

A PROPOSAL TO OBTAIN A LONG QUARTERLY CHILEAN GDP SERIES*

JUAN DE DIOS TENA

Universidad de Concepción y Universidad Carlos III, España

MIGUEL JEREZ

Universidad Complutense, España

SONIA SOTOCA

Universidad Complutense, España

NICOLE CARVALLO

Universidad de Concepción, Chile

An important limitation in order to specify and estimate a macroeconomic model that describes the Chilean economy resides in using variables with sufficient number of observations that allow for a reliable econometric estimation. Among these variables, the GDP constitutes a fundamental magnitude. Nevertheless, for this variable there is not quarterly information before 1980. This paper computes quarterly GDP series for the period 1965-1980 using the approach by Casals et al. (2000). As result, the new series incorporates the cyclical dynamic in the quarterly series later to 1979 respecting, in addition, all the annual existing information before the above mentioned period.

JEL: C22, E32

Keywords: Smoothing Algorithm, ARIMA Model, Transfer Function Model, Chilean GDP.

1. INTRODUCTION

An important limitation in order to specify and estimate a macroeconomic model that describes the Chilean economy resides in using variables with sufficient

*We are indebted to an anonymous referee for helpful comments. The usual qualifier applies. Sonia Sotoca and Miguel Jerez gratefully acknowledge financial support from Ministerio de Educación y Ciencia, DGCICYT, ref. SEJ2005-07388.

E-mail: juande@udec.cl, jtена@est-econ.uc3m.es

number of observations that allow for reliable econometric information. Among these variables, the GDP constitutes a fundamental magnitude. However, although there is annual information of this series from the XIX century, quarterly information is only available from 1980.

In this paper we reconstruct quarterly series of the Chilean GDP for the period 1966-1979 using the interpolation methodology advocated by De Jong (1989) and Casals *et al.* (2000) for the nonstationary case. This procedure, explained in a simple way, consists firstly of: 1) specifying and estimating a statistical model of the series of interest; 2) expressing the resulting model in state space form; and then 3) reconstructing this series backward using a fixed-interval smoothing algorithm.

As far as we are aware, in spite of its empirical relevance, there is only a previous attempt of reconstructing the Chilean GDP by Haindl (1986). An important difference between his paper and ours is that he uses the methodology proposed by Chow and Lin (1971). This methodology imposes a very restrictive functional form that cannot represent accurately the features of our data. More specifically, this procedure assumes that: 1) the dependent variable and the set of indicators are fully cointegrated; 2) there is not feedback between the variables in the model; and 3) the residual of the regression between the dependent variable and the set of indicators follows an ad-hoc functional form.

Here, instead of imposing any particular functional form on the variables, we specify models that are plausible given the statistical properties of the time series in our analysis. Two alternative statistical specifications are considered in our analysis: a univariate ARIMA model and a transfer function model. These two specifications are robust and filter the data to uncorrelated and normal residuals. The quarterly series obtained under both procedures incorporate the cyclical dynamic in the quarterly series later to 1979 respecting, additionally, all the previous annual information.

The remainder of the paper is organized as follows. In the next section, we specify and estimate an ARIMA model for the Chilean GDP. Section 3 discusses how to write the univariate model in the state space form and how to use this formulation to reconstruct the quarterly series backward using a fixed interval smoothing algorithm. In Section 4, we extend the analysis to consider the use of transfer function models. Quarterly GDP series obtained from the univariate and the transfer function model are shown and compared in Section 5. Some concluding remarks follow in Section 6.

2. UNIVARIATE ANALYSIS

In this section we describe some of the important features of the individual time series used in the analysis and then we specify and estimate an ARIMA model for the Chilean GDP that will be used to interpolate quarterly values of this series before 1980. Our main interest is in Chilean GDP in real terms. This series is freely available from the Central Bank of Chile at the following URL:

<http://www.bcentral.cl>. The Central Bank of Chile publishes GDP series on a quarterly basis since 1980. However, it is possible to find GDP information on annual basis before this period.

Additionally, we also consider other quarterly series that will be used in the subsequent analysis to specify an econometric model. More specifically, our intention is to study the properties of these series to use them as indicators to interpolate quarterly information of the Chilean GDP before 1980. Three series have received our attention in this respect: 1) the monetary aggregate M1; 2) the price of copper; and 3) the terms of trade.

