CIRANO

Centre interuniversitaire de recherche en analyse des organisations

Série Scientifique
Scientific Series

99s-17

Content Horizons for Forecasts of Economic Time Series

John W. Galbraith

Montréal
Avril 1999

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche. La *Série Scientifique* est la réalisation d'une des missions que s'est données le CIRANO, soit de développer l'analyse scientifique des organisations et des comportements stratégiques.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams. The Scientific Series *fulfils one of the missions of CIRANO: to develop the scientific analysis of organizations and strategic behaviour.*

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan Aluminium Ltée
- Banque Nationale du Canada
- Bell Québec
- Développement des ressources humaines Canada (DRHC)
- Egis
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Microcell Labs inc.
- Raymond Chabot Grant Thornton
- Téléglobe Canada
- Ville de Montréal

ISSN 1198-8177

Content Horizons for Forecasts of Economic Time Series*

John W. Galbraith†

Résumé / Abstract

Nous considérons la détermination de l'horizon après lequel les prévisions provenant des modèles des series chronologiques stationnares n'ajoutent rien à la valeur de la prévision implicite dans la moyenne. Nous appellons cette quantité le « content horizon » pour prévisions, et nous définissons la fonction de valeur ajoutée aux horizons $s = 1, \dots S$ par la réduction proportionnelle dans la moyenne des erreurs de prévisions carrées disponible en utilisant une prévision provenant d'un modèle formel relatif à la moyenne non-conditionelle. Cette quantité dépend de l'incertitude dans les estimés des paramètres du modèle, ainsi que des autocorrélations du processus considéré. Nous donnons une expression approximative – jusqu'à $o(T^{-1})$ – pour la fonction de valeur ajoutée à s pour les processus autorégressifs généraux, et nous démontrons par simulation que l'expression est bonne même dans les petits échantillons. Enfin nous considérons les estimés paramétriques et non-paramétriques (kernel) pour la fonction de valeur ajoutée empirique, en appliquant les résultats aux horizons de prévision pour le taux de croissance du PNB et le taux d'inflation, au Canada et aux États-Unis.

We consider the problem of determining the horizon beyond which forecasts from time series models of stationary processes add nothing to the forecast implicit in the conditional mean. We refer to this as the content horizon for forecasts, and define a forecast content function at horizons $s = 1, \dots S$ *as the proportionate reduction in mean squared forecast error available from a time series forecast relative to the unconditional mean. This function depends upon parameter estimation uncertainty as well as on autocorrelation structure of the process under investigation. We give an approximate expression – to* $o(T^{-1})$ *– for the forecast content function at* s *for a general autoregressive processes, and show by simulation that the expression gives a good approximation even at modest sample sizes. Finally we consider parametric and non-parametric (kernel) estimators of the empirical forecast content function, and apply the results to forecast horizons for inflation and the growth rate of GDP, in U.S. and Canadian data.*

---

* Corresponding Author : John W. Galbraith, CIRANO, 2020 University Street, 25th floor, Montréal, Qc, Canada H3A 2A5 Tel.: (514) 985-4008 Fax: (514) 985-4039 email: galbraij@cirano.umontreal.ca
Thanks to Christoph Schleicher for valuable research assistance, Lutz Kilian, Claude Nadeau, Greg Tkacz and Victoria Zinde-Walsh for insightful comments, and the Fonds pour la formation de chercheurs et l'aide à la recherche (Québec) and Social Sciences and Humanities Research Council of Canada for financial support of this research.

† McGill University and CIRANO

Mots Clés : Processus autorégressif, horizon de prévision, PNB, taux d'inflation

Keywords : Autoregressive Process, forecast horizon, GDP, inflation

JEL : C12, C22

1 Introduction

Short-range meteorological forecasts of daily deviations from seasonal norms of temperature and precipitation are often characterized as containing useful information up to horizons of approximately ten days into the future (somewhat less for precipitation). Longer-range forecasts of monthly or quarterly average temperatures are useful up to about three quarters (quarterly averages) or eleven months (monthly averages); see for example Wilks (1996) and Colucci and Baumhefner (1992). Such information is valuable as a guide to appropriate use of forecasts, as a benchmark for developers of refined forecasting methods, and as a check on charlatanism.

Little information of this type exists for economic variables, however. In the present paper we attempt a systematic study of the horizons within which we can usefully forecast stationary transformations of economic variables. We begin by defining forecast content for such variables relative to the information content of the estimated unconditional mean, and define a forecast content function using relative mean squared error as a function of the forecast horizon. We give approximate analytical expressions whereby the forecast content function can be computed for a general AR(p) process at a forecast horizon of s periods into the future. For comparison and evaluation of the analytical expression, we obtain the exact functions by simulation for particular AR processes and sample sizes.

Empirical estimation of the forecast content function, as opposed to its evaluation for a known process, can be carried out via substitution of estimated parameter values into

the parametric expressions just described, or by non-parametric regression. We describe both methods and apply them here to data from Canada and the U.S. on both real GDP growth (measured quarterly) and inflation (in the Consumer Price Index, measured monthly). We characterize the forecast content functions and forecast content horizons–spans of time beyond which we are not able to make useful forecasts of data at a given frequency beyond use of the unconditional mean–for each of these time series. We find content horizons of about two quarters for real GDP growth, and twenty-four (Canada) to forty (U.S.) months for inflation. [1] The functions and horizons obtained based on the analytical results are compared with those obtained through non-parametric

[1] Of course, these content horizons can in principle be extended by incorporating information on variables other than the past of the variable to be forecast, evaluating the chosen forecasting model using the non-parametric method, but we do not pursue such alternatives in the present paper.

estimation.

In section 2 we give definitions and analytical expressions for the time series forecast content function and corresponding content horizons. Section 3 provides estimates from simulations of the forecast content functions for a variety of time series processes, compares the results with the analytical results, and investigates robustness to heavy-tailed and skewed error distributions. Section 4 examines parametric (based on the analytical expressions of Section 2) and non-parametric estimators of the empirical forecast content function, and Section 5 applies these results to estimation of the forecast content functions and horizons for the growth rate of real GDP and inflation.

2 The forecast content function and content horizon

2.1 Definitions

Let $\{y_t\}_{t=1}^T$ be a sequence of T observations on a stationary, ergodic process y. All moments of y are assumed to be unknown. Our aim is to forecast the value $y_{T+s},\ s>0$, using the observations $\{y_t\}_{t=1}^T$. Let this forecast be $\tilde{y}_{T+s|T}$, so that the mean squared forecast error is $E(\tilde{y}_{T+s|T}-y_{T+s})^2$.

