PROGRAMA DE POSTGRADO EN ECONOMIA
ILADES / GEORGETOWN UNIVERSITY

Documento de Investigación #107
Diciembre de 1997

Nonlinearities in the Demand for Money: A Neural Network Approach

Raimundo Soto

Erasmo Escala 1835. Santiago de Chile. Tels: (562) 698-0046 / 696-8286. Fax: (562) 696-4880
e-mail: economia@uahurtado.cl
www.ilades.cl/economia/publi.htm

Nonlinearities in the Money Demand: A Neural Network Approach

Raimundo Soto
Graduate Program in Economics
ILADES - Georgetown University

December, 1997

I Introduction

A crucial element when undertaking monetary policies is to count on reliable projections regarding the likely effects of changes in income, interest rates, and other macroeconomic variables on monetary aggregates. Understandably, the estimation of money demand functions has been a dynamic field of econometric analysis. The frequently observed instability of estimated parameters and the tendency of money demand equations to systematically overestimate actual balances raises interesting questions concerning the correct model specification, the possible segmentation of monetary markets, and the role of financial intermediaries.[1]

The importance of estimating a reliable money demand equation is evident when monetary aggregates are the target of the authority, as opposed to real interest rates or real exchange rates. The issue becomes critical when the authority decides to switch from targeting real variables to nominal aggregates, as is usually the case in the context of countries with receding inflationary pressures.

Latinamerican countries which experienced acute inflationary problems in the 1980s have recently managed to reduce inflation to tolerable levels (e.g., Argentina, Brazil, Chile and Peru). In the Chilean case, annual inflation reduced from three-digit levels in the mid 1970s to around 25% in the early 1980s. However, despite draconian monetary policies

[1] This phenomenon is known as the case of "missing money" (Goldfeld, 1976; Goldfeld and Sichel, 1990).

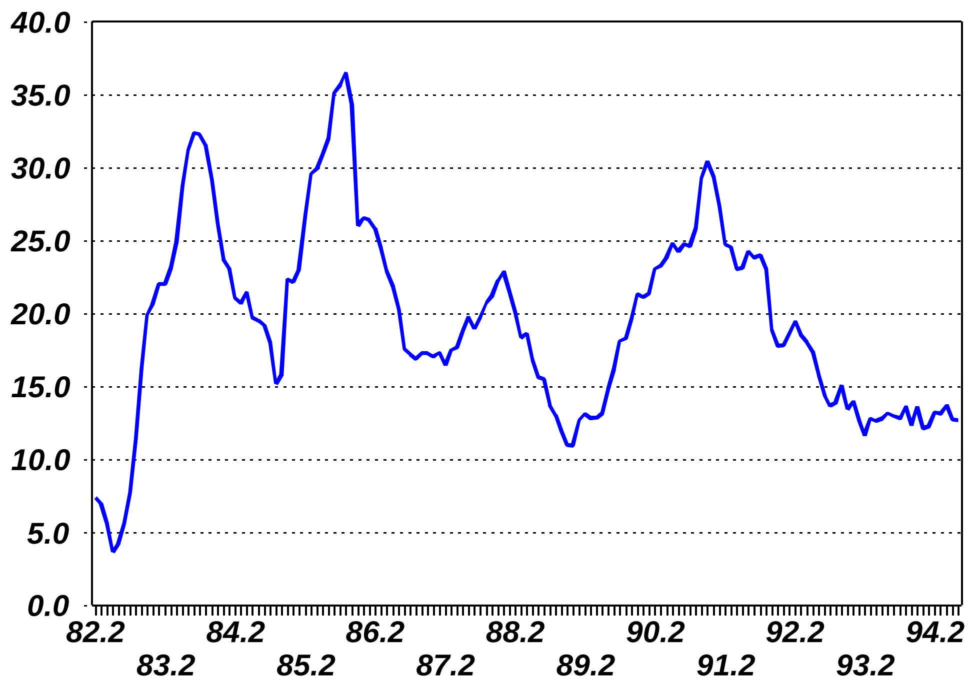
during most of the 1980s and early 1990s, inflation has proved extremely hard to reduce to single-digit levels (see Figure 1).

It has been argued that the inability of the Central Bank to reduce inflation below the annual mark of 10%, stems from the decision of the authorities of using the real interest rate as their preferred policy tool. This, in the context of an economy in which wages, prices, and the exchange rate are backwardly indexed, induces the economy to work without "nominal anchors" (Corbo and Piedrabuena, 1995). In defense of this real interest rate policy, the Central Bank remarks that targeting monetary aggregates cannot be successful because the demand for money in Chile is very unstable, even when considering extended horizons (Apt and Quiroz, 1992). Some empiricall studies tend to support such hypothesis; using a variety of techniques Matte and Rojas (1989), Labán (1991) and Herrera and Vergara (1992) show that estimated money demand equations for real balances lack stability and usually present parameters regarded as implausible for theoretical tastes.

Using newer and more sophisticated techniques, however, Apt and Quiroz (1992) and Martner et al (1995) found a stable long-run relationship between money and its fundamentals (income, interest rates, etc), which contradict previous studies and the Central Bank's conjecture. As discussed in section 3 of this paper, forecasting of these latter models display large out-of-sample errors, which casts doubts regarding their applicability for monetary policies of the sort described above. Note that in the context of reducing inflation from 15% per year to the range of 5 to 8% by targeting the growth of monetary aggregates, forecasting nominal aggregates requires far more accuracy than what was traditionally necessary in the

Chilean case, when the authorities were concerned with keeping inflation in the 20 to 30% range.

Figure 1
Annualized Rate of Inflation (%)



This discussion raises the point of the extent to which it is possible to obtain an accurate forecast of the demand for money *for the purpose of controlling inflation within narrow bands*, using the type of econometric techniques which has been used so far. All the above studies concentrate on linear models, or in models which display a long-run linear response to shocks to fundamentals, with disregard of the state of these and other economic variables. That is the case of classical OLS and uniequational and VAR time-series models as used by Matte and Rojas (1989), Lagos (1991), Garrido and Valenzuela (1991), Rosende and Herrera (1991) and García et al. (1995). Error-correction mechanisms, as those presented in Herrera and Vergara (1992), Apt and Quiroz (1992) and Martner et al (1995), are a step forward in the sense that the response of money demand to shocks in the short run is different than in the long-run or equilibrium. However, in this case the long-run relationship remains linear: a 2% increase in income will generate in the long run twice the effect of a 1% increase, though in the short run the responses might differ.

There are a number of theoretical considerations which suggest that specifying and estimating linear models can be, at best, a reasonable first approximation to the problem of the determinants of money demand, but a misleading tool when a more sophisticated analysis is required. Section 2 discusses briefly three sources of nonlinear response of the aggregate demand for money to changes in fundamentals: misspecification of the decision problem of economic agents; aggregation of individual demands; and financial intermediation. It is important to be precise about the concept of nonlinearity; throughout the paper we are concerned with "nonlinearity in the conditional mean" of a process, which is the relevant concept for forecasting purposes.[2] Nonlinearity in other moments of the distribution are not considered; nevertheless, some of them might still present linearity in the conditional mean (e.g., ARCH models).

---

[2]A process is linear in the mean if $P[E(y_t|x_t)=X'_t\theta] = 1$, for some $\theta \in \mathbb{R}$.

This paper explores empirically the role of nonlinearities in the demand for money using a nonlinear dynamic time-series model, known as a neural network structure.[3] An explicit model of the nonlinear structure of money demand is presented elsewhere (Soto, 1995); the discussion in section 2, however, indicates the type of non-linear behavior to be expected. In this paper, I focus on estimating a nonlinear reduced-form representation of the problem, with which explore the importance of nonlinear elements and their role in improving money demand forecastability.

The initial step in the analysis consists in evaluating the out-of-sample performance of estimated money demand equations, in the context of targeting monetary aggregates to reduce inflation levels of around 15% per year to single-digit levels (section 3). It is shown that even when uncertainty arises only from the estimated parameterization of the function, i.e., when money holdings are forecast out of the sample with the observed values for the fundamentals, the degree of inaccuracy in predicted money balances is too large to base anti-inflationary monetary policies mainly on targeting nominal monetary aggregates.

A second step, consists in testing the existence of nonlinear relationships between money holdings and its fundamentals. In section IV, I use the neural net framework suggested by Lee, White, and Granger (1993) to test monthly data in the 1982-1994 period. Rejection of the linear model suggests it is desirable to explore systematically the role of nonlinearities and that, to obtain better forecasts, the use of more complex models is inevitable. In section

---

[3] Nonlinear time-series models, such as neural networks, chaos models, and threshold models are relatively new in econometric theory (see Gallant and White, 1992, and Kuan and White, 1994); a brief exposition of neural networks is contained in section 4.

5, two types of non-linear dynamic times-series models are estimated (smooth-transition models and single hidden-layer neural networks)and their forecasts compared to linear predictions. Section 6 presents the conclusions.

II. Sources of Nonlinearities in Money Demand.

This section discusses three elements that may induce a nonlinear structure in the aggregate money-income-interest rate relationship: microeconomic specification, aggregation, and financial intermediation. For expository purposes, the analysis is undertaken using a highly stylized framework which omits some of the details of more developed theoretical models and focus on the role of the above mentioned topics.

II.a The Model

The problem of why in some circumstances money has a relative command over other goods in excess of its relative value as a security, i.e. the "liquidity preference" as denoted by Lucas (1984), remains a central issue in monetary theory. In the Arrow-Debreu framework of the decentralized economy with no uncertainty, money does not play a role because, as an asset, it is dominated by other assets which pay a higher return. Thus, the existence of money and why it is valued as it is by economic agents has to be discussed as a departure of the friction-free paradigm, either by including transaction costs or as a result of the time overlap of economic agents.[4] In the transaction costs literature there are two lines of analysis: one, in which money provides a direct service to agents, thus entering in productive or consuming activities (e.g. Sidrauski, 1967; Fischer, 1974); in the other, money is necessary because agents face cash-in-advance restrictions (Lucas, 1978; Svensson, 1985).

Overlapping generations models justify the existence of money as a Pareto-optimal solution to the problem of relocating wealth from one generation to the other (Samuelson, 1958); agents believe that essentially worthless pieces of paper (money) will be accepted by the next-period young generation as wealth and, consequently, they accept money in exchange for the commodity goods they provide to the unproductive old generation. Wilson (1989) shows that under certain conditions this solution is equivalent to the cash-restricted infinitely-lived agent solution.

