JOURNAL OF APPLIED ECONOMETRICS
*J. Appl. Econ.* **16**: 657–669 (2001)
DOI: 10.1002/jae.619

# MEASURING PREDICTABILITY: THEORY AND MACROECONOMIC APPLICATIONS

FRANCIS X. DIEBOLD[a]* AND LUTZ KILIAN[b]

[a] *Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297, USA, and National Bureau of Economic Research, Cambridge, MA, USA*
[b] *Department of Economics, University of Michigan, Ann Arbor, MI 48109-1220, USA, and Centre for Economic Policy Research, London, UK*

SUMMARY

We propose a measure of predictability based on the ratio of the expected loss of a short-run forecast to the expected loss of a long-run forecast. This predictability measure can be tailored to the forecast horizons of interest, and it allows for general loss functions, univariate or multivariate information sets, and covariance stationary or difference stationary processes. We propose a simple estimator, and we suggest resampling methods for inference. We then provide several macroeconomic applications. First, we illustrate the implementation of predictability measures based on fitted parametric models for several US macroeconomic time series. Second, we analyze the internal propagation mechanism of a standard dynamic macroeconomic model by comparing the predictability of model inputs and model outputs. Third, we use predictability as a metric for assessing the similarity of data simulated from the model and actual data. Finally, we outline several non-parametric extensions of our approach. 

## 1. INTRODUCTION

It is natural and informative to judge forecasts by their accuracy. However, actual and forecasted values will differ, even for good forecasts. To take an extreme example, consider a zero-mean white-noise process. The optimal linear forecast under quadratic loss in this case is simply zero, so the paths of forecasts and realizations will clearly look different. These differences illustrate the inherent limits to predictability, even when using optimal forecasts. The extent of a series' predictability depends on how much information the past conveys regarding future values of this series; as a result, some processes are inherently easy to forecast, and others are more difficult. In addition to being of interest to forecasters, predictability measures are potentially useful in empirical macroeconomics. Predictability provides a succinct measure of a key aspect of time series dynamics and is therefore useful for summarizing the behaviour of economic series, as well as for assessing agreement between economic models and data.[1]

Remarkably little attention has been paid to methods for measuring predictability. Existing methods include those based on canonical correlations between past and future, and those based on comparing the innovation variance and unconditional variance of stationary series (see Jewell

* Correspondence to: Francis X. Diebold, Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297, USA.

[1] We do not advocate comparing models to data *purely* on the basis of predictability. Rather, predictability simply provides an easily digested summary distillation of certain important aspects of dynamics. More complete frameworks for assessing agreement between models and data are developed in King and Watson (1996) and Diebold, Ohanian and Berkowitz (1998).



*Received 15 July 1998*
*Revised 5 February 2001*

and Bloomfield, 1983; Hannan and Poskitt, 1988; Granger and Newbold, 1986). Those methods, however, are inadequate in light of recent work stressing the possible presence of unit roots, rich and high-dimensional information sets, non-quadratic and possibly asymmetric loss, and variations in forecast accuracy across horizons (see Stock, 1995; Forni and Reichlin, 1998; Christoffersen and Diebold, 1996, 1997; Diebold and Mariano, 1995).

The lack of methodological development coincides with a lack of substantive exploration. Even for the major macroeconomic aggregates, little is known about comparative predictability. At first glance, the assertion that we know little about predictability seems exaggerated. We know, for example, that for the USA consumption is less volatile than output, and that investment is more volatile than output. Such statements, however, concern *unconditional* variances, whereas predictability concerns variances *conditional* on varying information sets. The concepts are fundamentally different.

In this paper, we contribute to the theory of predictability measurement and apply our results in several macroeconomic contexts. In Section 2, we discuss some of the difficulties involved in predictability measurement and propose a simple measure of relative predictability based on the ratio of the expected loss of a short-run forecast to the expected loss of a long-run forecast. Our measure allows for covariance stationary or difference stationary processes, univariate or multivariate information sets, general loss functions, and different forecast horizons of interest. In Section 3, we propose parametric methods for estimating the predictability of observed series, and we suggest using bootstrap methods for inference. We also illustrate the implementation of predictability measures based on fitted parametric models for several US macroeconomic time series. In Section 4, we illustrate the use of these parametric predictability measures in assessing the propagation mechanisms of economic models, and in assessing agreement between economic models and data.[2] In Section 5 we discuss alternative non-parametric measures of predictability and directions for future research.

## 2. POPULATION PREDICTABILITY MEASURES

The expected loss of an optimal forecast will in general exceed zero, which illustrates the inherent limits to predictability, even when using optimal forecasts. Put differently, poor forecast accuracy does not necessarily imply that the forecaster failed. The extent of a series' predictability in population depends on how much information the past conveys regarding the future; given an information set, some processes are inherently easy to forecast, and others are more difficult.

In measuring predictability it is important to keep three points in mind. First, the question of whether a series is predictable or not should be replaced by one of *how* predictable it is. Predictability is always a matter of degree. Second, the question of how predictable a series is cannot be answered in general. We have to be clear about the relevant forecast horizon and loss function. For example, a series may be highly predictable at short horizons, but not at long horizons. Third, to compare the predictability of several series we need a common numeraire. It may be tempting to simply compare the expected losses of forecasts for two series to assess

[2] The exercises in Section 4 are similar in spirit to Rotemberg and Woodford (1996). Rotemberg and Woodford's analysis of predictability is tightly and appropriately linked to the assumptions of the particular economic model they study. Our focus, in contrast, is on general macroeconometric methods of assessing predictability that can be applied in a variety of situations, under minimal assumptions.

