Abstract

This paper develops a comparative analysis of different forecasting approaches for an aggregate which is a main target of the monetary policy in Argentina. First we present the results of estimating a conditional equilibrium-correction model of money demand, which is theory consistent, accounts for the main features of the data and shows parameter constancy. Then we compare its forecasts with those obtained by other methods: VAR in differences, “naive” models, robustified devices, forecasting aggregates by disaggregates and pooling of forecasts using different models and windows. They are evaluated over an unstable period in which there was often uncertainty about the economic regime. For this aggregate it can be useful to apply several of these methods to complement the forecasts of the equilibrium-correction model, in particular the pooling of forecasts obtained by this conditional model with those forecasts of models that do not require the future path of the conditioning variables to be known.

Keywords: Forecasting – Equilibrium-correction – Monetary-aggregate

JEL: C53, E41, E47
1. Introduction

Parameter instability pervade empirical models, particularly those of monetary aggregates and interest rates (Stock and Watson, 1996, 2003). Breaks, both in-sample and out-of-sample but particularly the latter, are extremely critical to understand forecasting abilities. Thus, main developments in forecasting theory have been focussed on data generating processes subject to major and unanticipated shifts (Clements and Hendry, 1998, 1999). In this more realistic setting conventional results have been reviewed. Specifically, a model which approximates the conditional expectations of the data (“an econometric model”) does not need to produce the best forecasts. In turn, a simpler and non causal model (“a naive model”) “can” produce better forecasts as many of this kind of models are also more adaptive to the unanticipated changes.

Taking into account the assumption of intermittent and unanticipated shifts to understand economic forecasting is good news for unstable economies like Argentina, where empirical researchers are accustomed to face parameter instability and forecast failures. Argentine structural breaks and policy regime changes have not been exceptional but largely intermittent and mainly unanticipated. Thus, this environment provides a fruitful and defying field to study the relative performance of different forecasting methods. No previous work has been focussed on forecasting performance for Argentina’s monetary aggregates except Aguirre et. al. (2006). They considered the issue of forecasting a set of different monetary aggregates during 1993-2005 period.

Forecasting monetary aggregates has been a difficult task usually associated to the intrinsic instability of the money demand. Therefore, the aim of this paper is twofold. Firstly, we estimate a money demand for M2, which has become recently a main
target of the monetary policy. This aggregate was modelled by an equilibrium-correction model (EqC) which shows parameter stability “within-sample”. Secondly, its forecasting performance is compared with a wide range of forecasting alternative approaches which include VAR in differences, “naive” (univariate) models, robustified devices, forecasting aggregates by disaggregates and pooling of forecasts obtained using different models. The effect on accuracy of forecast pooling using different estimation windows for a given autoregressive model is also evaluated.

The forecasting sample is 2002-2006. This period seems very interesting to study because it includes the aftermath of the abandonment of the Convertibility regime, which lasted for more than ten years and a singular event, the default on the sovereign debt. In this period there was also a lot of uncertainty about stabilization results until they finally were obtained.

The next section presents a description of the data from a historical perspective. Section 3 shows the results of estimating an equilibrium-correction model of money demand. Section 4 analyses its forecast performance in comparison with the other approaches. Section 5 concludes.

2. Data Description.

The analysis is focussed on an Argentine monetary aggregate denoted as M2 (defined as narrow money, current account and saving deposits in pesos). Quarterly data over 1977-2006 are empirically studied, although different estimation windows are also used according to the purpose of the model. Figure 1 shows the time plot of real M2 (deflated by the consumer price index), expressed in logs (mp) and in log differences (Dmp), along with transactions (y) and the velocity (ymoon).
jointly with the nominal interest rate (i) for the whole sample. In this Figure two periods can be observed according to the underlying trend of real money: downwards until 1991 and upwards after this year.

Figure 1: Real M2, in log levels, in log differences and its long run determinants

The first period (1977-1991) was also characterised by an upward trend in inflation that accelerates in the mid-seventies when consumer prices passed the 50% annual rate and becomes a hyperinflation process in 1989 and 1990. The interest rate reflects this behaviour. However, the downward trend of real money was accompanied by several attempts to stabilize inflation. In 1985 a stabilization program known as “Plan Austral” led to a temporary decrease in inflation and to an increase of money holdings, but inflation soon accelerated and the reduction in real money holdings was dramatically during the hyperinflation process of the end of 1989 and the beginning of 1990.
The upward trend in real money holdings started in 1991 along with the Convertibility regime that backed the money base on external reserves to guarantee the one-peso to one-dollar rate of exchange. This monetary regime was undertaken at the same time that deep reforms were performed while a large growth in activity was experienced. This trend -both in real activity and real money- broke up in the second half of the nineties. The relative tranquillity of the first half of the nineties was temporarily interrupted in 1995 due to the regional consequences of the Mexican devaluation (known as "Tequila effect"). Although the Convertibility withstood these external shocks, it was a first evidence of the vulnerability of this monetary regime. The government external debt was increasing over time and began to be perceived as unsustainable once the economy entered a deep recession after the Russian (1998) and Brazilian (1999) crises.