The sample mean $\overline{y}=T^{-1}\sum_{t=1}^T y_t$ provides another possible forecast of any future value, with mean squared forecast error $E(\overline{y}_T-y_{T+s})^2$. For large s, information about the past of y will not improve on the forecast implicit in the unconditional mean. We will evaluate forecasts relative to this forecast from the unconditional mean, and will define a forecast $\tilde{y}_{T+s|T}$ as having positive *content* if $E(\tilde{y}_{T+s|T}-y_{T+s})^2 < E(\overline{y}_T-y_{T+s})^2$, or $MSE_{\tilde{y}} < MSE_{\overline{y}}$. [2] For a set of forecasts at different horizons, $\{\tilde{y}_{T+s|T}\}_{s=1}^S$, we define the *forecast content function* as the proportionate reduction in mean squared forecast error available relative to the unconditional mean forecast: that is,

$$C(s) = 1 - \frac{MSE_{\tilde{y}(s)}}{MSE_{\overline{y}(s)}}, \qquad s = 1,\ldots,S. \tag{2.1}$$

[2] Throughout the present paper we will confine ourselves to the quadratic loss case, although content could be defined relative to a general loss function, since any alternative loss function can be substituted for the MSE in (2.1). While a different loss function will in general imply a different optimal predictor, the content horizon may not be greatly affected, although numerical values of the forecast content function will change.

For the stationary, ergodic processes considered here, $C(s) \to 0$ as T and $s \to \infty$, for a forecast $\tilde{y}_{T+s|T}$ based on a correct specification of the forecasting model; our ability to improve upon the unconditional mean as a forecast disappears with lengthening forecast horizon. For fixed T, however, $C(s)$ may be less than zero for values of s such that extra parameters beyond the mean make negligible contribution to forecasting, but raise MSE through parameter uncertainty.

Related definitions have been made by numerous authors, including Granger and Newbold (1977), Box and Tiao (1977), Bhansali (1991), Diebold and Kilian (1997), and others. Bhansali (1991), Box and Tiao (1977), do so in defining a measure of R^2 for time series processes; the measure of Granger and Newbold (1977) is similar. However, here we do not measure unexplained variance relative to true variance of the process, but instead forecast MSE relative to the variance of the forecast implicit in the unconditional mean, which incorporates not only process variance and the variance of the estimate of the mean, but also parameter estimation uncertainty. In each case one obtains a measure which can in principle be computed *a priori* from the form of the process, for any forecast horizon. Diebold and Kilian (1997) offer a general framework for defining predictability measures from which the other definitions, including that which we will examine here, can be obtained. A survey of the various related definitions, as well as applications for the quadratic loss case such as computations of relative predictability at different horizons, can also be found in Diebold and Kilian.

Finally, we define the *content horizon* as the forecast horizon s_0 such that $C(s) \leq 0$ for $s \geq s_0$: that is, the point beyond which forecasts based on the unconditional mean are no worse than those obtained from explicit forecasting models. In cases the forecast content horizon may remain strictly positive even for large s, while approaching zero asymptotically in T and s, and in such cases it may be convenient to refer to the $\delta-$level content horizon, that is, s_δ such that $C(s) \geq \delta, s \leq s_\delta$ and $C(s) < \delta, s > s_\delta$.

It is important to note that these forecast content functions and content horizons are specific to a given time interval between observations. For example, a forecast content horizon of one period in annual data does not imply a forecast content horizon of twelve in monthly data; the two data generation processes must be examined separately.

In the next section, we examine the forecast content functions of some parametric time series models.

2.2 Analytical forecast content functions for autoregressive processes

Pure autoregressions serve as relatively good forecasting models for a variety of processes. It is often noted, in fact, that it can be difficult to better the forecast performance of an AR model even with much more elaborate forecasting structures; see for example Meese and Geweke (1984) or Stock and Watson (1998). Stock and Watson find the autoregression the best overall method within the set of AR, artificial neural network, smooth-transition autoregression and exponential smoothing methods, in a study of forecast performance on approximately two hundred macroeconomic time series. [3] Since (in both Stock and Watson and the present study) variable lag order is used to find an approximation, and chosen via an information criterion, the AR model is being used in what might be called a non-parametric fashion. Nonetheless, we will refer to forecast content functions from such AR models as parametric, reflecting the fact that our computation of the forecast content function uses the form of the AR structure.

Consider a stationary AR(p) process $\alpha(L)y_t = \varepsilon_t$, with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. With known parameters, forecast error variance at any horizon s depends only on the horizon and the parameters of the polynomial $\alpha(L)$. As $s \to \infty$, this variance (and therefore the forecast MSE) converges to the unconditional variance, which is the MSE of the forecast implied by the unconditional mean. When parameters are unknown, the forecast MSE depends as well on the uncertainty in parameter estimates, and therefore on sample size. [4] The same is true of the unconditional mean forecast, which depends on the uncertainty in a single estimated parameter.

In order to obtain the forecast content function analytically, we need expressions for the mean squared errors of both types of forecast. For the sample mean, we can obtain asymptotically the necessary expressions from a central limit theorem for processes with dependence of a form compatible with these processes. For forecasts based on the autoregressive model with normal errors, expressions for the forecast MSE accurate to $O(T^{-\frac{3}{2}})$ are given by Fuller and Hasza (1981). [5] Without the assumption of Normal errors, results accurate to $O(T^{-1})$ are available. In Proposition 1 we combine these to obtain an analytical ex-

[3] Unit root pre-tests are applied before AR model estimation, a step not explicitly used here because of our assumption that a stationary transformation has been found.

[4] See Sampson (1991) on the importance of parameter uncertainty in obtaining confidence intervals for forecasts of processes with deterministic or stochastic trends.