In this paper we use a stylized version of the generic transaction-cost model in Arrau and de Gregorio (1994) to show what are the potential sources of nonlinearities arising from the microeconomic behavior of agents. The model, which accommodates easily money-in-the-utility (MUF) and fixed-velocity cash-in-advance (CIA) models, is developed for the case of the consumer, but it can be extended straightforwardly to include the demand for money of productive agents (firms), as shown in the Appendix. Consider an infinitely-lived consumer

[4] We abstract here from "legalist" interpretations of the monetary phenomenon, as in Wallace (1983), which place the emphasis on the monopoly of the government in creating money.

which maximizes the expected present-value of utility derived from the consumption of the only good produced in the economy[5]:

$$max_{c_t, b_t, m_t} \qquad W = E_{\Omega_{t-1}} \left[\int_{s=t}^{\infty} U(c_s) e^{-\xi s} ds \right] \tag{1}$$

where ξ is the subjective discount factor, U(c) is a strictly concave utility function. The consumer can allocate his/her income on the two assets available in this economy: domestic bonds (b), with nominal return i, and money (m), with zero nominal return. Money is held because it reduces the cost of acquiring the consumption good. Consumption costs, thus, depend on the money-to-consumption ratio and also on a combination of idiosyncratic and aggregate shocks faced by the consumer, which are summarized in parameter ψ^i (superscript i represents the individual). Let us specify the total transaction-cost as:

$$G(c_t, m_t, \psi^i) = c_t \cdot g(\frac{m_t}{c_t}, \psi^i) \qquad\qquad g^{/}(\cdot \tag{2}$$

where g(.) is the unitary transaction-cost function for the household. The signs of the derivatives reflect the assumption that transaction costs reduce as more money balances are held from last period, and increases with the size of the idiosyncratic shocks faced by an agent.

[5] Note that, to maintain a simple set-up, we abstract from leisure-labor considerations. These can be easily included.

The consumer faces the following flow budget constraint (in real terms):

$$b_t + m_t + c_t[1 + g(\frac{m_t}{c_t}, \psi^i)] = b_{t-1}(1 + r_{t-1}) + \frac{m_{t-1}}{1 + \pi_{t\cdot}} \tag{3}$$

where b_t and m_t are government bonds and money respectively, y_t is income, π_t is inflation and r_t is the real interest rate. The analytical solution of the problem is presented in the Appendix. From the first order conditions of the problem we have:

$$\frac{\partial g(m_t/c_t, \psi_i)}{\partial m_t/c_t} = \frac{-\ i_t}{1\ +\ i_t} \tag{4}$$

which allows us to obtain a money demand equation once function g(.) is specified. Equation 4 states that the consumer will hold money balances until the reduction in the cost of transaction (a benefit for the consumer) equals the cost of holding money in terms of foregone interest. The latter includes both the (real) alternative cost and the expected depreciation of money due to inflation (the model can be extended to include exchange-market capital losses). In general, if g(.) is well behaved, its inverse exists. Then:

$$m_t^i = h(c_t, i_t, \psi^i) \qquad where \qquad h(.) = g(\tag{5}$$

II.b Nonlinearities Arising from the Specification of Microeconomic Behavior

In order to obtain a testable specification for the demand for money, explicit utility and transaction costs functions are to be assumed. Here lies the first source of potential nonlinearities in money demand because, to obtain linear closed-form solutions for the demand for money, very restrictive utility and transaction costs functions are required. Implicitly, we have assumed that the discount rate is constant and independent from wealth; moreover, to generate the standard money-demand equations which are usually estimated in the empirical work, we have to restrict the utility and transaction-cost functions to arbitrary cases. For example, the MUF model can be recovered only by allowing a separable utility function and specifying the following CES function:

$$g\left(\frac{m_t}{c_t}, \psi^i\right) = e^{\psi^i}\left[K - \frac{\sigma - 1}{\sigma}\left(\frac{m_t}{c_t}\right)^{\frac{\sigma - 1}{\sigma}}\right] \qquad \textbf{(6)}$$

where K is large enough to guarantee positive transaction costs. In this case, σ corresponds to the intratemporal elasticity of substitution between consumption and money. Using (5) and (6) we obtain the following money demand function for the individual:

$$log\ m_t^i = \psi^i + log\, c_t - \frac{1}{\sigma} log\left(\frac{i_t}{1 + i_t}\right) \qquad \textbf{(7)}$$

The fixed-velocity CIA model can be obtained by specifying the transaction cost function as:

$$g\left(m_t/c_t, \psi^i\right) = K\psi^i + \sigma\left[m_t/c_t \; log\left(\frac{m_t/c_t}{\psi^i}\right)\right] \tag{8}$$

which in turn would generate the following money demand:

$$log\, m_t^i = \psi^i + log\, c_t \tag{9}$$

In general, the microeconomic specification of utility functions and its use to explain the dynamics of macroeconomic variables has encountered problems, both on theoretical and empirical grounds. On one hand, standard utility functions do not always fit the strict mathematical requirements of the theory (e.g. the logarithmic function is unbounded); on the other hand, representative agent models with microfoundations do not always explain important characteristics observed in aggregate consumption or asset price data. Time separability has been rejected by several authors (e.g., Hansen and Singleton, 1982) which find that implausible elasticities of intertemporal substitution are required to explain the behavior of macroeconomic data. Mehra and Prescott (1985) show that the time-separable constant relative risk-aversion specification cannot explain the historical premium paid to holders of risky assets.[6] In general, all these departures from simple closed-form models generate non-linear structures are characterized, for example, by thresholds and/or time dependent behavior on the part of agents.

II.c *Nonlinearities Arising from Aggregation*

In order to obtain a specification of the aggregate demand for money, individual functions need to be aggregated. The simplest way to aggregate consists in assuming that the economy is inhabited by n-identical individuals, with *n* large enough to avoid non-competitive outcomes, and that a single representative-agent setup is sufficient to describe the aggregate behavior. Though in practical terms researchers do not pretend that all agents are identical, they expect idiosyncratic deviations to play a minor role in affecting individuals. However, as Caballero (1992 and 1994) notes, there are several cases in which the combined behavior of agents affected by idiosyncratic and collective shocks can deviate completely from the individual response, leading to the well known problem of fallacy of composition. In particular, given a degree of heterogeneity in the behavior of individual agents it is possible to obtain aggregate responses that are insensitive to this heterogeneity; on the other hand,

[6] Alternative specifications (e.g., Constantinides, 1991 and Epstein and Zin, 1991), developed to explain the observed volatility of consumption and the equity premium puzzle, cannot explain the demand for low return assets (e.g. Treasury bonds). This risk-free asset puzzle led researchers to focus on heterogeneous agents models (Heaton and Lucas, 1992).

given a rather homogeneous group of agents, it is possible to observe a non-linear response if shocks are strong enough to reshape the cross-section distribution of agents.

To portray this problem in the case of the demand for money, let us extend the model to the n-agent case to include aggregate and idiosyncratic shocks to income, each with a specific (and independent) probability function. Using a linear specification of the above set of equations we abstract from nonlinearities arising from the individual behavior of economic agents to concentrate on aggregation issues. Aggregation amounts to obtaining the cross-section distributional implications of individual's behavior in the presence of both types of shocks. That is, we would like to see whether the response of the combined agents to multiple shocks differs from the individual response to single shocks. The main reason to expect a different response between an individual and the aggregate in this context is the presence of transactions costs in consumption. In their absence, agents can adjust consumption to optimal levels after a each shock costlessly and instantaneously; with friction, the response of consumption, and hence the demand for money, will depend on the path of aggregate and idiosyncratic shocks.

If we assume that ψ^i comprises of idiosyncratic and aggregate shocks and denote f(ψ) the density function of individual shocks, the problem of determining the aggregate demand for money can be expressed as:

$$m_t = \int_{i=1}^{n} E_{\Omega_t} m_t^{\,i}\, di = \int_{i=1}^{n}\int_{\psi_0}^{\psi_1} h(c_t, i_t, \psi^i) f(\psi^i)\, d\psi^i\, di \qquad \textbf{(10)}$$

where m_t stands for the aggregate real money demand. Equation (10) shows that m_t corresponds to the sum over all individuals in the economy of their expected level of money balances, conditional on the probabilities of facing general and specific shocks.

A key element in simplifying the analysis is the application of the Glivenko-Cantelli theorem,[7] which states that as the number of agents becomes large, the cross-sectional distribution of outcomes will converge to the common probability function describing all possible outcomes for the individual (i.e. probabilities are stationary). The representative agent model, then, is a truthful representation of aggregate dynamics only if all agents have the same objective and transaction costs functions, they start from the same endowment (wealth), and face the same shocks. The latter case amounts to perfect shock correlation among agents (i.e., $cov(\psi^i,\psi^j) =1, \forall\ i\neq j$).

A second case of microeconomic behavior which is truthfully reflected in the aggregate arises when all shocks are completely uncorrelated among individuals ($cov(\psi^i,\psi^j) = 0, \forall\ i\neq j$); in this case, the cross-sectional distribution does not depend on the structure of individuals responses because of the stationarity of idiosyncratic shocks.

When shocks are imperfectly correlated ($1>cov(\psi^i,\psi^j)>0, \forall\ i,j$), the cross-sectional distribution is not stationary; however, it is stationary *conditional* over the entire path of aggregate shocks. That is, the response of the aggregate money demand depends on the

[7] See Caballero (1992).

history of realizations of aggregate shocks, on the degree of heterogeneity of agents, and on the correlation of shocks.

The literature on S,s rules has exploited this issue of aggregation asymmetries to show that only in the few cases described above, aggregate shocks does not induce the collective behavior to be different than that of individuals (see Caballero and Engel, 1991). In the context of monetary economics, Caplin and Spulber (1987) show that in a menu-costs framework the neutrality of money holds only in the unlikely event that money supply shocks are monotone and continuous. With non monotone money supply shocks, the "second positive shock may have very different effects on real variables than the first positive shock".

II.d Nonlinearities Arising from the Role of Financial Intermediation

A third source of nonlinear responses of monetary aggregates to shocks to the fundamentals arises from the presence of financial intermediaries. The Modigliani-Miller theorem, which states that financial variables and hence financial intermediaries do not affect decisions concerning "real variables", has been widely criticized at both the theoretical and empirical levels. In a comprehensive survey of the literature, Gertler (1988) divides the arguments used to justify the presence of intermediaries in two areas. First, financial intermediaries appear because of the existence of information costs that make too costly for depositors to monitor borrowers; asymmetric and nonlinear responses usually arise in this context. For example, Stiglitz and Weiss (1981) show that kinked credit supply functions (non-linear rationing) are consistent with competitive equilibrium outcomes, when banks cannot determine the idiosyncratic risk of individual borrowers, but only the aggregate level. Second, financial intermediaries appear when they can reduce transaction costs between depositors and borrowers. Boyd and Prescott (1986) show that intermediaries can effectively reduce the "lemon" risk problem described in Stiglitz and Weiss (1981) through an adequate portfolio diversification. Banks exist also because they can provide credit to firms which face uncertain liquidity needs as in the model of Bernanke and Gertler (1987).