Previous to the abandonment of the Convertibility regime it should be noted that in 2001 a financial and external crisis led to a reduction of real money holdings. The regime collapsed in January 2002 after the government announced the default on its sovereign debt and the abandonment of the currency board scheme. Before the crisis, the access to capital markets by Argentina was severely restricted ending the financial liberalization experienced during the 1990’s. Although financial flows to emerging countries had been decreasing since the Russian crisis (the “sudden stop” of Calvo, Izquierdo and Talvi, 2002), after the sovereign debt default, the Argentine economy faced further credit restrictions arising from both external and domestic sources. Not only did capital outflows accelerated but also, at the same time, there was a domestic credit disruption because of financial restrictions and the asymmetric pesification of bank deposits and loans which took place after devaluation (Miller, et.al., 2004). Although the devaluation provoked a jump in inflation rate, that reached a peak in the second quarter of 2002, it then returned to similar levels to those of the Convertibility. Argentine financial system tended to
recover and the real money holdings started to increase continuously, also
motivated since 2004 by the strong growth that the economy experienced after the
prolonged recession that had suffered for several years.

Understanding the behaviour of money demand in such a long period could be
useful to learn about the effects of different policies and to obtain lessons for the
future. In the next section the results of estimating an equilibrium-correction model
are presented and its parameter stability analysed. In order to test the out of sample
forecasting performance of the model the whole sample (1977:1–2006:4) was
divided into three samples that ended in: (i) 2001:4, the quarter previous to the
abandonment of the Convertibility regime; (ii) 2003:2, just before the period in which
policy makers recognised that the crisis finished and (iii) 2005:4, which leaves a
one year forecast horizon at the end of the sample.

3. A causal equilibrium-correction model

The demand for the aggregate M2 can be related to the transaction and
precautionary motives for holding money (see Baumol, 1952, Tobin, 1956, and
Friedman, 1956). In both cases, real money would depend on a measure of the
volume of real transactions and the opportunity cost of cash holdings. The
transactions elasticity is anticipated positive, taking values of 1 or 0.5, according to
the Cambridge interpretation of the Quantity Theory as a demand function or the
Baumol-Tobin hypothesis, respectively. We approximate transactions by aggregate
supply (GDP plus imports). Regarding the opportunity cost three alternatives could
be taken into account: inflation, exchange rate and the domestic interest rate. The
issue is whether or not they are substitutes or complement measures of the
opportunity cost of holding money. In the case of the interest rate it is worth noting
that it embodies an expected rate of inflation, which could be – in some periods – quite different from the actual ones\textsuperscript{iii}.

The estimation started with the analysis of cointegration\textsuperscript{iv} using the system-based procedure of Johansen (1988 and 1992) and Johansen and Juselius (1990). It also allowed us to evaluate exogeneity in order to estimate a conditional model of money demand. To study forecasting performance in different periods several estimations windows (1977:1–2001:4; 1977:1–2003:2 and 1977:1–2005:4) were used for the systems that included, initially, the money aggregate (m), the level of prices (p), the aggregate supply (y), inflation (\(\pi\)), the nominal interest rate of time deposits (i) and the nominal exchange rate peso-dollar (E)\textsuperscript{v}. The results showed that the variables have one long run (cointegration) relationship. A long run elasticity of prices equal to one was found and m was normalized as m-p (mp). Inflation and nominal exchange rate variables were not significant as long run determinants of mp for the three samples\textsuperscript{vi}. The results of the reduced system are summarized in Table 1. The nominal interest rate was the proxy of the opportunity cost that resulted as significant with a stable long run coefficient of approximately 0.70 for the three samples. The transactions elasticity equal to 1 was not rejected for the three samples. Also the Likelihood Ratio (LR) tests indicated that we can estimate a conditional model of mp on y and i for each sample at the traditional 5% significance level. Therefore, the relationship between these three variables was modelled as a conditional univariate equilibrium-correction model.
Table 1: Cointegration Analysis

<table>
<thead>
<tr>
<th>(\lambda_i)</th>
<th>(H_0: r=p)</th>
<th>Max (\lambda_i)</th>
<th>(Tr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.278</td>
<td>p = 0</td>
<td>32.6**</td>
<td>47.73*</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 1</td>
<td>12.7</td>
<td>15.13</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 2</td>
<td>2.908</td>
<td>2.908</td>
</tr>
</tbody>
</table>

Max \(\lambda_i\) is the Maximum Eigenvalue statistic and Tr is the Trace statistic, for each statistic the second column presents the adjusted by degrees of freedom and the third the 95% (Osterwald-Lenum,1992) critical values (See Hendry and Doornik, 1997).