[5] See Samaranayake and Hasza (1988), Lütkepohl (1993) for multivariate expressions; Ericsson and Marquez (1998) provide a general review of forecast MSE computation and related points such as potential non-monotonicity in the horizon.

pression for the forecast content function of a correctly- specified AR(p) model estimated on a sample of size T. Forecasts are made according to $\hat{y}_{T+s|T} = \hat{\alpha}_0 + \sum_{j=1}^{p} \hat{\alpha}_j \hat{y}_{T+s-j|T}$, with $\hat{y}_{T+s-j|T} = y_t$ if $s - j \leq 0$. Following Fuller and Hasza we define

$$Y_{T+s} = \begin{pmatrix} y_{T+s} \\ y_{T+s-1} \\ y_{T+s-2} \\ \vdots \\ y_{T+s-p} \\ 1 \end{pmatrix}.$$

Proposition 1. Let the AR(p) process $\alpha(L)y_t = \varepsilon_t$ be such that all roots of $\alpha(L)$ are outside the unit circle, and let $\varepsilon_t \sim IN(0, \sigma_\varepsilon^2)$. Then

$$C(s) = 1 - \frac{B_{[1,1]}}{\nu} + o(T^{-1}), \tag{2.2.1a}$$

where $B_{[1,1]}$ is the top-left element of the matrix

$$B = \sum_{j=0}^{s-1} A^j M A'^j + T^{-1} \sum_{j=0}^{s-1} \sum_{k=0}^{s-1} A^j M A'^k \cdot \mathrm{tr}[(A^{s-j-1},)'(, ^{-1} A^{s-k-1})], \tag{2.2.1b}$$

$$\nu = \sum_{i=0}^{\infty} a_i^2 \left(1 - 2T^{-1} \sum_{\ell=s}^{T+s-1} \rho(\ell)\right) + T^{-1} \left(\sum_{i=0}^{\infty} a_i\right)^2, \tag{2.2.1c}$$

and where M is the $(p+1) \times (p+1)$ matrix with 1 in the upper-left corner and zeroes elsewhere, , $= E(Y_t Y_t')$, $\rho(\ell)$ is the autocorrelation function at lag ℓ, and A is the $(p + 1) \times (p + 1)$ matrix such that $Y_t = AY_{t-1} + \epsilon_t$, with $\epsilon_t = (\varepsilon_t, 0, \ldots 0)'$. [6] The sequence of coefficients $\{a_i\}_{i=0}^{\infty}$ is the set of coefficients of the innovations representation of y, that is,

$$y_t = \sum_{i=0}^{\infty} a_i \varepsilon_{t-i}, \tag{2.2.2}$$

with $a_0 = 1$.

Proof.

The matrix B of (2.2.1) is the expression for the s-step-ahead forecast MSE of Y from the autoregressive predictor given by Fuller and Hasza

[6]The form of A is given explicitly in Fuller and Hasza (1981); the series $\sum_{i=0}^{\infty} a_i$ is convergent by the stationarity of y.

(1981, corollary 2.1), up to a factor of σ_ε^2; the top-left element, $B_{[1,1]}$, is therefore the corresponding forecast MSE for y_{T+s}. The scalar ν is the asymptotic mean squared error of the unconditional mean forecast, that is, $\overline{y}_T$ used as a forecast of y_{T+s}; therefore $\nu = E(\overline{y}_T - y_{T+s})^2 =$

$$E[(\overline{y}_T - \mu) - (y_{T+S} - \mu)]^2 = \ \mathrm{var}(\overline{y}_T) + \ \mathrm{var}(y_{T+s}) - 2\ \mathrm{cov}(\overline{y}_T, y_{T+s}),$$

where μ is the unconditional mean of Y. The second term in ν is the unconditional variance of the process y, equal to $\sigma_\varepsilon^2 \sum_{i=0}^{\infty} a_i^2$, and the first term is the asymptotic variance of the unconditional mean, $\sigma_\varepsilon^2[T^{-1}(\sum_{i=0}^{\infty} a_i)^2 + o(T^{-1})]$, from a central limit theorem for processes with dependence of the form in (2.2.2); see for example Theorem 6.3.3 of Fuller (1976) or Theorem 7.7.9 of Anderson (1971). The covariance term $-2\ \mathrm{cov}(y_{T+s}, \overline{y}_T)$ is equal to $-2T^{-1}\sum_{j=0}^{T-1}\gamma(j+s)$, where $\gamma(.)$ is the autocovariance function. Dividing by $\gamma(0)$ and removing the factor of σ_ε^2, the factors of σ_ε^2 in B and ν cancel from numerator and denominator of in (2.1) to give $C(s) = 1 - \frac{B_{[1,1]}+O(T^{-\frac{3}{2}})}{\nu+o(T^{-1})}$; removing a factor of ν^{-1} from the denominator and expanding, we obtain (2.2.1a). ■

The covariance term in ν converges to zero in both s and T, and can be omitted for moderately large samples. We examine the impact of omitting this term below.

The coefficients of the innovations representation of a general AR(p) or ARMA(p,q) process can be computed straightforwardly using the well-known recursions given by, for example, Fuller (1976: Theorems 2.6.1, 2.6.2). [7] The expressions (2.2.1) can then be used directly for computation, truncating infinite sums at a high value. In low-order cases it is also straightforward to simplify the expressions, as in the following.

Example. For the AR(1) process with mean zero and parameter α, $a_i = \alpha^i,\ A = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix},\ M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ ,\ = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & 1 \end{bmatrix}$, and from (2.2.1) the forecast content function to o(T^{-1}) is:

[7] Note that the forecast content function for a stationary, invertible ARMA process can also be obtained using an autoregressive approximation to the ARMA process, as in Galbraith and Zilde-Walsh (1997), in the numerator of (2.2.1).

$$
\begin{aligned}
C(s) &= 1 - \left[\frac{\sum_{j=0}^{s-1} \alpha^{2j} + T^{-1} \sum_{j=0}^{s-1} \sum_{k=0}^{s-1} \alpha^{j+k}(1+\alpha^{(2s-j-k-2)})}{(\sum_{i=0}^{\infty} \alpha^{2i})(1-2T^{-1}\sum_{\ell=s}^{T+s-1} \rho(\ell)) + T^{-1}(\sum_{i=0}^{\infty} \alpha^{i})^2)} \right] \\
&= 1 - \left[\frac{(1-\alpha^{2s})(1-\alpha^2)^{-1} + T^{-1} \sum_{j=0}^{s-1} \sum_{k=0}^{s-1} \alpha^{j+k}(1+\alpha^{(2s-j-k-2)})}{(1-\alpha^2)^{-1}\left[1-2T^{-1}\alpha^s(\frac{\alpha^T-1}{\alpha-1})\right] + T^{-1}(1-\alpha)^{-2}} \right].
\end{aligned}
\tag{2.2.3}
$$

For a model which is correctly specified in the sense that the true process is autoregressive of order $p^* \leq p$, consistency of the estimated forecast content function follows from consistency of the parameter estimates (from, for example, OLS or Yule-Walker estimators); however, it is the finite sample performance of the estimates of $C(s)$, embodying parameter estimation error, which are of greater interest. We examine this in section 3.

