \succ DM$. Given the computational burden of bootstrapping $ENC - F$, the use of HLN, based on Gaussian critical values, is an interesting simple way for comparing forecast accuracy in multi-step ahead predictions.

4 Monte Carlo evaluation for non-nested models

The same data generating process of section 3 is used to evaluate the properties of the tests of equal mean square prediction error and of forecast encompassing for non-nested models. In particular, we consider the VAR(1) process

[11] Contrary to the results of Clark and McCracken (2005a) we find that size distortions tend to vanish as the number of out-of-sample observations P increases. One difference, however, is that in their simulation experiment regressors are chosen according to information criteria: in finite samples, this may be an important source of additional noise and mis-specification of the restricted model.

(9), and (10) and the error-in-variables model (12). Let M_x denote the linear regression model of y_t on $(1, y_{t-1}, x_t)'$ and M_w that of y_t on $(1, y_{t-1}, w_t)'$. Then, if c and q_w are different from zero the two models M_x and M_w are non-nested. Of course, if $q_w = 0$ ($|\rho_{xw}| = 1$) the two models and forecast errors are identical, while if $c = 0$ the two regressions M_x and M_w are essentially the same.

The Monte Carlo simulations in this section aims at measuring the ability of the tests of MSPE and of FE towards rejecting the mis-specified model M_w in favor of M_x. The important parameter is the correlation $\rho_{xw} = 1/\left(1+q_w^2\right)$ between the regressors x_t and w_t. If ρ_{xw} is nearly one (q_w small) the two models produce very similar forecasts; on the other hand, the smaller ρ_{xw} the better the tests are likely to discriminate between the models.

We consider the following tests: (i) the DM test of equal MSPE, both one and two-sided, where the one-sided alternative corresponds to the hypothesis that M_x provides better predictions than M_w; (ii) the HLN test of FE, both one and two-sided[12], where the null hypothesis is that M_w encompasses M_x; (iii) one-sided DM tests comparing each model M_w and M_x with a combined forecast (with equal weights) that use predictions from both regressions, denoted as $DM\text{-}FCw$ and $DM\text{-}FCx$ respectively.

As in section 3.1, we present results where the parameters of the data generating process are set as follows, $\phi_y = \phi_x = 0.8$, $q = 1$, $\rho_{\varepsilon u} = 0$ (and, without loss of generality, we set $\mu_y = \mu_x = 0$). For each ρ_{xw}, the rejection probabilities of the tests depend on the magnitude of c; in the Monte Carlo simulations below (with $R = 200$) we have set $c = 0.2$, but qualitatively similar results would hold for other values of this parameter.

Figure 6 shows the rejection frequencies of the tests (against the correlation parameter ρ_{xw} taking decreasing values from 0.999 to 0) for $R = 200$ in-sample and $P = \pi R$ out-of-sample observations, where $\pi = (0.1, 0.25, 0.50, 1)$, for the case of one-step ahead predictions and tests run at 10% significance level. For ρ_{xw} near 1 all tests display rejection frequencies close to the nominal size, while the probability of rejection increases as ρ_{xw} becomes smaller. Clearly, the one-sided tests DM_1 and HLN_1 are more powerful than the corresponding two-sided version, DM_2 and HLN_2. Again, as for the case of nested models, the encompassing tests are significantly more powerful than the tests of equal MSPE: for example, for $\pi = .25$ and $\rho_{xw} = 0.5$ the simulated rejection probability of HLN_1 is 74%, against 37% of DM_1. While the ranking $HLN_1 \succ HLN_2 \succ DM_1 \succ DM_2$ applies for all π, it is interesting to observe that the good performance of $DM - FCw$ that dominates DM_1

[12] The two-sided HLN test is equivalent to the often used t-test of FE based on the regression of y_t on the forecasts from the two alternative models.

(and in some cases it also rejects more frequently than HLN_2). On the other hand, as expected, $DM - FCx$ is not able to reject the correct model M_x in favour of the forecast combination. The behaviour of $DM - FCw$ is an interesting result that is investigated analytically in a separate paper (Busetti and Marcucci, 2009).

Finally Figure 7 shows the corresponding results for two and four steps ahead forecasts, $h = (2, 4)$, and $\pi = (0.25, 0.50)$, where the statistics have been corrected for serial correlation by estimating long run variances with bandwidth parameter $m = 1.5h$, as done in the previous section. The ranking between the tests remains unchanged. Note that, as expected, rejection frequencies tend to decrease when we move from $h = 2$ to $h = 4$, while between $h = 1$ and $h = 2$ the difference in the ability of detecting the correct model is rather small.

5 Empirical applications: forecasting real GDP

In this section we apply the tests of equal MSPE and FE with empirical examples of nested and non-nested forecasting models of real GDP for Italy and the Euro area.

5.1 A *bridge model* versus nested competitors

We investigate the forecast accuracy of the so-called *bridge models*, that use monthly indicators to forecast quarterly GDP in Italy; see e.g. Parigi and Golinelli (2007). Here, for simplicity, we first aggregate these indicators to quarterly frequency, assuming that at each step all information is available for the quarter under prediction.

In practice, we augment a simple autoregressive model by a set of (timely) indicators. The models that we estimate have the general specification

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2' x_t + error,$$

where y_t is the growth rate in the Italian GDP, and x_t is one or more regressors chosen from the following set: industrial production growth (IP), an error correction term between industrial production and GDP (ECM), growth in net exports (NX), growth in car registrations (CAR), business climate in the construction sector (CC). A fully fledged bridge model contain all of the above indicators as well as the lagged dependent variable. The prediction sample runs from 1999Q1 to 2007Q4, while the estimation sample recursively expands one quarter at each step, starting from 1987Q1 until 1998Q4.

In the first comparison (Panel A of Table 3) the benchmark is the random walk forecast for the growth rate of GDP, while the nesting models are: an AR(1) model, an AR(1) augmented with one indicator at a time, the fully fledged bridge model. We find that while the forecasting performance of the random walk, evaluated in terms of RMSE, is indeed worse than any of the other models, the significance of this result is not necessarily supported by all the tests; in particular DM fails to reject the null hypothesis of equal predictive accuracy for the AR(1) model and when the autoregression is augmented by any single indicator, excluding the industrial production. We then compare the forecast accuracy of each regression against the bridge model (Panel B of Table 3). The finding is that the forecast accuracy of the bridge is superior to all alternatives. The failure to reject the null hypothesis for DM is in line with our previous simulation results that showed the lower power of this test.

5.2 *Eurocoin* versus non-nested competitors

Here we compare the real-time forecasting performance of Eurocoin, the real-time GDP monthly indicator for the Euro area, with that of univariate AR and ARMA models.[13] Eurocoin is derived from a large sized factor model (see Altissimo et al., 2009) to obtain each month an estimate of the medium-long run GDP growth in the euro area. We consider the indicator releases at the end of each quarter (i.e. March, June, September and December), and take the indicator figures as a forecast of GDP growth in that quarter. As Eurocoin aims to capture medium-long run fluctuations in GDP growth, we compare its value with the quarter-on-quarter growth rate as well as with the year-on-year rate one quarter ahead. The prediction sample runs from 1999Q1 to 2005Q4, while the estimation sample recursively expands one quarter at each step, starting from 1987Q1 until 1998Q4. As for the competitor models, we limit our analysis to simple univariate AR and ARMA models of quarterly GDP growth, which are not nested by Eurocoin. The AR and ARMA specifications are obtained at each step using the BIC in-sample information criteria.[14]

The RMSE of the ARMA and AR models are in general higher than the ones obtained using Eurocoin, both when the target is the quarter-on-quarter growth rate and more so when the target is the year-on-year rate. Panel C of Table 3 shows that HLN provides strong support in favour of Eurocoin

[13]Eurocoin is produced by the Bank of Italy and published monthly by Bank of Italy and the CEPR.

[14]This approach is often used when comparing factor models to simple benchmarks. See for example Stock and Watson (2002)

for both quarter-on-quarter and year-on-year euro area GDP growth, while according to DM Eurocoin has superior forecast ability only for quarter-on-quarter growth. Again, these results are consistent with our Monte Carlo simulations.

6 Concluding remarks

The performance of several tests for comparing out-of-sample forecasts between competing models has been evaluated. Overall, the tests of forecast encompassing seem preferable to those of equal mean square prediction error. In particular, for nested models the best properties are displayed by the $ENC-F$ test of Clark and McCracken (2001); its power advantage however tends to become smaller for longer prediction samples and for multi-step ahead forecasts. In these cases, a standard forecast encompassing test, based on Gaussian critical values, becomes relatively attractive. Moreover, as the issue surrounding the standard FE and equal MSPE tests in nested comparisons is that of undersizing and low power, it is clear that if these tests already reject the null hypothesis then the use of $ENC-F$ may become superfluous.

The simulation results presented, however, do not account either for structural breaks or for model uncertainty. In fact, Clark and McCracken (2005b, 2009) show that breaks significantly affect the properties of tests of predictive ability and thus they may render harder the task of discriminating between competing models, and they also argue that the choice of estimation window in rolling regressions becomes crucial given the bias variance tradeoff induced by parameter instability. As real world forecasts can never be generated by the underlying "true model", we believe that taking into account mis-specification and model uncertainty is an important direction for research.