 

their relative predictability, but that ignores the possibility that the two series may be measured on different scales.

Granger and Newbold (1986) therefore propose a natural measure of the forecastability of covariance stationary series under squared-error loss, patterned after the familiar regression $R^2$, $G = \text{var}(\hat{y}_{t+j,t})/\text{var}(y_{t+j}) = 1 - \text{var}(e_{t+j,t})/\text{var}(y_{t+j})$, where $\hat{y}_{t+j,t}$ is the optimal (i.e. conditional mean) forecast and $e_{t+j,t} = y_{t+j} - \hat{y}_{t+j,t}$. Related work includes Clements and Hendry (1998).

Our approach to predictability measurement is squarely in the tradition of Granger and Newbold, with the important difference that we relax several constraints that limit the broad applicability of their methods. Like Granger and Newbold, we assume that structural breaks in the stochastic process either have not occurred or have been appropriately modelled. The essence of the Granger–Newbold suggestion is that it is natural to base a measure of predictability on the difference between the conditionally expected loss of an optimal short-run forecast, $E(L(e_{t+j,t}))$, and that of an optimal long-run forecast, $E(L(e_{t+k,t}))$, $j \ll k$, where $E(.)$ denotes the mathematical expectation conditional on the information set $\Omega$. If $E(L(e_{t+j,t})) \ll E(L(e_{t+k,t}))$, we say that the series is highly predictable at horizon $j$ relative to $k$, and if $E(L(e_{t+j,t})) \approx E(L(e_{t+k,t}))$, we say that the series is nearly unpredictable at horizon $j$ relative to $k$. Thus, we define a general measure of predictability as $P(L, \Omega, j, k) = 1 - E(L(e_{t+j,t}))/E(L(e_{t+k,t}))$, where the information set $\Omega$ can be univariate or multivariate, as desired. The Granger–Newbold measure emerges when the series is covariance stationary, $L(x) = x^2$ (and hence the optimal forecast is the conditional mean), the information set is univariate, and $k = \infty$. The advantages of our generalization include:

(1) It is valid for both covariance stationary and difference stationary series, so long as $k < \infty$.
(2) It allows for general loss functions. The loss function $L(\cdot)$ need not be quadratic or even symmetric; we only require that $L(0) = 0$ and that $L(\cdot)$ be strictly monotone on each side of the origin. By the restrictions imposed on $L(\cdot)$, we have that for all covariance stationary or difference stationary processes $P(L(\cdot), \Omega, j, k) \in [0, 1]$, with larger values indicating greater predictability.
(3) It allows for univariate or multivariate information sets, and economic theory may suggest relevant multivariate information sets.
(4) It allows for flexibility in the choice of j and k and enables one to tailor the predictability measure to the horizons of economic interest.

Our predictability measure is closely related to Theil's (1966) $U$ statistic, which we define for the one-step-ahead horizon as $U = E(e_{t,t-1}^2)/E((y_t - y_{t-1})^2)$. To see this, specialize $P$ to the quadratic, univariate, $j = 1$ case and write it as $P(\text{quadratic, univariate}, 1, k) = 1 - E(e_{t,t-1}^2)/E(e_{t,t-k}^2)$, or $1 - P = E(e_{t,t-1}^2)/E(e_{t,t-k}^2)$. Thus, under certain conditions, $1 - P$ is similar in spirit to Theil's $U$. The key difference is that Theil's $U$ assesses one-step forecast accuracy relative to that of a 'naive' no-change forecast, whereas $P$ assesses one-step accuracy relative to that of a long-horizon ($k$-step) forecast. In the general case, $P(L(\cdot), \Omega, j, k) = 1 - E(L(e_{t,t-j}))/E(L(e_{t,t-k}))$. Thus, $P(L(\cdot), \Omega, j, k)$ is effectively one minus the ratio of expected losses of two forecasts of the same object, $y_t$. Typically, one forecast, $\hat{y}_{t,t-j}$, is based on a rich information set, while the other forecast, $\hat{y}_{t,t-k}$, is based on a sparse information set.

The formula for $P(L(\cdot), \Omega, j, k)$ also makes clear that the concept of predictability is related to, but distinct from, the concept of persistence of a series. Suppose, for example, that the series $y_t$ is a random walk. Then $P(e^2, \text{univariate}, j, k) = 1 - j/k$, as will be shown later. The corresponding

 

$j$-step variance ratio, a common persistence measure, is $V_j = \text{var}(y_t - y_{t-j})/\text{var}(y_t - y_{t-1}) = j$. It is clear, however, that although $P(e^2, \text{univariate}, j, k)$ and $V_j$ are deterministically related in the random walk case ($P = 1 - V/k$), they are not deterministically related in more general cases.

## 3. SAMPLE PREDICTABILITY MEASURES

Predictability is a population property of a series, not of any particular sample path, but predictability can be estimated from a sample path. We proceed by fitting a parametric model and then transforming estimates of the parameters into an estimate of $P$. To keep the discussion tractable, and in keeping with the empirical analysis of subsequent sections, we postulate a quadratic loss function $L(e) = e^2$ for estimation, prediction, model selection, and construction of predictability measures.