It is possible, then, to obtain nonlinear responses in the aggregate money demand to changes in fundamentals as a result of the behavior of financial intermediaries reacting in asymmetric environments to changes on those fundamentals. In fact, Goldfeld and Sichel (1990), among others, have argued that in the case of the US after 1974, "the period of missing money is largely attributed to the effects of financial innovation" (pp 305).

There have been several attempts to test for financial innovations, mostly for developed countries, with a variety of techniques. Justification and interpretation of the results, however, have been hampered by the lack of theoretical models and the implicit or explicit presumption that innovation is an exogenous event. Simple time trends have been used by Lieberman (1977) and Moore et al. (1988). Proxy variables, such as the ratio of nonbank financial assets to total financial assets and the ratio of currency to money have been employed by Bordo and Jonung (1987) and Soklos (1993). Virén (1994) tests the volume of credit card transactions. Cooper and Ejarque (1994) model financial intermediation as a pure random shocks at the individual level, while Arrau et al (1995) suggest intermediation can be described by a random walk process.[8]

II.e A Simple Model of Endogenous Financial Innovation

A microeconomic model of endogenous financial innovation is presented in Soto (1995), which develops a simple framework to obtain a money demand function similar to that in equation 7, but where the intercept is a nonlinear function of the investment in research and development undertaken by the financial sector.

The standard model presented above is extended to consider the role of financial intermediaries in creating mechanisms to economize on the cost of holding money balances. Considering technical innovation a non-negative random process controlled by a Poisson distribution,[9] the following aggregate money demand equation can be derived:

[8] Financial innovation presupposes that technical progress induces always a reduction in transaction costs; the random walk specification, however, allows for positive and *negative* financial innovation.

[9] This standard specification of innovation has been used in other contexts to discuss endogenous growth models (see for example, King and Robson, 1993).

$$
\begin{aligned}
\log m_t &= \theta_t + \beta_0 \log c_t + \beta_1 \log\left[\frac{i_t}{1+i_t}\right] \\
\theta_t &= \theta_{t-1} e^{\phi(\pi_t)} \mu_t \qquad \mu_t \sim N(O,\sigma_\mu^2) \qquad \textbf{(11)} \\
\phi(\pi_t) &= a\,[1 - e^{\frac{x}{1-x}(e^{-b}-1)}]
\end{aligned}
$$

where $\phi(\pi_t)$ is the expected benefit from innovation, i.e., the probability of finding an innovation (x/1-x) times the benefit of it $^{-b}$ – .

When there is no financial innovation (e.g., when costs outweigh its benefits), the standard linear specification of the money demand function describes adequately the behavior of monetary aggregates. However, when innovation is active (e.g., when a long-lasting inflationary shock makes profitable to invest in R&D even if the outcome of such investment is random). the non-linear component modifies in a significant way the response of monetary aggregates to changes in fundamentals. In fact, elasticities become time-dependent.

The empirical counterpart to the endogenous financial innovation model derived for the case of a single representative financial intermediary is the logistic smooth-transition model (LSTR). In the case of a competitive banking industry, several $\phi(\pi)$ functions may appear and single hidden-layer neural network would be a more appropriate specification. A thorough discussion of neural network and smooth transition specifications, as well as the empirical results for the demand for money in the case of Chile is presented below.

III. Estimating the Demand for Money in Chile

This section discusses the literature on money demand models applied to the Chilean economy. Each specification is subject to specification tests -in particular for nonlinearities- and their forecasting abilities evaluated. In replicating previous work, I do not use the original data for reasons of availability and comparability across studies; replication results based on information produced by the Central Bank of Chile, however, do not differ significantly from the original papers. Figures 2 and 3 presents the evolution of money balances and its fundamentals.[10]

[10]We use monthly data of the industrial sales index as a proxy for expenditures, the 30-to-90 day deposit interest rate, the average nominal exchange rate, and the end-of-month M1A (deflated by the CPI) as the measure of money balances.

Figure 2
Real Money Balances and Expenditures
3-month moving average

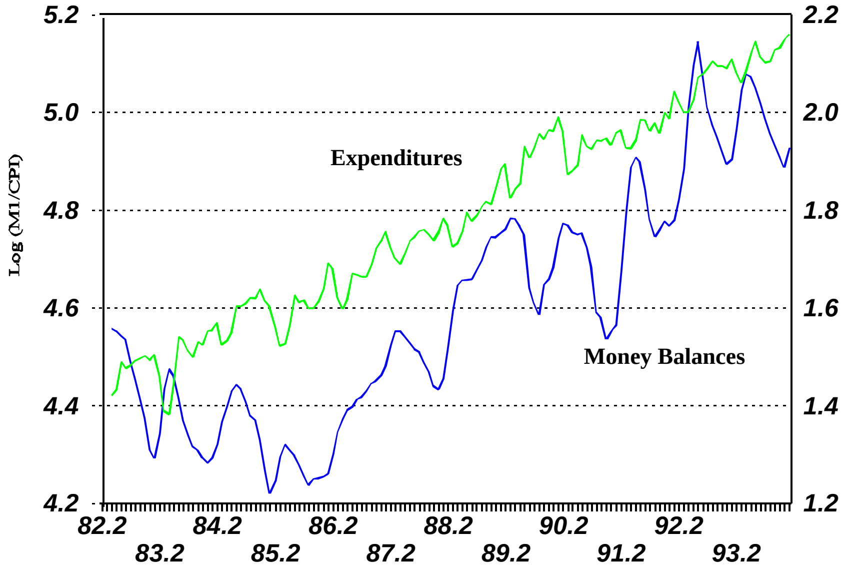
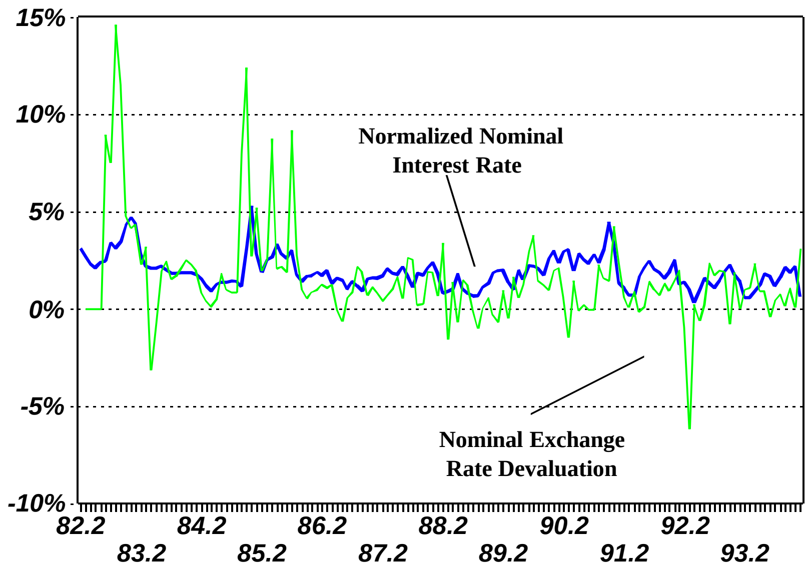


Figure 3
Nominal Interest Rate and Nominal Exchange Rate Devaluation



III.a Traditional specifications of the demand for money

Traditional models of money demand (Matte and Rojas, 1989; Rosende and Herrera, 1991) parameterize aggregate money balances and fundamentals by means of standard econometric procedures (OLS, IV, etc), which correspond to variations of the following static specification:

$$Log\ m_t = \beta' x_t + \mu_t \tag{12}$$

where m_t represents end-of-period real money balances, x_t is a vector of (weakly) exogenous variables, β is a vector of coefficients, and μ_t is an i.i.d. error term. In the case of Matte and

Rojas (1989) and Rosende and Herrera (1991) the specification corresponds to the Cambridge model, where vector *x* includes the log of real expenditures and the level of the nominal interest rate, supplemented by a multiplicity of dummies. Since currency substitution was a common phenomenon in the early 1980s, exchange rate devaluations are also included as an explanatory variable. It would be natural, then, to expect close correlation between devaluations and interest rates; empirically, the correlation in the 1982-1994 period is not very high (36%), probably because neither the exchange rate nor interest rates have been entirely market determined. The Central Bank managed the nominal exchange rate within a exchange band and followed a policy of tight control over the interest rate.

Table 1
Traditional Money Demand Estimations
Dependent Variable: Log M1A / CPI
1978.1 - 1992.8

	OLS	Cochrane-Orcutt Correction	OLS	Cochrane-Orcutt Correction
Log Expenditures	1.04 (20.74)	-0.0016 (-0.16)	1.020 (19.87)	-0.017 (-0.15)
Log Interest Rate	-0.126 (-5.77)	-0.117 (-6.19)	-0.115 (-5.11)	-0.117 (-6.19)
Nominal Devaluations	-	-	-0.78 (-1.77)	-0.005 (-0.06)
Constant	-3.97 (-16.30)	1.28 (2.22)	-3.81 (14.73)	1.28 (2.21)
ρ	-	0.97 (34.62)	-	0.97 (34.4)
Adj. R^2	0.821	0.923	0.825	0.923
DW	0.889	2.42	0.915	2.42

Note: t-statistiscs in parenthesis. Critical values for the null hypothesis of no autocorrelation of the DW test are: 1.69 and 1.77 at 5% and 10% confidence. Seasonal dummies, included in the estimation, are omitted from the table.