References

[1] Altissimo, F., Cristadoro, R., Forni, M., Lippi, M. and Veronese, G., (2009), "Eurocoin: tracking economic growth in real time", *Review of Economics and Statistics*, forthcoming.

[2] Busetti, F. and Marcucci, J., (2009), "MSE comparison of forecast combinations", mimeo.

[3] Chao, J., Corradi V., and Swanson N., (2001), "An out of sample test for Granger causality", *Macroeconomic Dynamics*, 5, 598-620.

[4] Chong, Y. Y., and Hendry D.F., (1986), "Econometric evaluation of linear macroeconomic models", *Review of Economic Studies*, 53, 671-690.

[5] Clark, T.E., and McCracken, M.W., (2001), "Tests of equal forecast accuracy and encompassing for nested models", *Journal of Econometrics*, 105, 85-110.

[6] Clark, T.E., and McCracken, M.W., (2005a), "Evaluating direct multistep forecasts", *Econometric Reviews*, 24, 369-404.

[7] Clark, T.E., and McCracken, M.W., (2005b), "The power of tests of predictive ability in the presence of structural breaks", *Journal of Econometrics*, 124, 1-31.

[8] Clark, T.E., McCracken, M.W. (2009), "Improving Forecast Accuracy by Combining Recursive and Rolling Forecasts", *International Economic Review*, 50, 363-395.

[9] Clark, T.E., and West, K.D., (2006), "Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis", *Journal of Econometrics*, 135, 155-186.

[10] Clark, T.E., and West, K.D., (2007), "Approximately normal tests for equal predictive accuracy in nested models", *Journal of Econometrics*, 138, 291-311.

[11] Clements, M.P., and Hendry, D.F., (1993), "On the limitations of comparing mean square forecast errors" (with discussion), *Journal of Forecasting*, 12, 617-676.

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

[13] Ericsson, N.R., (1992), "Parameter constancy, mean square forecast errors, and measuring forecast performance: an exposition, extensions, and illustration", *Journal of Policy Modeling*, 14, 465-495.

[14] Giacomini, R., and White, H., (2006), "Tests of conditional predictive ability", *Econometrica*, 74, 1545-1578.

[15] Granger, C.W.J., and Newbold, P., (1986), *Forecasting economic times series* (2nd ed.), New York: Academic Press.

[16] Hansen, P.R., (2005), "A Test for superior predictive ability", *Journal of Business & Economic Statistics*, 23, 365-380.

[17] Harvey, D.I., Leybourne, S.J., and Newbold P., (1998), "Tests of forecast encompassing", *Journal of Business & Economic Statistics*, 16, 254-259.

[18] Hubrich K., and West K.D., (2007), "Forecast Evaluation of small nested model sets", European Central Bank, *mimeo*.

[19] Kilian, L. (1998), Small-sample confidence intervals for impulse response functions, *Review of Economics and Statistics,* 80, 218–230.

[20] McCracken, M.W., (2007), "Asymptotics for out of sample tests of Granger causality", *Journal of Econometrics*, 140, 719-752.

[21] Nelson, C. R., (1972), "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy", *American Economic Review*, 62, 902-917.

[22] Newey, W.K., and West K.D., (1987), "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, 55, 703-708.

[23] Parigi G., and Golinelli R., (2007), The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries, *Journal of Forecasting*, 26, 77-94.

[24] Stock, J.H., and Watson M.W., (2002), Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20, 147-162.

[25] West, K.D., (1996), "Asymptotic inference about predictive ability" *Econometrica*, 64, 1067-1084.

[26] West, K.D., (2006), "Forecast Evaluation", 100-134, in *Handbook of Economic Forecasting*, Vol. 1, G. Elliott, C.W.J. Granger and A. Timmerman (eds), Amsterdam: Elsevier.

[27] White, H. (2000). "A reality check for data snooping", *Econometrica*, 68, 1097-1126.

Table 1: Empirical size of the tests of equal forecast accuracy for one-step ahead forecasts run at nominal 5 and 10% (Nested case).

	$R = 100$				$R = 200$			
π	0.10	0.25	0.50	1.00	0.10	0.25	0.50	1.00
				(A) Recursive 5%				
DM	0.06	0.03	0.02	0.01	0.04	0.02	0.01	0.01
$MSE-t$	0.11	0.09	0.07	0.06	0.08	0.08	0.06	0.06
$MSE-F$	0.07	0.07	0.06	0.06	0.06	0.06	0.05	0.05
HLN	0.07	0.05	0.04	0.03	0.05	0.04	0.04	0.03
$ENC-t$	0.10	0.08	0.07	0.06	0.08	0.07	0.06	0.06
$ENC-F$	0.08	0.07	0.07	0.06	0.07	0.06	0.06	0.06
CCS	0.09	0.07	0.07	0.06	0.07	0.06	0.06	0.06
				(B) Recursive 10%				
DM	0.10	0.06	0.04	0.02	0.08	0.05	0.03	0.02
$MSE-t$	0.17	0.14	0.12	0.11	0.14	0.13	0.11	0.11
$MSE-F$	0.12	0.11	0.11	0.10	0.11	0.10	0.10	0.10
HLN	0.12	0.09	0.08	0.07	0.10	0.08	0.07	0.06
$ENC-t$	0.16	0.14	0.12	0.11	0.14	0.13	0.11	0.11
$ENC-F$	0.13	0.12	0.12	0.11	0.12	0.11	0.11	0.11
CCS	0.15	0.13	0.12	0.12	0.13	0.11	0.11	0.11

Notes: Results from 50,000 Monte Carlo iterations. One-step ahead forecasts with recursive and rolling schemes. In sample sizes: $R = (100, 200)$.

Table 2: Empirical size of the tests of equal forecast accuracy for multi-step ahead forecasts run at nominal 10% (Nested case).

	$R = 100$				$R = 200$			
π	0.10	0.25	0.50	1.00	0.10	0.25	0.50	1.00
				(A) Recursive $h = 2$, 10%				
DM	0.17	0.09	0.05	0.03	0.13	0.07	0.05	0.02
$MSE-t$	0.25	0.15	0.14	0.12	0.20	0.15	0.13	0.11
$MSE-t^*$	0.10	0.09	0.10	0.10	0.11	0.10	0.11	0.10
$MSE-F^*$	0.10	0.09	0.09	0.10	0.11	0.10	0.10	0.10
HLN	0.21	0.12	0.10	0.08	0.16	0.11	0.09	0.08
$ENC-t$	0.25	0.18	0.15	0.13	0.21	0.16	0.13	0.13
$ENC-t^*$	0.10	0.09	0.09	0.10	0.10	0.10	0.10	0.11
$ENC-F^*$	0.10	0.09	0.09	0.10	0.10	0.10	0.10	0.10
CCS	0.30	0.20	0.15	0.16	0.21	0.16	0.15	0.13
CCS^*	0.10	0.10	0.10	0.11	0.10	0.09	0.10	0.09
				(B) Recursive $h = 4$, 10%				
DM	0.23	0.11	0.07	0.04	0.17	0.09	0.05	0.03
$MSE-t$	0.30	0.20	0.15	0.13	0.24	0.17	0.14	0.12
$MSE-t^*$	0.11	0.09	0.11	0.09	0.10	0.11	0.10	0.09
$MSE-F^*$	0.12	0.11	0.11	0.12	0.11	0.10	0.11	0.11
HLN	0.27	0.17	0.13	0.10	0.20	0.13	0.10	0.08
$ENC-t$	0.31	0.22	0.18	0.14	0.25	0.19	0.15	0.13
$ENC-t^*$	0.11	0.09	0.11	0.10	0.10	0.11	0.10	0.10
$ENC-F^*$	0.12	0.11	0.12	0.13	0.11	0.11	0.11	0.12
CCS	0.42	0.26	0.20	0.18	0.30	0.21	0.17	0.14
CCS^*	0.11	0.11	0.10	0.12	0.10	0.10	0.11	0.10

Notes: Results from 5,000 Monte Carlo iterations. h-step ahead forecsts computed with the direct method with recursive and rolling schemes. Forecast horizons: $h = (2, 4)$. In-sample sizes: $R = (100, 200)$. Starred (*) tests are bootstrapped.

Table 3: Empirical application: forecasting the Italian GDP (Nested and non-nested case).