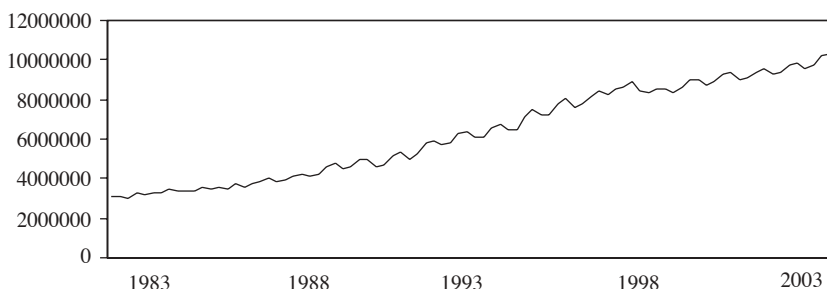
Series of the monetary aggregate M1 were collected from the Monthly Bulletin of the Central Bank of Chile. Series of the price of copper was obtained from “Informe Económico y Financiero”, also published by the Central Bank of Chile. Series of the terms of trade can be found in the paper by Bennett and Valdés (2001). These three series are available from 1966.

In order to specify an univariate ARIMA model for the GDP series we choose not to consider the first 10 quarterly observations as they are very erratic and generate an spurious AR(1) or AR(2) structure. That is, we only consider the period 1982:3-2004:2. The evolution of this series is shown in Figure 1. Inspection of the figure reveals that the GDP grows during the period under consideration and this growth is affected by the seasonal cycle.

Figures of the monetary aggregate M1, the price of copper and the terms of trade are not shown for the sake of brevity. However, M1 grows during the period and its growth is also affected by the seasonal cycle. Series of the terms of trade and the price of copper, on the other hand, do not grow but show little tendency to return to mean.

This and the additional information provided by correlograms suggests that M1 and GDP series require two differences to become stationary and one of them could be a seasonal difference. Also, terms of trade and price of copper should be stationary after one regular difference.

FIGURE 1
REAL GDP IN MILLIONS OF CHILEAN PESOS. SERIES IN LEVELS



More formally, we employ the Augmented Dickey-Fuller (ADF) test for unit roots on the series in levels, first and second differences. It is, of course, necessary to choose the number of augmentation lags to account for serial correlation and this is done using the sequential approach in Ng and Perron (1995). The results are shown in Table 1.

TABLE 1
AUGMENTED DICKEY-FULLER (ADF) TESTS

Series	Series in levels			Series in first differences			Series in second differences		
	Trend	Intercept	None	Trend	Intercept	None	Trend	Intercept	None
GDP	-2.52	0.02	2.28	-3.09	-3.09	-1.31	-28.19(**)	-28.37(**)	-28.52(**)
M1	0.41	2.60	3.07	-3.07	-1.63	-0.50	-13.63(**)	-13.68(**)	-13.69(**)
Copper	-2.36	-2.37	-0.07	-7.52(**)	-7.57(**)	-7.58(**)	-13.32(**)	-13.39(**)	-13.47(**)
Terms of trade	-1.98	-1.83	0.53	-7.87(**)	-7.92(**)	-7.91(**)	-14.17(**)	-14.24(**)	-14.33(**)

Trend, Intercept and None denote the Augmented Dickey Fuller tests applied to a regression with intercept and trend, intercept and no deterministic parameters respectively.

(**), (*) denote rejection of the null hypothesis at the 0.01, 0.05 significance levels respectively.

Estimation sample: 1982:III-2004:II.

For series in levels, the unit root null hypothesis cannot be rejected at conventional significance levels in any case. However, the unit root null can be rejected at the 0.01 level for first differences of the price of copper and terms of trade. For M1 and GDP it is necessary to take second differences in order to reject the null at the 0.01 level.

Once the number of unit roots has been determined for each of the series, the next step is to specify and estimate an ARIMA model for the Chilean GDP. In order to do this, we use the Box-Jenkins methodology based on the observation and interpretation of correlograms; see Box *et al.* (1994). The simple correlogram for the GDP series with one regular and one seasonal difference show that the series is stationary and it does not have any known structure in the regular part. However, the seasonal part suggests either a MA(1) or AR(2) specification. When the two models are estimated, the standard deviation of the estimated residuals is 0.12 in both cases. However, a MA(1) is more parsimonious than an AR(2) model and we choose the first option. This amounts to specifying the following model for the Chilean GDP (denoted by y_t):

$$(2.1) \quad \Delta_4 \Delta y_t = (1 - \theta L^4) a_t$$

where L is the lag operator; Δ and Δ_4 are operators for the regular and seasonal differences respectively; θ is a constant parameter; and $a_t \sim N(0, \sigma_a)$.

