Journal of Forecasting
J. Forecast. **20**, 37–45 (2001)

Empirical Analysis of Systematic Errors in Chilean GDP Forecasts

RÓMULO A. CHUMACERO*
Department of Economics, University of Chile

ABSTRACT
This paper presents a statistical comparison between the actual and predicted evolution of the Chilean GDP for the period 1986–1998 made by several forecasters. We show that the forecasters systematically underestimate the true growth rate of the economy. The magnitude of this bias tends to be correlated with the phase of the business cycle. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS forecasting; business cycles

INTRODUCTION

The Chilean economy presents one of the highest average growth rates in the Western Hemisphere. In fact, the average annual growth rate of the GDP between 1986 and 1997 was approximately 7.7%.

Accompanying this spectacular performance there has been a proliferation of projections of the growth of the economy by different sources. In this work we evaluate these projections by comparing them with the actual evolution of the rate of growth of the GDP for the period 1986 to 1997.

Although there is a long-standing tradition on evaluating the performance of economic forecasters (e.g. Zarnowitz and Lambros, 1987; McNees, 1989; Ito, 1990; Keane and Runkle, 1990; Bonham and Cohen, 1995; Lamont, 1995; Ehrbeck and Waldmann, 1996; Laster, Bennet, and Geoum, 1997; Stark, 1997) this is the first attempt to do so for the case of Chile.

We gathered 857 forecasts made by different sources and published in the financial newspaper *Estrategia* during the same period. These forecasts correspond to projections made by 43 individual forecasters, 28 organizations, 16 commercial banks, 23 private companies, 7 insurance companies, and 11 pension funds.[1]

The paper is organized as follows. The next section presents a comparison between the effective

* Correspondence to: Rómulo A. Chumacero, Department of Economics, University of Chile, Diagonal Paraguay, Torre 26, Santiago, Chile. E-mail: rchumace@econ.uchile.cl

[1] The 'lead time' of the forecasts varies because *Estrategia* collected the forecasts at different times in different years. On average, the 'lead time' is 9 months with a standard deviation of 4 months.

evolution of the growth rate of the GDP and the projections made by the forecasters and the third section presents some final comments.

HOW WELL DO FORECASTERS DO?

This section analyses the statistical properties of the forecast errors made by several forecasters during the period 1986 to 1997. Our objective is not to provide an explanation of why they incurred these errors but to stress some of the empirical regularities that can be associated with them.

How are they distributed?

In this paper we define the absolute (relative) forecast error ($e_{i,t}$ and $r_{i,t}$ respectively) made by forecaster i in period t as the difference (ratio) between the effective value of the growth rate of GDP in period t (g_t) and the forecast made by forecaster i for that period ($f_{i,t}$). That is:

$$e_{i,t} = g_t - f_{i,t} \quad r_{i,t} = \frac{g_t}{f_{i,t}}$$

According to our definition, an underestimate of the growth rate would lead to a positive value for e and a value of r greater than one. Conversely, an overestimate would lead to negative values for e and values of r smaller than one.

Let us begin by defining some of the most important properties that have these variables. For that purpose, we build a vector of absolute and relative errors to evaluate their unconditional distribution. As mentioned previously, 857 observations were available. Figure 1 shows the non-