Panel A: Forecasting Italian quarterly GDP one-quarter ahead
Nested case: random walk benchmark vs column model

	AR(1)	ARX_{car}	ARX_{CC}	ARX_{ECM}	ARX_{NX}	ARX_{IP}	Bridge
DM	-	-	-	-	-	***	***
$MSE-t$	**	*	**	*	**	***	***
$MSE-F$	***	***	***	***	***	***	***
HLN	***	***	***	***	***	***	***
$ENC-t$	***	***	***	***	***	***	***
$ENC-F$	***	***	***	***	***	***	***
CCS	***	**	***	**	**	***	***

Panel B: Forecasting Italian quarterly GDP one-quarter ahead
Nested case: column model vs Bridge benchmark

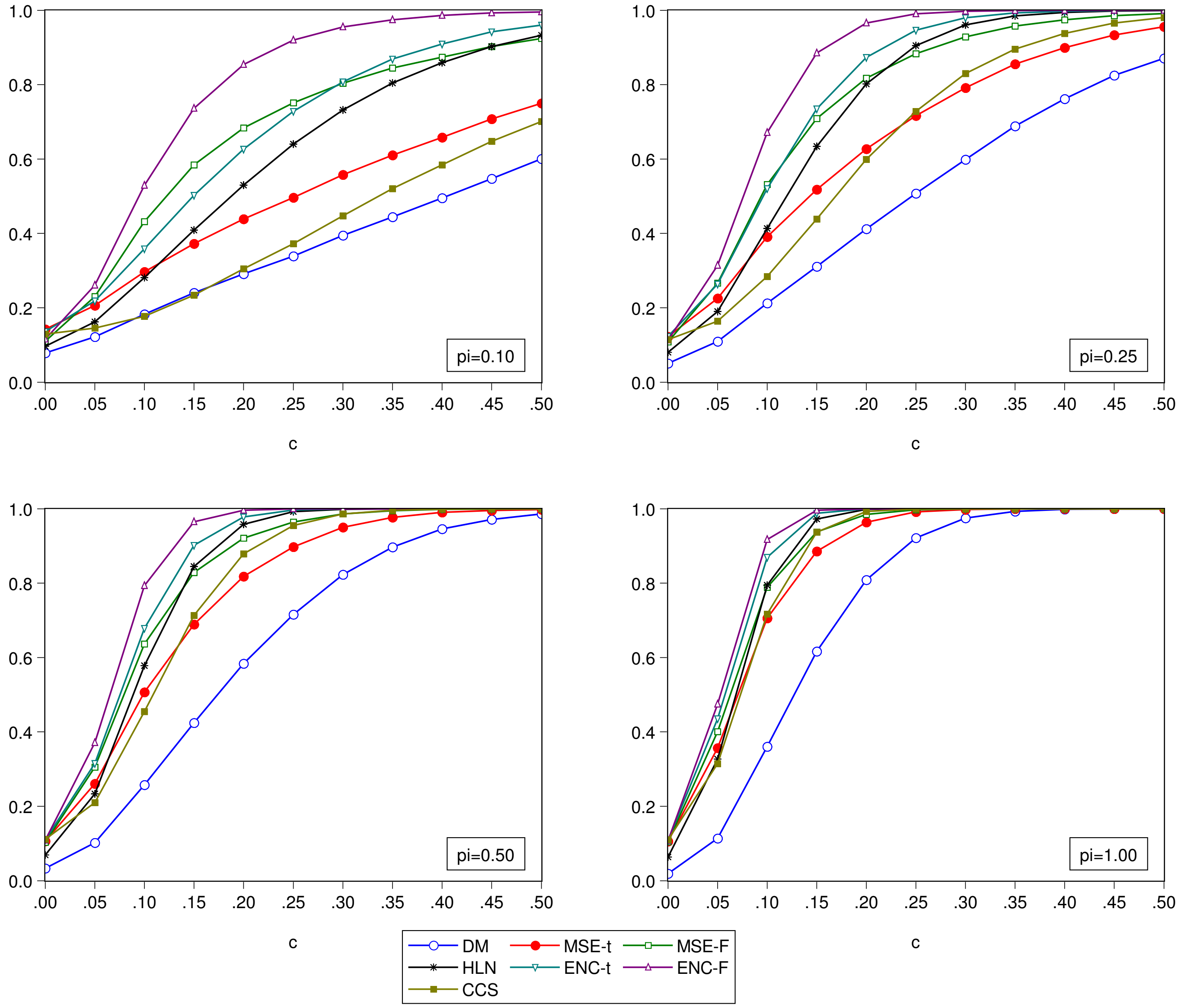
	RW	AR(1)	ARX_{car}	ARX_{CC}	ARX_{ECM}	ARX_{NX}	ARX_{IP}
DM	***	***	***	***	***	***	-
$MSE-t$	***	***	***	***	***	***	*
$MSE-F$	***	***	***	***	***	***	**
HLN	***	***	***	***	***	***	**
$ENC-t$	***	***	***	***	***	***	***
$ENC-F$	***	***	***	***	***	***	***
CCS	***	***	***	***	***	***	**

Panel C: Forecasting euro-area quarterly GDP one-quarter ahead (quarter-on-quarter and year-on-year)
Non-Nested case: Eurocoin benchmark vs column model

	$\Delta_1 \log(GDP_t)$			$\Delta_4 \log(GDP_t)$		
	RW	AR-BIC	ARMA-BIC	RW	AR-BIC	ARMA-BIC
DM	***	***	**	*	-	**
HLN	***	***	***	***	***	***

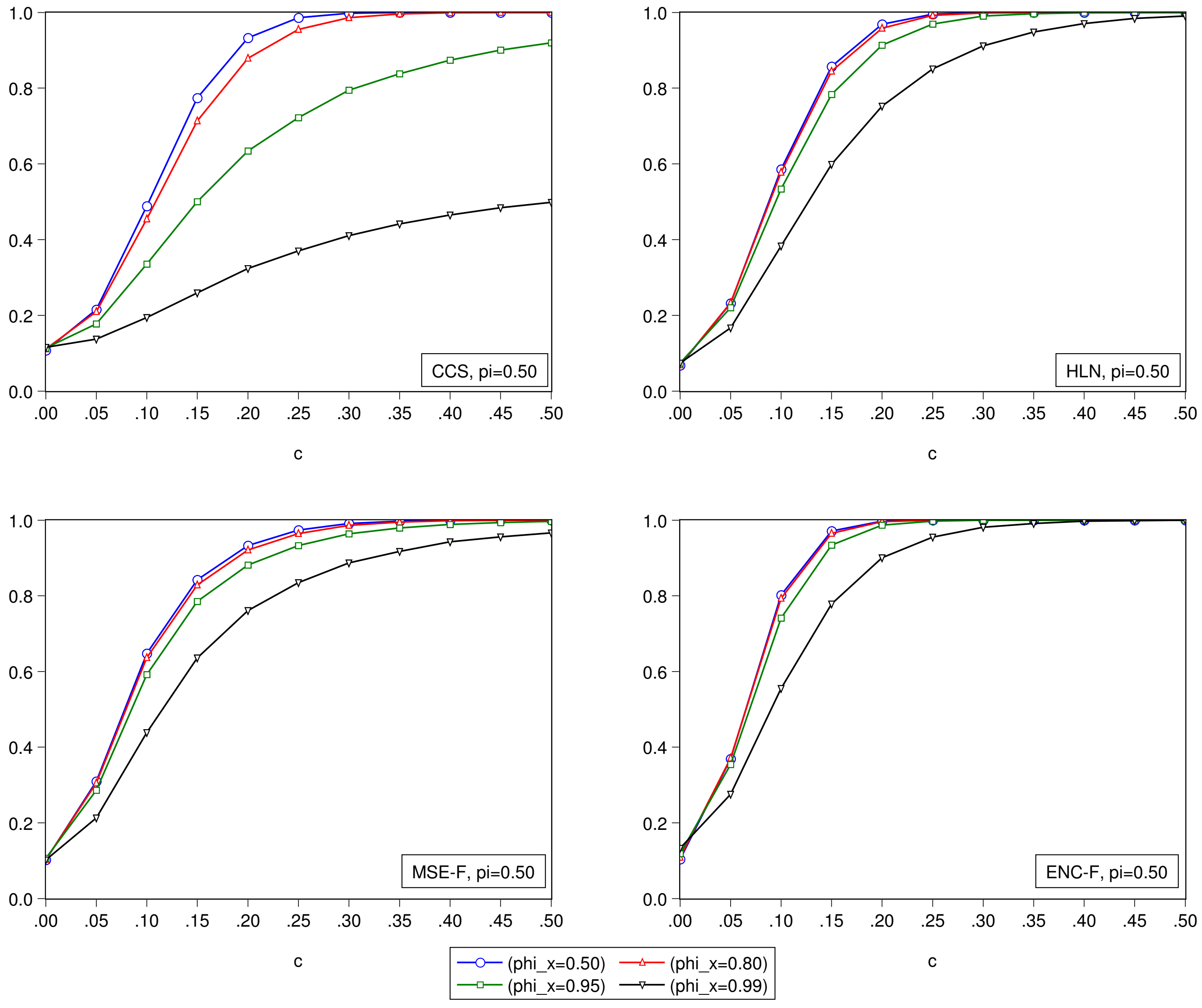
Notes: Panel (A) shows the results of different tests of equal forecast accuracy and FE at one-step ahed for the quarterly Italian GDP. The benchmark model is a random walk (RW) and the alternative models are AR(1), ARX(1) models and a Bridge model. The exogenous variables in the ARX(1) models are the new car sales ARX_{car}, the construction confidence indicator ARX_{CC}, an error correction term ARX_{ECM}, the growth rate of net exports ARX_{NX} and the growth rate of industrial production ARX_{IP}. The Bridge model combines the lagged dependent variable with all the exogenous indicators. Panel (B) reports the results of the same forecasting models as Panel (A) but here the benchmark is the Bridge model that nests all the others. Panel (C) displays the results of the DM and HLN tests in the non-nested case where the benchmark is the Eurocoin model. On the left side we forecast the quarter on quarter growth rate of GDP while on the right hand side we forecast its year on year growth rate one quarter ahead. In all panels ***, ** and * indicate rejection at 1, 5 and 10%, respectively, while '-' represents no rejection of the null hypothesis at 10%.

