Structural Time Series Modelling of Capacity Utilisation

Proietti, Tommaso

Università di Udine

6 June 1999

STRUCTURAL TIME SERIES MODELLING OF CAPACITY UTILISATION

Tommaso Proietti

Dipartimento di Scienze Statistiche,

Università di Udine

Via A. Treppo, 18, 33100 Udine, Italy

E-mail proietti@dss.uniud.it

Tel. +39 0432 249581; Fax: +39 0432 249595

ABSTRACT

In this paper we introduce a structural non-linear time series model for joint estimation of capacity and its utilisation, thereby providing the statistical underpinnings to a measurement problem that has received ad hoc solutions, often underlying arbitrary assumptions. The model we propose is a particular growth model subject to a saturation level which varies over time according to a stochastic process specified a priori. A bivariate extension is discussed which is relevant when survey based estimates of utilization rates are available. Illustrations are provided with respect to the US and the Italian industrial production.

Key words: Structural Time Series Models, Nonlinear models, Extended Kalman Filter, Interpolation.

1 Introduction

The measurement of productive capacity and its utilization has attracted a lot of attention in the economic and statistic literature. Since it deals with an artificial construct, its domain is rather controversial, and it is not surprising that different methods of measurement have been proposed, each appealing to economic theory in different degrees.

This article does not aim at reviewing this literature, mostly because excellent reviews are already available, such as Christiano's (1981). The recent additions mainly deal with the related issue of measuring potential output for the aggregate economy, and are paced to the developments in the econometric analysis of time series. For instance, DeSerres *et al.* (1995) measure potential output within a structural VAR framework, as the output dynamics produced by permanent supply and oil shocks; other bivariate work exploits either the Okun Law or the Phillips Curve relationships, according to which the output gap is related to the unemployment rate and to inflation, respectively; see Evans (1989), Clark (1989), Kuttner (1994) and Norden (1995). The latter uses an observable index model (Sims and Sargent, 1977), whereas the others treat potential output as a latent variable.

In this article we propose a measurement method that is quite neutral in that it does not necessarily assume a particular definition of capacity; perhaps it is closer to an engineering concept, for which capacity is defined as the as the maximum output that can be produced using a given plant and equipment under "realistic operating conditions", rather that to an economic one.

The main focus will be instead on the time series properties of the model used to tackle the measurement problem. Our can be considered as an attempt to bring the measurement problem back to an inferential framework, in which a model is formulated, its parameters estimated, and goodness of fit assessed.

The specific nature of the problem at hand will lead us to the introduction of a non linear unobserved components model which simultaneously extracts estimates of capacity and of its utilization. The former represents the saturation level of output which changes slowly over time as a result of shocks to technology. On the other hand, the extent to which the resources are employed results from the propagation of innovations that can have only transitory effects.

The plan of the paper is the following: the next section will introduce the univariate model, by which capacity is measured using the output series alone. Section 3 presents an application with respect to US industrial production and the derived measure of capacity utilization is compared with the estimates produced by the Federal Reserve (FED). The extension of the model to the bivariate framework, where the output series is used in conjunction with a survey based index of capacity utilization, is presented in section 4. The model is illustrated with respect to the Italian industry sector (section 5), and in the following section we tackle the important issue of interpolating quarterly figures at the monthly level. Section 7 concludes.

2 The Univariate Model of Capacity Utilization

In this section we introduce a model for the measurement of capacity exploiting the information contained in the output series alone. Later on we shall be able to argue that the model provides a rationalisation of the Wharton Index method (see Christiano, 1988, p.150, for a description), sharing the simplicity and computational ease (The Wharton Index is currently produced by the Bank of Italy).

Let $Y_t,\ t = 1, \ldots, T$, denote the level of industrial production, usually in index form; this can be thought of as characterised by an upper limit, namely capacity, which identifies the maximum output that can be produced, given the current state of technology and the availability of inputs; for the measurement problem at hand we are interested in decomposing the output at time t as the product of capacity and the utilisation rate:

$$Y_t = \alpha_t \upsilon_t, \tag{1}$$

where α_t denotes capacity, characterised by the property $\alpha_t \geq Y_t$; the utilization rate is given by the ratio of output to capacity and is denoted here by υ_t. The latter can be modelled in the following fashion:

$$\upsilon_t = \frac{1}{1 + \exp(-\phi_t)}.$$

We further assume that ϕ_t is generated by a linear stationary process admitting, say, an ARMA(p, q) representation, with innovations $\kappa_t \sim$ WN$(0, \sigma^2_\kappa)$, and mean E$[\phi_t] = b$. Therefore, in the long run υ_t tends to $[1 + \beta]^{-1}$, with $\beta = \exp(-b)$, and this is interpreted as the utilization rate that would be observed at time t in the absence of shocks on ϕ_t, and as a first order approximation to E$[\upsilon_t]$. Notice that when $\beta = 0$, $y_t = \mu_t$, i.e. production takes place at full capacity. The logistic transformation of bounded variables is also considered in Wallis (1987), although with reference to observables.

Hence, Y_t lies below the saturation level α_t, and the range of υ_t is the interval (0,1). Rewriting $\phi_t = b + \psi_t$, where $\psi_t = \phi_t - b$ is a zero mean process, and substituting into (1), we get:

$$Y_t = \frac{\alpha_t}{1 + \beta \exp(-\psi_t)};$$

then, taking natural logarithms of both sides,

$$y_t = \mu_t - \ln[1 + \beta \exp(-\psi_t)],$$

with $y_t = \ln Y_t$ and $\mu_t = \ln \alpha_t$.

The model implies that the logarithms of capacity and output are non linearly cointegrated, in the sense given by Granger (1991), as the difference $\mu_t - y_t$ is *short memory in mean*; this is so since $E(\mu_{t+h} - y_{t+h}|\mathcal{F}_t) \to \ln(1 + \beta)$, where $\mathcal{F}_t$ denotes information up to time t.

As far as μ_t is concerned, we adopt the local linear model specification used by Harvey (1989) to model a trend component:

$$
\begin{aligned}
\mu_t &= \mu_{t-1} + \gamma_{t-1} + \omega_t, \qquad \omega_t \sim \mathrm{WN}(0, \sigma^2_\omega),\\
\gamma_t &= \gamma_{t-1} + \zeta_t, \qquad \zeta_t \sim \mathrm{WN}(0, \sigma^2_\zeta),
\end{aligned}
$$

with $\mathrm{E}(\omega_t \zeta_t) = 0$.