Table 1 presents the results of estimating equation 14 using OLS and the Cochrane-Orcutt procedure to correct for autocorrelation. Both OLS equations, including or excluding exchange rate devaluations, present obvious econometric problems. Autocorrelation of the residuals, as expressed in DW tests well below the critical value of 1.46 for a sample of 100 observations, suggests either that errors are truly autocorrelated (in which case the Cochrane-Orcutt technique would be useful) or that there is a problem of spurious correlation. The latter case looks as the more probable, since the DW is similar in size to the adjusted R^2

(Granger and Newbold, 1974). The DW can also be used as a test of stationarity for the residuals; in this case, the above results suggest that errors are probably non stationary. Moreover, Cochrane-Orcutt results show that the estimated ρ is suspiciously close to 1 so as to suggest the presence of a unit root, in addition to coefficients for expenditures that are non-significant and negative.[11]

There is, however, an alternative explanation for this evidence, which is explored below, namely that the relationship may contain, in addition to the standard linear determinants of money balances, additional nonlinear components. If the relationship between money and income, interest and exchange rates is truly nonlinear, a linear approximation would probably display several of the above symptoms: residual autocorrelation (as it is probable that the non-linear component may act for periods of time), lack of significance of parameters (which average over periods of different intensity of response) and relatively high fit (when the linear component is important).

III.b Cointegration and dynamic money demand functions

Arrau and de Gregorio (1994) use Dickey-Fuller tests to present evidence that most variables in vector x and real money balances may contain a dominant unit-root. If that is the case, parameters in equation 14 do not have standard probability distributions which, in turn,

[11] No major differences would arise if the estimation period includes only the post Debt-crisis years (1984-1993).

invalidates inferences upon them (although they remain unbiased).[12] Engle and Granger (1987) show that when each variable in model is I(1), there may be a linear combination of them which yields stationary residuals, i.e., an I(0) error term. In such case, there is a long-run relationship among the variables and, in the time-series jargon, the model *cointegrates*. The parameters in cointegrating regressions are superconsistent as they converge to their asymptotic values at a much faster rate than in standard models (T instead of $T^{1/2}$), which helps identifying long-run relationships in rather reduced samples. Nevertheless, the covariance matrix of errors does not have the standard properties, which in turn complicates inferences upon the parameters.

Arrau and de Gregorio (1994) also found that money, income and interest rates failed to cointegrate in the 1975-89 period. They interpret this as evidence that an important determinant of money balances has been left out of the estimation, which they label "financial innovation" and propose to capture this phenomenon by adding a random-walk component to the standard money demand specification:

$$\begin{aligned} \log m_t &= \alpha_t + \beta' x_t + \mu_t & \mu_t &\sim N(0,c \\ \alpha_t &= \alpha_{t-1} + \eta_t & \eta_t &\sim N(0,\sigma^2_r \end{aligned} \qquad \textbf{(13)}$$

where η is an i.i.d. error-term. By construction, the model in equation (13) should generate a stationary error-term (the sum of η_t and μ_t) and very high fit, and thus the variables should

[12] There is some concern regarding the ability of the unit root tests to distinguish between stationary series with long memory from truly non-stationary variables (Cochrane, 1988). More important in the Chilean context, DF tests are very sensitive to structural breaks (Hendry and Neale, 1991; León and Soto, 1994).

look as if they were cointegrated. The problem of the specification is that by introducing a random-walk as an explanatory variable we turn the problem into one of trivial integration; the random walk will capture all the permanent shocks of real money and vector x_t will pick up only transitory shocks.

From a policy point of view, the main disadvantage of this specification is that the model is useless for predictive purposes because, being financial innovation represented by a random-walk, the best prediction of future financial innovation available is its current level. Moreover, the variance of predictions, itself a function of the variance of both types of shock (η and μ), increases linearly as it is dominated by the random-walk process. Hence, the forecast in the medium to long term has nearly infinite variance and it becomes useless for monetary policy.

Of the several techniques available to model dynamic money demand equations, the Chilean case has been analyzed mostly using VARs (García et al., 1995), cointegration-error correction specifications (Labán, 1991; Herrera and Vergara, 1992; Apt and Quiroz, 1992), and multivariate error-correction models (Martner et al, 1995). Since traditional VARs exclude cointegrating vectors thus missing the long-run relationship of the variables and, in addition, present severe econometric limitations (see Granger, 1989), this section concentrates on error-correction models (ECM).

The simplest specification of this type of model is:

$$\begin{aligned} Log\,m_t &= \beta^{/}x_t + \mu_t & (a) \\ \Delta Log\,m_t &= \alpha^{/}\Delta x_t - \gamma\,[Log\,m_{t-1} - \beta^{/}x_{t-1}] + \nu_t & (b \end{aligned} \tag{14}$$

equation (a) is usually called "cointegrating vector", while equation (b) is referred as the "error-correction mechanism" (Engle and Granger, 1987). Table 2 presents the results of estimating the error-correction model using monthly data for the 1983.1-1992.8 period; we reserve the data on the 1992.9-1994.06 period for out-of-sample forecasting comparisons.

Apt and Quiroz (1992) prefer to use the Engle-Granger two-step procedure to estimate the model. In the first stage, the cointegrating vector is estimated by OLS and the residuals are computed; in the second step, the error-correction model is estimated using lags of those residuals. Our results match that of Apt and Quiroz for the cointegrating vector, although we obtain a smaller elasticity for expenditures (0.90 versus 1.07). The estimation of the error-correction model also matches their results in general terms, though some parameters are less significant than in the original study. There are, however, technical reasons of efficiency and feedback effects which suggest using a one-step nonlinear least squares estimation of the model (see Phillips and Loretan, 1991; Stock and Watson, 1993); column 3 presents the results of the estimation. The structure of the long-run relationship between money, expenditures and the alternative cost of holding money do not change significantly and, in fact, are closer to the original Apt-Quiroz results. Nevertheless, the speed of adjustment implied by the one-step specification is considerably faster than that of the 2-

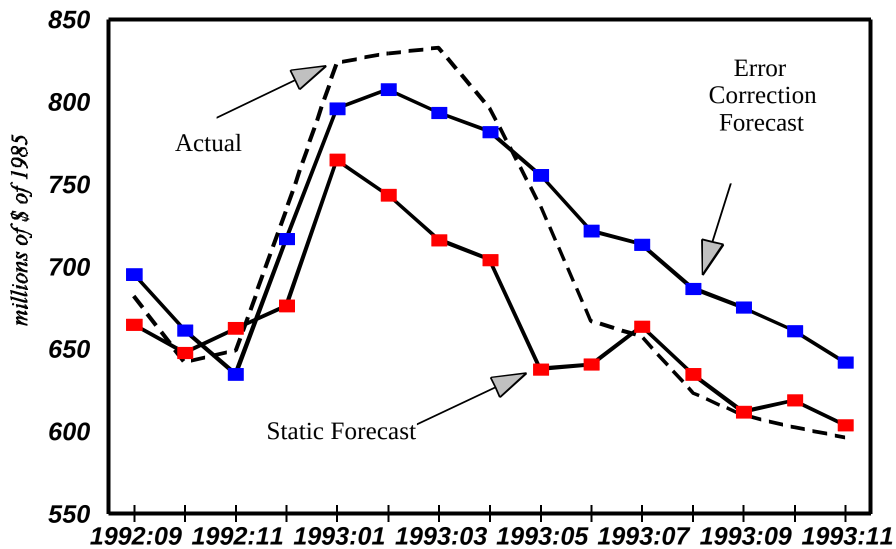
step model (0.234 vs 0.134): in practical terms, 95% of a shock would dissipate in 21 months according to the 2-step model, while it would only take 11 months in the single-step model.

Apt and Quiroz (1992) show that an important advantage of their model with regards to previous estimations is that it presents a smaller forecast error. Their within-sample mean-squared forecast-error of the estimated error-correction model was in the 1.5 to 3% range, while that of the static model in Herrera and Vergara (1992) was in the 2.5-6% range. Nevertheless, when out-of-sample forecast errors are computed the model does not perform as impressive as within the sample. Figure 4 presents the dynamic forecast of money demand for the 1992:9 to 1993:11 period. It can be seen that, although the error-correction model performs satisfactorily in the short-run (6 to 9 months), in a longer horizon it deviates considerably more than a naive forecast based on the static models estimated in table 1.

Table 2
Error Correction Money Demand Estimations, 1983:1 1992:8
Dependent Variable: Log M1A/CPI

Exogenous Variables	Two-Step Procedure		Single Step Procedure
	Cointegration	Error Correction	
Constant	-3.250 (-11.3)		-4.569 (-7.10)
Expenditures	0.906 (15.7)		1.06 (9.14)
Normalized Interest Rates	-0.125 (-4.92)		-0.314 (-4.87)
Nominal Exchange Rate Devaluation	-1.686 (-3.44)		-1.742 (1.45)
Lagged (1) Error		-0.134 (-2.96)	-0.234 (-5.34)
Lagged (1) ΔM1A		-0.347 (-4.07)	
Lagged (2) ΔM1A		-0.167 (-3.66)	
Δ Expenditures		0.095 (1.32)	
Lagged (1) ΔExpenditures		0.67 (0.91)	-0.128 (-1.64)
Lagged (7) ΔExpenditures		-0.025 (-0.38)	
Δ Interest rate		-0.097 (-7.99)	-0.128 (-8.47)
Lagged (1) Δ Int. rate		-0.080 (-5.52)	
Lagged (2) Δ Int. rate		-0.056 (-3.82)	
Lagged (5) ΔDevaluation			0.305 (1.86)
Lagged (1) Δ Devaluation		-0.165 (-0.95)	
ADF test	-5.73	-	-8.25
adj R^2	0.753	0.621	0.635
DW	0.829	1.97	2.62

Figure 4
Out-of-sample forecast of money demand



A more important issue for the purpose of using estimated equations to target nominal aggregates is that the magnitude of the forecasting error is too large to be confident that monetary policies based on the estimated model would be successful. In a 6-month ahead forecast, money balances predicted by static models fails in the order of 10% of actual balances; the error-correction model performs better in 6 months horizons (2.7%) but for a one-year horizon the latter increases to 9%. To give an idea of why these magnitudes are too large for reducing inflation from 15% to single digit levels, it suffices to think in terms of the quantitative theory of money (MV=PY); with income growing at 5% in real terms and velocity approximately constant in the short-run, a range of ± 10% of error in setting nominal money levels would induce inflation to fluctuate in the 5 to 15% range.[13]