Figure 1: Empirical power functions for the case of one-step ahead forecasts under correct specification (R=200, recursive regressions)



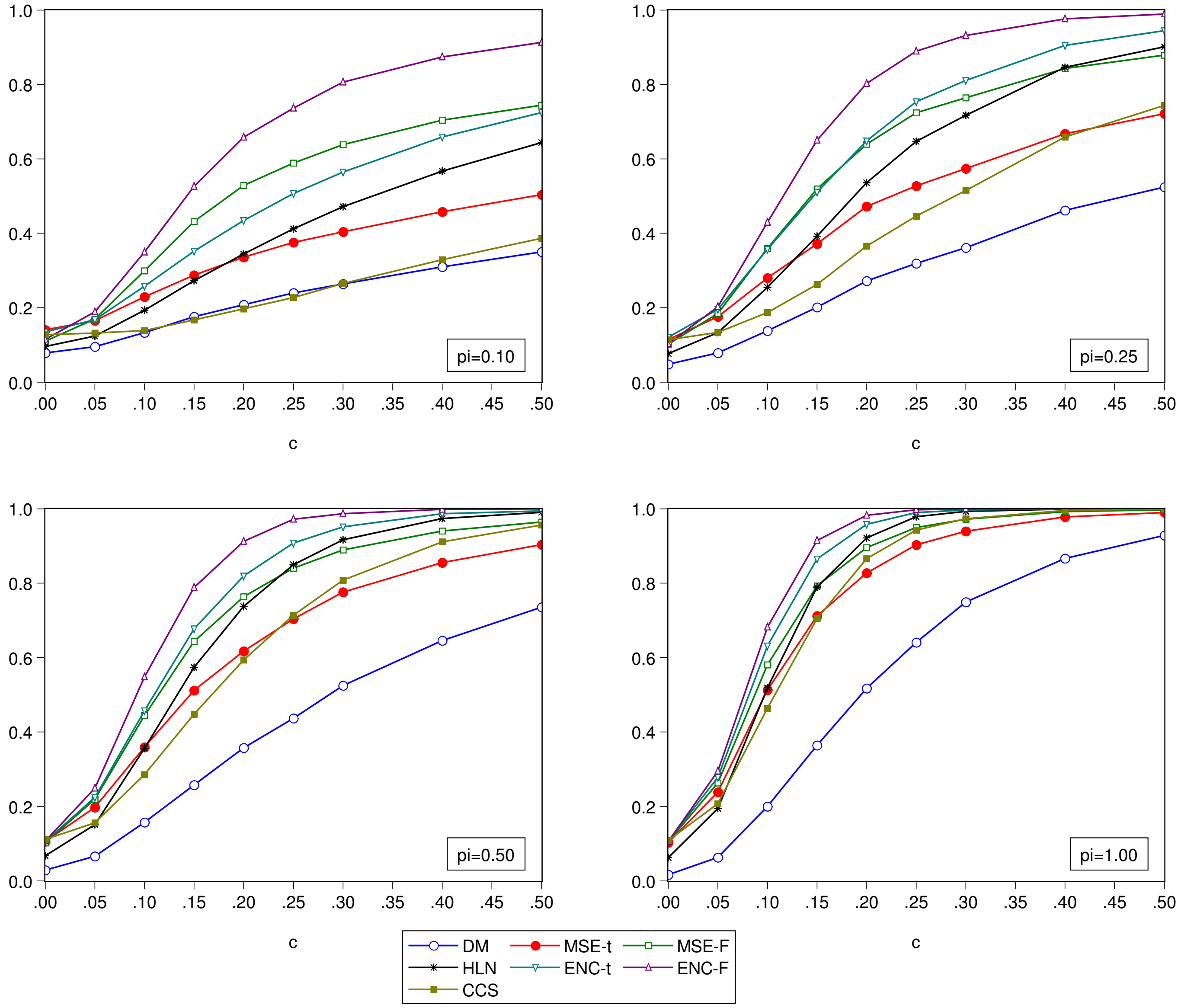
Notes: Results from 50,000 Monte Carlo simulations of one-step ahead forecasts. Recursive scheme with in-sample $R = 200$.

Figure 2: Empirical power functions for the case of one-step ahead forecasts under different degrees of persistence (R=200, π = .50, recursive regressions)



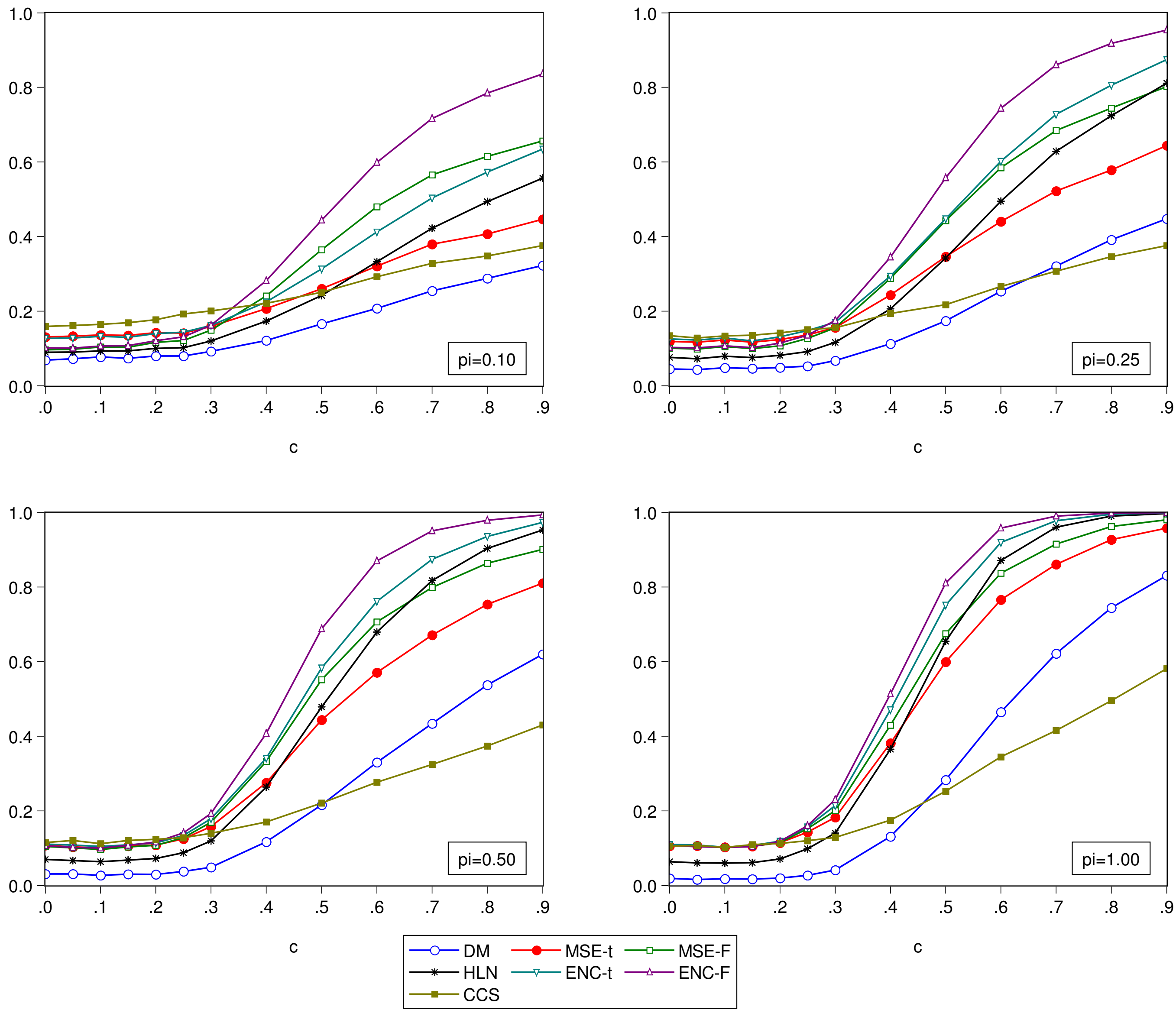
Notes: Results from 50,000 Monte Carlo simulations of one-step ahead forecasts. Recursive scheme with in-sample $R = 200$.