If it is deemed that Y_t is affected by a multiplicative measurement error, model (1) can be correspondingly extended so as to include it:

$$
Y_t = \alpha_t \upsilon_t \epsilon_t, \tag{2}
$$

so that, rewriting $\varepsilon_t = \ln \epsilon_t$, the logarithmic version of the model becomes:

$$
y_t = \mu_t - \ln[1 + \beta \exp(-\psi_t)] + \varepsilon_t. \tag{3}
$$

The model (2) admits the following non linear state-space representation:

$$
\begin{aligned}
y_t &= z_t(x_t) + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2_\varepsilon)\\
x_t &= T_t x_{t-1} + R_t \eta_t, \qquad \eta_t \sim \mathrm{WN}(0, \Sigma_\eta)
\end{aligned} \tag{4}
$$

and for its statistical treatment we shall make reference to the extended Kalman filter. For a general treatment see Harvey (1989, sec. 3.7).

The state vector can be partitioned as: $x_t = [x'_{\alpha t},\ x'_{\psi t}]'$, with $x_{\alpha t} = [\mu_t,\ \gamma_t]'$, and $x_{\psi t}$ being an $m \times 1$ vector containing the elements of the markovian representation of ψ_t, so that $z' x_{\psi t} = \psi_t$, for $z = [1, 0, \ldots, 0]'$; R_t is a suitable selection matrix acting on the innovations η_t. The covariance matrix of the innovations, Σ_η, is assumed diagonal, since the innovations to the different components are taken to be mutually uncorrelated.

Furthermore, let $x_{t|t-1}$ denote the expectation of x_t conditional on the information available at time $t-1$, $\mathrm{E}[x_t|\mathcal{F}_{t-1}]$; the first order Taylor expansion of $z_t(x_t) = \mu_t - \ln[1 + \beta \exp(-\psi_t)]$ about $x_{t|t-1}$ is:

$$
z_t(x_t) = z_t(x_{t|t-1}) + \frac{\partial z_t(x_{t|t-1})}{\partial x'_t}(x_t - x_{t|t-1}), \tag{5}
$$

where

$$\frac{\partial z_t(x_{t|t-1})}{\partial x_t'} = \left[1, 0, \frac{\beta \exp(-\psi_{t|t-1})}{1+\beta \exp(-\psi_{t|t-1})}, 0, \ldots, 0\right].$$

Hence, it is possible to rewrite the approximate model:

$$y_t = z_t(x_{t|t-1}) + \frac{\partial z_t(x_{t|t-1})}{\partial x_t'}(x_t - x_{t|t-1}) + \varepsilon_t,$$

which is *conditionally Gaussian*. Pseudo-maximum likelihood estimates of the hyperparameters can be obtained via the Kalman filter. However, the presence of nonstationary components in the state vector poses an additional difficulty concerning initial conditions; recently, De Jong (1991a, 1991b) has proposed an extended algorithm, the *diffuse* Kalman filter, that overcomes the problem. Smoothed estimates of the unobserved components, $x_{t|T}$ can be obtained by related algorithms. A Gauss program implementing the diffuse Kalman filter and smoother is available from the author.

As hinted at the beginning of the section, our model based approach shares with the Wharton Index method the feature of relying solely on the information provided by the output series. Nevertheless, it differs from it in many respects, and in particular it does not suffer from the following criticism, which applies to the Wharton Index: (i) it is not assumed that each major peak underlies the same intensity of resource utilisation; (ii) it is not assumed that capacity grows linearly from peak to peak; rather, technology shocks occur with continuity and not only at each cyclical peaks; (iii) the amount of revision as new data become available is not related to the location of a new peak. The most recent estimates of capacity are updated as soon as new observations become available.

Up to now, we have abstracted from the presence of seasonality in the output series, and in fact model (2) is suitable either for annual data or for subannual data that are seasonally adjusted. However, seasonality deserves further consideration, since the output in most sectors of the economy is seasonal. In the case of multiplicative seasonality a seasonal component, S_t, can be brought into the model in the following fashion:

$$Y_t = Y_t^{ns} \times S_t,$$

where Y_t^{ns} is modelled as in (2). The notion of capacity is replaced by that of *capability* adopted by the FED (1985), and the output series can now reach up to a level above capability due to seasonal peaks. The seasonal component can be modelled by a model of stochastic seasonality, such as the trigonometric model (see Harvey, 1989), so there is no real need for prior adjustment of the series, which represents, in our opinion, an additional advantage of structural modelling.

3 Illustration: US Industrial Production

In order to illustrate the performance of the model proposed in the previous section, we consider the problem of estimating capacity and its utilization for the US industry sector. The information set consists of the monthly index of industrial production, seasonally adjusted, for the sample period 1967:1-1996:7 (the series, whose code is B50001, has been retrieved from the Internet at the URL `http://www.bog.frb.fed.us`).

The results can be compared with those produced by the US Federal Reserve Board: this exercise is carried out only in order to certify the reliability of the model, since we must bear in mind that the procedure used by the FED is much more complex; furthermore, it is based on a larger information set. The main steps are described by Raddock (1985) and will be summarised in the sequel. (i) Preliminary end of year estimates of capacity are obtained at the industry level by dividing the production index by an utilization rate obtained from an external source (BEA, McGraw-Hill, Census Bureau), usually in the form of a business survey; (ii) the preliminary estimates are found overly procyclical and are corrected into a refined capacity estimate using the capital stock or capacity data in phisical units measures as additional information. The rationale underlying the refinement is that it is possible to get rid of the short run fluctuations in capacity exploiting the cointegration with the capital stock. Thus the predicted values from a regression model where the logarithm of ratio of the preliminary estimate to capital stock is regressed on a deterministic polynomial trend. (iii) The monthly capacity series is obtained by linear interpolation of the end of year refined estimate. (iv) A further adjustment called annual capability adjustment is made to obtain more appropriate levels of utilization. (v) The individual series are aggregated into market and industry groups applying value-added weights. (vi) The utilization rates for each individual series and groups are calculated by dividing the pertinent production index by the related capacity index.

We are now going to estimate model (2) under two scenarios. In the first, we shall assume that the average utilisation rate is known a priori, and β will be set equal to $(1/m-1)$, where m is the FED average utilisation rate in the sample period. In the second, we will treat β as an unknown parameter, and estimate it via maximum likelihood.