It is clear that parametric measures of predictability in general will depend on the specification of the parametric model. Here we focus on univariate autoregressive models, although one could easily generalize the discussion to other parametric models, such as vector ARMA models. We construct $P$ by simply reading off the appropriate diagonal elements of the forecast MSE matrices for forecast horizons $j$ and $k$. To build intuition, consider a univariate AR(1) population process with innovation variance $\Sigma_u$: $y_t = A_1 y_{t-1} + u_t$. Then for $A_1 = 0$ the model reduces to white noise, and short-run forecasts are just as accurate as long-run forecasts. As a result, relative predictability is zero: $P(j, k) = 1 - \Sigma_u/\Sigma_u = 0$ for all $j$. In contrast, for $A_1 = 1$ the model becomes a random walk, and relative predictability steadily declines as the forecast horizon increases: $P(j, k) = 1 - (j\Sigma_u)/(k\Sigma_u) = 1 - j/k$.

Forecast errors from consistently estimated processes and processes with known parameters are asymptotically equivalent. In practice, we estimate $P$ by replacing the underlying unknown parameters by their least squares estimates. The legitimacy of this predictability measure is invariant to the possible presence of unit roots in the autoregressive lag order polynomial. To determine the autoregressive lag order we use the Akaike Information Criterion (AIC) with a suitable upper bound on the admissible lag orders. The AIC is less likely to underestimate the lag order in small samples than alternative criteria. The latter property is crucial in preserving the higher-order dynamics implicit in $P$ (see Kilian, 2001). Throughout this paper, inference is based on the bias-corrected bootstrap method proposed by Kilian (1998a,b). We account for lag order uncertainty along the lines suggested by Kilian (1998c). Recent theoretical results by Inoue and Kilian (2000a) suggest that—with few exceptions (such as most random walk processes)—the method of bootstrap inference we propose is asymptotically valid both for stationary and for unit root processes. Some Monte Carlo evidence of the small-sample accuracy of the proposed bootstrap method can be found in Diebold and Kilian (1999).

We now turn to an illustrative example. Figure 1 shows point and interval estimates for the predictability of several post-war quarterly US investment series. For expository purposes, we use univariate information sets and fix $L(e) = e^2$ and $k = 40$, as we vary the near-term forecast horizon $j$. Note that the univariate autoregressive representation of the series may be interpreted without loss of generality as a marginalized reduced form of a more general vector autoregressive model. We plot $P$ (in per cent) against near-term forecast horizons $j = 1, \ldots, 20$. Higher values of $P$ indicate greater predictability. All data are in logs and seasonally adjusted. We model the data as level autoregressions and, with the exception of the inventory series, allow for a linear time trend.

Private investment spending is one of the least predictable components of real expenditures. Although predictability is close to 80% at a horizon of one quarter, it drops sharply at higher

 

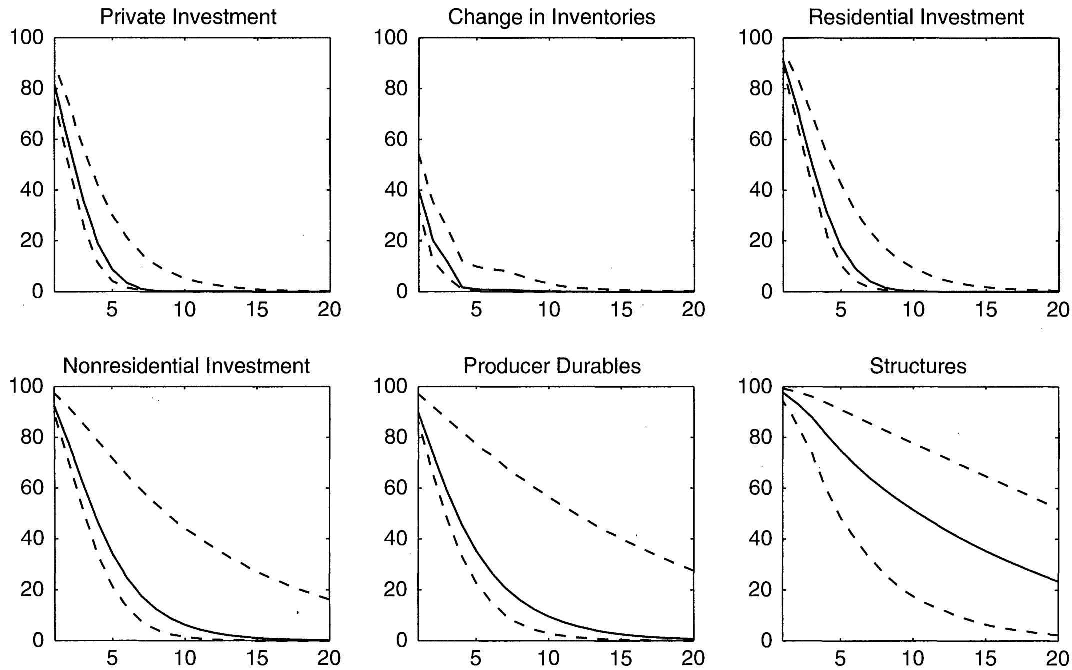


Figure 1. Predictability estimates with 90% confidence intervals. Notes: All data are from Citibase. The sample period is 1947.II–1994.III. For a description of the methodology see text. The plots are based on $k = 40$ and $j = 1, \ldots, 20$

horizons. Our estimates suggest that its one-year predictability is just 20%. Beyond a horizon of two years it is essentially zero. Broken down by component, changes in inventories of nonfarm businesses are least predictable, followed by residential and non-residential investment, in that order. Within non-residential investment spending, however, structures are much more predictable than producer durables. This example highlights the fact that there are important differences in predictability across expenditure components that are potentially useful for macroeconomic modelling. For example, changes in inventories are virtually unpredictable for horizons in excess of one year. In contrast, the predictability of investment in structures remains in excess of 20% even after five years.