Figure 3: Empirical power functions for the case of one-step ahead forecasts under error-in-variable mis-specification (R=200, recursive regressions)



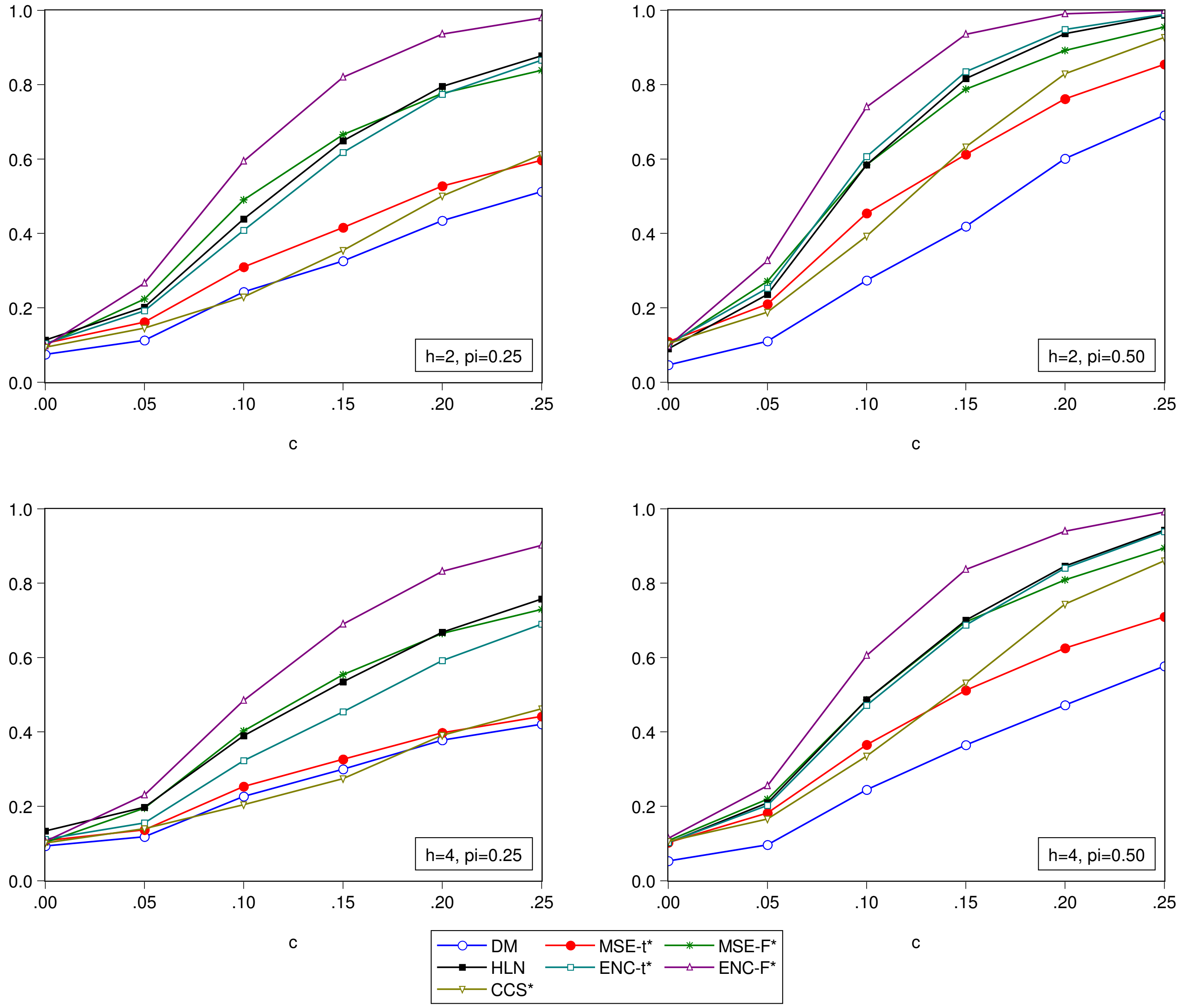
Notes: Results from 50,000 Monte Carlo simulations of one-step ahead forecasts. Recursive scheme with in-sample $R = 200$.

Figure 4: Empirical power functions for the case of one-step ahead forecasts under mis-specification ($AR(p)$ model selected by BIC, R=200, recursive regressions)



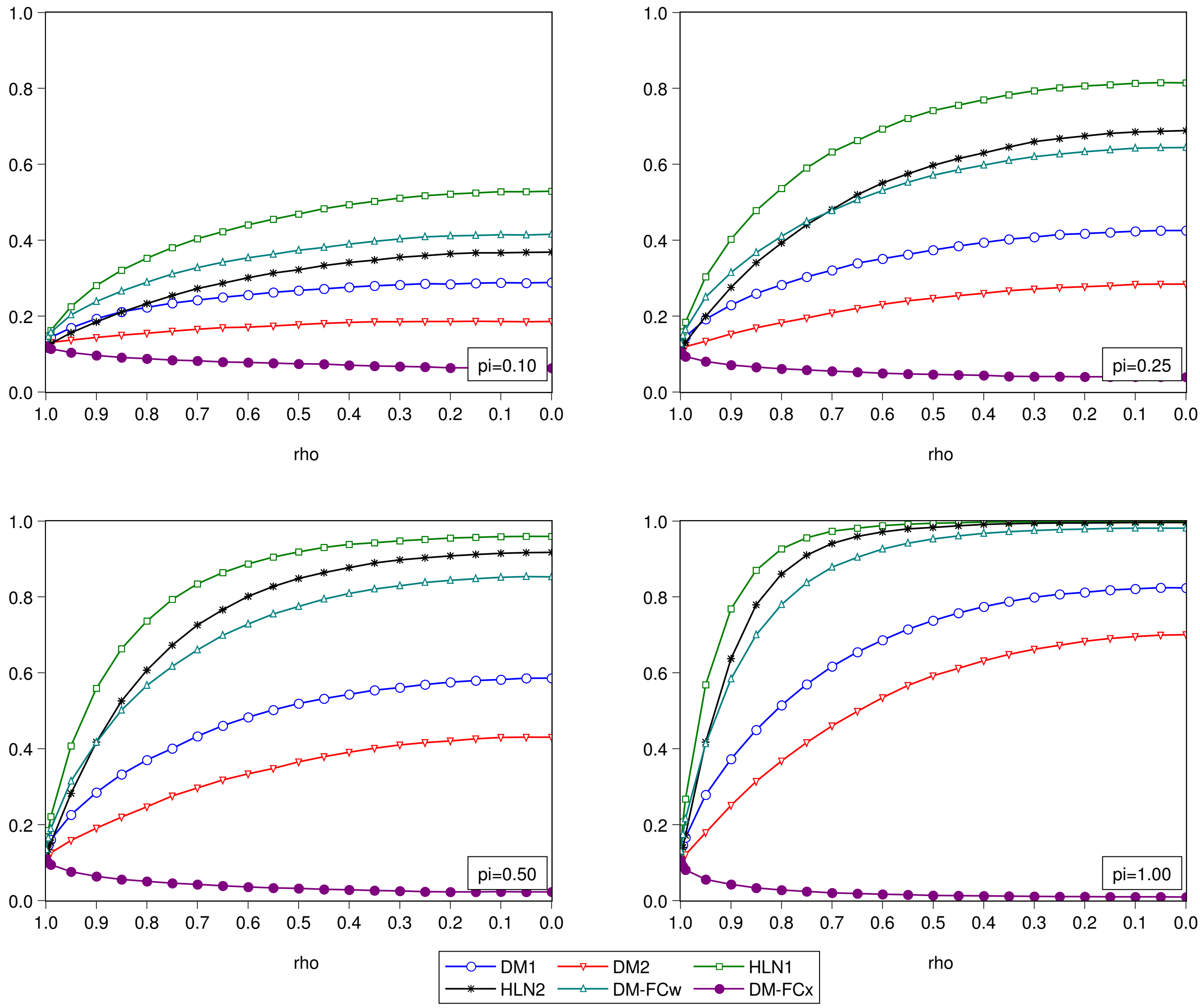
Notes: Results from 50,000 Monte Carlo simulations of one-step ahead forecasts. Recursive scheme with in-sample $R = 200$.