In both cases the logarithm of the capacity index, μ_t, is modelled as indicated in the previous section, with the restriction $\sigma^2_\zeta = 0$ (μ_t is a random walk with drift), whereas for ψ_t we have adopted the trigonometric specification:

$$\begin{aligned} \psi_t &= \quad \rho\cos\lambda_c\psi_{t-1} + \rho\sin\lambda_c\psi^*_{t-1} + \kappa_t, \qquad \kappa_t \sim \mathrm{WN}(0,\sigma^2_\kappa) \\ \psi^*_t &= - \ \rho\sin\lambda_c\psi_{t-1} + \rho\cos\lambda_c\psi^*_{t-1} + \kappa^*_t, \qquad \kappa^*_t \sim \mathrm{WN}(0,\sigma^2_\kappa) \end{aligned}$$

which has been introduced by Harvey (1989) to model an economic cycle. Thus, ψ_t has a restricted ARMA(2,1) reduced form representation; the restrictions imply, amongst other things, that the roots of the autoregressive polynomial are a pair of complex conjugates with modulus ρ^{-1} and phase λ_c. The model is stationary if $|\rho| < 1$. We further assume that the disturbances κ_t and κ^*_t are uncorrelated with each other and with the disturbances driving the other components.

The parameter estimates are reproduced in table 1, along with diagnostics and goodness of fit. In both cases the irregular component is absent and the model for μ_t is constrained to be a random walk with constant drift. Notice also that the value of $\hat{\rho}$, very close to one, implies that ψ_t is close to the nonstationarity region at the frequency $\hat{\lambda}_c$, corresponding to a period of about 57 months (4 years and 9 months). These estimates are comparable to those presented by Harvey and Jaeger (1993, table 1, p. 236) for quarterly S GNP: namely, the period of the cycle is virtually the same (22.2 quarters), and the estimated damping factor is lower (.92) as should be expected since the data considered are quarterly, rather than monthly.

The diagnostic quantities highlight significant residual correlation at lag 1, which is responsible for the significant Ljung-Box statistic. Yet the model is satisfactory as it reproduces the main stylised facts concerning capacity and its utilisation as we are going to argue.

Figure 1 displays the smoothed estimates of the logarithm of capacity, $\exp(\mu_{t|T})$, for the case $\beta = 1/m - 1$; these estimates do not appear unduly procyclical and appear to capture well the notion of capacity as a potential series whose variations are due solely to long run shocks. In figure 2 we compare the implied utilization rates (calculated as $[1 + \beta \exp(-\psi_{t|T})]^{-1}$), with those calculated by the FED, according to the procedure outlined above.

The comparison reveals that the profile of the two measures is pretty much the same, although the FED estimates are somewhat slightly trending downwards, whereas our estimates hover around a constant level. The main message conveyed by the plot is that the univariate model, despite its simplicity and the little information requirements, does a good job in replicating the dynamics of the utilisation rates.

When we treat β as an unknown parameter, the estimated variances of the disturbances to both components are somewhat larger. Since the ML estimate of β is less than the value imposed for the previous model ($\hat{\beta} = .06$), and corresponds to an average utilisation rate of 0.94, the capacity series gets closer to the production index (see figure 3); further, its dynamics are somewhat rougher.

Estimates of the utilization rates, displayed in figure 4, reproduce the dynamics of figure 2, in the sense that the alternation of low and high capacity utilization regimes coincides with that highlighted by the FED estimates. However, they oscillate around a higher value and are smoother than in the previous case. More generally, there is a tradeoff between the smoothness of capacity and of the utilization rates; usually it is believed that capacity should be slowly changing, whereas the utilization rates, which are affected by short run variations, ought to be more volatile.

As far as the average level of capacity utilization is concerned, is a well known fact that "data based" utilization rates tend to be systematically higher on average than those computed employing survey based data (Christiano, 1981). This simple example conveys the message that the main uncertainty concerning capacity measurement concerns the the average level of capacity, rather than its dynamics.

Likelihood contours for the problem at hand, considered as a function of β and the innovation ratio $\sigma^2_\omega/\sigma^2_\kappa$ are plotted in figure 5. The cycle parameters were fixed at $\rho = .98$ and $\lambda_c = .11$. The picture shows on the y axis the utilization rate, $(1 + \beta)^{-1}$, and the innovation ratios $\sigma^2_\omega/\sigma^2_\kappa$ on the horizontal axis, and confirms that the likelihood is well behaved, although relatively flat going in the top-left bottom-right direction: in general, for lower average utilisation rates we expect smoother capacity estimates and rougher utilization rates.

4 The Bivariate Model

In some countries direct measures of capacity utilization are elicited by business surveys, asking a sample of companies the percentage at which the company operates in the reference period. In Italy, for instance, a judgmental survey is conducted quarterly by ISCO (*Istituto Nazionale per lo Studio della Congiuntura*) for the manifacturing sector. The individual responses are then aggregated with weights proportional to a measure of size of the firm, such as value-added or the number of employees. The respondent is offered no formal definition, apart from a generic reference to maximum capacity as an upper limit.

The resulting estimates are affected by both sampling errors and non sampling errors, in the form of sample selection bias, since the judgemental selection of the sampling units usually leads to a gross underrepresentation of small productive units. We shall not deal with this source of systematic error in the sequel, firstly because it is not easily quantified, secondly because the output measure, at least in Italy, suffers from the same sample selection bias.

Furthermore, there are two main sources of ambiguity that harm the interpretation of the results; the first arises as a consequence of the fact that the definition of capacity is left to the respondent, who may alternatively refer to either a particular productive factor, namely capital, or to all resources. The second element is "the time horizon that businesses have in mind in evaluating their capacity" (Christiano, 1981, p.171).

Nevertheless, these measures cohere with the information coming from other sources and display dynamics reflecting the stage of the business cycle, and what is more, they can make a significant contribution to the information set for estimating capacity. In this section we shall make an attempt to incorporate this additional information into a suitable multivariate structural time series model.