<table>
<thead>
<tr>
<th>(\lambda_i)</th>
<th>(H_0: r=p)</th>
<th>Max (\lambda_i)</th>
<th>(Tr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.279</td>
<td>p = 0</td>
<td>34.7**</td>
<td>52.57**</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 1</td>
<td>12.7</td>
<td>17.87</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 2</td>
<td>5.161</td>
<td>5.161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\lambda_i)</th>
<th>(H_0: r=p)</th>
<th>Max (\lambda_i)</th>
<th>(Tr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.225</td>
<td>p = 0</td>
<td>29.59*</td>
<td>49.09**</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 1</td>
<td>16.73</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>p &lt;= 2</td>
<td>2.768</td>
<td>2.768</td>
</tr>
</tbody>
</table>

The econometric analysis continued with the estimation for each sample of an equilibrium-correction model that included the equilibrium-correction term of each sample and 4 lags of the log differences of each variable; also the proxies of opportunity cost that did not enter the long run relationship were re-considered but...
as part of the dynamics. The restricted model which has homoscedastic white-noise and normal residuals is presented for the different periods in Table 2.

Table 2: Equilibrium-correction model for mp

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.750</td>
<td>-0.809</td>
<td>-0.257</td>
</tr>
<tr>
<td></td>
<td>0.173</td>
<td>0.167</td>
<td>0.076</td>
</tr>
<tr>
<td>Dy_2</td>
<td>0.204</td>
<td>0.225</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>0.081</td>
<td>0.074</td>
<td>0.067</td>
</tr>
<tr>
<td>Di</td>
<td>-0.317</td>
<td>-0.319</td>
<td>-0.303</td>
</tr>
<tr>
<td></td>
<td>0.028</td>
<td>0.027</td>
<td>0.024</td>
</tr>
<tr>
<td>Di_1</td>
<td>-0.172</td>
<td>-0.175</td>
<td>-0.188</td>
</tr>
<tr>
<td></td>
<td>0.034</td>
<td>0.032</td>
<td>0.029</td>
</tr>
<tr>
<td>Dπnetynthia_2</td>
<td>-0.029</td>
<td>-0.027</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td>EqC_1</td>
<td>-0.074</td>
<td>-0.075</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>0.017</td>
<td>0.016</td>
<td>0.010</td>
</tr>
<tr>
<td>R^2</td>
<td>0.76</td>
<td>0.83</td>
<td>0.84</td>
</tr>
<tr>
<td>σ</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>AR 1- 1</td>
<td>[0.1290]</td>
<td>AR 1- 1</td>
<td>[0.2377]</td>
</tr>
<tr>
<td>AR 1- 5</td>
<td>[0.6025]</td>
<td>AR 1- 5</td>
<td>[0.5341]</td>
</tr>
<tr>
<td>ARCH 1</td>
<td>[0.8562]</td>
<td>ARCH 1</td>
<td>[0.4831]</td>
</tr>
<tr>
<td>ARCH 4</td>
<td>[0.1988]</td>
<td>ARCH 4</td>
<td>[0.5479]</td>
</tr>
<tr>
<td>Normality</td>
<td>[0.8076]</td>
<td>Normality</td>
<td>[0.6069]</td>
</tr>
<tr>
<td>Xi^2</td>
<td>[0.8089]</td>
<td>Xi^2</td>
<td>[0.8651]</td>
</tr>
<tr>
<td>RESET</td>
<td>[0.6971]</td>
<td>RESET</td>
<td>[0.8118]</td>
</tr>
</tbody>
</table>

LM statistics of Autocorrelation (AR) and Heteroscedasticity (ARCH); Normality, White’s heteroscedasticity (squares, Xi^2) and Specification (RESET) tests are reported (see Hendry and Doornik, 1996).

In the model presented in Table 2 the equilibrium-correction term (EqC_1) is significant and about 0.07 (for the first two samples) and 0.03 (for the last one) of the disequilibria is corrected in the first quarter in order to adjust the long run relationship between mp, y and i. There is also a short run lagged effect of aggregate transactions (Dy_2) of 0.20 approximately. In addition, the rate of growth of nominal interest rate of time deposits has a contemporaneous (Di) and one lag effect (Di_1) on the rate of growth of real money holdings; the total short run effect is negative and approximately 0.5. The inflation entered this equation, expressed as differences (only positive changes), two period lagged and net from hyperinflation outbreaks. The delay in this effect could be due to the period of time money holders
need to adapt their decisions to changes in the opportunity costs apart from the interest rate.

The dummy variables included in the three samples were not shown for simplicity but coincide with periods of crises or monetary regime changes: 1982:2 with the Falklands/Malvinas conflict, 1983:3-4 and 1988:3 with periods of inflation acceleration before the first democratic elections and a stabilization plan, respectively, 1989:2-3 with the hyperinflation period, 1991:2 with the beginning of the Convertibility regime, 2001:3-4 and 2002:1-2 with the external and financial crises of 2001, the abandonment of the Convertibility regime and the announcement of the default on sovereign debt.