3 Simulated forecast content functions

In this section we compare the forecast content functions from the analytical expressions above with simulated functions, and also consider simulated functions for a number of cases not directly covered by Proposition 1, including heavy-tailed and skewed error distributions.

3.1 Analytical results vs. finite-sample results from simulation

Figure 1 (a/b/c/d) compares the forecast content function for $s = 1, \ldots 20$ computed from expression (2.2.1), or (2.2.3) in AR(1) cases, with a simulation estimate of the exact function, for specific parameter values in AR(1) (Figure 1 a/b) and AR(2) (Figure 1 c/d) processes. [8] The simulated forecast content functions are obtained from 50,000 replications of experiments with normally-distributed pseudo-random errors, in which mean squared errors for forecasts based on both the AR model and the in-sample mean are computed for each s and for sample sizes $T = \{100, 400\}$. In the AR(1) case, we use roots of 0.4 and 0.8. In the AR(2) case, we use roots of $0.8, 0.3$ and $0.8, -0.5$, corresponding to parameters

[8] The infinite sums in (2.2.1c) are truncated at 50.

(α_1, α_2) of $1.1, -0.24$ and $0.3, 0.4$. Results for AR(1) roots are approximately, but not exactly, symmetric around zero. A further approximation is examined in which the covariance term $-2\ \mathrm{cov}(y_{T+s}, \overline{y}_T) = -2T^{-1}\sum_{j=0}^{T-1}\gamma(j+s)$, which appears in (2.2.1c) scaled by the process variance, is omitted; This term, while only $O(T^{-1})$, tends to be small at moderately large sample sizes.

There are several noteworthy features. First, the analytical results from (2.2.1/2.2.3), while exact only asymptotically, provide a very good approximation even at $T = 100$; the analytical and simulated forecast content functions are difficult to distinguish visually. This holds true in both AR(1) and AR(2) cases. With the covariance term omitted ("ex-cov" in the Figures), the analytical form gives, a good approximation at $T = 100$, but the approximation is clearly distinguishable visually from the 'true' function. At $T = 400$ the simulated and analytical functions are difficult to distinguish visually whether or not $\mathrm{cov}(y_{T+s}, \overline{y}_T)$ is incorporated; for clarity only the "ex-cov" form is recorded.

Second (a feature not easy to read from the graphical results presented here), the true forecast content functions obtained by simulation can take on negative values at long horizons. This possibility arises because the forecast based on the sample mean requires estimation of only one parameter, while a model using an autoregression or other multiparameter model for forecasting estimates the additional parameters of the model as well. The extra parameter or parameters lead to slightly higher variance, which is more than offset by the information content in the past of the process of interest for small s; however, for large s, the information content can be sufficiently low that the extra parameters dominate and leave the forecast content function slightly negative. For example, in the AR(2) case with roots of 0.8 and 0.3, the forecast content function at $T = 100,\ s = 20$ is -9×10^{-3}; at $T = 400,\ s = 20$ it is -2×10^{-4}.

It is easy to see that the forecast content is roughly zero by lag 10 for each of the cases considered here, despite the fact that three of the four examples contain a fairly large root, of 0.8. The 5% forecast content horizons, that is, the horizons beyond which forecast content is strictly less than .05, are (for $T = 400$) $s = 1$ and $s = 6$ for AR(1) roots of 0.4 and 0.8 respectively, $s = 7$ and $s = 5$ for AR(2) roots of $0.8, 0.3$ and $0.8, -0.5$ respectively.

3.2 Impact of some non-Normal error distributions

We now consider the adequacy of the forecast content functions obtained via the expression (2.1.1) using the Normal distribution, where errors are in fact heavy-tailed or skewed. To do so we repeat the simulations of section 3.1, making a comparison of the simulated forecast content functions with Normal pseudo-random errors (well approximated by (2.1.1)) with those applying where errors have $t-$ or χ^2- distributions. The simulations again use 50,000 replications.

For heavy-tailed errors we draw from the t_4 distribution. The t_4 has variance of 2; its 99th and 99.5th percentiles are 3.747 and 4.604. For the Normal distribution with mean zero and variance of 2, the 99th and 99.5th percentiles are 3.289 and 3.643. For skewed errors, we use the χ^2_8, which has median 7.34 (versus mean of 8), with 1st and 99th percentiles of 1.65 and 20.1; the χ^2_8 errors are re-scaled to mean zero here. [9]

The resulting forecast content functions are in some cases difficult to distinguish visually from those pertaining to Normal errors, and so are presented in Table 1 rather than graphically, using the AR(2) models of Figure 1c/d for comparison; AR(1) results show similarly small discrepancies. Since the discrepancies relative to the Normal case are similar for $T = 100$ and 400, we present only $T = 100$.

Table 1

Simulated Forecast Content Functions, Normal and non-Normal errors
AR(2) cases, $T = 100$

α_i	s	$\mathrm{N}(0,\sigma^2)$	t_4	χ^2_8
0.8,0.3	1	0.794	0.796	0.789
0.8,0.3	2	0.545	0.548	0.534
0.8,0.3	3	0.351	0.353	0.341
0.8,0.3	4	0.214	0.218	0.211
0.8,0.3	5	0.129	0.132	0.127
0.8,0.3	6	0.075	0.077	0.079
0.8,-0.5	1	0.352	0.355	0.344
0.8,-0.5	2	0.303	0.306	0.296
0.8,-0.5	3	0.149	0.150	0.148
0.8,-0.5	4	0.103	0.107	0.106
0.8,-0.5	5	0.060	0.061	0.058
0.8,-0.5	6	0.036	0.036	0.039

[9] Results are similar for neighbouring values of these degrees of freedom parameters.

Clearly the forecast content functions computed via (2.1.1) are robust to substantial departures from Normality and symmetry. Of the two alternative distributions, the χ^2_8 produces the larger discrepancies.