IV. Using Artificial Neural Networks to Test for Neglected Nonlinearities

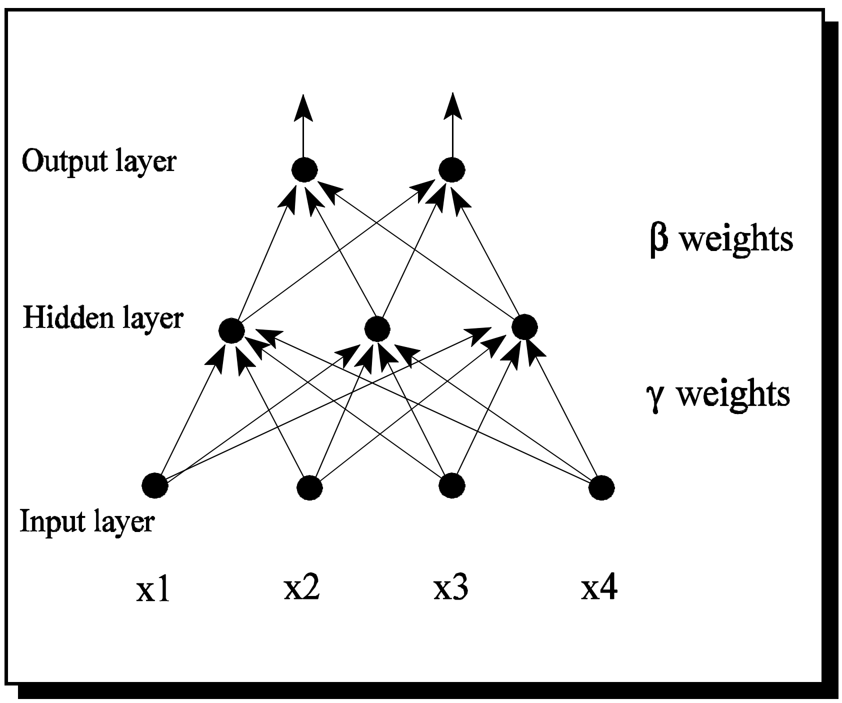
The previous section presented evidence that linear money demands approximate the true relationship between money and its determinants in an unsatisfactory way for the purpose of doing fine adjustments to monetary policy. This section revises the extent to which potential nonlinearities discussed in section 2 might be the cause of these problems. We focus on neural networks as they are able to model highly complex dynamic models in simple but powerful manners (Kuan and White, 1994).

Artificial neural networks (ANN) are a class of input-output models developed to understand the way in which the human brain processes information, characterized by their ability to learn from trial and error procedures. The ANN is based on four principles described below: massive parallelism, nonlinear neural unit response to neural unit input, processing in multiple layers, and dynamic feedback among units. In econometric terms, ANN are a particular group of nonlinear parametric models, where "learning" corresponds to the statistical estimation of model parameters.

[13] Two other specifications are not included in the analysis for lack of replicability. Labán (1991) estimates a traditional specification with time-varying parameters, using Kalman filter algorithms. Easterly et al (1994) test Cagan's specification with a variable interest rate semi-elasticity.

Figure 5 presents a schematic representation of the simplest ANN. This single hidden-layer ANN is composed of an input layer, a single intermediate layer (hence the name) and the output layer, linked by γ and β functions. The first layer uses as input a vector of data $X_t=\{x1,x2,x3,x4\}_t$ which are processed typically in simple ways, using a γ-weight function. The outcome of the operation is processed by the hidden layer using the β-weight functions to yield the output response.

Figure 5
Single Hidden-Layer Network



This simple structure portrays some of the features of ANN, and helps us establish a link with standard econometric practices. Parallelism in processing information, though not as massive as it is usually encountered in neural biology for example, is a standard feature of most econometric work. A widely utilized example of this feature is the simultaneous equation model, in which the information of the so-called exogenous variables (vector *x* in figure 5) is often used in several equations. A second element, nonlinear response to unit inputs, is at the heart of ANN. The simplest form of nonlinear behavior is a zero-one response to inputs. Processing units activate only when input activity {x} reaches a certain threshold я, such that G(x)=1 if x>я and G(x)=0 otherwise. A smoother version, used in the next section, consists of using a sigmoid function, in particular the cumulative logistic distribution, that would create an S-shaped response: unit response turns gradually on as input activity increases, but beyond certain limits (superior and inferior) the response is attenuated. The network in Figure 5 can be easily adapted to multiple layer processing, by introducing additional layers between the hidden and output layers, whilst feedback mechanisms among units, from equal or different layers, can be also easily included.

The mathematical representation of the ANN model belongs to the family of the flexible functional form models (Elbadawi et al, 1989; Gallant and White, 1992) and can be written as:

$$m_t = f_h(x_t, \theta) = F\left(\beta_{h_0} x_t + \sum_{j=1}^{q} G(x_t^{\prime} \gamma_j) \beta_{hj} \right) \qquad h = 1, \ldots, v \qquad \textbf{(15)}$$

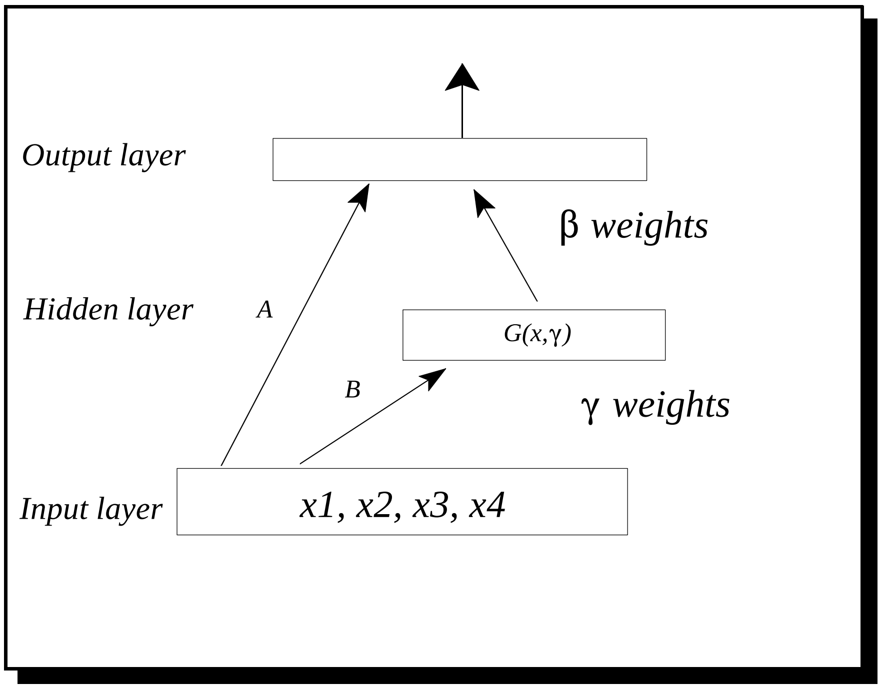
Equation (17) says that output y_t is a nonlinear function of inputs $\{x_t\}$ and a series of parameters $\theta=\{\beta'_1, \beta'_{2,\ldots,} \beta'_v, \gamma'_{1\ldots}\gamma'_q\}$ with $\beta'_h=\{\beta_{h0\ldots}\beta_{hq}\}$ belonging to the *v* inputs in vector *x* and the *q* different units in the hidden and input layers. F(.), which maps from $\mathbb{R}\rightarrow\mathbb{R}$, is usually called the "output activation function", while G(.) is the hidden layer specification. A well known example of this type of specification in standard nonlinear procedures is the logit model; in such case, F(.) is the identity function and G(.) is the cumulative normal distribution.

In the empirical analysis we use an augmented version of equation (15) which allows for a combination of linear and nonlinear processes in determining the output, called the augmented single hidden layer network. Figure 6 sketches the structure of the neural net in this case. This specification, which is the base for the neural net test for nonlinearity, retains the standard linear component of money demands, labeled A, but incorporates an additional nonlinear component, labeled B. Mathematically:

$$m_t = f_h(x,\theta) = \beta_0^{/} x_t + \sum_{j=1}^{q} G_j(x_t^{/}\gamma_j)\beta_j \qquad \textbf{(16)}$$

where F function is assumed to be the identity, and the weight of input in each hidden layer unit is given by β_{0}.. The where the null hypothesis of linearity is $\beta_j=0$, $\forall$ j. Note that $\beta_0'x_t$ is the optimal linear predictor of y given x, and corresponds, for example, to the classical money demand specifications, as in Matte and Rojas (1989), or to an error-correction model, as in Apt and Quiroz (1992).

Figure 6
Augmented Hidden-Layer Network



The power of the test relies on the ability of the set of G functions to extract structure from the residual $e^*_t = y_t - x'_t\beta_0$. Stinchcombe and White (1991) suggest that when the G function is the logistic, the terms $G(x'\gamma)$ are, generically in γ, able to extract such structure. Implementing the tests as a Lagrange multiplier test correspond to testing:

$$H_0: \quad E(\psi_t e_t^*) = 0 \qquad versus \qquad H_1: \quad E(\psi_t e_t^*) \tag{17}$$

where $\psi_t \equiv (\psi(x_t'\Gamma_1), \psi(x_t'\Gamma_2), ..., \psi(x_t'\Gamma_q))$ and $\Gamma = (\Gamma_1, \Gamma_2, ..., \Gamma_q)$ is chosen a priori, independently of the sequence $\{x_t\}$ for a given $q \in \mathbb{N}$. Following Lee et al. (1993), in the empirical application vector T will be chosen randomly from a probability distribution.

In constructing the test, e* is replaced by the sample residuals ($\hat{e}$). The statistic has the form:

$$M_n = \left(n^{-1/2} \sum_{t=1}^{n} \psi_t \hat{e}_t \right)' \hat{W}_n^{-1} \left(n^{-1/2} \sum_{t=1}^{n} \psi_t \hat{e}_t \right) \qquad \textbf{(18)}$$

where $\hat{W}_n$ is a consistent estimator of the variance of the term in parenthesis. Under the null hypothesis of linearity, M_n distributes as a $\chi^2(q)$ as $n \to \infty$. There are, however, two practical problems in computing the test: (1) elements of ψ_t tend to be collinear with x_t and with themselves, and (2) computation of $\hat{W}_n$ can be tedious. The solution to the first problem consists in using $q < q^*$ principal components; for the second problem an equivalent test can be utilized, which avoids the explicit computation of $\hat{W}_n$:

$$nR^2 \quad \to \quad \chi^2(q*) \qquad \textbf{(19)}$$

where R^2 is the uncentered squared multiple correlation from a standard linear regression of $\hat{e}$ on ψ_t and x_t. Lee et al (1993) show that this test is, in general, more or equally powerful than other tests (e.g. Ramsey's Reset, White's, Tsay, etc.) when the hidden-layer function G is modeled as a cumulative logistic function. Teräsvirta et al. (1993), however, show that a test based on the Lagrange multiplier of a Volterra expansion of the series (such as Ramsey's

RESET test used below) performs as well or better than the neural network test when the activating unit lacks the intercept.