Although, as a whole it was a period of great macroeconomic variability, the first two samples (1977:1–2001:4 and 1977:1–2003:2) showed parameter stability. Instead, some differences in the estimated coefficients, the Constant$^a$ and the adjustment coefficient of EqC_1 were found for the last sample (1977:1–2005:4). However, stability was not rejected at traditional levels by their recursive estimation, as can be observed in the next graphics (the recursive estimates of the main coefficients are within the previous 2 times standard errors intervals and the N-descendant Break-Chow test shows values below the 5% significance critical value).
These estimations seem to indicate that this equilibrium-correction model can be used to explain the M2 monetary aggregate (in real terms) despite the changing environment in which the Argentine economy was immersed during the last three decades.

The forecast performance of the model will be evaluated in the next section.

4. A comparative analysis of alternative forecasting methods

The model presented in Section 3 can be considered as a congruent representation of the Argentine money demand over near three decades, accordingly to its within-sample properties. In particular the analysis of recursive estimates do not reject “ex-post” constancy of the model conditionally to the model uncertainty (see Clements
and Hendry, 1999, Ch. 2). However, the actual forecasting performance of an econometric model may be different “because of the things we don’t know we don’t know” and in this context forecasting failure may not be considered as a fatal flaw. Therefore, in this section we investigate the forecasting performance of the money demand model presented in the previous section, in relation to other approaches, which are explained in 4.1. We apply them to the Argentine rate of growth of real money in 4.2.

4.1 A review of the literature

The traditional theory of forecasting has been based on two key assumptions: i) the empirical model is a good representation of the economy and ii) the structure of the economy will remain relatively unchanged. Given these assumptions it has been proved that the best in-sample model produces the best forecasts, that adding causally relevant variables should improve forecast and it should not pay to pool forecast across models (once encompassing has been evaluated), among other properties.

However, empirical evidence has shown how inadequate these assumptions can be. In particular “since the future is rarely like the past in economics, forecast failure has been all too common” (Hendry and Nielsen, 2007, p.325).

Clements and Hendry (1998, 1999) reconsider the theory of forecasting relaxing the above mentioned assumptions and their results show that the properties based on them do not hold. A main point they addressed is which of the components of the econometric model is the responsible for forecast failure. They found that the main “culprit” is the behaviour of determinist components (intercepts, linear trends, etc.). When they allowed for parameter changes, they showed that the best model in-
sample need not produce the best forecasts, models with no causally relevant variables can outperform and so can do forecast pooling. These properties are examined for an Argentine monetary aggregate, by studying the performance of different forecasting approaches as explained below.

As the major source of changes in the determinist components are derived from location shifts, particularly, long run means (see Clements and Hendry, 1999 and 2005 for the detailed reasons and Hendry and Nielsen, 2007 for a simple example), the Vector equilibrium-correction model (VEqCM) is the more affected one. The VEqCM is an appropriate re-parameterization of a VAR model in levels (assumed of first order for simplicity) for cointegrated I(1) variables (see Johansen, 1988). In “mean-deviations” the system becomes,

\[
(\Delta x_t - \gamma) = \alpha(\beta' x_{t-1} - \mu) + \varepsilon_t, \quad \varepsilon_t \sim IN[0, \Omega] \quad (1)
\]

where \( x_t \) is a vector of \( n \) variables and \( \beta \) and \( \alpha \) are \( nxr \) matrices when there are \( r \) cointegration vectors.

Suppose that at \( T+1 \) the DGP suffered a break in the long run mean and became

\[
\Delta x_t = \gamma + \alpha(\beta' x_{t-1} - \mu^*) + \varepsilon_t, \quad t > T + 1 \quad (2)
\]

If forecasts (denoted by \(^\hat{}\)) are performed based on (1), the model prior the mean shift, then for the 1-step ahead forecast, the expected forecast error is

\[
E[\Delta x_{T+1} - \hat{\Delta x}_{T+1} | \cdot] = -\alpha' \mu^* \quad (3)
\]
The same expected forecast errors would result for h-periods ahead as they are based on model (1). If the model were estimated and not assumed to be the DGP, it could take several periods until the new long run mean were approximated by re-estimation.

In this case a possible solution to avoid systematic errors is differencing the VEqC (DVEqC). To see this, first it can be noted that when $\mu$ does not change,

$$
\Delta^2 x_t = \alpha \beta^t \Delta x_{t-1} + \Delta \epsilon_t \\
\Delta x_t = \Delta x_{t-1} + \alpha \beta^t \Delta x_{t-1} + \Delta \epsilon_t = (I + \alpha \beta^t) \Delta x_{t-1} + u_t , \quad (4)
$$

which is a restricted version of a VAR in differences: DVAR (a preferred model by many users).