4 Parametric and non-parametric estimation of forecast content functions

For an AR process with given parameters, the forecast content function can be obtained from (2.1.1). In an empirical context where the parameters are unknown, the function may be estimated parametrically if an AR (or, by approximation, ARMA) process is used, or non-parametrically. It is important to distinguish estimation of the forecast content function from estimation of the underlying forecasting model; non-parametric estimation of the forecast content function can be carried out for a forecasting model which will itself usually be parametric. This will be the case in the empirical study of section 5, where the forecast content function for AR(p) forecasting models will be estimated both parametrically and non-parametrically.

Parametric estimation may be carried out by substituting consistent estimates of the $p+1$ parameters (mean and autoregressive terms), $\hat{\alpha}$, into the matrix A of (2.1.1), and proceeding as with known autoregressive parameters to compute C(s). Whether or not the forecasting model is autoregressive, the forecast content function can also be evaluated without imposing the form of forecast content function implied by an autoregressive forecasting model. We refer to such estimation as non-parametric, and will carry it out using kernel regression techniques (see, for example, Härdle, 1990, for an introduction); other non-parametric techniques, as well as purely unsmoothed estimates, may also be used.

The estimation problem here has the feature that the function of interest (the forecast content function) will have a non-zero slope for low values of s, except in the degenerate case in which forecast content is zero at all horizons. For this reason, the problem of bias near a boundary that occurs in standard kernel regression estimators, such as that of Nadaraya and Watson, will be important. We therefore do not use such locally-constant kernel regression estimators, but instead the locally-linear kernel estimator examined by, among others, Fan (1992, 1993). This kernel regression estimator in effect extends the traditional locally-constant approach by incorporating an additional term in the numerator and denominator of the expression which yields the estimated regression function, with the result that the bias is not dependent on the derivatives of the marginal density.

To obtain these non-parametric forecast content function estimates, we use the following sequence of steps:

(i) estimate the parameters of the forecasting model on the full sample;

(ii) using the parameterized forecasting model, forecast each in-sample observation $t > t_0 = p + s + 1$, at each horizon $s = \{1, \ldots S\}$, (where p is the number of lags required by the forecasting model);

(iii) from the set of forecasts, obtain the sequence of pairs $\{z(s,t),\ s\}$, $s = 1, \ldots S,\ t = t_0, \ldots T$, where

$$z(s,t) = \left(\frac{(\hat{y}_{t|t-s} - y_t)^2}{(T - t_0 + 1)^{-1} \sum_{t=t_0}^{T} (\overline{y} - y_t)^2} \right); \qquad (4.1.1)$$

(iv) compute the locally linear regression of the set of first elements z_i on the fixed grid of regressors given by the set of second elements $(s = 1, \ldots S)$, and subtract from a corresponding $S \times 1$ vector of ones.

The term $z(s,t)$ represents the squared error of the model-based forecast for s and t relative to the overall MSE of mean-based forecasts; the final step performs a non-parametric regression of this quantity on the forecast horizon. The denominator of the expression (4.1.1), the average MSE from forecasts based on the mean given the same set of observations used in computing the numerator, is used in order to avoid occasional division by very small values, and the attendant variability.

Full-sample parameter estimates are used since we are evaluating the forecast content function at the existing full sample, not at points in the past. In the limit as bandwidth approaches zero (no smoothing is applied), step (iv) is equivalent to plotting the average over t at each s of the first terms in (4.1.1), independently of the averages at other values of s. We refer to the latter estimate of the function as the unsmoothed non-parametric estimate. In other words, we may estimate the conditional expectation of the forecast content at a given forecast horizon without smoothing by averaging sample values of forecast content at that horizon; alternatively, we may smooth the estimates by taking into account the estimated forecast content at neighbouring values of s as well.

In the next section, we estimate forecast content functions for some monthly and quarterly time series using both parametric and non-parametric methods.

5 Forecast content functions for quarterly real GDP growth and inflation

We now consider estimated forecast content functions for four data series which can be transformed to approximate stationarity by computing growth rates: price level and real GDP, in Canada and the U.S. [10]

For each of the four transformed time series (inflation and real GDP growth) we estimate a sequence of autoregressive models, and select a model using the Schwarz information criterion; this model is then used directly for forecasting, and its parameters are used to produce the forecast content function by (2.2.1). The forecast content function obtained from the parameters is then compared with the same function estimated non-parametrically using a locally-linear kernel method with Gaussian kernel. [11] Again, because we are interested in obtaining the forecast content function that applies at the size of sample now available, full-sample parameter estimates are used in the computations. The results appear in Figures 2a/b and 3a/b; the model orders selected by the SIC are 2 and 1 for real GDP growth in Canada and the U.S., 7 and 10 for inflation in Canada and the U.S.

Figure 2 describes real GDP growth forecasts. In the U.S. data, parametric (from (2.2.1)) and non-parametric forecast content functions correspond very closely; in Canadian data, the direct estimates become slightly negative beyond two quarters, but are more substantial in the first quarter. On both Canadian and U.S. data, forecast content is virtually zero beyond two quarters.

Figure 3 describes monthly inflation forecasts. On both samples there is a fairly close match between parametric and non-parametric estimates, although the non-parametric estimates tend to be somewhat more optimistic. The forecast content horizon extends to about twenty-four months in Canada, and to about forty (only twenty-four quarters are reproduced in Figure 3b) in the U.S.

In comparing the parametric and non-parametric estimates, note that

[10] Quarterly Canadian data on seasonally adjusted real GDP are available from 1947:1 as series D20463 in CANSIM; US seasonally-adjusted real GDP are available from 1947:1 from the Survey of Current Business, Table 2A. Monthly Canadian CPI data (all items, not seasonally adjusted) are available from 1914:1 as P700000 in CANSIM; US CPI data from 1946:1 are taken from the Federal Reserve Bank of St. Louis FRED database.

[11] Forecast content functions estimated by the standard Nadaraya-Watson (locally constant) kernel regression are not shown because of the boundary problem noted in section 3: in each of the cases examined here, they differ from the locally-linear estimates in that the estimated regression function lies substantially below the locally linear estimate for small s.

the parametric estimates are based on the smaller number of parameters (the autoregressive parameters estimates alone) than the non-parametric estimates, for which the forecast content at each horizon must be estimated as a separate quantity based on a separate sequence of computed forecasts. The parametric quantities show a correspondingly greater smoothness (visible in Figure 3), are much less cumbersome to compute, and, where the model class used in Proposition 1 provides a good approximation to the process, will be more efficient. The non-parametric estimates, however, do not use information about model form and may be computed for any forecasting model, parametric or non-parametric, requiring only a realized set of in-sample forecasts. In these examples the two types of estimate produce similar results, although on inflation data the non-parametric estimates of forecast content tend to be somewhat more optimistic at intermediate values of s.