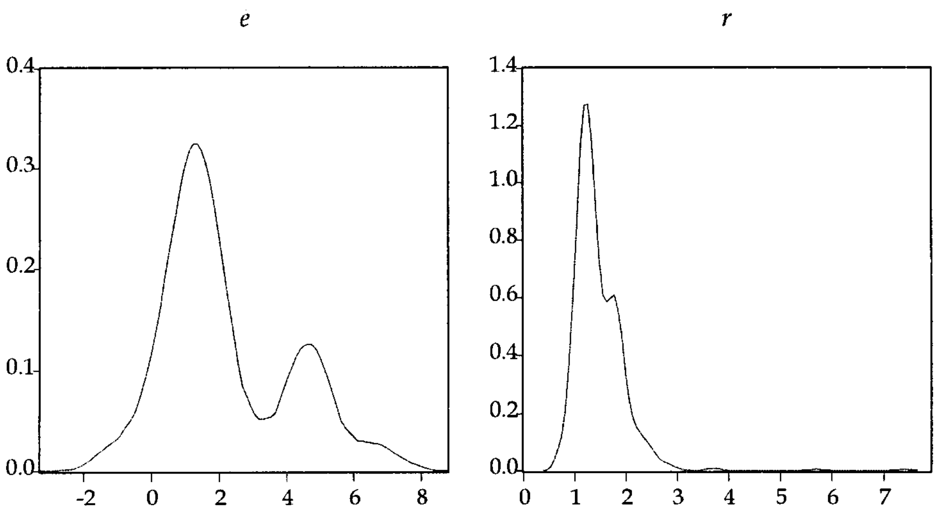

Figure 1. Unconditional densities of e and r. (*Note*: the unconditional densities were estimated using the Epanechnikov kernel and the bandwidth selection proposed by Silverman, 1986)

Table I. Descriptive statistics of e and r

	e		r	
Statistic	Value	p-value	Value	p-value
Mean	2.165	0.000	1.458	0.000
Median	1.600	0.000	1.325	0.000
Standard deviation	1.892		0.477	
CV	0.874		0.327	
S	0.727	0.000	3.751	0.000
K	2.821	0.269	38.171	0.000
JB	76.651	0.000	46180.890	0.000

Mean = the p-values correspond to the nulls that the mean of e is 0 and of r is 1. Median = the p-values correspond to the null that the median of e is 0 and of r is 1. CV = coefficient of variation. S = Skewness. The p-value corresponds to the null that S is 0. K = Kurtosis. The p-value corresponds to the null that K is 3. JB = Jarque and Bera Normality test. The p-value corresponds to the null of $S = 0$ and $K = 3$.

parametric estimators of the unconditional distributions of both series, while Table I displays some of their descriptive statistics.

As can be seen the unconditional distributions of both variables show strong departures from normality. Both distributions are bimodal and asymmetric. Notice that in both cases there is an important bias towards underestimation. That is, e is biased towards positive values and r is biased towards values exceeding 1. In fact, simple tests show that the null of unbiased forecast errors is strongly rejected in both cases.[2] More importantly, on average, the forecasters underestimated the growth rate of GDP by more than 2 points.

An important characteristic that a forecast error should have is that it should not be systematic (unpredictable).[3] However, as Figure 1 shows, and formal predictability tests would confirm, this is hardly the case. In fact, as we will show later, the underestimation bias is always present.

Are they all the same?

Given that in our database we can follow different individuals through time and through affiliation, we can evaluate whether there is any systematic difference among groups. This will enable us to verify whether there is a group of forecasters that dominates another group.

Figure 2 and Table II show the estimation of the unconditional density functions and equality tests for means among six groups. As can be observed, all the densities (with the sole exception of insurance companies) present evidence of bimodality.[4] Table II also shows a test for equality of the means among groups (both for absolute and relative errors). As the null cannot be rejected at standard levels of significance, for every practical purpose, there is no statistical difference

[2] Despite being asymptotically valid, the tests of equalities (in mean and median) assume independence. More formal tests will be developed later.

[3] There is, however, a large body of empirical and theoretical literature that advances some ideas on why forecasts may be biased (see e.g. Stark, 1997).

[4] As Table II shows, there were only ten forecasts made by insurance companies in the whole sample. Thus, as will be shown later, this group does not refute the existence of bimodality in the other groups and on the aggregate.