Table 3
Testing for nonlinearities in the demand for money in Chile
(1983.1-1992.8)

	Traditional		Error Correction-Cointegration	
Tests#	Without Seasonal Dummies	With Seasonal Dummies	Without Seasonal Dummies	With Seasonal Dummies
Keenan (univariate)				
6 lags, 2 fitted terms	4.10*	3.61*	-	-
12 lags, 2 fitted terms	5.21*	1.51	-	-
RESET (multivariate)				
2 fitted terms	51.5*	42.9*	3.9**	5.17*
4 fitted terms	53.5*	48.6	5.47*	6.68*
Neural Network I				
q = 3	21.9*	19.8*	21.9*	49.8*
q = 5	31.7*	48.9*	23.0*	51.8*
Neural Network II				
q = 3	8.4**	1.49	8.4**	37.4*
q = 5	34.2*	27.3*	10.14*	31.8*

Note: (#) Keenan's and RESET estimators are tested against an F(k-1,n-k) where k is the number of fitted terms and n is the number of observations. The neural-network test distributes as a $\chi^2(q)$. (*) Rejects the null hypothesis of no-nonlinearities at the 1% level; (**) rejects the null hypothesis of no-nonlinearities at the 5% level.

Table 3 presents the results of applying three nonlinearity tests to money demands. We focus on the two models discussed in section 3 (classical specifications and error correction-cointegration models). Keenan's and the RESET tests are based on using transformations of the residuals from estimated money demand functions to verify whether there remains some information that would improve forecasts.[14]

[14] Both tests regress the residuals of a linear regression between money and its fundamentals against those fundamentals and powers of the forecasted money demand. Keenan includes as fundamentals only lags of money balances, while RESET includes also income, interest rates and devaluations.

Two versions of the neural-net test are used in the analysis. The first model (ANN-I) considers the following specification for the case of the traditional money demand (an equivalent expression is used in the case of the error-correction model):

$$Log\, m_t = \theta_0 + \theta_1 Log(Exp_t) + \theta_2 Log\left[\frac{i_t}{1+i_t}\right] + \theta_3 \frac{\hat{e}}{1+\hat{e}} + \sum_{j=1}^{q} G_j([1, Exp_t, i_t, \hat{e}_t]'\gamma_j)\beta_j \quad \textbf{(20)}$$

This specification is consistent with the type of non-linear response discussed in section 2. If the G functions are able to improve forecasts, the source of that improvement cannot be assigned to changes in a particular fundamental. An alternative specification is to test a specific model of nonlinearity such as that proposed in Soto (1995) and summarized in section 2, where financial innovation is driven by financial intermediaries responding optimally to the costs of inflation. In that case, equation (20) is modified to consider a simpler model:

$$Log\ m_t = \theta_0 + \theta_1 Log(Exp_t) + \theta_2 \frac{i_t}{1+i_t} + \theta_3 \frac{\hat{e}}{1+\hat{e}} + \sum_{j=1}^{q} G_j([1,\pi_t]'\gamma_j)\beta_j \quad \textbf{(21)}$$

Both specifications were estimated using a logistic specification for G functions and considering two cases for the number of hidden layers (q=3 and 5) to control for the fact that the number of G functions is arbitrary. To implement the test, γ coefficients were randomly drawn from a uniform distribution in the [-2,2] interval. A large number of repetitions (not reported) confirmed that these results are not due to singularities of the starting values used in the estimation.

The results are remarkably consistent. When considering the traditional specification all tests reject the null hypothesis of linearity, regardless of the inclusion of seasonal dummies or the number of fitted terms in Keenan and RESET tests. The null hypothesis of linearity is also rejected by neural network tests, although when including only inflation in the specification of the sigmoid, the test is less able to extract structure from the residuals. A similar phenomenon is observed in the case of the cointegration error-correction model, for which tests in the majority of cases reject linearity at 1%.

V. Estimating Non Linear Specifications of the Demand for Money

The first part of this section analyzes a simplified version of neural network specification, called *smooth-transition* model, which corresponds to the case where there is only one G-function in a single hidden-layer model. Estimating a smooth transition model for money demand is useful as a first step toward the estimation of the neural network model, as it helps characterize the parameter space and solve minor technical issues. In addition, it provides important insights concerning the relationship between money and its fundamentals. The second part of this section presents the estimation of the neural network models. Given the limitations of the data, we focus on a single hidden-layer model with five G-functions which, in the most complicated model, implies estimating recursively 34 parameters with 127 degrees of freedom.

Smooth Transition Models

Two specifications for the sigmoid activation function are tested; one including all fundamentals as activating variables and a restricted version, consistent with the last model in section 2, which concentrates only on inflation as an activator. For the traditional specification of the demand for money, we then have:

$$\begin{aligned} &(a) \quad Log\ m_t = \theta^{\prime}[Exp_t, i_t, \hat{e}_t] + G_1([1, Exp_t, i_t, \hat{e}_t]^{\prime}\gamma_1)\beta_1 \\ &(b) \quad Log\ m_t = \theta^{\prime}[Exp_t, i_t, \hat{e}_t] + G_2([1, \pi_t]^{\prime}\gamma_2)\beta_2 \end{aligned} \tag{22}$$

while for the error-correction model the following equations are estimated:

$$\begin{aligned} &(a) \quad \Delta Log\, m_t = \alpha^{\prime}\Delta[Exp_t, i_t, \hat{e}_t] - \phi\left(m_{t-1} - \theta^{\prime}[Exp_{t-1}, i_{t-1}, \hat{e}_{t-1}, G_1([1, Exp_{t-1}, i_{t-1}, \hat{e}_{t-1}]^{\prime}\gamma_1)\beta_1\right) \\ &(b) \quad \Delta Log\, m_t = \alpha^{\prime}\Delta[Exp_t, i_t, \hat{e}_t] - \phi\left(m_{t-1} - \theta^{\prime}[Exp_{t-1}, i_{t-1}, \hat{e}_{t-1}, G_1([1, \pi_{t-1}]^{\prime}\gamma_1)\beta_1\right) \end{aligned} \tag{23}$$

note that in the error-correction model the sigmoid function is considered part of the long-run relationship of the variables.

As is the case in all nonlinear estimations, starting values and estimation procedures are of importance. I used as starting values those obtained from the linear estimation; this is intuitive since we already know that the linear specification is a reasonable first approximation to the data. Regarding the estimation strategy, I first use a quasi-Newton method (e.g., Davidon-Fletcher-Powell or Broyden-Fletcher-Goldfarb-Shanno) to approach quickly a solution and narrow the parameter space; to obtain final estimates, Newton-Raphson methods are used.

Table 4
Smooth Transition Models of the Demand for Money in Chile
(1983.1-1992.8)

	Traditional Specification			Error-Correction Specification		
Parameters[#]	Linear	Smooth I All Fundamentals	Smooth II Inflation Only	One Step Estimation	Smooth I All Fundamentals	Smooth II Inflation Only
θ_1	-0.035	-0.120	3.430	-0.013	0.577	-0.044
θ_2	0.924	1.393	0.916	1.066	1.280	1.096
θ_3	-0.123	-0.157	-0.097	-0.303	-0.276	-0.230
θ_4	-1.617	-3.618	-1.700	-0.852	-1.581	-0.819
β_1		2.703*	-3.473		-0.621	-2.046
γ_1		-3.961	4.402		6.646	1.331
γ_2		-5.330			16.761	
γ_3		0.359**			-1.851	
γ_4		12.123			-13.092	
γ_5			2.470			-0.436
α_1				0.115	0.131	0.051
α_2				-0.134	-0.133	-0.118
α_3				-0.189	-0.186	-0.235
ϕ				0.231	0.357	0.307
R^2 first diff				0.433	0.495	0.475
R^2 level	0.772	0.867	0.790	0.636	0.791	0.767
DW	0.99	1.48	1.15	2.32	2.26	2.43

Note: (#) Contrary to the standard practice (*) non-significant at 5%, (**) non-significant at 10%, the rest are significant at 1%. Seasonal dummies, included in the estimation, are omitted from the table.

The results, which are presented in table 4, show several interesting results. First, the smooth transition model improves considerable the estimation of the traditional specification of the demand for money. The goodness of the fit improves substantially when considering the case of a sigmoid function driven by all fundamentals (higher R^2) and residuals tend to exhibit less serial correlation (as measured by Durbin-Watson statistics), though in the case of the inflation activated sigmoid, they do not signal unequivocally non-correlated residuals. Nevertheless, in both cases DW statistics support the notion that, if variables are integrated, they cointegrate. Second, when considering the error-correction model, both smooth-transition models are superior to the standard specification in terms of fit and dynamic properties. The speed of convergence in the latter models is markedly faster than that of the one-step model (itself faster than the two-step procedure in Table 2); a coefficient in the range of 0.30-0.35 implies that, on average, 95% of a shock will dissipate in a period of 6 to 8 months.

It is interesting to note that when including the non-linear structure in the estimation the point estimates of the parameters of the linear component do not change markedly. Scale elasticities, for example, remain in the neighborhood of one. This would suggest that the sigmoid function is effectively extracting structure from the residuals of the linear component, as discussed above when testing for non-linearities, and is not too colinear with fundamentals. Moreover, this is consistent with the theoretical discussion in section 2 which hypothesized that as an economy moves toward non-extreme situations (e.g., high and volatile inflation to low and stable), non-linearities are likely to become important. In the Chilean case, the sharp decline in the volatility of inflation in the early 1990s changed the structure of the demand for money in a way such that linear models cannot apprehend, but that the smooth-transition model is able to uncover.

The estimated models were used to calculate the out of sample mean-squared forecast-error for the 1992.9-1994.6 period. The simulation was undertaken using the observed out-of-sample values for fundamentals. The results are reported in table 5.[15] The main conclusion, which concerns both the traditional and error-correction specifications, is that the inflation driven sigmoid presents the lower forecast error as the horizon increases (15-21 months), coinciding with the period in which inflation reduced markedly (see figure 1). This suggests that in models which include activation functions, it may not be necessarily important to obtain a good fit within the sample to generate successful forecasts, but rather to approximate the data with a flexible specification able to accommodate changes in the behavior of exogenous variables in ways that are not contained in the estimation sample. In our case, it can be argued that the stabilization of inflation and its reduction in the post estimation period encouraged agents to maintain money balances beyond what had been traditional for the volume of transactions they undertake and the alternative cost they face.

[15] Note that the error-correction model present a higher mean-squared forecast-error than the static models. This is a consequence of estimating money balances dynamically, i.e., forecasted money balances are used to generate the next-period forecast of money balances, thus accumulating deviations on time.