After the break in $T+1$,

$$
\Delta x_{T+2} = \Delta x_{T+1} + \alpha (\beta^t \Delta x_{T+1} - \nabla \mu^*) + \Delta \epsilon_{T+1} \quad (5)
$$

and the 1-step forecast error is,

$$
\Delta x_{T+1} - \tilde{\Delta} x_{T+1|T} = -\alpha \nabla \mu^* + \Delta \epsilon_{T+1} \quad (6)
$$

But as $\nabla \mu^* = 0$ after that period,

$$
\Delta x_{T+j} - \tilde{\Delta} x_{T+j|T+j-1} = \Delta \epsilon_{T+j} \quad j > 1 \quad (7)
$$
For $j > 1$, forecasts are unbiased. However, there is a cost in terms of variance (it doubles the forecast variance of the VEqC).

For the same reasons the DVAR model is another alternative to explore for forecasting since by construction it does not include the EqC terms.

Alternatively, the EqC model can be adjusted after the break occurs by residual adjustments or intercept corrections (IC). This can be done by putting the forecast “back on track” when forecast errors are correlated, as follows

$$\Delta x_t^{IC} = \Delta x_{t-1} + (\Delta x_{t-1} - \Delta x_{t-1})$$  \hspace{1cm} (8)

Moreover, a quite simple model that removes deterministic components is the second differences models, DDV. Since many economic variables do not continuously accelerate, then,

$$E[\Delta x_t^2] = 0$$  \hspace{1cm} (9)

And therefore, a “minimal information”, “constant change” forecasting rule (see Clements and Hendry, 2006) is,

$$\tilde{\Delta} x_{T+j}|T+j-1 = \Delta x_{T+j-1}$$  \hspace{1cm} (10)

It reduces the impact of the breaks offsetting breaks in intercepts and trends as well.

The case of M2 allows us to evaluate the forecasting approach of estimating an aggregate by including disaggregates (DISAGG), as M2 is the sum of narrow
money, current account and saving deposits in pesos. Hendry and Hubrich (2006) found that including disaggregate information can improve forecasts as opposed to forecasting first disaggregates variables and then aggregating those forecasts or using only lagged aggregate information in the aggregate model. They found such a property for the population but empirically results may be different due to uncertainty and changing collinearity, among other factors.

Finally, pooling of forecasts has been shown as a successful approach in empirical work. Timmerman (2006) enumerates the following reasons for gains due to forecast combinations: a simple portfolio diversification, different misspecifications of unknown form, different loss functions and different effects of the structural breaks. In this respect, when no model coincides with a non-constant DGP, Clements and Hendry (2004) show that better forecasts can be obtained by forecast pooling using forecasting models that are different mis-specified, particularly because of location shifts. They also found that just averaging may be the dominant strategy over estimated weights in the combination since weights are obtained on the basis of past performance for processes subject to unanticipated breaks.

It has also been suggested that pooling of forecasts obtained for different estimation windows of the same model can help when breaks have occurred within-sample (Pesaran and Timmerman, 2007). When breaks are not well defined it can be better to combine forecast based on different window sizes instead of choosing a single window, which is supposed to exploit the trade off between biases and forecast error variances.
4.2 Forecasting real M2 in the Argentine case

The forecasting models described in section 4.1 are employed to forecast the log differences of real M2. Their individual performances are evaluated in 4.2.1 in comparison with the EqC of section 3. The results of pooling are commented in 4.2.2.

4.2.1. Individual models

To start with, we report in Table 3 the forecast results using the conditional EqC model, which is also used for computing forecasts by DEqC.

For the DVAR model two kinds of estimation are explored. On the one hand, the equation of a VAR(4) for the log differences of real M2 and the log differences of the same set of variables entering the long run solution (transactions and interest rates) plus seasonals has been restricted (r) automatically using PcGets (Hendry and Krolzig, 2001). It is denoted as DVARr. On the other hand, an unrestricted (u) closed system for the same variables is directly employed for forecasting the log differences of real M2. It is denoted as DVARu.

In the case of the univariate model, an AR(4) for the log differences of real M2 (also restricted by PcGets) is estimated for two different samples. The longest sample (l) denoted by DARl and a shorter (s) sample DARs that started in 1991:2 when the Convertibility regime began.

For the second difference model the log difference of the same quarter in the last year, was considered given the forecast horizon later explained. It is reported as D1D4, the corresponding double differences are supposed to have an expected value equal to zero (see equation (9)).
DISAGG is the model selected by PcGets from the regression on the 1 to 4-quarter lags of the log differences of each component (real narrow money, current account and saving deposits) plus seasonalsxii.

As many of the monetary authority projections are one year ahead, forecasts horizon of four quarters are studied for three different cases. The first forecast period started just at the end of the Convertibility regime and the announcement of sovereign debt default, it is the more difficult situation (see Figure 1) since it includes 2002:1, a huge “ex-post” outlier the forecaster had to face followed by another outlier of the converse sign. The second forecast period began in the quarter that the economy seems to have been stabilised from the policy makers perspective. The third forecast period corresponds to the last year of the sample. In order to evaluate the relative performance for a different horizon, for the second estimation sample, forecasts are also evaluated until the end of the sample (long horizon).