Estimation of model (2.1) by exact maximum likelihood gives the following results (with standard errors between brackets):¹

$$(2.2) \quad \Delta_4 \Delta y_t x 10^{-6} = (1 - \underset{(0.10)}{0.47 L^4}) \hat{a}_t$$

$$(2.3) \quad \hat{\sigma}_a = 0.12$$

The correlation between the two estimated parameters ($\hat{\sigma}_a$ and $\hat{\theta}$) is always below 0.01. Besides, the Jarque-Bera statistic on the estimated residuals is 1.80 indicates that they can be considered as normal. Also, the correlogram of the residuals do not show any significant correlation at any peak and its structure could be regarded as a white noise process.

From model (2.1), it is straightforward to describe a univariate model for the accumulated yearly GDP, (denoted by Y_t) as:

$$(2.4) \quad Y_t = y_t + y_{t-1} + y_{t-2} + y_{t-3} = (1 + L + L^2 + L^3) y_t.$$

Notice that

$$(2.5) \quad (1 - L^4) = (1 - L)(1 + L + L^2 + L^3),$$

then, using expressions (2.2), (2.3) and (2.5) one can obtain a model for Y_t as:

$$(2.6) \quad \Delta^2 Y_t x 10^{-6} = (1 - 0.47 L^4) \hat{a}_t$$

This last model is used for the backward interpolation of Y_t . An explanation of this process follows in the next section.

3. AN ALGORITHM FOR THE INTERPOLATION OF THE CHILEAN GDP

Once we have defined a model for Y_t , the quarterly series is constructed using a fixed interval smoothing algorithm where for each four values one is fixed. For a brief description of this process, notice that every linear econometric model with fixed parameters can be represented in the state space form as:²

¹In the estimation we use the econometric software E4 (<http://www.ucm.es/info/icae/e4/>). One important advantage of this software compared with other more conventional ones is that E4 estimates parameters by exact maximum likelihood. This method is more efficient than the estimation by conditional maximum likelihood. Moreover, exact maximum likelihood is especially convenient when initial observations in the sample are very erratic; see Hamilton (1994), Chapter 5.

²See Terceiro (1990) for a more detailed and formal description of state space models.

$$(3.1) \quad Z_t = H\alpha_t + Du_t + cv_t$$

$$(3.2) \quad \alpha_{t+1} = \Phi\alpha_t + \Gamma u_t + Ew_t$$

where Z_t is a vector of observed variables (in our case $Z_t = \Delta^2 Y_t$); u_t is a vector of exogenous variables; α_t is a vector of nonobservable state variables; w_t and v_t are white noise processes; and H, D, C, Φ, Γ and E are known parameter matrices.

Thus, for example, a possible representation of a seasonal MA(1) process in the state space form could be:

$$(3.3) \quad \alpha_{t+1} = -\theta a_{t-3}$$

$$(3.4) \quad \Delta^2 Y_t = \alpha_t + a_t$$

A fixed interval smoothing algorithm consists on obtaining an estimation of the state variable and its variance from the available sample information. We denote these estimations as $\alpha_{t/N}$ and $P_{t/N}$, $P_{t/N} = E[(\alpha_t - \alpha_{t/N})(\alpha_t - \alpha_{t/N})' / \Omega_N]$.³ A detailed description of the computational procedure of interpolation proposed by De Jong (1989) and extended by Casals *et al.* (2000) is confined to the appendix. Intuitively, the algorithm departs from some initial conditions α_1 and P_1 and then estimates the state variable using a Kalman filter and reconstructs the series backward using a smoothing algorithm.

An important drawback in the algorithm advocated by De Jong (1989) lies in the fact that the initial value of the covariance of the state value, P_1 , is arbitrarily close to infinite for all the states. Casals *et al.* (2000) extend this previous methodology by proposing a fixed interval algorithm that allows for both, stationary and unit roots, and also treats the case of exogenous inputs. This last procedure is used here to interpolate nonexistent values of y_t . Thus, from 1980 backward, for each 4 values, 3 nonexistent values are interpolated using E4.⁴ The algorithm gives the interpolated values of y_t and the estimated standard deviations of the interpolations.