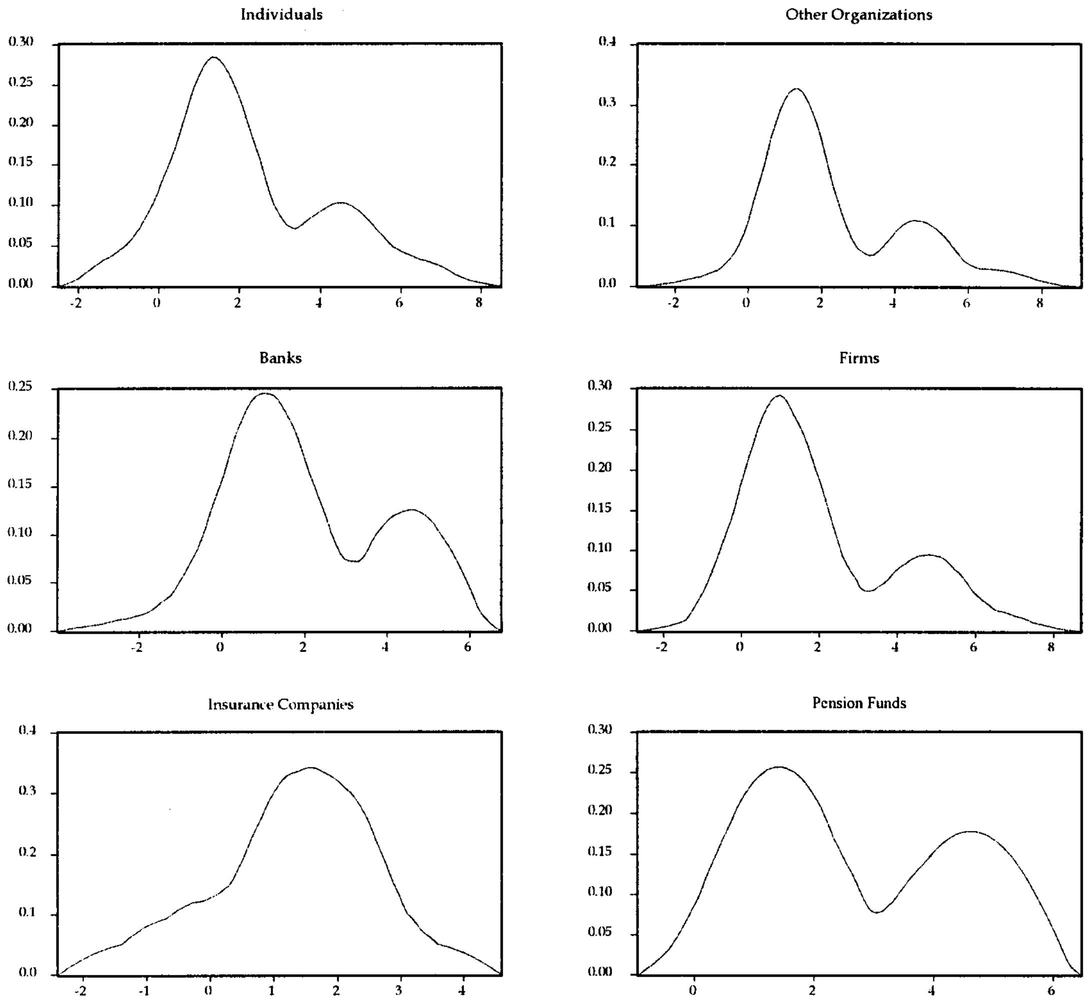

Figure 2. Function of density and of group. (*Note*: the unconditional densities were estimated using the Epanechnikov kernel and the bandwidth selection proposed by Silverman, 1986)

among the groups. Thus these forecasters tend to commit the same type of errors both in terms of direction as well as magnitude.[5]

When do they make fewer mistakes?

As the evidence shows, the forecasters not only make systematic mistakes but the unconditional distribution of the forecast errors is bimodal. This fact can be easily explained when observing the forecast errors by year. If we associate a 'contraction' with a year in which the GDP grew less than average (7.7%) and an 'expansion' with its complement, we observe that the forecast

[5] Although not reported, there is weak evidence of Granger causality from individual forecasters (mostly academics in different universities) to the forecasters in other organizations (mostly producer organizations).

Table II. Descriptive statistics of *e* and *r* by group

		e		*r*	
Group	Observations	Mean	Deviation	Mean	Deviation
Individuals	353	2.183	0.101	1.469	0.030
Other organizations	269	2.249	0.115	1.474	0.026
Banks	58	1.999	0.247	1.439	0.057
Firms	139	1.986	0.163	1.399	0.031
Insurance companies	10	1.325	0.373	1.386	0.123
Pension funds	28	2.675	0.310	1.514	0.047
Total	857	2.165	0.065	1.457	0.016
		Test	*P*-value	Test	*P*-value
ANOVA		1.252	0.283	0.668	0.648

Deviation = standard error of the mean. ANOVA = test of equality of means whose asymptotic distribution is an *F*-test with 5 degrees of freedom in the numerator and 851 in the denominator.

errors are, in addition to systematic, asymmetric. That is, the absolute (relative) errors, despite being positive (exceeding 1) in all the phases of the cycle, are smaller in the contractions than in the expansions.