Table 5
Mean Squared Forecast Error
(in percentage)

Year / Month	Linear Models		Smooth Transition All Fundamentals		Smooth Transition Inflation		Neural Network All Fundamentals		Neural Network Inflation	
	Traditional	Error Correction	Traditional	Error Correction	Traditional	Error Correction	Traditional	Error Correction	Traditional	Error Correction
92.09	3.8	1.3	3.1	2.4	2.6	3.6	2.7	0.8	1.2	0.8
92.10	3.5	1.1	3.0	2.2	2.5	3.2	1.4	0.7	0.6	0.7
92.11	3.2	1.1	2.8	2.1	2.3	3.0	1.0	0.5	0.4	0.6
92.12	2.9	1.0	2.6	2.0	2.2	2.8	0.9	0.8	0.4	0.5
93.01	3.0	1.0	2.6	1.4	2.3	2.7	0.7	1.0	1.1	0.5
93.06	3.1	1.3	2.1	0.8	1.8	0.9	1.1	1.3	1.5	0.3
93.12	2.4	3.7	2.3	5.1	1.8	2.7	0.9	1.0	1.0	0.6
94.06	2.4	8.8	2.9	9.7	1.8	6.0	1.1	1.1	1.0	1.4

Neural Network Models

When considering the full estimation of a neural-net specification for the demand for money the following identification problem arises. Let us focus on a reduced version of the traditional specification, as in equation (22), with only 2 G-functions in the hidden layer:

$$Log\ M_t \ = \ \theta^{/}[Exp_t, i_t, \hat{e}] \ + \ G_1([1,\pi_t]^{/}\gamma_1)\beta_1 \ + \ G_2([1,\pi_t]^{/}\gamma_2)\beta_2 \tag{24}$$

Clearly, switching γs and βs would not affect the estimation of vector θ, nor any feature of the estimation procedure. Note that this problem does not invalidate linearity testing procedures as specified in the previous section, because the test was based on the ability of G(.) functions to extract structure from the residuals of the linear model and not on an adequate parameterization of the neural net. This is why we could draw the γ parameters from a random distribution without estimating them.

A simpler solution to this problem, and also a more robust procedure, consists in expanding the Munro-Robbins recursive estimation algorithm (Kuan and White, 1994) to consider the following updating scheme for the parameters in equation (26):

$$\Omega_{T+1} \ = \ \Omega_T \ + \ \eta_{T+1}\nabla\tilde{f}_{T+1}(m_{T+1} \ - \ \tilde{f}_{T+1}) \tag{25}$$

where $\Omega = \{\theta', \gamma_1 \ldots \gamma_n, \beta_1 \ldots\beta_q\}$ and T represents a sufficient number of observations that make a first iteration of the recursive estimation feasible:

$$
\begin{aligned}
(a) \quad & f_T(x_T, \Omega) = \sum_{j=1}^{q} \beta_j \phi(\gamma_j, x_T) \\
(b) \quad & \tilde{f}_T = f_T(x_t, \Omega_{T-1}) \\
(c) \quad & \nabla \tilde{f}_T = \left[\frac{\partial}{\partial \Omega} f(x_T, \Omega) \right]_{\Omega = \tilde{\Omega}_{T-1}}
\end{aligned}
\qquad \textbf{(26)}
$$

Equation (28a) corresponds to the activation units in the hidden layer of the neural net, while equation (28b) is the forecast value of m_T based on the recursively estimated parameters as of time T-1. The last equation correspond to the actualization of the parameters, using the gradient method and based on the distance between the forecasted money balances and the actual value.

A key parameter in the estimation is the convergence rate, η_T, which is time-dependent. White (1989) and Granger and Teräsvirta (1993) suggest that specifying it to be proportional to sample size is an adequate procedure ($\eta_T = \propto T^{-1}$). White shows that, if *m* and *x* are random vectors and η is a declining rate, then the estimated vector θ converges to a local minimum with probability 1 or diverges to infinity with probability 1.

This "backward propagation" algorithm, when used in large samples, should be able to attain convergence with a single pass through the data, using quasi-Newton methods at the start of the estimation to reduce the parameter space and Newton-Raphson techniques for the final iterations. In small samples, however, bootstrapping is necessary; in our case (127 observations) the final estimation of the model required the latter procedure. To avoid falling into sub-optimal parameter regions, the starting values for the grid search are those of smooth-transition models for the linear components and the first G-function, while the rest are drawn randomly from a standard normal distribution.

Table 6 presents the results of the estimation of the neural network with 5 units in the hidden layer for both specifications of the activation unit. It can be seen that the estimation does not differ markedly from that of the smooth transition models in terms of the degree of fit and the size of the parameter of the linear component, except that scale parameters are slightly smaller in the case of the traditional specification. When considering the error-correction model, we note that the parameters of the linear component resemble those estimated by Apt and Quiroz (1992) and others. Scale parameters are very close to unity and, in the case of the sigmoid activated by all fundamentals, the interest rate elasticity and the semi elasticity of nominal devaluations are also very close to the single-step error-correction model in table 2. Nevertheless, the degree of fit of the long-run model is far superior. In contrast, the inflation-activated sigmoid model fits the data better in the case of the first differences model. As we will see, this will have implications for the forecasting ability of each model. Only the estimated weights of the sigmoid functions (β parameters) are reported for space limitations.

From the estimated parameters it would seem, then, that the neural network model has not been able to extract more structure from the data. Nevertheless, when contrasting the out-of-sample forecasting performance of these models, neural networks projections present a considerably lower mean-squared forecast-error (MSFE) than smooth transition and linear models (see table 5). When considering the traditional static specification, neural networks present a MSFE between 33 and 50% lower than linear models and smooth transition models; in the case of dynamic error-correction models, the MSFE of the neural network is one half that of linear models in short horizon forecasts and one fifth in a 22-month prediction. Note, in addition, the ability of the inflation-activated sigmoid model to predict money balances, despite the fact that the fit within the sample is worse than that of linear and smooth transition models.

Table 6
Neural Network Models of the Demand for Money in Chile
(1983.1-1992.8)

	Traditional Specification		Error-Correction Specification	
Parameters#	Neural I All Fundamentals	Neural II Inflation Only	Neural I All Fundamentals	Neural II Inflation Only
θ_1	-0.246	-0.446	-0.614	-0.535
θ_2	0.891	0.860	1.057	1.098
θ_3	-0.045	-0.111	-0.521	-0.203
θ_4	-2.146	-1.478	-1.769	-0.476
β_1	-20.738	-0.982	0.944	-1.287
β_2	20.797	0.480	-0.370	1.209
β_3	28.016	-1.205	-1.026	-1.711
β_4	-4.951	1.034	0.566	0.821
β_5	-0.125	2.316	-0.975	2.339
α_1			0.064	0.189
α_2			-0.113	-0.146
α_3			-0.089	-0.159
ϕ			0.164	0.281
R^2 first diff R^2 level	 0.825	 0.708	0.334 0.813	0.457 0.575
DW	1.11	0.78	2.31	2.39

Note: (#) Contrary to the standard practice, (*) non-significant at 5%, (**) non-significant at 10%, the rest are significant at 1%. Seasonal dummies are omitted from the results. Standard errors are computed using White's heteroskedasticity robust procedure.

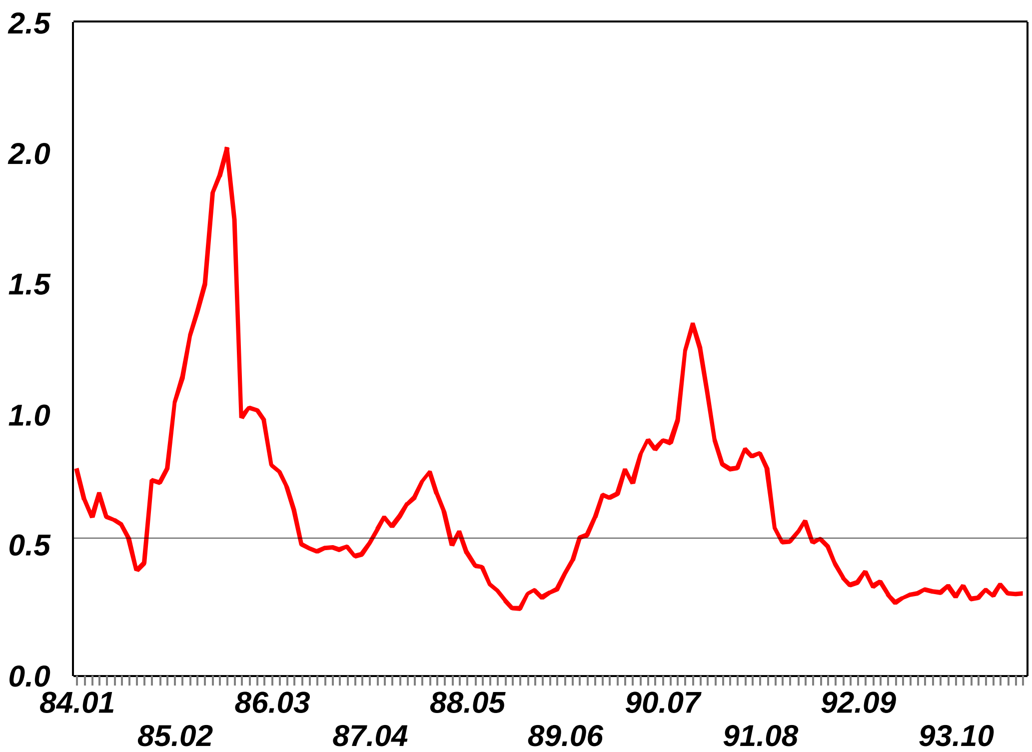
It is also interesting to note that long-run parameters in the error-correction version of the neural network are of similar size to those of the standard error-correction model. This is consistent with the endogenous financial innovation model described in section 2, which suggests that innovation is more likely undertaken when the economy moves toward moderate and stable inflation levels. In that case, shocks to the fundamentals have an additional effect on money demand and a highly nonlinear structure arises. The neural network model encompasses these two situations: parameters of the linear component, which is dominant when the state of the economy does not give rise to innovation, are very similar to those of standard error-correction models estimated in the period of high and volatile inflation, while the estimated G-functions incorporates the non-linear effect. It is precisely this last element which explains the improvement in forecastability in these latter models. Without the non-linear device, money demand forecasts obtained from error-correction models would deviate from observed values by as much as ±10% in a 15-month forecast (table 5), while forecasts including the nonlinear component would fail by 2.5 to 5%. Hence, for the purpose of controlling inflation within a narrow band (e.g., a target of 7% inflation with a ±2 percentage points of deviation), standard error-correction models would yield very inaccurate forecasts. Neural net models, on the contrary, would forecast real money balances with more accuracy and, provided a good estimate of real GDP growth, would generate targets for nominal monetary aggregates which are consistent with desired inflation levels.