4. MULTIVARIATE ANALYSIS

One potential caveat in the previous analysis is that a univariate ARIMA model only considers information referred to its own lagged values. Although our quarterly series respect the annual information provides by the Central Bank of Chile before 1980, however one could argue that some exogenous shocks could marginally affect the GDP series at specific moments. For example, the

³ N denotes the total number of observations.

⁴ E4 codes can be obtained from the authors upon requests. Some practical examples of interpolation methods can be found in the E4 manual from its webpage (<http://www.ucm.es/info/icae/e4>).

political crisis in the early seventies, the coup d'état in September 1973, the downfall in the price of copper in 1975 and the revaluation of the Chilean peso in 1976. In order to take into account some of these facts, we specify an econometric model to interpolate the Chilean GDP before 1980 using as indicators the quarterly series presented in Section 2.

The use of these series can be justified as follows. The monetary aggregate M1 series can be considered as a potential indicator to measure the relation between money and output suggested by monetarists. The use of the price of copper can be justified as there is a general agreement among economists on the influence of this mineral on the Chilean economy, see Meller (2002). The terms of trade is also an important variable to consider as it can take into account open market considerations in a small open country as Chile.

Typically, the methodology proposed by Chow and Lin (1971) has been used to estimate quarterly series from their annual values; see Pons and Surinach (1997). The procedure consists on relating the unknown quarterly series to a set of quarterly variables or indicators. When the model is expressed in terms of annual accumulated values, it can be written as:

$$(4.1) \quad Y_t = I_t' \beta + \varepsilon_t$$

where I_t is a vector of explanatory variables; β is a vector of parameters; and ε_t follows an ad-hoc univariate process. For example, Chow and Lin (1971) assume to be a white noise or a stationary AR(1) process. Fernández (1981), on the other hand, assumes that it follows a random walk.

An obvious drawback of this procedure is that it imposes a functional form that could be inaccurate given the features of the time series in our analysis. More specifically, expression (4.1) imposes the following assumptions:

- Series Y_t and I_t are cointegrated, unless we specifically assume a functional form similar to the one proposed by Fernández (1981).
- Y_t does not Granger cause any of the variables in I_t . That is, there is no reason to prefer a VAR system rather than a single transfer function equation.
- Even in the unlikely case that the two previous assumptions are true, an additional assumption imposes ε_t to follow an ad-hoc specification.

In our particular case, given the analysis in Section 2, it is clear that GDP cannot be cointegrated neither with the series of price of copper nor with terms of trade as they have different integration orders. Also, Granger causality tests, not explicitly reported here, clearly indicate that the null of no causality between the GDP and the two series, the price of copper and the terms of trade, cannot be rejected in any direction at conventional significance levels.

In order to test cointegration between GDP and M1 we use the tests proposed by Engle and Granger (1997) as this procedure is simple, intuitive and it is especially suitable to be used with only two series. Thus, first we run a regression between the two variables (GDP and M1). Then, in a second step, we

test for the presence of a unit root in the residuals of the regression using an ADF test. Results of the test do not indicate the presence of a cointegration relationship between these two variables. More specifically, the values of the ADF statistic for a regression with intercept and trend, intercept and no deterministic components are 2.07, -0.65 and -0.75 respectively.

Given these results, we include in our multivariate model the GDP and M1 series with one regular and one seasonal difference in order to ensure that they are stationary. In a first attempt, not reported here, we considered a VARMA model. However, no evidence of causation from M1 to GDP was found in that model. Then, we estimate the following model (standard deviations are between brackets):

$$(4.2) \quad (1 + L + L^2 + L^3)y_t \times 10^{-6} = \underset{(0.15)}{(0.43)} + \underset{(0.15)}{(0.28)} i_t + L + L^2 + L^3 i_t \times 10^{-6} + \frac{(1 - 0.40 L^4)}{\Delta^2} \hat{a}_t, \\ \hat{\sigma}_a = 0.11,$$

where i_t is the quarterly indicator (M1).

In this regression, the correlations among the estimated parameters are always below 0.33. Also, the residuals can be assumed to be normal with a Jarque Bera statistic of 2.09 and correlograms of the estimated residuals do not show any known structure and they are not serially correlated.

At this stage, it is important to emphasize that model (4.2) is only a reduced form statistical model and it is beyond purpose of this paper to develop any structural analysis on the relation between GDP and M1.

Using this econometric model, we interpolate the quarterly GDP series before 1980 using the smoothing algorithm described in the previous section.