Let U_t denote a survey based measure of capacity utilization, taking values within the range (0,1] and u_t its *logit* transformation, namely:

$$u_t = \ln\left(\frac{U_t}{1-U_t}\right).$$

We are now in a position to introduce the following bivariate model:

$$\begin{aligned} y_t &= \mu_t - \ln[1+\beta\exp(-\psi_t)] + \varepsilon_{1t} \\ u_t &= b + \psi_t + \varepsilon_{2t}. \end{aligned} \tag{6}$$

The second equation captures the idea that the survey based measure has extra variability due to a measurement error. Since the data sources are independent, it is quite natural to assume that the measurement noises ε_{1t} and ε_{2t} are mutually and serially uncorrelated. It should also be noticed that ψ_t enters both equations: as a matter of fact, it is the latent factor of the logit transformation of U_t, devoid of measurement noise, and it enters non linearly in the equation for y_t, being the source of the short run fluctuations in the utilization rates.

Whereas in the univariate case we assumed the stationarity of ψ_t, the increase in the information set can allow more precise statements on the time series properties of the data generating process for the utilization rate. More precisely, a particular specification for it may be suggested by the univariate analysis of u_t. Finally, the parameter $\beta = \exp(-b)$ can be concentrated out of the likelihood function and estimated in terms of $\hat{b} = N^{-1} \sum u_t$.

As far as the statistical treatment is concerned, the only source of non linearity is the first measurement equation; the model can be linearised by a first order Taylor expansion as was done in section 2. The resulting approximated model is conditionally gaussian and likelihood evaluation and prediction can be performed by means of the extended Kalman filter with diffuse initial conditions.

As a further extension, it would be interesting to explore the possibility to employ the information arising from capital stock estimates, which are built in some countries according to the perpetual inventory method. Following the point (ii) of the FED procedure, one can suitably assume that capacity and capital stock are cointegrated with cointegrating vector $[1 \quad -1]$. This enhances the smoothness of the capacity estimates since the perpetual inventory method *per se* results in a one sided MA filter applied to the investment series, smoothing out the high frequency components (although inducing a phase shift). The model (6) would be amended so as to contain a common trend which enters log-capacity and the capital stock equation with loadings matrix $[1\ 1]'$.

5 Illustration: Capacity Estimates for the Italian Industrial Sector

In this section we derive capacity estimates for the Italian industrial sector, by applying the model (6) to a bivariate system made up of the index of industrial production (IP) and the index of capacity utilisation produced by ISCO. The former is available on a monthly basis and we consider the seasonally adjusted series with trading days correction. The latter is available only quarterly, and the most straightforward solution is to aggregate the IP series to the same observation interval.

Time aggregation is an issue, since for the ISCO series there exists some uncertainty surrounding the time horizon considered by the respondent in practice; in particular, it is unclear whether reference is made to the end of the quarter or to the average capacity utilization over the quarter; the survey question is formulated in terms of the latter, asking for an assessment “in the course of the quarter”. On the contrary, the entire questionnaire makes explicit reference to the situation at the end of the quarter; moreover, the respondent is more likely to refer to the end of period situation in the capacity assessment.

Therefore, we assumed that the quarterly utilization rates are end of period estimates and we derived a quarterly production series taking a systematic sample of the series values referring to the last month of each quarter (March, June, September, December). The sample period goes from the first quarter of 1970 to the fourth quarter of 1993.

Univariate structural time series modelling of u_t suggested a stationary AR(1) plus noise specification for ψ_t; table 2 reproduces the maximum likelihood estimates obtained via the diffuse Kalman filter for the unrestricted case and when the smoothness prior, $\sigma^2_\omega = 0$, is imposed.

Figure 6 shows the smoothed estimates of capacity, $\exp(\mu_{t|T})$, obtained from the restricted model, which, although is not the preferred model, according to the information criteria reported in table 2, provides the smoother capacity estimates. The figure highlights that at the beginning of the 80's a slowdown affected capacity growth. In figure 7 the implied capacity utilization rates, $[1+\hat{\beta}\exp(-\psi_{t|T})]^{-1}$, are compared with the ISCO series, showing that the noise in the latter is somewhat smoothed out.

6 Interpolation

Usually, since the output series are monthly, it is desirable to estimate capacity with the same periodicity. This need justifies the interpolation technique adopted by the FED and the peak to peak interpolation in the Wharton Index methodology. The techniques adopted are quite elementary and result in linear interpolation. The interpolation problem can actually be solved within the models we have proposed in the previous sections, by a suitable modification of the measurement equation.

We thus turn our attention to the problem of estimating the model (6) using the monthly index of production and the quarterly utilization rate series. The latter is subject to missing values in a systematic fashion, since the first and the second month of each quarter are not observed. As the unobserved components are stock variables, time disaggregation is less of a problem and the model (6) holds at the monthly level without any modification.

A possible strategy is illustrated in De Jong (1991b), and amounts to setting the missing values to zero and zeroing out the elements of the measurement equation system matrices corresponding to the series. This produces a singularity in the covariance matrix of the Kalman filter innovations, which is remedied upon replacing its inverse by a generalised inverse, using a selection matrix picking up a suitable basis for this matrix.

Preserving the AR(1) specification for ψ_t, the maximum likelihood estimates of the hyperparameters in the unrestricted case are as follows ($\ln\mathcal{L} = 1355.76$):

$$\psi_t = 0.96\psi_{t-1} + \kappa_t, \quad \kappa_t \sim \text{WN}(0, .00181);$$

$\hat{\sigma}^2_\omega = 537 \times 10^{-7}$, $\hat{\sigma}^2_\zeta = .4 \times 10^{-7}$, $\hat{\sigma}^2_{\varepsilon_1} = 2169 \times 10^{-7}$, $\hat{\sigma}^2_{\varepsilon_2} = 5555 \times 10^{-7}$.

When the smoothness prior is imposed on the model for capacity, the hyperparameter estimates are the following ($\ln \mathcal{L} = 1349.66$): $\hat{\sigma}^2_\zeta = 13 \times 10^{-7}$, $\hat{\sigma}^2_{\varepsilon_1} = 2488 \times 10^{-7}$, $\hat{\sigma}^2_{\varepsilon_2} = 7039 \times 10^{-7}$. The model for ψ_t is unchanged:

$$\psi_t = 0.96\psi_{t-1} + \kappa_t, \quad \kappa_t \sim \mathrm{WN}(0, .00182);$$

The smoothed estimates of productive capacity obtained from the second model are graphed in figure 8. The series reproduces the behaviour of capacity observed at a quarterly reference period with the monthly variability being absorbed by the utilization rate, reported in figure 9.