We also show the corresponding nominal 90% bootstrap confidence intervals. The confidence intervals often are wide, reflecting the limited information available in macroeconomic data, but not so wide as to render the exercise futile. Moreover, the degree of sampling uncertainty may differ across investment components. For example, the predictability of structures is rather uncertain, whereas that of inventories is more precisely estimated.

## 4. EVALUATING MACROECONOMIC MODELS BASED ON PREDICTABILITY

Predictability measures are potentially useful for assessing the internal propagation mechanisms of economic models, and for assessing agreement between models and quarterly data. We illustrate both uses with the indivisible labour model of Hansen (1985), in which a representative agent chooses labour input, $h_t$, and next period's capital stock, $k_{t+1}$, to maximize expected lifetime utility, $E_0 \sum_{t=0}^{\infty} \beta^t (\log c_t + B(1 - h_t))$, subject to the constraints $c_t + i_t \leq z_t k_t^{\theta} h_t^{1-\theta}$, $k_{t+1} = (1 - \delta)k_t + i_t$, and $z_{t+1} = (1 - \gamma) + \gamma z_t + \varepsilon_{t+1}$, where $k_0$ and $z_0$ are given, $0 < \beta < 1$, $0 \leq \delta \leq 1$, and $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$. We parameterize the model as in Hansen (1985) and solve for the associated linear decision rules for $h_t$ and $k_{t+1}$ in terms of the current period states $k_t$ and $z_t$. The data source is Hansen (1985).

To assess the internal propagation mechanism of Hansen's model, we first compare the predictability of the model input (the exogenous technology shock) and the model outputs (the endogenous model variables). A weak propagation mechanism, assessed in terms of predictability, is associated with nearly identical input and output predictability. We calculate the predictability of the technology shock analytically, based on population parameter values, and we calculate the predictability of the model outputs numerically. The predictability of the model outputs can be calculated to any desired degree of accuracy by simulating a long enough realization from the model (the 'model data') and then estimating predictability by fitting a univariate autoregressive model and using our substitution estimator. We compute the predictability of the model data on output, consumption, investment, productivity, capital stock, and hours using a simulated realization of length 10,000.

The results appear in Figure 2(a), in which we show the comparative predictability of the technology shock and the model outputs. We again fix $L(e) = e^2$ and set $k = 40$, as we vary the near-term forecast horizon $j = 1, \ldots, 20$. Figure 2(a) makes clear that many of the model outputs have predictability patterns very different from the model input. Model real GNP is about as predictable as the technology shock. In contrast, model investment and hours are less predictable than the technology shock, and model consumption, productivity and the capital stock are more predictable than the technology shock. Figure 2(a) indicates that the indivisible labour model has a strong internal propagation mechanism, in the sense that the predictabilities of model outputs are distinctly different from that of the technology shock and different from one another. Some model series are more predictable than the technology shock, some are less predictable, and the differences arise endogenously from the model's internal propagation mechanism. This result is surprising, insofar as other studies using other criteria have concluded that models such as Hansen's have weak internal propagation mechanisms (e.g. Cogley and Nason, 1995a).

Thus far, neither the model input nor the model outputs have been logged or filtered, which effectively amounts to population linear detrending of both the input and output series, because no trends are operative. In our judgement that is the right way to proceed if we are interested in assessing the propagation mechanism of the model. In many applications, however, the focus is on how well the dynamics in simulated model data match those of the cyclical component of the actual data (expressed in percent deviations from a smooth trend). A common approach is to log and HP-filter the model outputs. We show the results of doing so in Figure 2(b). All model outputs are now much less predictable than the technology shock, which is not surprising, because the HP filtering removes highly predictable low-frequency components. Thus it is nonsensical to compare model input and output predictabilities when only the outputs have been HP filtered. We can, however, still compare the predictability of the various model outputs: with the possible exception of the capital stock, all become strikingly similar after HP filtering. To an unsuspecting observer this finding may seem to suggest that the technology shock imparts a common pattern of predictability.

A natural conjecture, however, is that common predictability pattern of HP-filtered model outputs is an artifact of HP filtering. A number of earlier studies have documented such effects, from

 