5. COMPARING ALTERNATIVES

This section compares the quarterly GDP series obtained under both the ARIMA and the transfer function model. The main advantage of using a univariate model is based on its simplicity. In general, parsimonious models are especially accurate to forecast values out of the sample. Transfer function models, on the other hand, are useful in this context as they can measure the impact of other variables for which quarterly values before 1980 can be obtained. However, one can argue that there have been important structural changes in the Chilean economy from 1967 and it is very risky to assume that a set of economic indicators may have a constant effect on Chilean GDP through the whole sample. Following this argument, to test how robust our estimations are, we estimate models (2.6) and (4.2) using annual data for the whole sample (1967-2004) but results were very similar to those obtained from the sample (1980-2004).

Figure 2 shows the quarterly, y_t , and the accumulated annual, Y_t , GDP series obtained under both procedures. The figure describes the different economic

FIGURE 2
GDP SERIES IN MILLIONS OF CHILEAN PESOS

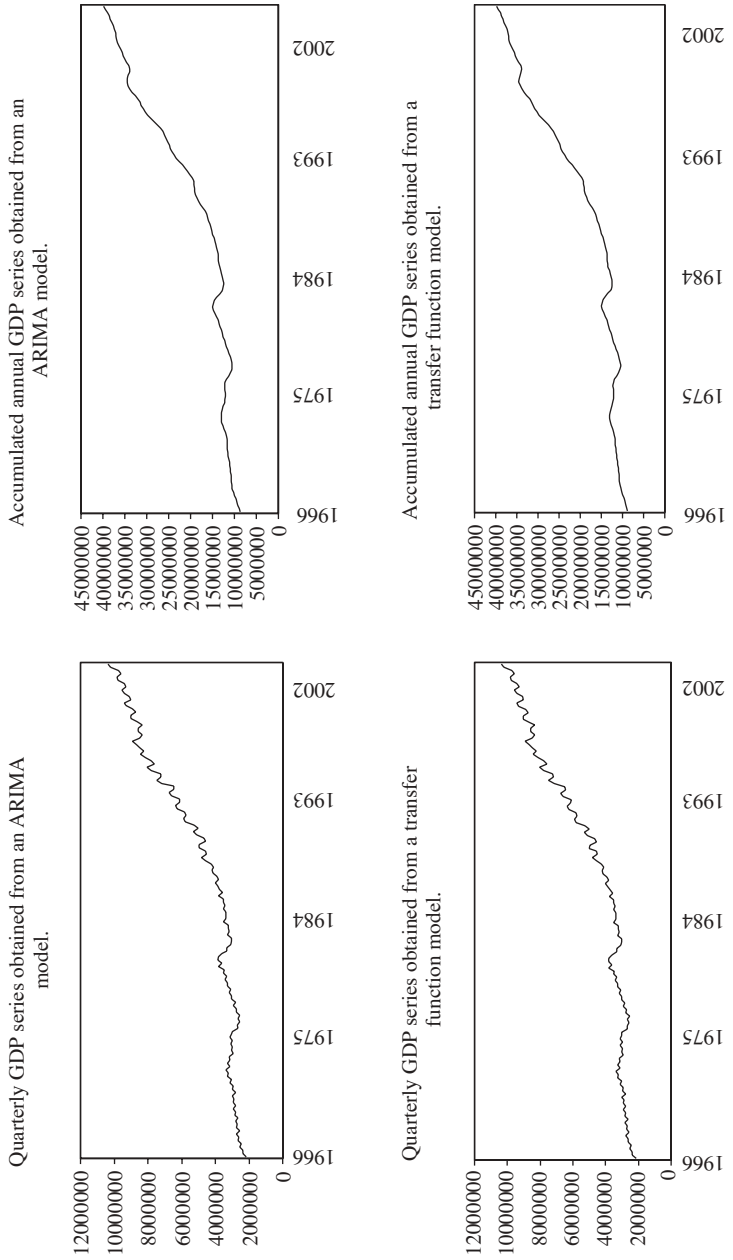


TABLE 2
 QUARTERLY GDP SERIES IN MILLIONS OF CHILEAN PESOS OBTAINED
 FROM A UNIVARIATE AND A TRANSFER FUNCTION MODEL

Year	Univariate model	Transfer function model	Year	Univariate model	Transfer function model
1966-I	2156613.7	2153258.1	1973-I	3149329.7	3147011.9
1966-II	2369677.9	2364634.6	1973-II	3271616.2	3266622.3
1966-III	2355350.1	2355745.8	1973-III	3144042.8	3143944.2
1966-IV	2534668.7	2542671.8	1973-IV	3187647.8	3195058.1
1967-I	2436189.5	2434117.0	1974-I	2930987.3	2928686.0
1967-II	2656683.8	2651701.5	1974-II	3047950.9	3042976.0
1967-III	2621572.1	2621357.4	1974-III	2963959.5	2963859.1
1967-IV	2751892.9	2759162.5	1974-IV	3100051.3	3107427.9
1968-I	2576202.0	2573821.2	1975-I	2984782.3	2982475.5
1968-II	2743786.2	2738789.1	1975-II	3148519.9	3143518.3
1968-III	2680066.5	2679998.5	1975-III	3016684.9	3016566.2
1968-IV	2806081.0	2813526.9	1975-IV	3010315.6	3017742.7
1969-I	2650385.7	2648079.0	1976-I	2657967.9	2655794.2
1969-II	2834982.8	2829989.3	1976-II	2700608.3	2695667.8
1969-III	2785293.0	2785189.7	1976-III	2563657.7	2563525.6
1969-IV	2922354.5	2929758.0	1976-IV	2668154.3	2675400.5
1970-I	2774723.5	2772399.2	1977-I	2542654.1	2540420.9
1970-II	2954458.3	2949464.1	1977-II	2759415.8	2754825.4
1970-III	2886979.7	2886884.9	1977-III	2743860.2	2743927.7
1970-IV	2993326.0	3000739.2	1977-IV	2917025.6	2923781.7
1971-I	2802053.1	2799732.1	1978-I	2807468.0	2805173.0
1971-II	2977271.7	2972277.3	1978-II	3032764.2	3028582.3