When possible we compute dynamic forecasts. For the multivariate variable models (except the unrestricted DVAR) we assume that policy makers “know” the true values of the variables on which real M2 are conditioned, therefore results would be biased to obtain better forecasts by this kind of models. It should be noted that the variables involved are the level of activity and the nominal interest rates, the first one is often conjectured by policy makers and the second is basically an instrument under their control. However, the effects of “estimating” instead of “knowing” the variables taken as given should also be investigated in a future researchxiii.

The models using disaggregates begins in 1994 after the reforms associated with a new regime. It gave a “transactional nature” to the saving accounts when wages and salaries started to be paid through deposits in such accounts. The model using
disaggregates would also be biased to obtain more accurate forecast as it supposes that the components are “known” in each of the four lagged periods (indeed they would be static forecasts).

The next table reports the corresponding root mean squared error (RMSE) and the mean absolute percentage error (MAPE) for each case. We can anticipate differences between these measures not only because of the normalization of the MAPE but also RMSE gives a relatively high weight to large errors, a possible outcome in the period analysed.

Table 3: Forecasting performance of alternative models

<table>
<thead>
<tr>
<th>Forecast Period</th>
<th>EqC</th>
<th>DEqC</th>
<th>DVAr</th>
<th>DVARu</th>
<th>D1D4</th>
<th>DAR</th>
<th>DARS</th>
<th>DISAGG</th>
<th>POOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002:1-2002:4</td>
<td>0.31</td>
<td>0.49</td>
<td>0.35</td>
<td>0.35</td>
<td>0.57</td>
<td>0.37</td>
<td>0.37</td>
<td>0.48</td>
<td>0.23</td>
</tr>
<tr>
<td>2003:3-2004:2</td>
<td>0.10</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.11</td>
<td>0.14</td>
<td>0.09</td>
</tr>
<tr>
<td>2006:1-2006:4</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Period</th>
<th>EqC</th>
<th>DEqC</th>
<th>DVAr</th>
<th>DVARu</th>
<th>D1D4</th>
<th>DAR</th>
<th>DARS</th>
<th>DISAGG</th>
<th>POOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003:3-2004:2</td>
<td>0.55</td>
<td>3.14</td>
<td>0.94</td>
<td>0.80</td>
<td>1.55</td>
<td>0.99</td>
<td>0.92</td>
<td>1.08</td>
<td>0.54</td>
</tr>
<tr>
<td>2006:1-2006:4</td>
<td>3.57</td>
<td>4.41</td>
<td>10.20</td>
<td>8.13</td>
<td>1.15</td>
<td>7.75</td>
<td>11.47</td>
<td>3.54</td>
<td>2.77</td>
</tr>
</tbody>
</table>

The results show that for individual forecasting methods there is no “winner” for all situations and overall they illustrate the difficulties to forecast the rate of change of Argentine M2 (all of the MAPE are higher than 0.5) probably because of the behaviour of this aggregate (See Figure 1). However, the EqC is not too far from
the “best” competitor particularly in terms of RMSE. The EqC wins in the first forecast period even after the abandonment of the Convertibility regime and the default occurred. For the last one year forecast period the EqC shares the best forecasting performance with DEqC, DVARu, D1D4 and DARl. DEqC wins in the 2003:3 – 2004:2 period. Results are different when we consider the MAPE, EqC wins in the 2002:1 – 2002:4 period, DVARr in the post–crisis period (2003:3 – 2004:2) and D1D4 is the better one in the last year.

For the long horizon EqC forecasts are the best ones in terms of RMSE and the second ones in terms of MAPE; D1D4 is better considering MAPE but when the horizon is higher than 4 periods xv. These results are similar to those obtained by Eirtheim, Husebo and Nymoen (1999) who found that at short horizon (up to 4 quarters) simple devices are better than econometric models estimated by the Norges Bank but the latter win for horizon longer than 12 quarters ahead. It would be due to increasing forecast error variances that offset smaller biasesxvi.

Disaggregates models are worse than aggregates models in all cases (except in the last year considering MAPE) even though the former are static (disaggregates are supposed to be known) and the second are dynamic. In spite of the population advantages of using disaggregates, it is not the case here found, as it was in the empirical cases studied by Hendry and Hubrich (2006).

There are no clear differences between the DVARu and the DVARr. It is important because being DVARu a dynamic closed system no conjectures about the interest rate and transactions are needed.

Simple devices can work. D1D4 has the best forecasting performance in terms of RMSE and MAPE for the last period studied.
4.2.2 Pooling

Given the previous results, we compare the relative performance of different pooling procedures. The linear combination of forecasts obtained by three models (EqC, DVARu and DARs) with reasonable forecasting performance is calculated using different techniques to obtain the weights. The EqC model is selected because it is the most useful for policy makers. The DVARu and the DARs models are included in the pooling since they do not require information about conditioning variables and both are more adaptive approaches to out of sample breaks than the EqC model.