Figure 5: Empirical power functions for the case of multi-step ahead forecasts under correct specification (R=200, recursive regressions)



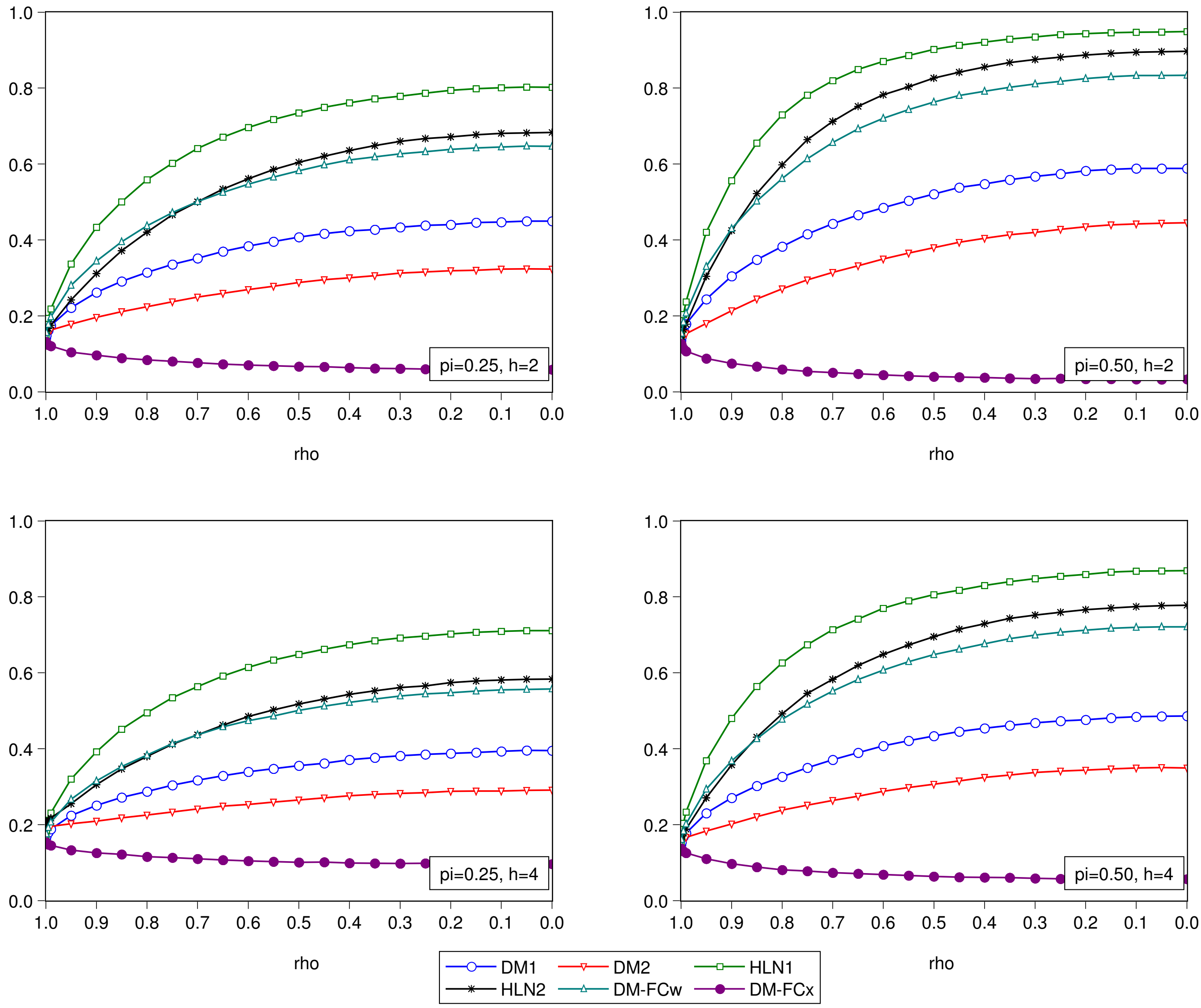
Notes: Results from 5,000 Monte Carlo simulations of multi-step ahead forecasts. Recursive scheme with in-sample $R = 200$.

Figure 6: Empirical power functions for the non-nested case of one-step ahead forecasts (R=200, recursive regressions)



Notes: Results from 10,000 Monte Carlo simulations of one-step ahead forecasts. Recursive scheme with in-sample $R = 200$. Non-nested case.

Figure 7: Empirical power functions for the non-nested case of multi-step ahead forecasts (R=200, recursive regressions)



Notes: Results from 10,000 Monte Carlo simulations of h-step ahead forecasts. Recursive scheme with in-sample $R = 200$ and $h = (2, 4)$. Non-nested case.

RECENTLY PUBLISHED “TEMI” (*)

N. 696 – *Reservation wages: explaining some puzzling regional patterns*, by Paolo Sestito and Eliana Viviano (December 2008).

N. 697 – *Technological change and the demand for currency: an analysis with household data*, by Francesco Lippi and Alessandro Secchi (December 2008).

N. 698 – *Immigration and crime: an empirical analysis*, by Milo Bianchi, Paolo Buonanno and Paolo Pinotti (December 2008).

N. 699 – *Bond risk premia, macroeconomic fundamentals and the exchange rate*, by Marcello Pericoli and Marco Taboga (January 2009).

N. 700 – *What determines debt intolerance? The role of political and monetary institutions*, by Raffaela Giordano and Pietro Tommasino (January 2009).

N. 701 – *On analysing the world distribution of income*, by Anthony B. Atkinson and Andrea Brandolini (January 2009).

N. 702 – *Dropping the books and working off the books*, by Rita Cappariello and Roberta Zizza (January 2009).

N. 703 – *Measuring wealth mobility*, by Andrea Neri (January 2009).

N. 704 – *Oil and the macroeconomy: a quantitative structural analysis*, by Francesco Lippi and Andra Nobili (March 2009).

N. 705 – *The (mis)specification of discrete duration models with unobserved heterogeneity: a Monte Carlo study*, by Cheti Nicoletti and Concetta Rondinelli (March 2009).

N. 706 – *Macroeconomic effects of grater competition in the service sector: the case of Italy*, by Lorenzo Forni, Andrea Gerali and Massimiliano Pisani (March 2009).

N. 707 – *What determines the size of bank loans in industrialized countries? The role of government debt*, by Riccardo De Bonis and Massimiliano Stacchini (March 2009).

N. 708 – *Trend inflation, Taylor principle and indeterminacy,* by Guido Ascari and Tiziano Ropele (May 2009).

N. 709 – *Politicians at work. The private returns and social costs of political connection*, by Federico Cingano and Paolo Pinotti (May 2009).

N. 710 – *Gradualism, transparency and the improved operational framework: a look at the overnight volatility transmission*, by Silvio Colarossi and Andrea Zaghini (May 2009).

N. 711 – *The topology of the interbank market: developments in Italy since 1990*, by Carmela Iazzetta and Michele Manna (May 2009).

N. 712 – *Bank risk and monetary policy*, by Yener Altunbas, Leonardo Gambacorta and David Marqués-Ibáñez (May 2009).

N. 713 – *Composite indicators for monetary analysis*, by Andrea Nobili (May 2009).

N. 714 – *L'attività retail delle banche estere in Italia: effetti sull'offerta di credito alle famiglie e alle imprese,* by Luigi Infante and Paola Rossi (June 2009)

N. 715 – *Firm heterogeneity and comparative advantage: the response of French firms to Turkey's entry in the European Customs Union,* by Ines Buono (June 2009).

N. 716 – *The euro and firm restructuring,* by Matteo Bugamelli, Fabiano Schivardi and Roberta Zizza (June 2009).

N. 717 – *When the highest bidder loses the auction: theory and evidence from public procurement,* by Francesco Decarolis (June 2009).

N. 718 – *Innovation and productivity in SMEs. Empirical evidence for Italy,* by Bronwyn H. Hall, Francesca Lotti and Jacques Mairesse (June 2009).

N. 719 – *Household wealth and entrepreneurship: is there a link?,* by Silvia Magri (June 2009).

(*) Requests for copies should be sent to:
Banca d’Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2006

F. BUSETTI, *Tests of seasonal integration and cointegration in multivariate unobserved component models,* Journal of Applied Econometrics, Vol. 21, 4, pp. 419-438, **TD No. 476 (June 2003).**

C. BIANCOTTI, *A polarization of inequality? The distribution of national Gini coefficients 1970-1996*, Journal of Economic Inequality, Vol. 4, 1, pp. 1-32, **TD No. 487 (March 2004).**

L. CANNARI and S. CHIRI, *La bilancia dei pagamenti di parte corrente Nord-Sud (1998-2000),* in L. Cannari, F. Panetta (a cura di), Il sistema finanziario e il Mezzogiorno: squilibri strutturali e divari finanziari, Bari, Cacucci, **TD No. 490 (March 2004).**

M. BOFONDI and G. GOBBI, *Information barriers to entry into credit markets,* Review of Finance, Vol. 10, 1, pp. 39-67, **TD No. 509 (July 2004).**

W. FUCHS and LIPPI F., *Monetary union with voluntary participation,* Review of Economic Studies, Vol. 73, pp. 437-457 **TD No. 512 (July 2004).**

E. GAIOTTI and A. SECCHI, *Is there a cost channel of monetary transmission? An investigation into the pricing behaviour of 2000 firms*, Journal of Money, Credit and Banking, Vol. 38, 8, pp. 2013-2038 **TD No. 525 (December 2004).**

A. BRANDOLINI, P. CIPOLLONE and E. VIVIANO, *Does the ILO definition capture all unemployment?,* Journal of the European Economic Association, Vol. 4, 1, pp. 153-179, **TD No. 529 (December 2004).**

A. BRANDOLINI, L. CANNARI, G. D'ALESSIO and I. FAIELLA, *Household wealth distribution in Italy in the 1990s,* in E. N. Wolff (ed.) International Perspectives on Household Wealth, Cheltenham, Edward Elgar, **TD No. 530 (December 2004).**

P. DEL GIOVANE and R. SABBATINI, *Perceived and measured inflation after the launch of the Euro: Explaining the gap in Italy,* Giornale degli economisti e annali di economia, Vol. 65, 2 , pp. 155-192, **TD No. 532 (December 2004).**