1971-III	2944402.6	2944306.0	1978-III	3018335.1	3018327.3
1971-IV	3124483.9	3131895.9	1978-IV	3185218.9	3191703.7
1972-I	3046071.8	3043752.4	1979-I	3061971.6	3059268.5
1972-II	3274930.8	3269937.8	1979-II	3278958.2	3274143.5
1972-III	3236481.6	3236383.8	1979-III	3261599.6	3261299.6
1972-IV	3351762.6	3359172.7	1979-IV	3430933.9	3438751.7

cycles in Chile during the period 1965-1971. Thus, there is a moderate growth in the years 1965-1971 motivated by a strong public investment. After this period, there is stagnation phase in 1973-1982 characterized by tight monetary policies aiming to control a high rate of inflation. There is a strong recovery in 1983-1998 stimulated mainly by foreign trade. Finally, the Asian crisis affected the Chilean economy at the end of the nineties slowing down economic growth from 1998 up to the present date.

From this figure, an important feature to remark is that the two new quarterly series describe the same economic trend as the accumulated annual series from the Central Bank of Chile. Besides, the two quarterly series reflect the seasonal behaviour inherent in the Chilean economy.

When the two different models are compared, it is clear that the series interpolated under the univariate and the transfer function model are almost identical. This is also corroborated by a Pearson's correlation value of 0.9998 between the quarterly series interpolated under both procedures for the period 1967:01-1979:04. Given this result, it is not clear which alternative is superior. Therefore, we leave to the analyst's criteria to choose between the two options. This election will not have dramatic consequences in any case as both series respect the annual GDP information provided by the Central Bank of Chile. The numerical values obtained from the two interpolations are shown in Table 2.

6. CONCLUSIONS

In this paper we obtain quarterly series of the Chilean GDP for the period 1966-1979 using the fixed interval smoothing algorithm proposed by De Jong (1989) and extended by Casals *et al.* (2000) for the nonstationary case. We propose two alternatives for the interpolation based on a univariate and a transfer function model. Quarterly series obtained in both cases are consistent with the annual information provided by the Central Bank of Chile before 1980.

A new line of research opened by this paper is based on the use of the quarterly GDP series to specify and estimate more efficient econometric models that could be considered to describe the relation of the Chilean GDP with other economic (national and international) variables for which quarterly information is already available for long periods.