Figure 3 and Table III summarize this evidence. When we condition the estimation of the densities of the forecast errors to the 'phase of the cycle' we now obtain unimodal distributions, although in both cases they are biased towards underestimation (as the simple mean tests suggest). Despite this, the forecast errors are asymmetric, in the sense that the underestimation is smaller in a contraction than in an expansion. Thus, on average, the forecasters underestimate the growth rate of the economy by more than one point during the contractions and by close to five points during the expansions.

Should this give us any comfort? Very little, because from the results reported it is easy to verify that the variation coefficient of *e* is 2.3 times greater during a contraction than in an expansion (this coefficient is 3.6 times greater in the case of *r*). Summarizing, the forecasters are unnecessarily pessimistic in all the phases of the cycle, but particularly so during the expansions. The forecast errors (and therefore their projections) are more volatile (in relative terms) during a contraction than during an expansion.

Do they learn from their mistakes?

Even though the forecasts tend to present systematic biases towards underestimation, we could ask ourselves if the forecast errors tend to diminish when the forecasts are made closer to the period of projection. Given that *Estrategia* conducts several surveys during a given year, we construct a series that measures the distance (in months) between the period where the forecast was made and the period for which that forecast was made. Denoting the resulting variable by L, Table IV shows the results of a regression between the forecast errors and L.

The regression displays several interesting features. Given that the forecast errors are systematic and asymmetric (with respect to the phase of the cycle) we include a dummy variable that controls for this factor. As seen, this variable is highly significant for both absolute and relative errors. If the forecasters learned from their mistakes, we would expect the coefficient associated with L to be positive given that further away from the realization of the series (the greater the value of L)

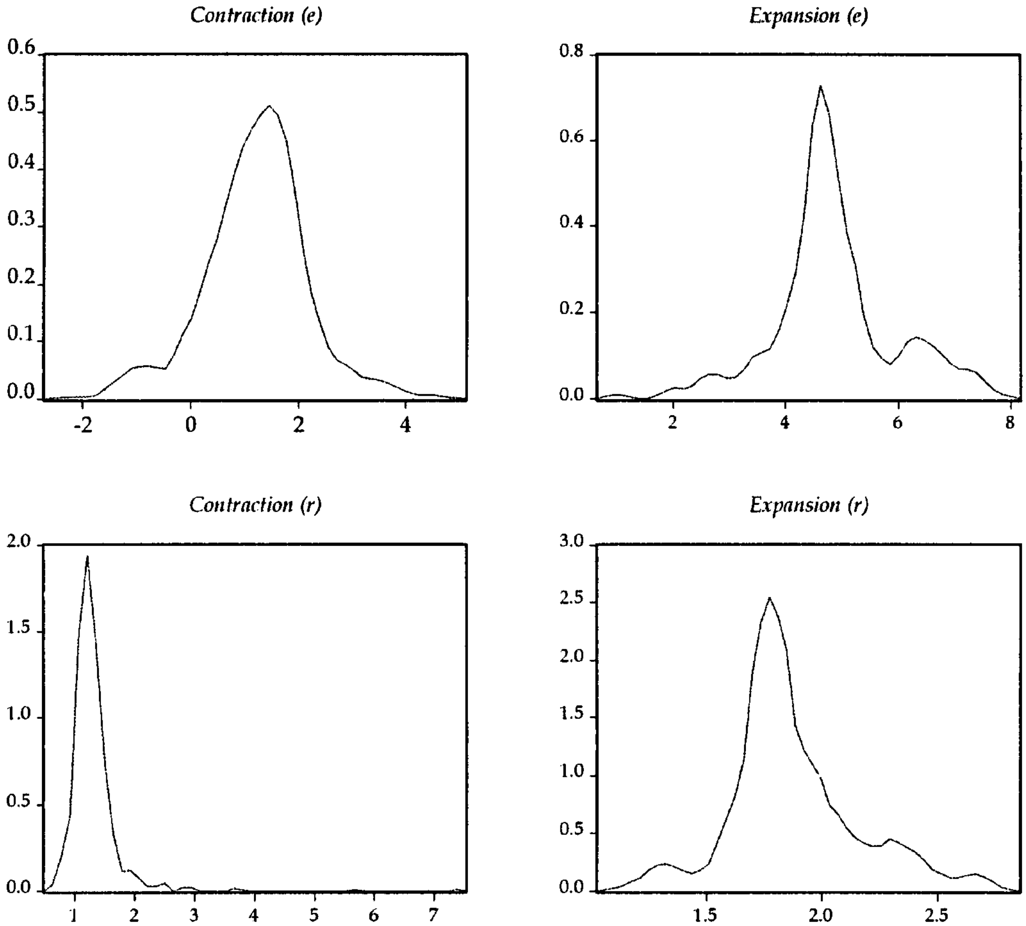

Figure 3. Density functions of e and r by phase of the cycle. (*Note*: the unconditional densities were estimated using the Epanechnikov kernel and the bandwidth selection proposed by Silverman, 1986)