The first pooling is made considering a simple average. The weights are 1/3 for each model: EqC, DVARu and DARs. It is reported in Table 3 and Table 4 for comparison with individual methods and other ways of pooling, respectively.

The second pooling method in Table 4 estimates the weights by ordinary least squares, regressing realizations of the estimated variable $y_\tau$ on the N-vector of forecast, $\hat{y}_{\tau}$, using data over the period $\tau = 1, \ldots, t-h$,

$$\hat{w}_{t+h,j} = \left( \sum_{k=1}^{t-h} \hat{y}_{t+k,\tau} \hat{y}_{t+k,\tau}^\prime \right)^{-1} \sum_{k=1}^{t-h} \hat{y}_{t+k,\tau} y_{t+k}$$

Then, the least squares projection used is (following Granger and Ramanathan, 1984, see Timmermann, 2006 and also Marcellino, 2002 ),

$$y_{t+h} = w^\prime_h \hat{y}_{t+h,j} + \xi_{t+h} \quad s.t. \quad \sum_{t=1}^{l} w_{t} = 1$$
The weights are computed for two estimation samples of 16 quarters, one of them ends in 1999:4 and the other, in 2005:4.xvii

Besides, a linear combination of forecasts was performed but with the weights calculated as follows,

\[
\hat{y}_{t+h} = \sum_{m=1}^{M} k_{m,h,t} \hat{y}_{t+h,m}, \quad k_{m,h,t} = \left( \frac{1}{RMSE_{m,h,t}^v} \right) / \sum_{j=1}^{M} \left( \frac{1}{RMSE_{j,h,t}^v} \right)
\]

where \( m \) indexes the models, \( k_{m,h,t} \) indicates the weighting factors. The weights for each model are inversely proportional to their RMSE when \( v=1 \) in the previous equation. We also compute this measure but using another loss function, MAPE, to obtain the weights.

<table>
<thead>
<tr>
<th>Forecast Period</th>
<th>Simple average</th>
<th>w - 1999*</th>
<th>w - 2005**</th>
<th>w - inv. RMSE***</th>
<th>BP</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002:1-2002:4</td>
<td>0.23</td>
<td>0.30</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003:3-2004:2</td>
<td>0.09</td>
<td>0.11</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006:1-2006:4</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Forecast Period</th>
<th>Simple average</th>
<th>w - 1999*</th>
<th>w - 2005**</th>
<th>w - inv. MAPE***</th>
<th>BP</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002:1-2002:4</td>
<td>0.54</td>
<td>0.55</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003:3-2004:2</td>
<td>0.79</td>
<td>0.90</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006:1-2006:4</td>
<td>2.77</td>
<td>2.64</td>
<td>2.50</td>
<td>2.57</td>
<td>3.41</td>
<td>3.92</td>
</tr>
</tbody>
</table>

*weights based on estimations until 1999
**weights based on estimations until 2005
***weights inverse loss function using RMSE/MAPE

When a pooling of EqC, DVARu and DARs (the ones in bold in Table 3) is calculated the forecasting performance is the best compared with these three individual alternatives (it is never worse)xviii. Simple averaging is also the best in
Table 4 for all forecasting periods when comparing RMSE. In terms of MAPE, for 2003:3-2004:2, the inverse MAPE is the best and for the last year, forecast using 2005 estimated weights wins. Therefore, forecast pooling using different models appears to be a recommendable alternative for forecasting real M2 in the Argentine case. Moreover simple averaging seems a very useful method because of its simplicity and the net gain of using other ways of pooling appears to be not too large.

Finally, an exploratory analysis of forecast pooling using different windows was performed for a simple AR(1) model of the differences to evaluate forecasts of the last year of the sample. Estimation windows were selected by using Bai-Perron (BP) statistics (Bai Perron, 1998, 2003) and according to the period in which an institutional change took place (DR).

BP indicated a break in the last quarter of 1989 (and a less significant break in 2002:1) and therefore, a full sample and a window starting in 1990:1 were used for estimation. For the case of the different monetary regime the window started in 1991:2 at the beginning of the Convertibility regime. Simple averages were used to pool windows determined by BP and DR, respectively. Window pooling of an AR(1) model for the differences is better than individual AR models for fixed windows of Table 3 (DARI and DARs). Comparing with pooling of models RMSE measures are lower than those obtained with the other pooling approaches but MAPE are higher.