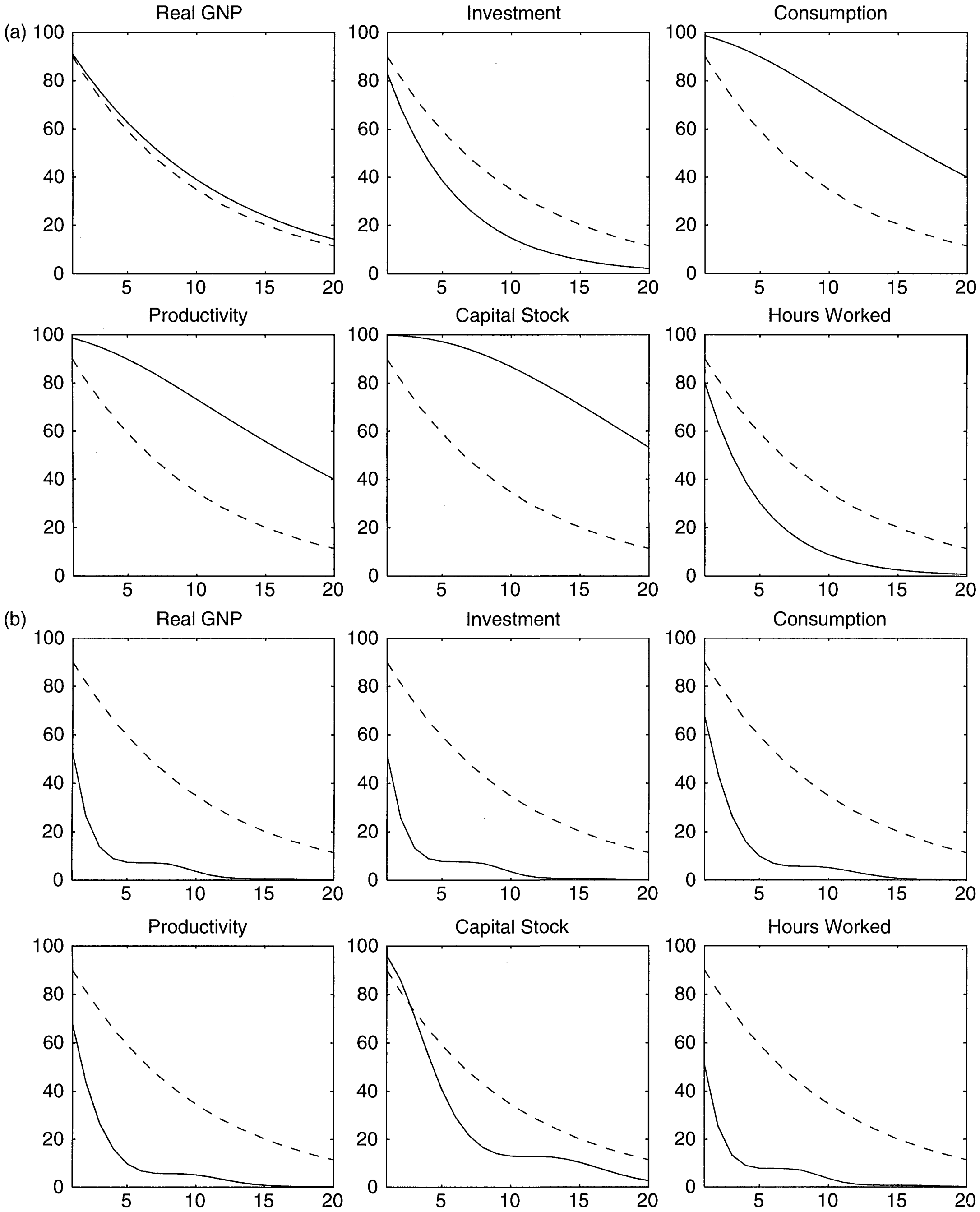


Figure 2. The propagation mechanism of the Hansen (1985) model. Model input (dashed line), model output (solid line). (a) Predictability in levels without filtering. (b) Predictability after logging and HP-filtering the model output data. Notes: For a description of the model and data see text. The plots are based on $k = 40$ and $j = 1, \ldots, 20$

 

perspectives different from predictability (e.g. King, Plosser and Rebelo, 1988; Singleton, 1988; Harvey and Jaeger, 1993; Cogley and Nason, 1995b). To explore that conjecture, we compute the predictability of the logged and HP-filtered technology shock, whose dynamics, of course, cannot have anything to do with the model's propagation mechanism. For illustrative purposes we consider six alternative values for the persistence of the technology shock, including the value of $\gamma = 0.95$ used in Hansen (1985). The result appears in Figure 3: For all realistic values of $\gamma$, the predictability of the HP-filtered technology shock looks the same as the predictability of the typical HP-filtered model output! We conclude that HP filtering tends to make predictabilities look similar, and thereby masks the strong propagation mechanism in Hansen's model that is revealed in unfiltered data.

We now compare the predictability of model data to that estimated using actual US data, 1955:III-1984:I. Computation of the predictability of the model data is not subject to sampling error, while the estimation of predictability of the real data is; thus, we compute interval estimates only for the latter. We consider the model to be consistent with the data if the model measure of predictability is contained in the confidence bands estimated from the U.S. data.

The real business cycle school favours logging and HP filtering both the model data and the real data, so we begin with that strategy. In Figure 4(a) we show the predictability of model data and actual data, where both have been HP filtered and logged. Disregarding some minor discrepancies, the model data and real data predictabilities generally agree. Only at very short time horizons the model data are not contained in the confidence bands for the US data. Of course, the relative success of the model may simply reflect the small sample size of the US data and the large degree