Table III. Descriptive statistics of e and r by phase of the cycle

		e		r	
Group	Observations	Mean	Deviation	Mean	Deviation
Contraction	629	1.195	0.038	1.307	0.018
Expansion	228	4.842	0.072	1.872	0.018
Total	857	2.165	0.065	1.458	0.016
		Test	P-value	Test	P-value
ANOVA		2266.175	0.000	323.296	0.000

Deviation = standard error of the mean. ANOVA = test of equality of means whose asymptotic distribution is an F-test with 1 degree of freedom in the numerator and 855 in the denominator.

Table 4. Linear regression models for *e* and *r*

	e			*r*		
	Parameter	Deviation	*p*-value	Parameter	Deviation	*p*-value
Constant	0.866	0.213	0.001	1.335	0.120	0.001
L	0.035	0.023	0.130	−0.003	0.011	0.787
D	3.645	0.172	0.000	0.565	0.048	0.001
	$R^2 = 0.729$	SER = 0.985		$R^2 = 0.273$	EER = 0.407	

Deviation = standard error of the parameter computed using the HAC variance–covariance matrix. L = months between the projection and the realization of the series. D = dummy variable that adopts the value of 1 in expansion and 0 in contractions. R^2 = adjusted R^2. SER = standard error of the regression.

we would expect the forecasters to be less accurate. It turns out that once heteroscedasticity and autocorrelation consistent (HAC) estimates of the variance of the parameters are used this variable is not statistically significant for standard levels of significance (the associated *p*-values for *e* and *r* are 0.13 and 0.79 respectively). The inclusion of HAC estimates for the variance–covariance matrix is justified because the forecast errors display persistence and because White's heteroscedasticity tests (not reported) suggest its presence.[6]. On the other hand, the coefficient associated with *L* for the regression on *r* is not significant and has the 'wrong' sign.

Why do they err that much?

So far we have shown that the forecast errors are systematic and asymmetric. Here, we present a tentative explanation of why this phenomenon may occur. A reasonable search of variables that may explain the asymmetry found in the forecast errors is to see if there is any variable that has a different behaviour according to the phase of the cycle. A natural candidate is, of course, the interest rate.

From an intertemporal perspective, the real interest rate is simply a relative price (between consumption today and tomorrow). Thus, if the economy is in a contraction that the agents perceive as transitory, their willingness to smooth their consumption stream would (generally) create pressure for the interest rates to rise. Thus, it is not uncommon to find (weak) negative contemporary correlations between the growth rate of the economy and interest rates.

Even though the previous discussion applies to real interest rates, it is not uncommon for it to be translated to nominal or imperfectly indexed interest rates. In Chile, the Central Bank has a short-term instrument (an imperfectly indexed interest) called the PRBC that is usually taken to consider the stance of the monetary authority. In fact, this interest rate presents a mild negative contemporary correlation with the growth rate of GDP.

[6] In fact, the *p*-value associated with White's heteroscedasticity test for the regressions on *e* and *r* are 0.002 and 0.010 respectively. It is worth noting that *L* turned out to be significant in 'explaining' the squared residuals (with a positive sign). Estimations with weighted least squares (using *L* as the weight) were also performed without changing the results of Table IV significantly. This would imply that even though they tend to make the same mistakes regardless of how near the effective realization of the variable is, at least their forecast errors tend to appear similar due to the reduction in variance.