7 Conclusions

In this paper univariate and bivariate models for measuring capacity and its utilization were introduced, that are consistent with the recent developments of the econometric analysis of time series. They are nonlinear structural time series models that estimate capacity as the saturation level of output.

The performance of the models was illustrated with respect to the US and the Italian industrial production, and our conclusion is that they qualify as an useful addition to the currently available methods. Not only do they provide a rationalisation of the ad hoc procedures used by statistical agencies, but they also allow important issues such as interpolation to be treated within the same model based framework.

Furthermore, suitable extensions can be envisaged that can deal with related measurement issues, arising when a series is subject to an upper or lower bound, e.g. in the extraction of the natural rate of unemployment, and so forth.

AKNOWLEDGEMENTS

Financial support from MURST, Funds ex-60%, is gratefully acknowledged. I am indebted to Guido Pellegrini at the Bank of Italy for kindly providing the data.

REFERENCES

Christiano L.J. (1981), "A Survey of Measures of Capacity Utilization", *IMF Staff Papers*, 28, 1, 144-198.

DeJong P. (1991a), "The Diffuse Kalman Filter", *Annals of Statistics*, 19, 2, 1073-1083.

DeJong P. (1991b), "Stable Algorithms for the State Space Model", *Journal of Time Series Analysis*, 12, 2, 143-157.

DeSerres A., Guay A., and St.Amant, P. (1995), "Estimating and Projecting Potential Output using Structural VAR Methodology: The Case of the Mexican Economy", Mimeo *Bank of Canada*.

Evans, G.E. (1989), "Output and Unemployment Dynamics in the United States: 1950:1985", *Journal of Applied Econometrics*, 4, 213-237.

Granger, C.W.J. (1991), "Some Recent Generalizations of Cointegration and the Analysis of Long-Run Relationships", in Engle, R.F. and Granger, C.W.J. (eds.) *Long-Run Economic Relationships. Readings in Cointegration*, Advanced Texts in Econometrics, Oxford University Press, 277-287.

Harvey A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

Kuttner K.N. (1994), "Estimating Potential Output as a Latent Variable", *Journal of Business and Economic Statistics*, 12, 3, 361-368.

Raddock R.D. (1985), "Revised Federal Reserve Rates of Capacity Utilization", *Federal Reserve Bulletin*, 71, pp. 760-766. w

Sargent T.J., and Sims C.A. (1977), “Business Cycle Modeling without Pretending to Have Too Much *A Priori* Economic Theory”, in *Mew Methods in Business Cycle Research* (C. A. Sims, ed), Minneapolis, Federal Reserve Bank.

Wallis, K.F. (1987), “Time Series Analysis of Bounded Economic Variables”, *Journal of Time Series Analysis*, 8, 1, 115-124.

Table 1: *US Industrial Production, 1967:1.1996.6. Parameter estimates, Diagnostics and Goodness of Fit.*

	$\beta = 1/m - 1$	β unrestricted
$\hat{\sigma}^2_\epsilon$	0	0
$\hat{\sigma}^2_\omega$	83	256
$\hat{\sigma}^2_\kappa$	11952	56113
$\hat{\rho}$	.98	.99
$\hat{\lambda}_c$	.11	.11
$\hat{\beta}$		.06
$\ln \mathcal{L}$	1683.96	1686.94
AIC	-3361.92	-3363.88
BIC	-3342.41	-3346.37
R^2_d	.13	.11
r_1	.23	.24
$Q(12)$	39.93	48.65
G_1	2.67	0.56
G_2	64.23	43.19
G	66.90	43.75

NOTES: All variances are multiplied by $\times 10^{-7}$. r_1 denotes the residual correlation at lag 1. $Q(12)$ is Ljung-Box statistic based on 12 residual autocorrelations. G_1 is a test for residual skewness based on the standardised third moment of the residuals about the mean (see Harvey, 1989, 5.4.2.); G_2 is a test for residual kurtosis and $G = G_1 + G_2$ is the Bowman and Shenton test for non-normality. $R^2_D = 1 - SSE/SSD$, where SSE is the residual sum of squares and SSD is the sum of squares of the centered first differences. The Akaike Information Criterion is computed as AIC$= -2\ln\mathcal{L} + 2h$, where h is the number of hyperparameters, and the Bayes Information Criterion as $BIC = -2\ln\mathcal{L} + h\ln N$.

Table 2: *Italy, Bivariate Model for Capacity Estimation, 1970:1.1993.4.*

	Unrestricted	$\sigma^2_\omega = 0$
$\hat{\sigma}^2_{\epsilon 1}$	2247	3422
$\hat{\sigma}^2_{\epsilon 2}$	6206	5286
$\hat{\sigma}^2_{\omega}$	1640	0
$\hat{\sigma}^2_{\zeta}$	9	97
$\hat{\sigma}^2_{\kappa}$	51507	53083
$\hat{\rho}$	.89	.89
$\hat{\beta}$	.33	.33
$\ln \mathcal{L}$	590.31	586.33
AIC	-1170.61	-1164.66
BIC	-1157.79	-1154.40
$Q(8)$	57.88	71.90
G_1	2.44	1.04
G_2	4.45	2.92
G	6.90	3.97

NOTES: see notes at table 1. The statistics $Q(8), G_1, G_2$ and G are the multivariate Ljung-Box and normality tests.

Figure 1: US Capacity, $\exp(\mu_{t|T})$, and Industrial Production. The parameter β is set equal to $(1/m - 1)$, where m is the FED average utilisation rate.