M. CARUSO, *Monetary policy impulses, local output and the transmission mechanism*, Giornale degli economisti e annali di economia, Vol. 65, 1, pp. 1-30, **TD No. 537 (December 2004).**

L. GUISO and M. PAIELLA, *The role of risk aversion in predicting individual behavior, In P. A. Chiappori e C. Gollier (eds.) Competitive Failures in Insurance Markets: Theory and Policy Implications, Monaco, CESifo*, **TD No. 546 (February 2005).**

G. M. TOMAT, *Prices product differentiation and quality measurement: A comparison between hedonic and matched model methods*, Research in Economics, Vol. 60, 1, pp. 54-68, **TD No. 547 (February 2005).**

L. GUISO, M. PAIELLA and I. VISCO, *Do capital gains affect consumption? Estimates of wealth effects from Italian household's behavior*, in L. Klein (ed), Long Run Growth and Short Run Stabilization: Essays in Memory of Albert Ando (1929-2002), Cheltenham, Elgar, **TD No. 555 (June 2005).**

F. BUSETTI, S. FABIANI and A. HARVEY, *Convergence of prices and rates of inflation*, Oxford Bulletin of Economics and Statistics, Vol. 68, 1, pp. 863-878, **TD No. 575 (February 2006).**

M. CARUSO, *Stock market fluctuations and money demand in Italy, 1913 - 2003,* Economic Notes, Vol. 35, 1, pp. 1-47, **TD No. 576 (February 2006).**

R. BRONZINI and G. DE BLASIO, *Evaluating the impact of investment incentives: The case of Italy's Law 488/92.* Journal of Urban Economics, Vol. 60, 2, pp. 327-349, **TD No. 582 (March 2006).**

R. BRONZINI and G. DE BLASIO, *Una valutazione degli incentivi pubblici agli investimenti,* Rivista Italiana degli Economisti , Vol. 11, 3, pp. 331-362, **TD No. 582 (March 2006).**

A. DI CESARE, *Do market-based indicators anticipate rating agencies? Evidence for international banks,* Economic Notes, Vol. 35, pp. 121-150, **TD No. 593 (May 2006).**

R. GOLINELLI and S. MOMIGLIANO, *Real-time determinants of fiscal policies in the euro area*, Journal of Policy Modeling, Vol. 28, 9, pp. 943-964, **TD No. 609 (December 2006).**

2007

S. SIVIERO and D. TERLIZZESE, *Macroeconomic forecasting: Debunking a few old wives' tales*, Journal of Business Cycle Measurement and Analysis , v. 3, 3, pp. 287-316, **TD No. 395 (February 2001).**

S. MAGRI, *Italian households' debt: The participation to the debt market and the size of the loan*, Empirical Economics, v. 33, 3, pp. 401-426, **TD No. 454 (October 2002).**

L. CASOLARO. and G. GOBBI, *Information technology and productivity changes in the banking industry*, Economic Notes, Vol. 36, 1, pp. 43-76, **TD No. 489 (March 2004).**

G. FERRERO, *Monetary policy, learning and the speed of convergence*, Journal of Economic Dynamics and Control, v. 31, 9, pp. 3006-3041, **TD No. 499 (June 2004).**

M. PAIELLA, *Does wealth affect consumption? Evidence for Italy*, Journal of Macroeconomics, Vol. 29, 1, pp. 189-205, **TD No. 510 (July 2004).**

F. LIPPI. and S. NERI, *Information variables for monetary policy in a small structural model of the euro area*, Journal of Monetary Economics, Vol. 54, 4, pp. 1256-1270, **TD No. 511 (July 2004).**

A. ANZUINI and A. LEVY, *Monetary policy shocks in the new EU members: A VAR approach*, Applied Economics, Vol. 39, 9, pp. 1147-1161, **TD No. 514 (July 2004).**

D. JR. MARCHETTI and F. Nucci, *Pricing behavior and the response of hours to productivity shocks*, Journal of Money Credit and Banking, v. 39, 7, pp. 1587-1611, **TD No. 524 (December 2004).**

R. BRONZINI, *FDI Inflows, agglomeration and host country firms' size: Evidence from Italy*, Regional Studies, Vol. 41, 7, pp. 963-978, **TD No. 526 (December 2004).**

L. MONTEFORTE, *Aggregation bias in macro models: Does it matter for the euro area?*, Economic Modelling, 24, pp. 236-261, **TD No. 534 (December 2004).**

A. NOBILI, *Assessing the predictive power of financial spreads in the euro area: does parameters instability matter?*, Empirical Economics, Vol. 31, 1, pp. 177-195, **TD No. 544 (February 2005).**

A. DALMAZZO and G. DE BLASIO, *Production and consumption externalities of human capital: An empirical study for Italy*, Journal of Population Economics, Vol. 20, 2, pp. 359-382, **TD No. 554 (June 2005).**

M. BUGAMELLI and R. TEDESCHI, *Le strategie di prezzo delle imprese esportatrici italiane*, Politica Economica, v. 23, 3, pp. 321-350, **TD No. 563 (November 2005).**

L. GAMBACORTA and S. IANNOTTI, *Are there asymmetries in the response of bank interest rates to monetary shocks?*, Applied Economics, v. 39, 19, pp. 2503-2517, **TD No. 566 (November 2005).**

P. ANGELINI and F. LIPPI, *Did prices really soar after the euro cash changeover? Evidence from ATM withdrawals*, International Journal of Central Banking, Vol. 3, 4, pp. 1-22, **TD No. 581 (March 2006).**

A. LOCARNO, *Imperfect knowledge, adaptive learning and the bias against activist monetary policies*, International Journal of Central Banking, v. 3, 3, pp. 47-85, **TD No. 590 (May 2006).**

F. LOTTI and J. MARCUCCI, *Revisiting the empirical evidence on firms' money demand*, Journal of Economics and Business, Vol. 59, 1, pp. 51-73, **TD No. 595 (May 2006).**

P. CIPOLLONE and A. ROSOLIA, *Social interactions in high school: Lessons from an earthquake*, American Economic Review, Vol. 97, 3, pp. 948-965, **TD No. 596 (September 2006).**

L. DEDOLA and S. NERI, *What does a technology shock do? A VAR analysis with model-based sign restrictions*, Journal of Monetary Economics, Vol. 54, 2, pp. 512-549, **TD No. 607 (December 2006).**

F. VERGARA CAFFARELLI, *Merge and compete: strategic incentives for vertical integration*, Rivista di politica economica, v. 97, 9-10, serie 3, pp. 203-243, **TD No. 608 (December 2006).**

A. BRANDOLINI, *Measurement of income distribution in supranational entities: The case of the European Union*, in S. P. Jenkins e J. Micklewright (eds.), Inequality and Poverty Re-examined, Oxford, Oxford University Press, **TD No. 623 (April 2007).**

M. PAIELLA, *The foregone gains of incomplete portfolios*, Review of Financial Studies, Vol. 20, 5, pp. 1623-1646, **TD No. 625 (April 2007).**

K. BEHRENS, A. R. LAMORGESE, G.I.P. OTTAVIANO and T. TABUCHI, *Changes in transport and non transport costs: local vs. global impacts in a spatial network*, Regional Science and Urban Economics, Vol. 37, 6, pp. 625-648, **TD No. 628 (April 2007).**

M. BUGAMELLI, *Prezzi delle esportazioni, qualità dei prodotti e caratteristiche di impresa: analisi su un campione di imprese italiane*, v. 34, 3, pp. 71-103, Economia e Politica Industriale, **TD No. 634 (June 2007).**

G. ASCARI and T. ROPELE, *Optimal monetary policy under low trend inflation*, Journal of Monetary Economics, v. 54, 8, pp. 2568-2583, **TD No. 647 (November 2007).**

R. GIORDANO, S. MOMIGLIANO, S. NERI and R. PEROTTI, *The Effects of Fiscal Policy in Italy: Evidence from a VAR Model*, European Journal of Political Economy, Vol. 23, 3, pp. 707-733, **TD No. 656 (January 2008).**

B. ROFFIA and A. ZAGHINI, *Excess money growth and inflation dynamics*, International Finance, v. 10, 3, pp. 241-280, **TD No. 657 (January 2008).**

G. BARBIERI, P. CIPOLLONE and P. SESTITO, *Labour market for teachers: demographic characteristics and allocative mechanisms*, Giornale degli economisti e annali di economia, v. 66, 3, pp. 335-373, **TD No. 672 (June 2008).**

E. BREDA, R. CAPPARIELLO and R. ZIZZA, *Vertical specialisation in Europe: evidence from the import content of exports*, Rivista di politica economica, numero monografico,**TD No. 682 (August 2008).**

2008

P. ANGELINI, *Liquidity and announcement effects in the euro area*, Giornale degli Economisti e Annali di Economia, v. 67, 1, pp. 1-20, **TD No. 451 (October 2002).**

P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, International Journal of Central Banking, v. 4, 3, pp. 1-28, **TD No. 457 (December 2002).**