These forecast content functions are based purely on information which can be extracted from the past of the process of interest by linear models. Of course, we would expect that in many cases the forecast content can be raised, especially at short horizons, by incorporating appropriate additional explanatory variables. The forecast content from the pure time series model provides an overall lower bound on forecast content and a base against which to evaluate the potential of alternative models to extend the forecast horizon.

6 Concluding remarks

Forecast content and the horizon beyond which there is no content–that is, the model produces forecasts no better than the unconditional mean–can be estimated for purely autoregressive forecasting models using only the AR coefficients, and for general forecasting models can be estimated by non-parametric methods.

Such results provide benchmarks for interpretation of published forecasts; the application of the methods here to macro-economic time series suggest forecast horizons which are shorter than the maximum horizon provided in some published forecasts, suggesting that such forecasts may be of little value.

References

[1] Anderson, T.W. (1971) *The Statistical Analysis of Time Series.* Wiley, New York.

[2] Bhansali, R.J. (1991) Autoregressive estimation of the prediction MSE and an R^2 measure: an application. In Brillinger et al., eds., *New Directions in Time Series Analysis, Part I.* Springer-Verlag, New York, 9–24.

[3] Box, G.E.P. and G.C. Tiao (1977) A canonical analysis of multiple time series. *Biometrika* 64, 355–365.

[4] Colucci, S.J. and D.P. Baumhefner (1992) Initial weather regimes as predictors of numerical 30-day mean forecast accuracy. *Journal of the Atmospheric Sciences* 49, 1652–1671.

[5] Diebold, F.X. and L. Kilian (1997) Measuring predictability: theory and macroeconomic applications. NBER technical working paper 213.

[6] Ericsson, N..R. and J. Marquez (1998) A framework for economic forecasting. *Econometrics Journal* 1, C228–C266.

[7] Fan, J. (1992) Design-adaptive nonparametric regression. *Journal of the American Statistical Association* 87, 998–1004.

[8] Fan, J. (1993) Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics* 21, 196–216.

[9] Fuller, W.A. (1976) *Introduction to Statistical Time Series.* Wiley, New York.

[10] Fuller, W.A. and D.P. Hasza (1981) Properties of predictors for autoregressive time series. *Journal of the American Statistical Association* 76, 155–161.

[11] Galbraith, J.W. and V. Zinde-Walsh (1997) On some simple, autoregression-based estimation and identification techniques for ARMA models. *Biometrika* 84, 685–696.

[12] Granger, C.W.J. and P. Newbold (1977) *Forecasting Economic Time Series. (1st ed.)* Academic Press, New York.

[13] Härdle, W. (1990) *Applied Nonparametric Regression.* Cambridge University Press, Cambridge.

[14] Lütkepohl, H. (1993) *Introduction to Multiple Time Series Analysis.* (Revised ed.) Springer-Verlag, Berlin.

[15] Meese, R. and J. Geweke (1984) A comparison of autoregressive univariate forecasting procedures for macroeconomic time series. *Journal of Business and Economic Statistics* 2, 191–200.

[16] Samaranayake, V.A. and D.P. Hasza (1988) Properties of predictors for multivariate autoregressive models with estimated parameters. *Journal of Time Series Analysis* 9, 361–383.

[17] Sampson, M. (1991) The effect of parameter uncertainty on forecast variances and confidence intervals for unit root and trend stationary time-series models. *Journal of Applied Econometrics* 6, 67–76.

[18] Stine, R.A. (1987) Estimating properties of autoregressive forecasts. *Journal of the American Statistical Association* 82, 1072–1078.

[19] Stock, J.H. and M.W. Watson (1998) A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. NBER working paper no. 6607.

[20] Wilks, D.S. (1996) Statistical significance of long-range "optimal climate normal" temperature and precipitation forecasts. *Journal of Climate* 9, 827–839.

Figure 1a: AR(1) processes, T=100
Analytical and Simulated Forecast Content Functions

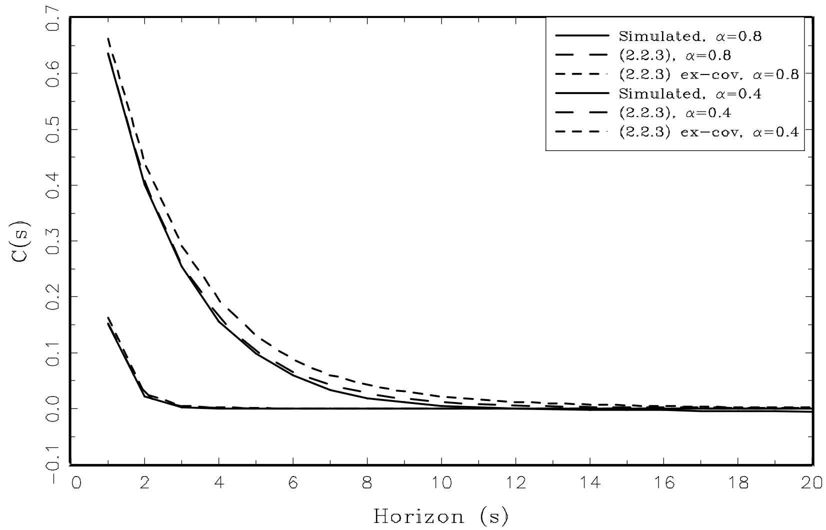


Figure 1b: AR(1) processes, T=400
Analytical and Simulated Forecast Content Functions