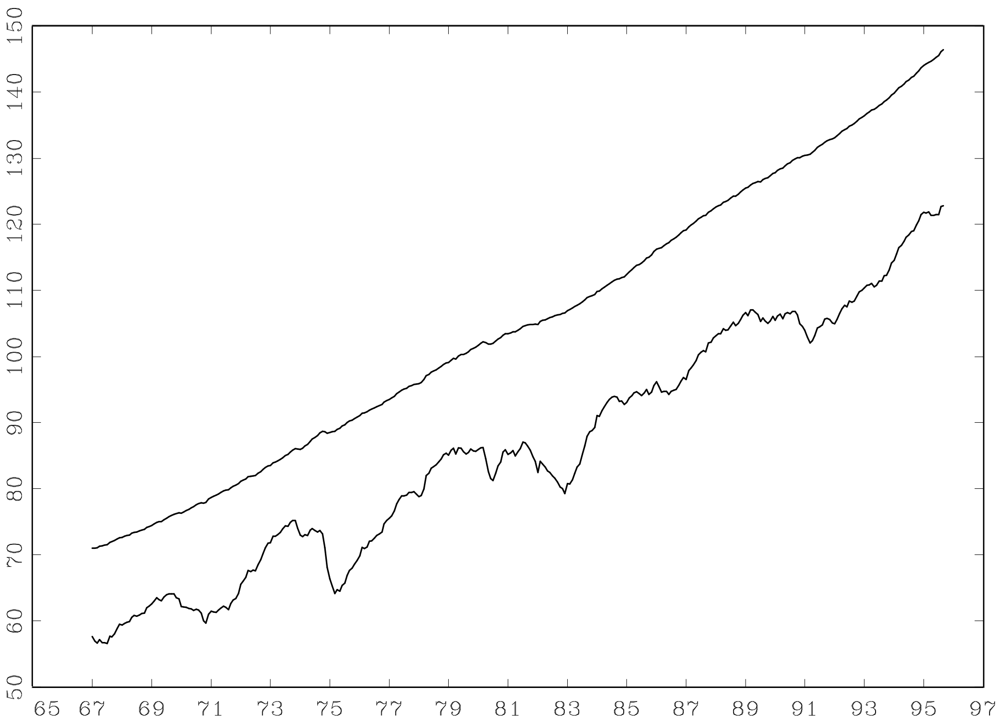

Figure 2: US Index of Capacity Utilisation. The parameter β is set equal to $(1/m - 1)$, where m is the FED average utilisation rate.

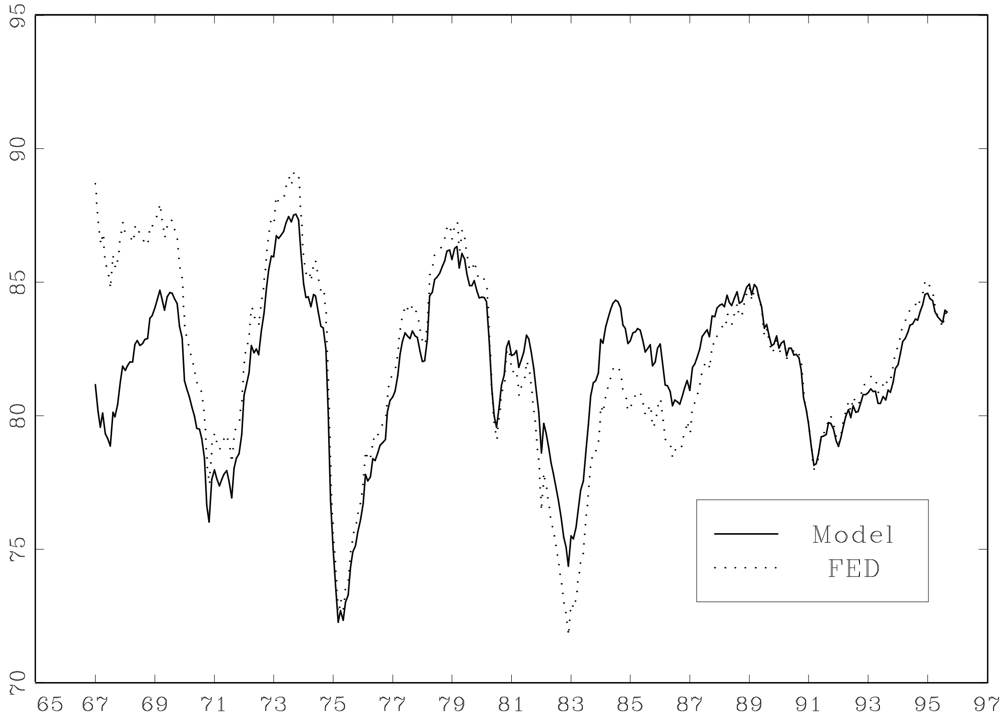

Figure 3: US Capacity (upper line), $\exp(\mu_{t|T})$, and Industrial Production (lower line). The parameter β is unconstrained.

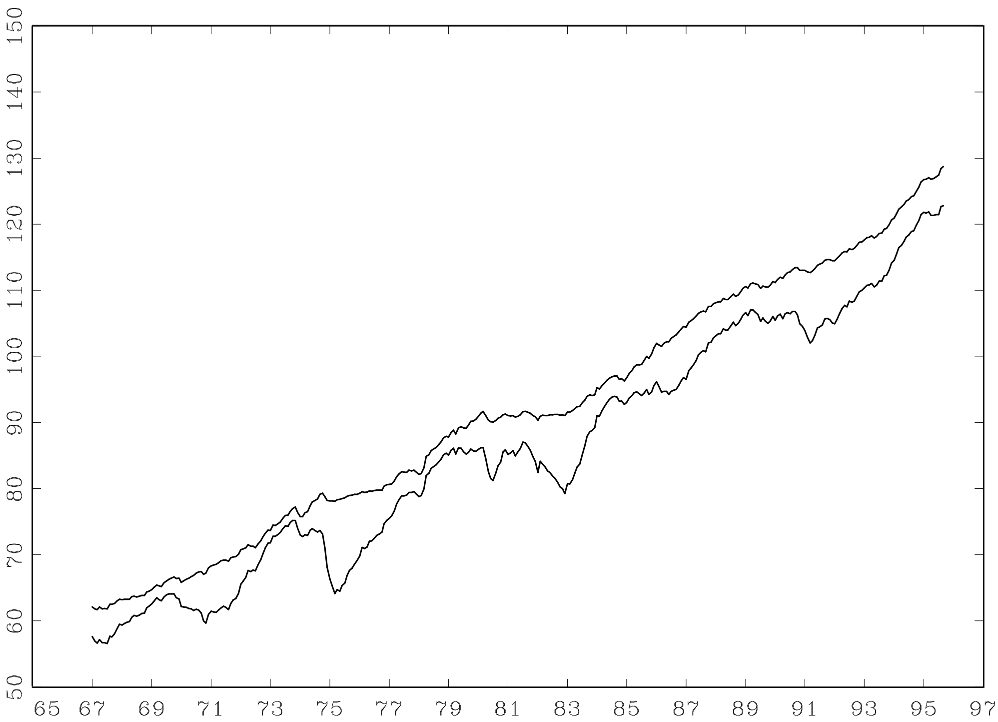

Figure 4: US Index of Capacity Utilisation. The parameter β is unconstrained.