REFERENCES

- Bennett, H. and R. Valdés (2001). "Series de Términos de Intercambio de Frecuencia Mensual para la Economía Chilena: 1965:1999", Working paper N° 98, Central Bank of Chile.
- Box, G.E.P., G.M. Jenkins and G.C. Reinsel (1994). *Time Series Analysis: Forecasting and Control*. (Third Edition), Englewood Cliffs, HJ: Prentice Hall.
- Casals, J., M. Jerez and S. Sotoca (2000). "Exact Smoothing for Stationary and Nonstationary Time Series", *International Journal of Forecasting*, 16, 59-69.
- Chow, G. and Lin, A.L. (1971). "Best Linear Unbiased Distribution and Extrapolation of Economic Time Series by Related Series", *The Review of Economics and Statistics*, 53, 471-76.
- De Jong, P. (1989). "Smoothing and Interpolation with the State-Space Model", *Journal of the American Statistical Association*, 84, 1085-88.
- Engle, R.F. and Granger, C.W.J. (1987). "Co-Integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-76.
- Fernández, R.B. (1981). "Methodological Note on the Estimation of Time Series", *The Review of Economic and Statistics*, 63, 471-478.
- Haindl, R. (1986). "Trimestralización del Producto Geográfico Bruto por Origen y Destino", *Estudios de Economía*, Universidad de Chile, 13, 119-153.

- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Meller, P. (2002). *Dilemas y Debates en Torno al Cobre*. Dolmen.
- Ng, S. and P. Perron (1995). "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag", *Journal of the American Statistical Association*, 429, 268-81.
- Pons, E. and J. Surinach (1997). "Trimestralización y Conciliación de Magnitudes Económicas: una Ampliación del Método de Chow-Lin", *Working Paper* N° 20, Universidad de Barcelona.
- Terceiro, J. (1990). *Estimation of Dynamic Econometric Models with Errors in Variables*. Berlin: Springer-Verlag.

APPENDIX
INTERPOLATION PROCEDURE

This appendix describes the procedure used for interpolation of the Chilean GDP before 1980. The methodology follows the lines stated by De Jong (1989) and extended by Casals *et al.* (2000) for the nonstationary case.

Let assume the following state-space model:

$$(A1) \quad Z_t = H\alpha_t + Du_t + Cv_t$$

$$(A2) \quad \alpha_{t+1} = \Phi\alpha_t + \Gamma u_t + Ew_t$$

where (A1) is the observation equation that generates the $(m \times 1)$ vector of measures Z_t , $t = 1, 2, \dots, N$, u_t is a $(r \times 1)$ vector of inputs and the state equation (A2) describes the evolution of the $(n \times 1)$ state vector α_t .

We make the following assumptions about (A1)-(A2):

- i) $w_t \sim IID(0, Q)$, $v_t \sim IID(0, R)$, $cov(w_t, v_t) = S$, for all $t = 1, 2, \dots, N$.
- ii) The initial state is independent of w_t and v_t , and such that $\alpha_1 / u_1, \dots, u_N \sim (\bar{\alpha}_1, P_1)$.
- iii) The matrices $H, D, C, \Phi, \Gamma, E, Q, R$ and S have been estimated previously whereas $\bar{\alpha}_1$ and P_1 are unknown.

We also denote the information available up to $t = j$ by: $\Omega_j = \{Z_1, Z_2, \dots, Z_j, u_1, u_2, \dots, u_j\}$ and the first and second-order conditional moments of the state vector by: $\alpha_{t|j} = E(\alpha_t / \Omega_j)$ and $P_{t|j} = E[(\alpha_t - \alpha_{t|j})(\alpha_t - \alpha_{t|j})' / \Omega_j]$.

Following De Jong (1989), based on (A1) and (A2), one can interpolate Z_t , u_t or α_t onto the space $[Z_1, Z_2, \dots, Z_{t-1}, Z_{t+1}, \dots, Z_N]$ using a fixed interval smoothing algorithm. This algorithm consists of a forward step given by a standard Kalman filter and a backward recursion that, for the case of fixed-interval smoothing, takes the form:

$$(A3) \quad \alpha_{t|N} = \alpha_{t|t-1} + P_{t|t-1} r_{t-1}$$

$$(A4) \quad P_{t|N} = P_{t|t-1} - P_{t|t-1} R_{t-1} P_{t|t-1}$$

$$(A5) \quad r_{t-1} = H' B_t^{-1} \tilde{Z}_t + \bar{\Phi}_t^T r_t, \text{ with } r_N = 0$$

$$(A6) \quad R_{t-1} = H' B_t^{-1} H + \bar{\Phi}_t^T R_t \bar{\Phi}_t, \text{ with } R_N = 0$$

$$(A7) \quad \bar{\Phi}_t = \Phi - K_t H$$

where $\alpha_{t|t-1}$ and $P_{t|t-1}$ were computed in a forward step, $\tilde{Z}_t = Z_t - Z_{t-1}$ is the sequence of Kalman filter innovations corresponding to (A1)-(A2); B_t is the covariance matrix of \tilde{Z}_t ; and K_t is the Kalman filter gain.