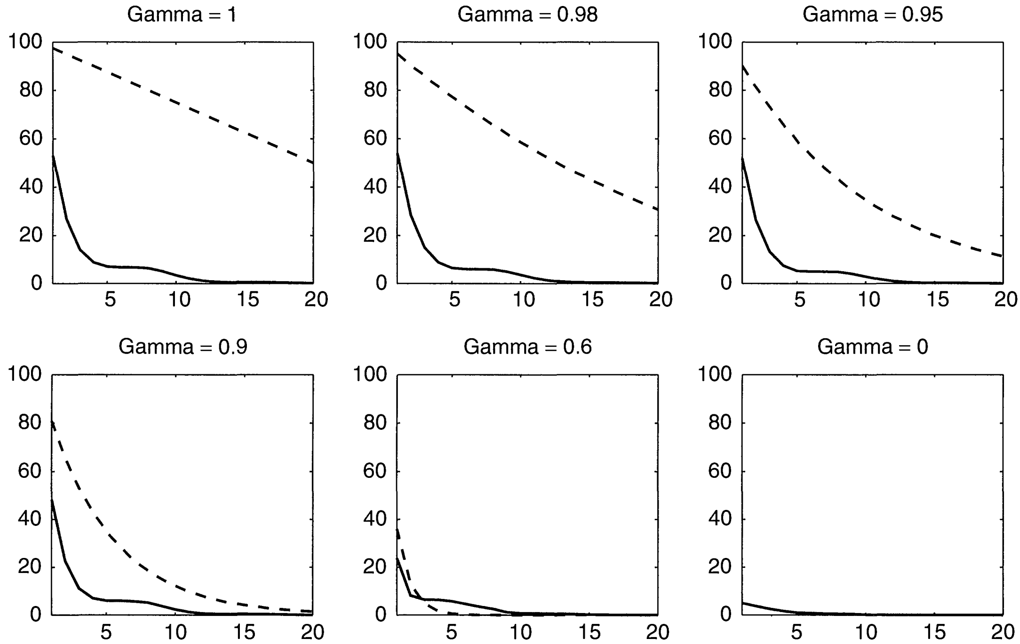


Figure 3. Predictability of the AR(1) technology shock raw data (dashed line), HP-filtered data (solid line): alternative degrees of persistence. Notes: For a description of the model see text. The plots are based on $k = 40$ and $j = 1, \ldots, 20$

 

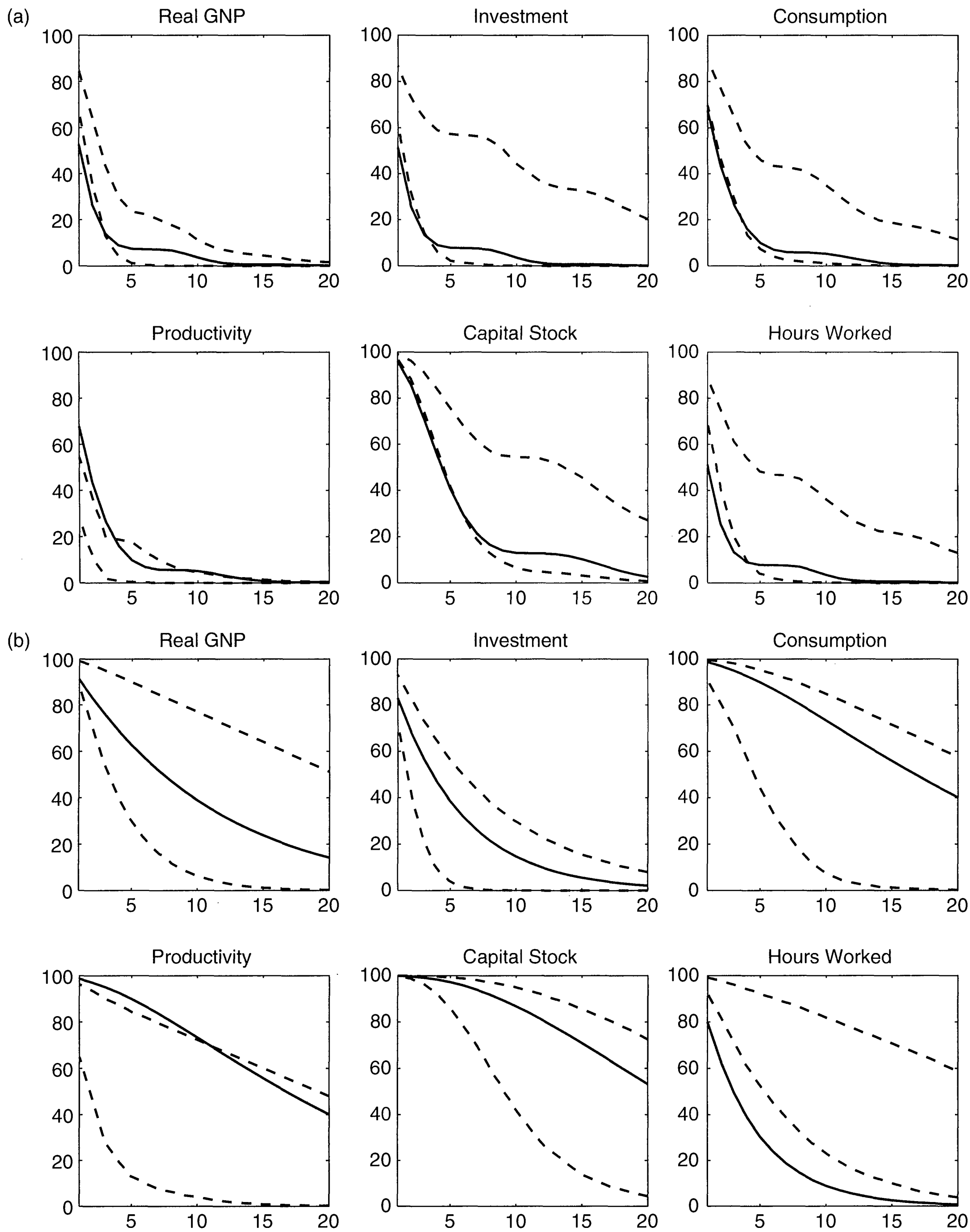


Figure 4. Model performance in terms of predictability. US data (dashed lines), model data (solid line). (a) All data logged and HP-filtered. (b) All data logged and linearly detrended. Notes: For a description of the model and data see text. The plots are based on $k = 40$ and $j = 1, \ldots, 20$. The dashed lines are the 90% confidence intervals for the US estimates

of sampling uncertainty. Furthermore, note that although the predictabilities of model data and actual data generally agree, they always have the same humped shape that the HP filter tends to impart, as we showed earlier. Thus we are naturally suspicious that the results may be an artifact of HP filtering. To address this possibility, we continue to treat the model data and real data symmetrically by using identical detrending procedures, but we use standard linear detrending instead of HP filtering.