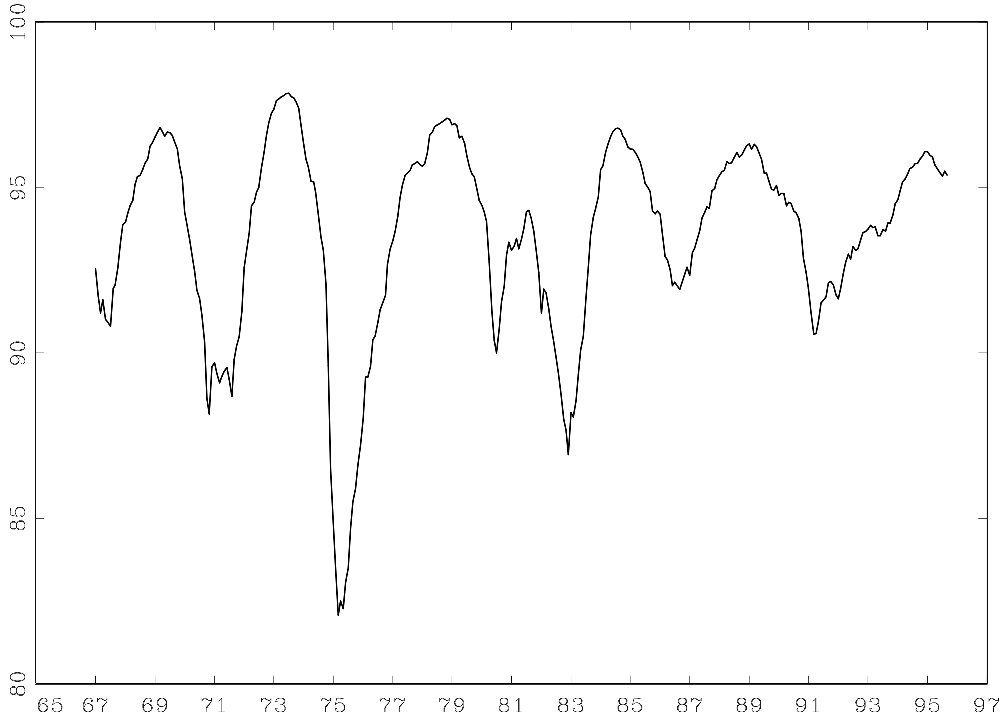

Figure 5: Likelihood Contours. The vertical axis measures the utilization rate, $(1+\beta)^{-1}$, and the horizontal axis the ratio $\sigma_\omega^2/\sigma_\kappa^2$. The small picture is a perspective plot of the likelihood function.

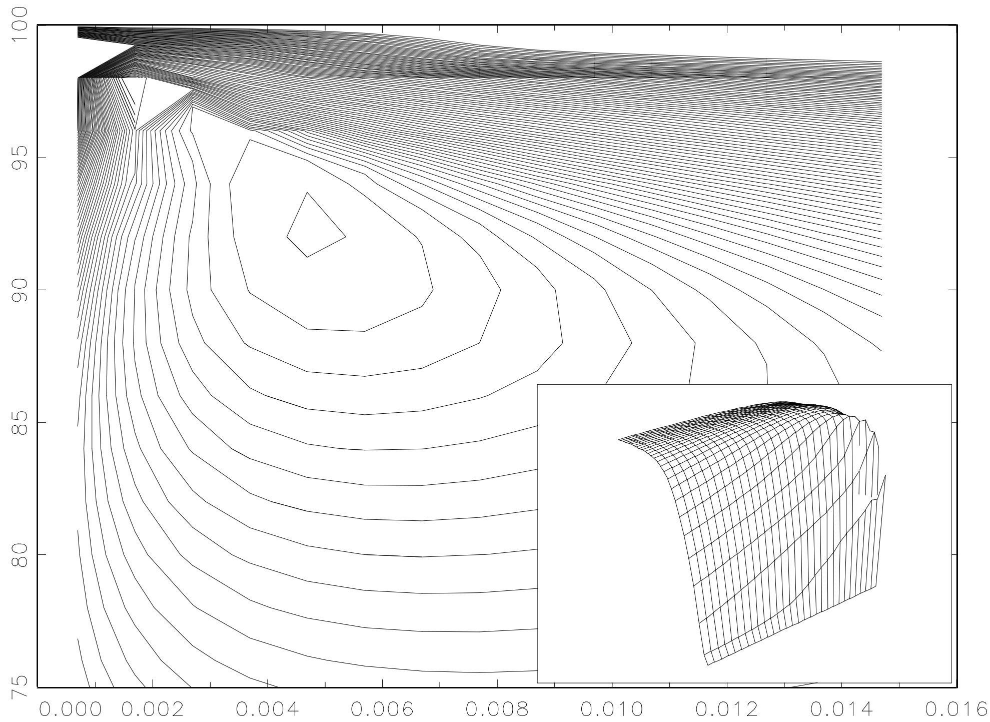

Figure 6: Italy: Quarterly Capacity (upper line) and Industrial Production (lower line). Capacity is estimated by a bivariate quarterly model.