Model (A1)-(A2) can be stationary, nonstationary or partially stationary, depending on the eigenvalues of Φ . De Jong (1989) emphasize the importance of initialization in the forward filtering phase and propose adequate solutions for stationary systems. However, in the nonstationary case the initial state covariance, P_1 , is arbitrarily close to infinity and, therefore, these initialization criteria cannot be used.

In a pure nonstationary framework, a common practice consists of approximating the diffuse initial conditions by $P_1 = kI$, where k is an arbitrary big value. Frequent 'rule of thumb' values for k may vary between $k = 10^2$ and $k = 10^7$. But, although this initialization allows one to keep using standard algorithms, however it commonly generates biased results when k is not adequately chosen and induces numerical error.

Casals *et al.* (2000) derive an exact algorithm that can be applied to stationary, nonstationary and partially nonstationary systems. This algorithm is obtained from the combination between exact fixed-interval smoothed moments and those obtained from an arbitrarily initialized smoother. To see this, consider the state space model:

$$(A8) \quad Z_t^* = H\alpha_t^* + Du_t + cv_t$$

$$(A9) \quad \alpha_{t+1}^* = \Phi\alpha_t^* + \Gamma u_t + Ew_t$$

where the states and measures correspond to (A1)-(A2) with the initial conditions $\alpha_1^* = 0$ and $P_1^* = 0$.

It is straightforward to see that propagating the state equations (A2) and (A9), it follows that

$$(A10) \quad \alpha_{t|N} = \Phi^{t-1}\alpha_{1|N} + \alpha_{t|N}^*$$

Also, from (A1)-(A2) and (A8)-(A9), Casals *et al.* (2000) prove that

$$(A11) \quad \tilde{Z}_t = H\bar{\bar{\Phi}}_{t-1}\alpha_1 + \tilde{Z}_t^*$$

where $\tilde{Z}_t^* = Z_t^* - Z_{t|t-1}^*$ are the innovations resulting from a Kalman filter applied to (A8)-(A9), hereafter $KF(0,0)$; and the matrices $\bar{\bar{\Phi}}_{t-1}$ are given by $\bar{\bar{\Phi}}_t = (\Phi - K_t H)\bar{\bar{\Phi}}_{t-1}$ with $\bar{\bar{\Phi}}_1 = I$.

Equation (A11) can be written for all the sample as:

$$(A12) \quad \tilde{Z}_t = X\alpha_1 + \tilde{Z}_t^*$$

where X is the block-diagonal matrix whose t -th block is $H\bar{\bar{\Phi}}_{t-1}$ and the $(m \times N) \times 1$ vectors \tilde{Z} and \tilde{Z}^* contain the $KF(0,0)$ innovations \tilde{Z}_t and \tilde{Z}_t^* respectively. Note that \tilde{Z}^* is independent of α_1 .

Then, the problem consists of obtaining the conditional expectations in the right-hand-side of (A10), taking into account the relationship (A12). The solution is given by the following expressions:

$$(A13) \quad \alpha_{t/N} = \left\{ \Phi^{t-1} - E \left[\alpha_t^* (\tilde{Z}^*)' \right] B^{-1} X \right\} \alpha_{t/N} + E \left[\alpha_t^* (\tilde{Z}^*)' B^{-1} \tilde{Z} \right]$$

$$(A14) \quad P_{t/N} = \left\{ \Phi^{t-1} - E \left[\alpha_t^* (\tilde{Z}^*)' \right] B^{-1} X \right\} P_{t/N} + \left\{ \Phi^{t-1} - E \left[\alpha_t^* (\tilde{Z}^*)' \right] B^{-1} X \right\} + P_{t/N}^*$$

An important point to be noted about these two expressions is that they can be applied to stationary, nonstationary and partially nonstationary systems, as the only term affected by P_1 are $Z_{1/N}$ and $P_{1/N}$ and this dependence occurs through P_1^{-1} , which is finite.

Algorithms for the computation of $E \left[\alpha_t^* (\tilde{Z}^*)' \right] B^{-1} X$ and $E \left[\alpha_t^* (\tilde{Z}^*)' \right] B^{-1} \tilde{Z}$ can be found in Casals *et al.* (2000), Section 3.