Table 5. Linear regression models for e and r

	e			r		
	Parameter	Deviation	p-value	Parameter	Deviation	p-value
Constant	3.765	0.610	0.001	1.608	0.220	0.001
PRBC	−39.352	9.168	0.001	−4.606	3.577	0.198
D	3.423	0.157	0.001	0.539	0.043	0.001
	$R^2 = 0.764$	SER = 0.920		$R^2 = 0.281$	EER = 0.404	

Deviation = standard error of the parameter computed using the HAC variance–covariance matrix. L = months between the projection and the realization of the series. D = dummy variable that adopts the value of 1 in expansion and 0 in contractions. R^2 = adjusted R^2. SER = standard error of the regression.

Table V shows the results from incorporating this variable into a regression for e and r. As can be observed, this variable is highly significant (at least in explaining the variation of the absolute errors), displaying a negative coefficient. This means that when this instrument increases, the forecast errors tend to diminish. This is congruent with the asymmetry found previously.

A misguided interpretation of these results would attribute some type of economic cause from PRBC to the growth rate of the economy. This interpretation is not correct because the PRBC (as any other interest rate) is a good leading indicator of the expectations of the growth of the economy; this fact does not bring any economic causation from one variable to the other. Thus, statistical precedence does not necessarily imply economic causation (see Chumacero, 1998, unpublished manuscript, for a detailed discussion).

The more reasonable explanation of these results is precisely the converse. Recalling that the dependent variable is the forecast error (not the growth rate of the economy), and that we already controlled for the phase of the cycle, the results tend to show that forecasters tend to attribute an unjustified influence to the monetary authority's stance in having real effects.

FINAL COMMENTS

This paper presents statistical evidence that shows that forecasters systematically underestimate the true growth rate of the Chilean economy for the period 1986 to 1997. The theoretical and empirical literature on the theory of forecasting advances some rationalizations of why forecasters may not be solely interested in minimizing their forecast errors. Strategic interactions and reputational incentives, among other factors, may provide an explanation of why forecasts may be biased.

Even though these explanations may help to justify some of the results of this study, there are some regularities that cannot. They may help to explain why forecasters commit the same mistakes in a given period, but not why they always underestimate the growth rate of the economy.

This study shows that, at least for the Chilean case, some additional factors have to be considered. The presence of bimodality in the unconditional distribution of forecast errors; its association with the 'phase of the cycle'; the fact that, independently of the phase of the cycle, forecasters of the sample always underestimated the growth rate of the economy; and that in addition to being correlated with the phase of the cycle, the magnitude of the bias in forecast errors is also associated with the monetary authority's stance. This regularity appears to present a more promising avenue for further research.

ACKNOWLEDGEMENTS

I would like to thank Rodrigo Fuentes, Osvaldo Larrañaga, Ricardo Paredes, José Miguel Sánchez and an anonymous referee for helpful comments. Carlos Oyarzún and Patricia Toledo provided able research assistance. The usual disclaimer applies.

REFERENCES

Bonham C, Cohen R. 1995. Testing the rationality of price forecasts: Comment. *American Economic Review* **85**: 284–289.

Ehrbeck T, Waldmann R. 1996. Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics* **CXI**: 21–40.

Ito T. 1990. Foreign exchange rate expectations: Micro survey data. *American Economic Review* **80**: 434–449.

Keane MP, Runkle DE. 1990. Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review* **80**: 714–735.

Lamont O. 1995. Macroeconomic forecasts and microeconomic forecasters, National Bureau of Economic Research Working Paper #5284.

Laster D, Bennett P, Geoum IS. 1997. Rational bias in macroeconomic forecasts, Federal Reserve Bank of New York Staff Reports Number 21.

McNees SK. 1989. Why do forecasts differ? *Federal Reserve Bank of Boston New England Economic Review*: 42–54.

Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall: London.

Stark T. 1997. Macroeconomic forecasts and microeconomic forecasters in the survey of professional forecasters, Federal Reserve Bank of Philadelphia Working Paper No. 97-10.

Zarnowitz V, Lambros LA. 1997. Consensus and uncertainty in economic prediction. *Journal of Political Economy* **95**: 591–621.

Author's biography:
Rómulo A. Chumacero is an Assistant Professor at the Department of Economics of the University of Chile. He holds a B.A. in Economics from the Catholic University of Bolivia, a M.A. in Economics from ILADES/Georgetown University and a Ph.D. in Economics from Duke University. His research interests are in Econometrics and Macroeconomics.

Author's address:
Rómulo A. Chumacero, Department of Economics, University of Chile, Diagonal Paraguay, Torre 26, Santiago, Chile.