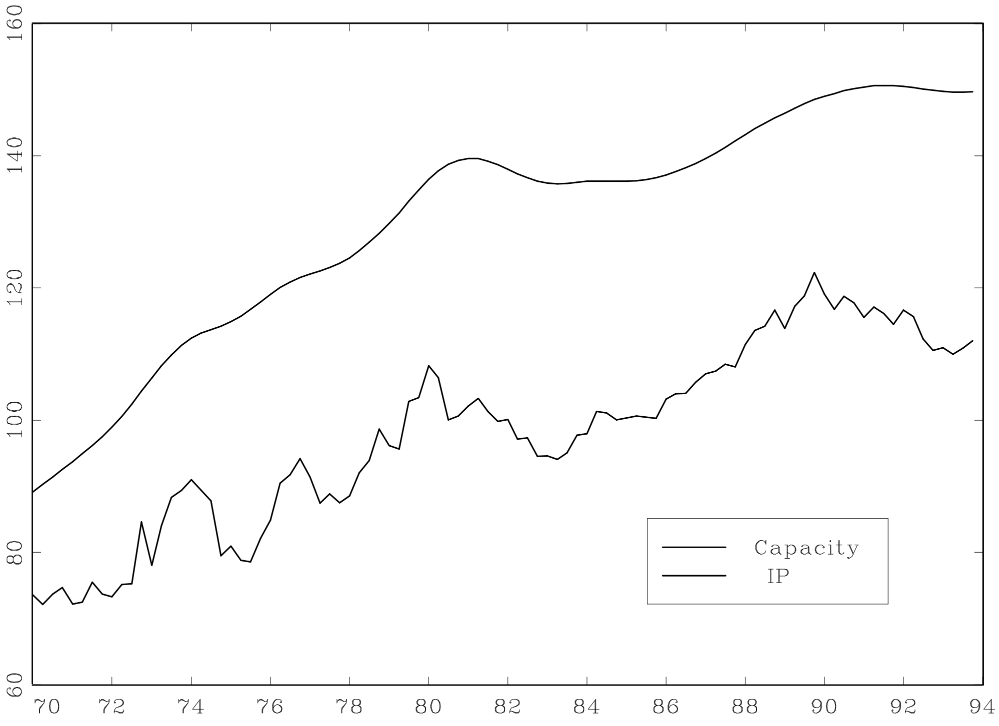

Figure 7: Italy: Quarterly Utilisation Rates and Comparison with ISCO.

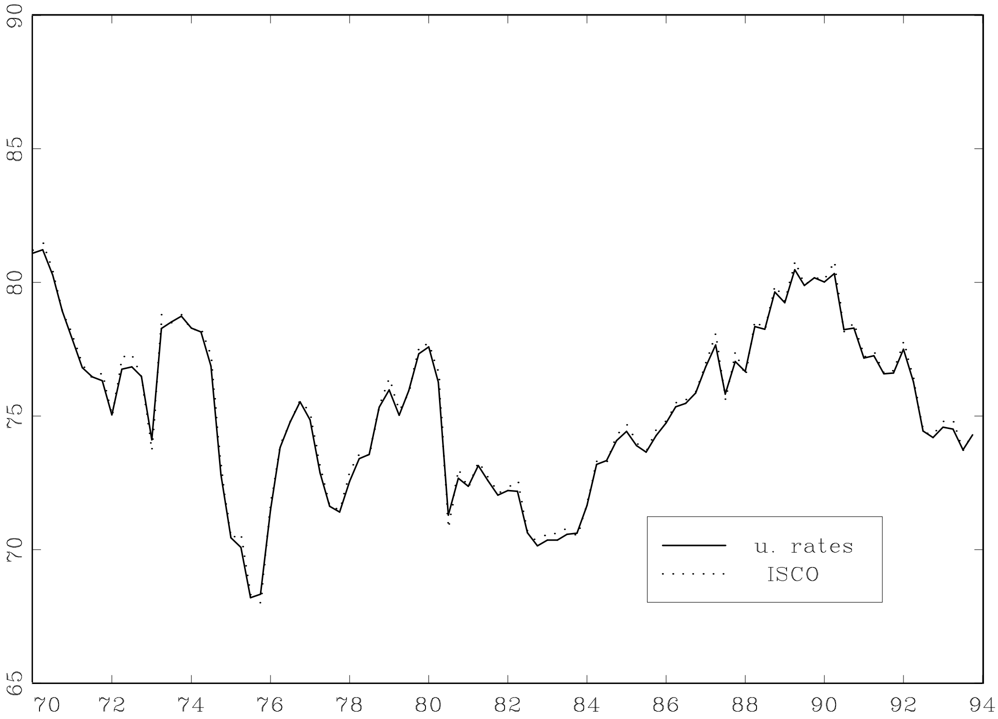

Figure 8: Italy: Monthly Capacity and Industrial Production.

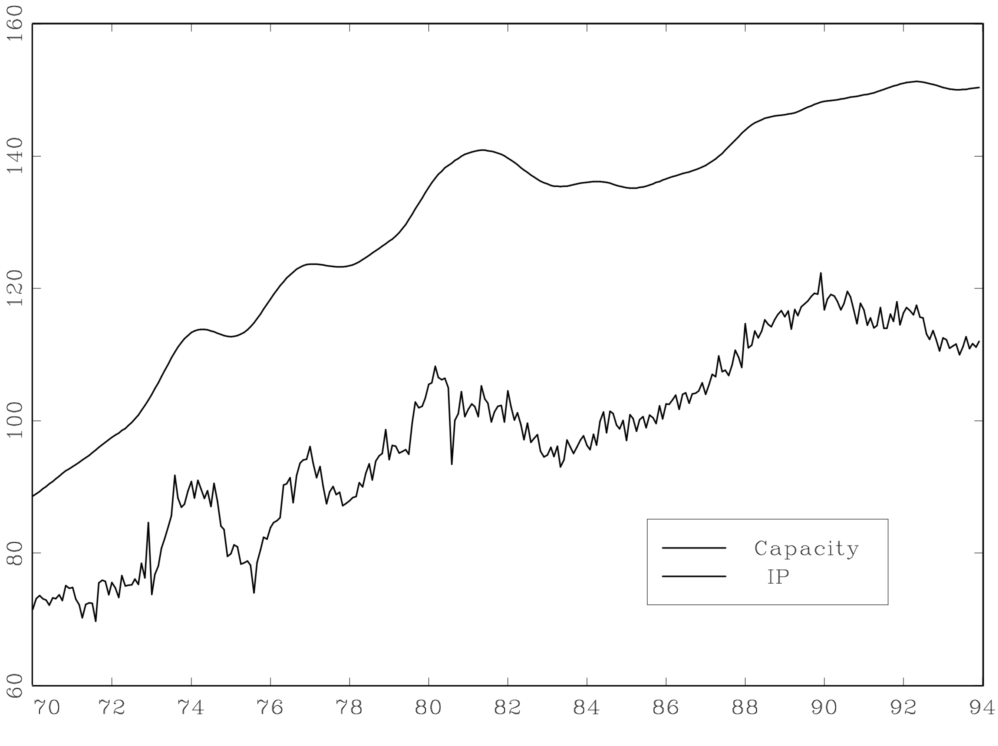

Figure 9: Italy: Monthly Utilisation Rates.