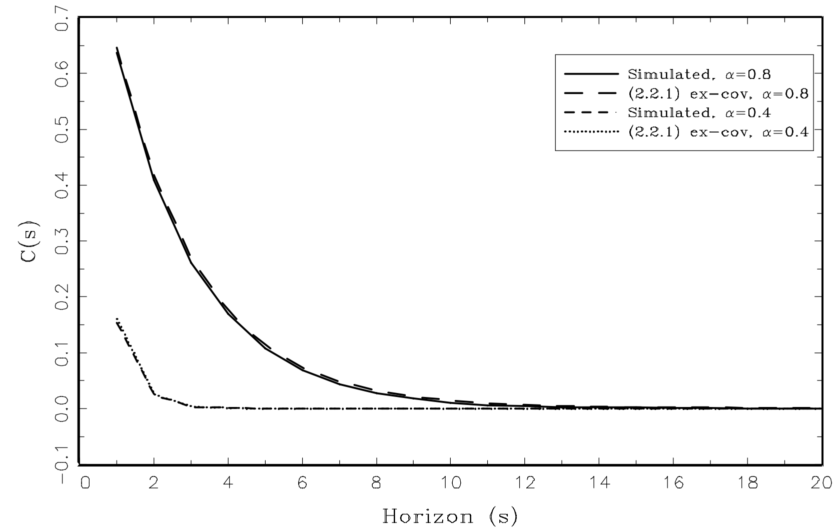


Figure 1c: AR(2) processes, T=100
Analytical and Simulated Forecast Content Functions

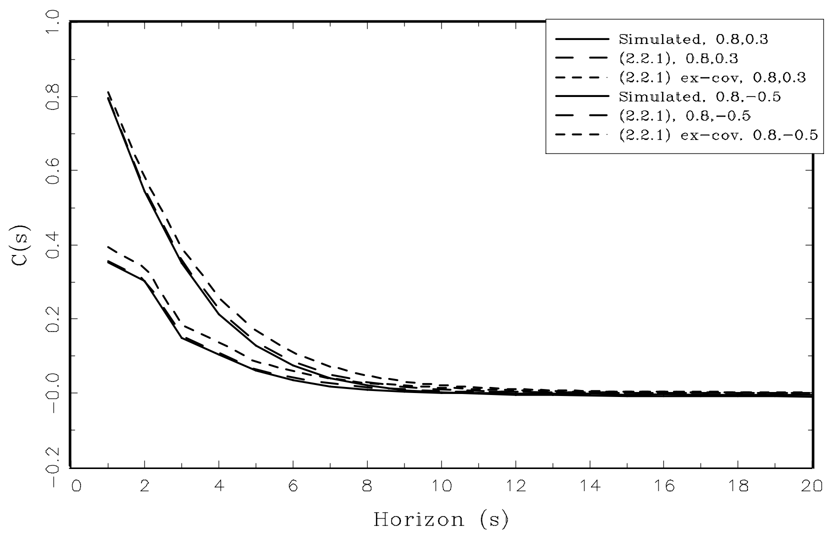


Figure 1d: AR(2) processes, T=400
Analytical and Simulated Forecast Content Functions

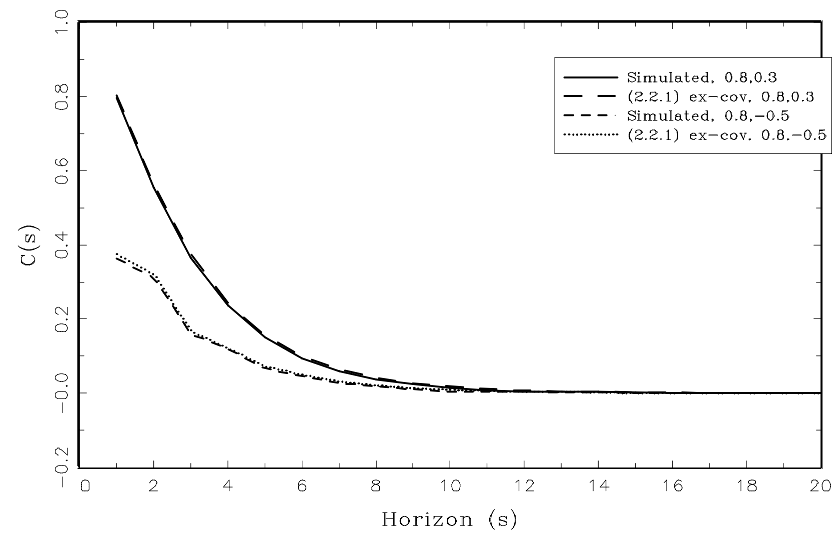


Figure 2a:
Analytical vs direct estimates of forecast content
Quarterly real GDP growth, Canada

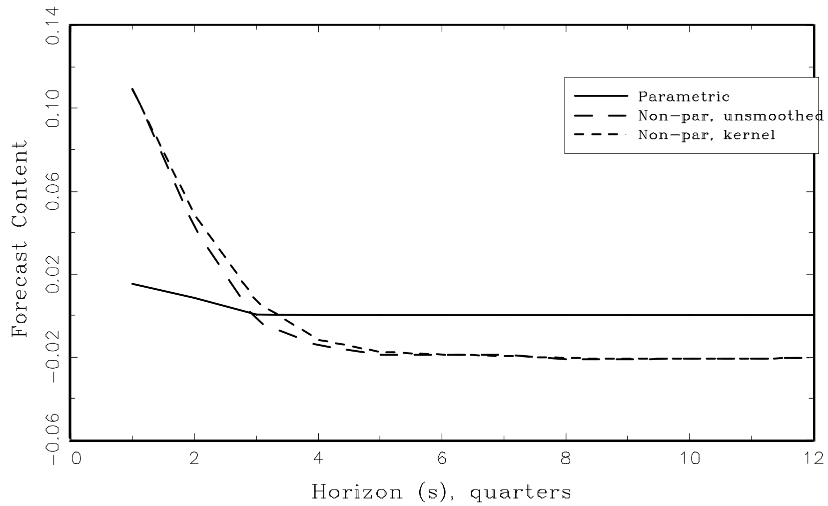


Figure 2b:
Analytical vs direct estimates of forecast content
Quarterly real GDP growth, USA