In Figure 4(b) we show estimates of the predictability of model data and US data based on log-linear detrending. The predictabilities of model data and actual data summarized in Figure 4(b) show a pronounced divergence: only model output, investment, consumption, and capital stock match the data; hours worked in the model are not nearly predictable enough, whereas productivity in the model is too predictable for the first three years. This result is robust to whether a trend is fitted to the model data, or the population trend of zero is imposed. We conclude that the HP filter is a likely source of spurious fits between model data and actual data.

## 5. CONCLUDING REMARKS AND DIRECTIONS FOR FUTURE RESEARCH

We have proposed operational methods for measuring predictability and put them to work in several contexts—measuring the predictability of a variety of US macroeconomic series, assessing the internal propagation mechanism of a simple macroeconomic model, and assessing agreement between the model and the data. Our main intent is the introduction and illustration of an approach to predictability measurement, not the provision of a complete evaluation of a particular macroeconomic model. Nevertheless, our results reveal some successes and some failures of the model.

There are many useful directions for future research. Some are obvious, but nevertheless important, variations on the applications reported here. To take one example, in our applications we estimate predictability on the basis of univariate information sets, whereas the theory allows for multivariate information sets. Empirical predictability measurement on the basis of multivariate information sets, and comparison of the univariate estimates to various multivariate estimates, will be of interest. To take a second example, although deterministic linear trends often compare favorably to competitors in forecasting and macroeconometric studies (see DeJong and Whiteman, 1994; Diebold and Senhadji, 1996), it will be of interest to explore models with non-linear deterministic trends and models that allow for stochastic trends only. Other directions for future research are more wide-ranging and fundamental. We shall briefly discuss three.

### 5.1. Nonparametric Predictability Estimation

We presented our autoregressive modelling approach as a parametric method. This convention facilitated the exposition and allowed us to draw on established results for bootstrap inference. However, in general, we need not assume that the fitted autoregression is the true data-generating process; rather, it may be considered an approximation, the order of which can grow with sample size. Thus the autoregressive model can be viewed as a sieve in the sense of Grenander (1981), so our approach actually is nonparametric. For the stationary case, the asymptotic validity of bootstrapping predictability measures based on the sieve approximation follows from Inoue and Kilian (2000b).

 

Nevertheless, the sieve approach has a parametric flavour. For any fixed sample size, we assess predictability through the lens of a particular autoregressive model. In the future, it may be of interest to develop an approach with a more thoroughly non-parametric flavour by exploiting Kolmogorov's well-known spectral formula for the univariate innovation variance,

$$\sigma^2 = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln 2\pi f(\omega)\mathrm{d}\omega\right)$$

where $f$ is the spectral density function. Kolmogorov's result has been extended to univariate $h$-step-ahead forecast error variances by Bhansali (1992), and to multivariate $h$-step-ahead forecast error variances by Mohanty and Pourahmadi (1996). Several technical problems remain, however, before we can operationalize those methods in our context.

### 5.2. Predictability of Financial Asset Returns and Volatilities

We have focused on the application of predictability measurement to macroeconomics. Predictability statistics should also prove useful in finance, in which the predictability of asset returns at various horizons is a central concern. In fact, if non-predictability was arguably the central concern of the 1960s and 1970s literature (e.g. Fama, 1970), precisely the opposite is true of the more recent literature (e.g. Fama, 1991). We have reserved application to finance for a separate paper, in order to devote the necessary attention to multivariate information sets, long-horizon predictability, conditional heteroscedasticity, and possibly non-quadratic loss functions.[3]

Of equal, and perhaps even greater, importance is measuring the predictability of asset return *volatility* across various horizons. Tracking and forecasting time-varying volatility is central to financial risk management. Little attention, however, has been given to assessing volatility predictability patterns, and in particular the speed and pattern with which volatility predictability decays as the horizon grows.

### 5.3. Survey-based Predictability Estimation

We have taken a *model-based* approach to predictability measurement. Conditional upon a particular fitted model, we make inferences about predictability. Alternatively, we could take a *survey-based* approach, based on the predictions of competitive professional forecasters. Conditional upon the assumption that the reported forecasts are optimal, those data can be used for inferences about predictability. The survey-based approach is of interest because the information sets used by actual forecasters are likely much richer than simple univariate histories. They are surely multivariate, for example, and they also contain hard-to-quantify subjective information. The survey-based approach does rely on a crucial and disputable assumption (optimality of reported forecasts), but so too does the model-based approach (adequacy of the fitted model). The key point is that the assumptions made by the two approaches are different, and that the approaches therefore naturally complement one another.

A number of relevant surveys exist, including the former NBER-ASA Quarterly Economic Outlook Survey dating back to 1968 (and continued as the Survey of Professional Forecasters

[3] Mean absolute error (MAE), for example, may be more useful for measuring the accuracy of forecasts of financial series. In high-frequency financial data, for example, fat-tailed distributions are common, in which case MSE may be infinite but MAE often remains finite.