Therefore, the main results of forecasting the rate of growth of Argentine real M2 showed that: (i) the accuracy of different methods depends on the horizon and the evaluation measure (RMSE or MAPE); (ii) a causal model as an equilibrium-correction of money demand, which can be used for economic purposes, can be useful for long horizon forecasts, also it was the best just after a major change in
the economic regime and its performance is not too far from best forecast approach in the rest of periods; (iii) simple devices can be used, as D1D4 in the last period analysed for the 4-quarter ahead forecasts and iv) forecast pooling (of both models and windows) can be better than individual forecasts, particularly, simple averages given the simplicity of this approach.

5. Conclusions

Different approaches to forecast Argentine real M2 (expressed as log differences) have been evaluated in line with the theoretical developments in forecasting theory that allow for processes subject to breaks. When they are unanticipated, the forecasting ability of causal models can be jeopardised. In particular these breaks can seriously affect equilibrium-correction models, which are more prone to suffer long run mean shifts. In this context devices that rapidly adapt to changes can be better.

For the case of the Argentine economy during an unstable period a conditional equilibrium-correction model of money demand has been analysed both within and out of sample. The model stability was not rejected for the whole period 1977-2006. Comparatively its forecasting performance has been satisfactory regarding the turbulent period analysed. It was the most accurate for long run horizons and its evaluation measures were not too far from those of the best forecast methods in the rest of horizons studied. Although for forecasting this equilibrium-correction model has the cost of having to assume the path of the conditioning variables, it is the more appropriate to understand the economic relationship the policy makers are interested in. It has also been shown how useful for forecasting can result to complement this model with other approaches like double differencing. Apart from this simple device, pooling has been very helpful to improve forecasts. Pooling of
forecasts obtained by AR models estimated for different windows are better than forecasts obtained with fixed samples. Better accuracy was also achieved when the equilibrium-correction was combined with a VAR and a simple AR, both in log differences. These two models are not only more adaptive to a changing environment but also they do not need conjectures about the future path of variables. The pooling of conditional and unconditional models appears as a practical route but there may be others worth exploring in future research that alleviate forecasting problems without losing the economic insights the equilibrium-correction model provides.

**Appendix 1: Data Definitions and Sources**

*M2:* Narrow money, current account and saving deposits in pesos of private sector at the end of period. Banco Central de la República Argentina. B.C.R. A.

*Aggregate Supply:* Gross Domestic Product plus Imports. ECLAC Bs.As. and Dirección Nacional de Cuentas Nacionales (INDEC).

*Nominal Exchange Rate:* Peso/Dollar. B.C.R.A.

*Interest Rate:* 30-59 day time deposits interest rates. B.C.R.A.

*Inflation:* \((p_t - p_{t-1})\) being \(p_t\) the log of general level of consumer prices. INDEC.
References


---

i See also Ericsson (1998) for the main issues related to money demand modelling.


iii Although saving accounts pays interest, this rate has not been considered because of the following reasons: (i) it has been very low in comparison to the rate of time deposits and (ii) its low participation in the aggregate at least until the nineties.

iv Unit-root tests indicate that the variables considered are I(1), they are available on request.

v All variables are expressed in logs, \( \pi \) as log differences and \( i \) as the log of one plus the rate. Unrestricted dummies were also included to get residual normality. A restricted trend was included but it was not significant.

vi Results not reported are available upon request.

vii This deterministic component is more prone to show parameter instability as Clements and Hendry (1998,1999) showed (see Section 4).


ix Clements and Hendry (1999) show that IC works in a similar way to differencing.

x It can run against forecast encompassing as it cannot be proved that only non-encompassed devices should be retained in the pooling.

xi For a previous analysis of monetary shocks using a VEqCM see Utrera (2002).

xii As the aggregate is the sum of the components using log differences as proxies for % variations implies that the coefficient estimates not only includes those of the parameters of the VAR representations of the components but also the shares of each component in the aggregate for the corresponding lags.

xiii Greene, 2000, considers the uncertainty of this mis-measurement source in terms of the deviations from mean of these variables (see also Clements and Hendry, 2006). Pesaran and Timmermann (2007) suggest doing so by simulation.

xvi The root mean squared error (RMSE) is 
\[ \sqrt{\frac{1}{T-h} \sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h} \]
and the mean absolute percentage error (MAPE) is 
\[ \frac{1}{T-h} \sum_{t=T+1}^{T+h} \left| \frac{\hat{y}_t - y_t}{y_t} \right| / h \] when \( \hat{y}_t, y_t \) are forecasted and actual values for \( t=T+1,\ldots, T+h \).

xiv D1-D4 requires us to know the 4- quarter lags of the log differences of real M2.

xv Another result we found is that non significant improvement has been achieved when the forecast was put back on track; this IC correction was calculated only for the EqC model in the long horizon showing nearly the same RMSE (not reported) as EqC without correction.

xvi The size of the estimate sample to compute weights is large enough to follow a mix strategy (not only simple averages) according to the shrinkage rule of Granger and Ramanathan (1984), discussed in Stock and Watson (2004b) and Timmermann (2006).

xvii It is also the best of all individual methods except in the last two period for MAPE. It is the second best measure in all cases.