F. SCHIVARDI and R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, Labour Economics, v. 15, 3, pp. 482-511, **TD No. 504 (June 2004).**

L. GUISO and M. PAIELLA,, *Risk aversion, wealth and background risk*, Journal of the European Economic Association, v. 6, 6, pp. 1109-1150, **TD No. 483 (September 2003).**

C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, Review of Income and Wealth, v. 54, 3, pp. 466-493, **TD No. 520 (October 2004).**

S. MOMIGLIANO, J. HENRY and P. HERNÁNDEZ DE COS, *The impact of government budget on prices: Evidence from macroeconometric models*, Journal of Policy Modelling, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004).**

L. GAMBACORTA, *How do banks set interest rates?,* European Economic Review, v. 52, 5, pp. 792-819, **TD No. 542 (February 2005).**

P. ANGELINI and A. GENERALE, *On the evolution of firm size distributions*, American Economic Review, v. 98, 1, pp. 426-438, **TD No. 549 (June 2005).**

R. FELICI and M. PAGNINI, *Distance, bank heterogeneity and entry in local banking markets*, The Journal of Industrial Economics, v. 56, 3, pp. 500-534, **No. 557 (June 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Labour Economics, v.15, 5, pp. 1040-1061, **TD No. 570 (January 2006).**

M. PERICOLI and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia,* Journal of Money, Credit and Banking, v. 40, 7, pp. 1471-88, **TD No. 580 (February 2006).**

E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector*, Labour Economics, v. 15, 6, pp. 1200-1222, **TD No. 594 (May 2006).**

S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy*, Review of World Economics, Journal of Money, Credit and Banking, v. 144, 2, pp. 295-324, **TD No. 613 (February 2007).**

F. BUSETTI and A. HARVEY, *Testing for trend*, Econometric Theory, v. 24, 1, pp. 72-87, **TD No. 614 (February 2007).**

V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for prices and the Euro cash changeover: an analysis for cinema prices in Italy,* In P. Del Giovane e R. Sabbatini (eds.), The Euro Inflation and Consumers' Perceptions. Lessons from Italy, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007).**

B. H. HALL, F. LOTTI and J. MAIRESSE, *Employment, innovation and productivity: evidence from Italian manufacturing microdata*, Industrial and Corporate Change, v. 17, 4, pp. 813-839, **TD No. 622 (April 2007).**

J. SOUSA and A. ZAGHINI, *Monetary policy shocks in the Euro Area and global liquidity spillovers,* International Journal of Finance and Economics, v.13, 3, pp. 205-218, **TD No. 629 (June 2007).**

M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms,* Journal of Regional Science, v. 48, 1, pp. 97-129, **TD No. 635 (June 2007).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers*, in P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 655 (January 2008).**

B. BORTOLOTTI, and P. PINOTTI, *Delayed privatization,* Public Choice, v. 136, 3-4, pp. 331-351, **TD No. 663 (April 2008).**

R. BONCI and F. COLUMBA, *Monetary policy effects: New evidence from the Italian flow of funds,* Applied Economics , v. 40, 21, pp. 2803-2818, **TD No. 678 (June 2008).**

M. CUCCULELLI, and G. MICUCCI, *Family Succession and firm performance: evidence from Italian family firms,* Journal of Corporate Finance, v. 14, 1, pp. 17-31, **TD No. 680 (June 2008).**

A. SILVESTRINI and D. VEREDAS, *Temporal aggregation of univariate and multivariate time series models: a survey*, Journal of Economic Surveys, v. 22, 3, pp. 458-497, **TD No. 685 (August 2008).**

2009

F. PANETTA, F. SCHIVARDI and M. SHUM, *Do mergers improve information? Evidence from the loan market*, Journal of Money, Credit, and Banking, v. 41, 4, pp. 673-709, **TD No. 521 (October 2004).**

P. PAGANO and M. PISANI, *Risk-adjusted forecasts of oil prices*, The B.E. Journal of Macroeconomics, v. 9, 1, Article 24, **TD No. 585 (March 2006).**

M. PERICOLI and M. SBRACIA, *The CAPM and the risk appetite index: theoretical differences, empirical similarities, and implementation problems*, International Finance, v. 12, 2, pp. 123-150, **TD No. 586 (March 2006).**

S. MAGRI, *The financing of small innovative firms: the Italian case*, Economics of Innovation and New Technology, v. 18, 2, pp. 181-204, **TD No. 640 (September 2007).**

S. MAGRI, *The financing of small entrepreneurs in Italy*, Annals of Finance, v. 5, 3-4, pp. 397-419, **TD No. 640 (September 2007).**

F. LORENZO, L. MONTEFORTE and L. SESSA, *The general equilibrium effects of fiscal policy: estimates for the euro area*, Journal of Public Economics, v. 93, 3-4, pp. 559-585, **TD No. 652 (November 2007).**

R. GOLINELLI and S. MOMIGLIANO, *The Cyclical Reaction of Fiscal Policies in the Euro Area. A Critical Survey of Empirical Research*, Fiscal Studies, v. 30, 1, pp. 39-72, **TD No. 654 (January 2008).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "Inflation Perceptions"? A survey-based analysis of Italian consumers*, Giornale degli Economisti e Annali di Economia, v. 68, 1, pp. 25-52, **TD No. 655 (January 2008).**

F. MACCHERONI, M. MARINACCI, A. RUSTICHINI and M. TABOGA, *Portfolio selection with monotone mean-variance preferences,* Mathematical Finance, v. 19, 3, pp. 487-521, **TD No. 664 (April 2008).**

M. AFFINITO and M. PIAZZA, *What are borders made of? An analysis of barriers to European banking integration*, in P. Alessandrini, M. Fratianni and A. Zazzaro (eds.): The Changing Geography of Banking and Finance, Dordrecht Heidelberg London New York, Springer, **TD No. 666 (April 2008).**

L. ARCIERO, C. BIANCOTTI, L. D'AURIZIO and C. IMPENNA, *Exploring agent-based methods for the analysis of payment systems: A crisis model for StarLogo TNG*, Journal of Artificial Societies and Social Simulation, v. 12, 1, **TD No. 686 (August 2008).**

A. CALZA and A. ZAGHINI, *Nonlinearities in the dynamics of the euro area demand for M1*, Macroeconomic Dynamics, v. 13, 1, pp. 1-19, **TD No. 690 (September 2008).**

L. FRANCESCO and A. SECCHI, *Technological change and the households' demand for currency*, Journal of Monetary Economics, v. 56, 2, pp. 222-230, **TD No. 697 (December 2008).**

M. BUGAMELLI, F. SCHIVARDI and R. ZIZZA, *The euro and firm restructuring*, in A. Alesina e F. Giavazzi (eds): Europe and the Euro, Chicago, University of Chicago Press, **TD No. 716 (June 2009).**

B. HALL, F. LOTTI and J. MAIRESSE, *Innovation and productivity in SMEs: empirical evidence for Italy*, Small Business Economics, v. 33, 1, pp. 13-33, **TD No. 718 (June 2009).**

FORTHCOMING

L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts*, Applied Economics, **TD No. 458 (December 2002).**

M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006).**

G. DE BLASIO and G. NUZZO, *Historical traditions of civicness and local economic development*, Journal of Regional Science, **TD No. 591 (May 2006).**

R. BRONZINI and P. PISELLI, *Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure,* Regional Science and Urban Economics, **TD No. 597 (September 2006).**

E. IOSSA and G. PALUMBO, *Over-optimism and lender liability in the consumer credit market,* Oxford Economic Papers, **TD No. 598 (September 2006).**

U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and the business cycle,* Journal of Financial Stability, **TD No. 601 (September 2006).**

A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, Journal of International Financial Markets, Institutions & Money, **TD No. 637 (June 2007).**

V. DI GIACINTO and G. MICUCCI, *The producer service sector in Italy: long-term growth and its local determinants,* Spatial Economic Analysis, **TD No. 643 (September 2007).**

Y. ALTUNBAS, L. GAMBACORTA and D. MARQUÉS, *Securitisation and the bank lending channel,* European Economic Review, **TD No. 653 (November 2007).**

F. BALASSONE, F. MAURA and S. ZOTTERI, *Cyclical asymmetry in fiscal variables in the EU*, Empirica, **TD No. 671 (June 2008).**

M. BUGAMELLI and F. PATERNÒ, *Output growth volatility and remittances*, Economica, **TD No. 673 (June 2008).**

M. IACOVIELLO and S. NERI, *Housing market spillovers: evidence from an estimated DSGE model,* American Economic Journal: Macroeconomics, **TD No. 659 (January 2008).**

A. ACCETTURO, *Agglomeration and growth: the effects of commuting costs,* Papers in Regional Science, **TD No. 688 (September 2008).**

L. FORNI, A. GERALI and M. PISANI, *Macroeconomic effects of greater competition in the service sector: the case of Italy,* Macroeconomic Dynamics, **TD No. 706 (March 2009).**

Y. ALTUNBAS, L. GAMBACORTA, and D. MARQUÉS-IBÁÑEZ, *Bank risk and monetary policy,* Journal of Financial Stability, **TD No. 712 (May 2009).**