 

by the Federal Reserve Bank of Philadelphia) and the Blue Chip Indicators available from the early 1980s onward (see Croushore, 1993). These surveys focus on the major macroeconomic aggregates, such as real GDP growth. An interesting extension of this paper will be to use these forecasts to compute survey-based estimates of predictability, and to compare the survey-based and model-based estimates.

## ACKNOWLEDGMENTS

For helpful discussions and comments we thank the editor, three anonymous referees, Bob Barsky, Dean Croushore, Eric Ghysels, Gary Hansen, David Hendry, Houston McCulloch, Garey Ramey, Julio Rotemberg, Mark Watson, and Mike Woodford.

## REFERENCES

Bhansali RJ. 1992. Autoregressive estimation of the prediction mean squared error and an $R^2$ measure: an application. In *New Directions in Time Series Analysis I*, Brillinger D *et al.* (eds). Springer-Verlag: New York; 9–24.

Christoffersen PC, Diebold FX. 1996. Further results on forecasting and model selection under asymmetric loss. *Journal of Applied Econometrics* **11**: 561–572.

Christoffersen PC, Diebold FX. 1997. Optimal prediction under asymmetric loss. *Econometric Theory* **13**: 808–817.

Clements MP, Hendry DF. 1998. *Forecasting Economic Time Series*. Cambridge University Press: Cambridge.

Cogley T, Nason JM. 1995a. Output dynamics in real business cycle models. *American Economic Review* **85**: 492.

Cogley T, Nason JM. 1995b. Effects of the Hodrick–Prescott filter on trend and difference stationary time series: implications for business cycle research. *Journal of Economic Dynamics and Control* **9**: 253–278.

Croushore D. 1993. The survey of professional forecasters. *Business Review*, Federal Reserve Bank of Philadelphia, November–December.

DeJong D, Whiteman CH. 1994. The forecasting attributes of trend- and difference-stationary representations for macroeconomic time series. *Journal of Forecasting* **13**: 279–298.

Diebold FX, Kilian L. 1999. *Measuring predictability: theory and macroeconomic implications*. Working Paper No. 99-10, Department of Economics, University of Michigan.

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–265.

Diebold FX, Ohanian LE, Berkowitz J. 1998. Dynamic equilibrium economies: a framework for comparing models and data. *Review of Economic Studies* **65**: 433–451.

Diebold FX, Senhadji A. 1996. The uncertain unit root in real GNP: Comment. *American Economic Review* **86**: 1291–1298.

Fama EF. 1970. Efficient capital markets: a review of theory and empirical work. *Journal of Finance* **25**: 383–417.

Fama EF. 1991. Efficient capital markets: II. *Journal of Finance* **46**: 1575–1617.

Forni M, Reichlin L. 1998. Let's get real: a dynamic factor analytical approach to disaggregated business cycles. *Review of Economic Studies* **65**: 453–473.

Granger CWJ, Newbold P. 1986. *Forecasting Economic Time Series* (2nd edn). Academic Press: Orlando, FL.

Grenander U. 1981. *Abstract Inference*. John Wiley: New York.

Hannan EJ, Poskitt DS. 1988. Unit canonical correlations between future and past. *Annals of Statistics* **16**: 7894–790.

Hansen GD. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics* **16**: 309–327.

 

Harvey A, Jaeger A. 1993. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics* **8**: 231–248.
Inoue A, Kilian L. 2000a. Bootstrapping autoregressive processes with possible unit roots. *Econometrica* (forthcoming).
Inoue A, Kilian L. 2000b. Bootstrapping smooth functions of slope parameters and innovation variances in VAR($\infty$) models. *International Economic Review* (forthcoming).
Jewell NP, Bloomfield P. 1983. Canonical correlations of past and future for time series: definitions and theory. *Annals of Statistics* **11**: 837–847.
Kilian L. 1998a. Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics* **80**: 218–230.
Kilian L. 1998b. Confidence intervals for impulse responses under departures from normality. *Econometric Reviews* **17**: 1–29.
Kilian L. 1998c. Accounting for lag order uncertainty in autoregressions: the endogenous lag order bootstrap algorithm. *Journal of Time Series Analysis* **19**: 531–548.
Kilian L. 2001. Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting* **20**: 161–179.
King RG, Plosser CI, Rebelo ST. 1988. Production, growth and business cycles. I. The basic neoclassical model. *Journal of Monetary Economics* **21**: 195–232.
King RG, Watson MW. 1996. Money, prices, interest rates and the business cycle. *Review of Economics and Statistics* **78**: 35–53.
Mohanty R, Pourahmadi M. 1996. Estimation of the generalized prediction error variance of a multiple time series. *Journal of the American Statistical Association* **91**: 294–299.
Rotemberg JJ, Woodford M. 1996. Real business-cycle models and the forecastable movements in output, hours, and consumption. *American Economic Review* **86**: 71–89.
Singleton KJ. 1988. Econometric issues in the analysis of equilibrium business cycle models. *Journal of Monetary Economics* **21**: 361–386.
Stock JH. 1995. Unit roots and trend breaks. In *Handbook of Econometrics Volume 4*, Engle R, McFadden D (eds). North-Holland: Amsterdam.
Theil H. 1966. *Applied Economic Forecasting*. North-Holland: Amsterdam.