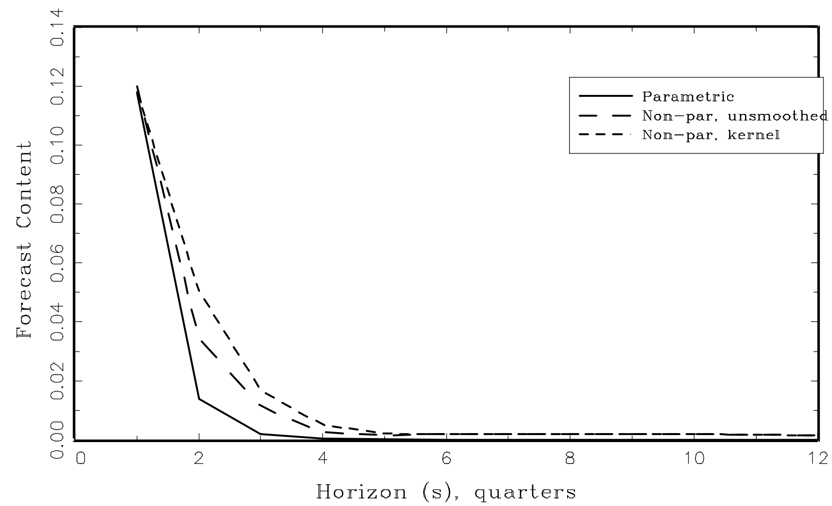


Figure 3a:
Analytical vs direct estimates of forecast content
Monthly inflation, Canada

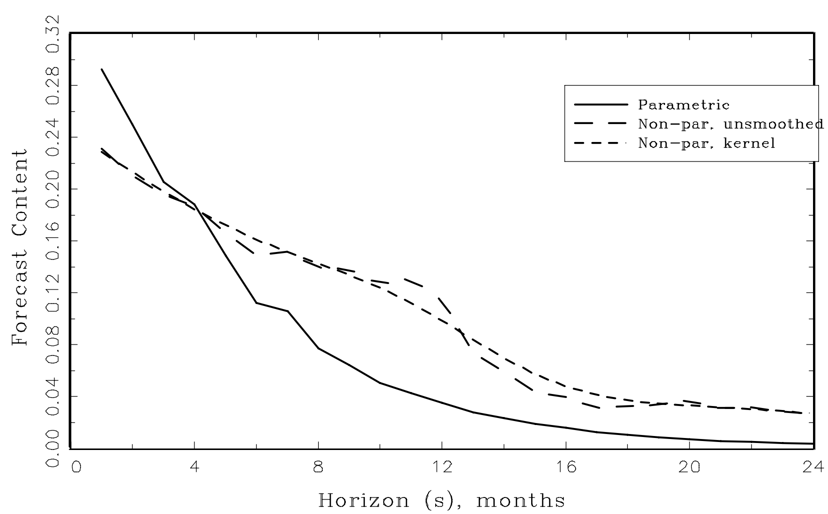
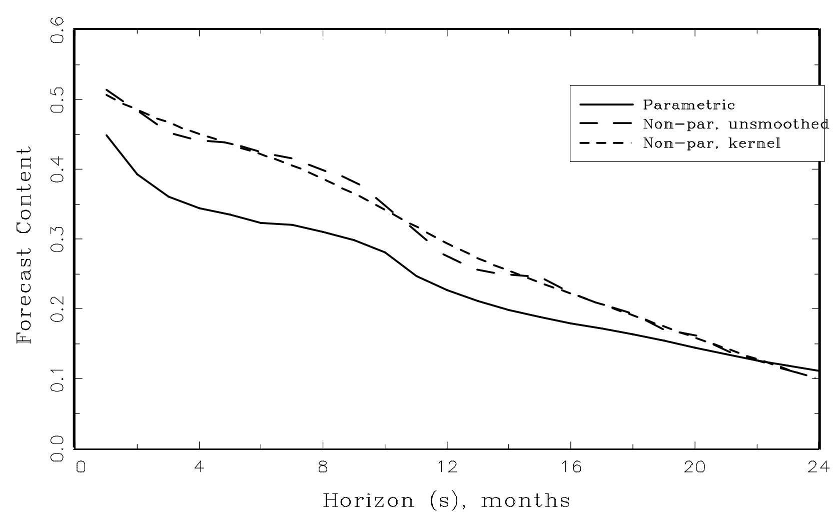


Figure 3b:
Analytical vs direct estimates of forecast content
Monthly inflation, USA



Liste des publications au CIRANO *

Cahiers CIRANO / *CIRANO Papers* (ISSN 1198-8169)

99c-1 Les Expos, l'OSM, les universités, les hôpitaux : Le coût d'un déficit de 400 000 emplois au Québec — Expos, Montréal Symphony Orchestra, Universities, Hospitals: The Cost of a 400,000-Job Shortfall in Québec / Marcel Boyer

96c-1 Peut-on créer des emplois en réglementant le temps de travail ? / Robert Lacroix

95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret

95c-1 La réglementation incitative / Marcel Boyer

94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette

94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel

94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

99s-16 Modelling the Role of Organizational Justice: Effects on Satisfaction and Unionization Propensity of Canadian Managers / Michel Tremblay et Patrice Roussel

99s-15 Pricing Discretely Monitored Barrier Options by a Markov Chain / Jin-Chuan Duan, Evan Dudley, Geneviève Gauthier et Jean-Guy Simonato

99s-14 Shame and Guilt in Lancashire: Enforcing Piece-Rate Contracts / Michael Huberman

99s-13 Cost Manipulation Games in Oligopoly, with Costs of Manipulations / Ngo Van Long et Antoine Soubeyran

99s-12 Using Employee Level Data in a Firm Level Econometric Study / Jacques Mairesse et Nathalie Greenan

99s-11 Incentives for Poluution Control: Regulation or (and?) Information / Jérôme Foulon, Paul Lanoie et Benoît Laplante

99s-10 Le coût du capital des entreprises à base de connaissance au Canada / Jean-Marc Suret, Cécile Carpentier et Jean-François L'Her

99s-09 Stratégies de financement des entreprises françaises : Une analyse empirique / Cécile Carpentier et Jean-Marc Suret

99s-08 Non-Traded Asset Valuation with Portfolio Constraints: A Binomial Approach / Jérôme Detemple et Suresh Sundaresan

99s-07 A Theory of Abuse of Authority in Hierarchies / Kouroche Vafaï

99s-06 Specific Investment, Absence of Commitment and Observability / Patrick González

99s-05 Seasonal Nonstationarity and Near-Nonstationarity / Eric Ghysels, Denise R. Osborn et Paulo M. M. Rodrigues

99s-04 Emerging Markets and Trading Costs / Eric Ghysels et Mouna Cherkaoui

99s-03 Sector-Specific Training and Mobility in Germany / Lars Vilhuber

---

* Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site World Wide Web à l'adresse suivante :
http://www.cirano.umontreal.ca/publication/page1.html