Can we avoid using a non-linear model to forecats money demand?

The previous analysis has pointed to the fact that inflation -among other variables- could have a non-linear impact on money demand as financial intermediaries react to changes in inflation in non-symmetric ways. Even if accepting that inflation may affect money demand beyond the effect of nominal interest rates, it is natural to be skeptic regarding the need of developing a complicated neural-network model. An indirect way of checking the need of using non-linear models consists of comparing inflation with the estimated impact of the non-linear component of one of the estimated money demand equations. To keep the comparison simple, we use the estimated neural network model for the traditional specification in table 6 and isolate the fitted non-linear component using the estimated parameters. This, in turn, is normalized by inflation for each point in the sample. The results are presented in figure 5. If the relationship between these two variables remain relatively flat of a period of time, we could argue that a non-linear dynamic model is unnecessary, as the same information could be obtained from a much simpler linear structure.

When analyzing the evolution of this ratio two elements are striking. First, it is evident that inflation cannot be used as a proxy of the sigmoid as the relationship among them is far from being linear (flat). The presence of large deviations from the mean ratio around the end of 1985 and in 1989-90 suggests that there are periods in which the sigmoid is very active while in other it remains rather inactive (e.g., post 1992). Second, it should be noted that in those periods in which there is activity in the sigmoid the economy faced a sudden increase in inflation -to almost 30% per year in each occasion- although the reasons for these phenomena differ; in the former, the debt crisis and its sequel of large nominal devaluations generated high, though short lived, inflationary pressures. In the latter, the political cycle induced unsustainable expansions in public expenditure and investment. It is, in fact, remarkable how the evolution of this the ratio matches with what we should expect from the theoretical model regarding the response of agents to the evolution of inflation. Note that, due to the above normalization, the ratio measures the "excess" response of the non-linear component when compared to inflation. It is precisely in those periods when there is high inflation that the ratio exhibits sudden jumps in activity, reflecting the behavior of agents avoiding the costs of those inflationary shocks.

Figure 5
Ratio of the Estimated Marginal Contribution of the Non-linear Component in the Forecast of Money Balances to Annual Inflation



6. Conclusions.

An important element when implementing monetary policies is to count on reliable forecasts of the behavior of monetary aggregates in response to shocks to fundamental variables, such as interest rates, the nominal exchange rate or the level of consumption. When estimating and forecasting money demand the Chilean experience mirrors the international case: frequently, parameters are unstable and appear distorted when confronted to theoretical priors, while forecasts tend to overestimate actual levels (the case of "missing money") or to fall short of actual demands (the "excess monetization" case).

Empirical studies have applied increasingly sophisticated econometric specifications to the demand for money, including Kalman filtering, error-correction models with cointegrating relationships, VARs, etc. Results, however, are mixed: although the within-sample estimation has improved considerably when compared to traditional static models, modern techniques have not been successful in reducing the out-of-the-sample forecast error, particularly at long horizons.

A common feature to these techniques is the assumption of a linear long-run relationship between money balances and its determinants. This paper contends that it is this particular structure which causes the predictive failure of estimated models, in special in the case of Chile. During the last years, monetary authorities have allowed an increasing deregulation of domestic and foreign financial markets. This more competitive environment led financial intermediaries to undertake a series of important financial innovations (e.g., automatic teller machines, electronic interbank transactions, etc) which, in turn, have affected the long-run determinants of the demand for money balances. Linear models have trouble in tracking these changes as their rigid parametric structure is not well suited for what is more likely a time-dependent relationship.

Non linear techniques, as those proposed in this paper, can tackle this issue in a direct way. The effect of nonlinearities in the estimation and forecast of the demand for money is analyzed using recently developed time-series techniques based on the architecture of neural networks. Neural networks have been developed to mimic the way in which information is usually processed in the human brain, allowing for massive parallelism, nonlinear neural unit response to neural unit input, processing in multiple layers, and dynamic feedback among units. In econometric terms, this framework corresponds to highly non-linear time series models with adaptive update of parameters.

The main advantages of these new techniques rely on exploiting the richness of their non-linear structure and the ability to learn in an adaptive way the underlying data generating process of the data. As more information is available and can be supplied to the model, the estimated structure is easily updated using an optimizing algorithm which improves the fit and forecastability of the model. In addition, neural networks are designed to process information in multiple layers, thus able to model and explain highly complex structures.

The methodology proceeded in three stages. First, standard models of the demand for money are estimated to provide a benchmark for the assessment of the forecasting abilities of different techniques. We include static and dynamic error-correction models. Second, we test for non-linearities in the data, both using univariate and multivariate tests; in general, the null hypothesis on linearity is amply rejected by the data, even for sophisticated specifications. Third, we estimate two of the simplest non linear dynamic models: the smooth-transition model and the augmented single-hidden layer specification of a neural network model.

The results can be grouped in three categories. First, there is strong evidence that the relationship between money balances and its determinants (expenditures, interest rates and nominal exchange rates devaluations) presents important nonlinear components, which, if omitted, have adverse effects when forecasting money balances with linear specifications. The out-of-sample mean-squared forecast-error of linear models increases drastically in the 1992.9-1994.6 period, when the economy moved out of a high uncertainty scenario (repressed growth and high and volatile inflation) towards a more stable regime (rapid growth with stable and slowly declining inflation).

Second, simple nonlinear structures, such as the smooth transition model and the single-layer neural network model, can improve significantly the quality and out-of-sample forecastability of static and dynamic models of the demand for money. Non linear models, which use a sigmoid activation function to capture potential non linear elements in addition to standard linear components, improve the within-sample performance and display better fit, less correlation of residuals and stable parameters. The latter, in general, also fit theoretical priors.

Third, forecasts from linear models of money balances tend to display large errors as the horizon lengthens which make them useless for undertaking monetary policies aiming at controlling nominal aggregates. On the contrary, neural network forecast money balances with very low forecast error at long horizons and in particular when fundamentals present large changes for sustained periods of time. For the period 1992.9-1994.6, neural networks present a mean-squared forecast-error in the order of 1%, while static linear models present more than 3% and dynamic error-correction models around 6%. To a large extent. the permanent reduction in inflation cause the large error in the latter.

Among the most interesting results it is important to note that neural network models estimated in this paper encompass previous estimates of linear models. Since the sigmoid activation function attenuate the effect of nonlinearities when the activating variables cross certain upper and lower thresholds, it should be expected that in those periods the estimated neural network model matches the structure of linear models, and that when activating variables activate the sigmoid income and interest rate elasticities deviate from the linear model. In both cases this is verified, but the inflation activated sigmoid is more interesting because inflation per-se is not a determinant of money balances, acting only through the sigmoid. When considering the recent macroeconomic experience in Chile in which inflation stabilized (1993) and reduced to single digit levels (mid 1994), it is reasonable to expect that before 1993 nonlinearities captured by the sigmoid function had a negligible effect when predicting money balances because during unstable periods the sigmoid component is not active. However, after price stabilization was achieved the contrary should be observed.

Finally, there remains to be explored the improvements that may arise from using multiple-layer neural network models applied to this case. In principle, multiple-layer neural nets could approach the data with arbitrary closeness, though this may not necessarily translate into improved forecasts. Nevertheless, it would be reasonable to expect an improvement in predictions, in particular in the short run, as the dynamic nature of the relationship between money and its determinants is better parameterized.

APPENDIX

A.1 Solution to the Maximization Problem of Consumers

The setup given by equations (1) to (3) implies the following Hamiltonian:

$$\mathcal{L} = e^{-\theta t}\left[U(c_t) - \lambda_t \left(b_t + m_t + c_t [1 + g(\frac{m_t}{c_t}, \psi_i)] \right.\right. \qquad \textbf{(1)}$$

which in turn implies the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial c_t} = e^{-\theta t}\left[\frac{\partial U(c)}{\partial c_t} - \lambda_t \left(1 + g(\frac{m_t}{c_t}, \psi_i) + \frac{m_t}{c_t} g'_r \right.\right.$$

$$\frac{\partial \mathcal{L}}{\partial m_t} = e^{-\theta t}\left[-\lambda_t \left(1 + g'_{m_t/c_t}(\frac{m_t}{c_t}, \psi_i) \right) + \frac{\lambda_{t+1}}{1 + \pi_t} \right] = 0 \qquad \textbf{(2)}$$

$$\frac{\partial \mathcal{L}}{\partial b_t} = e^{-\theta t}\left[-\lambda_t + \lambda_{t+1}(1 + r_t) \right] = 0$$

$$\lim \; b_t e^{-\theta t} = 0$$

from the bonds condition we obtain the relationship between λ_t and λ_{t+1} which in turn is used in the money holdings condition to obtain equation (4) in the text.

A.2 Specification and Solution to the Maximization Problem of Firms

Firms maximize the present value of its cash flows to be paid as dividends. Output is produced, with constant returns to scale, using two factors: capital (K_t) and labor (L_t):

$$\begin{aligned} Y_t &= F(K_t, L_t) = K_t^{\alpha} L_t^{1-\alpha} \\ y_t &= F(k_t, 1) = k_t^{\alpha} \end{aligned} \tag{3}$$

hence, y_t represents output per worker and k_t is capital stock per worker. Since labor is supplied inelastically by households we can normalize it, without loss of generality, at 1. The labor market, then, clears at wage w_t.

The firm's crucial variable of decision in this set-up is investment, i_t, which is undertaken facing transaction and installment costs, $H(i_t,m_t,k_t)$. Transaction costs arise from the need of obtaining information regarding profitable opportunities in the market; maintaining liquidity (money balances as a fraction of investment) is a cost-reducing strategy. Installment costs, on the other hand, represent the direct cost in terms of output which has to be devoted to produce capital goods; it is increasing in the amount of investment and decreasing in the stock of capital. Hence:

$$\begin{gathered} H(m_t, i_t, k_t, \psi_i) = i_t \cdot h\left(\frac{i_t}{k_t}, \frac{m_t}{i_t}, \psi_i\right) \\ h_{i_t/k_t} > 0 \qquad h_{\psi_i} > 0 \qquad h_{m_t/i_t} < 0 \end{gathered} \tag{4}$$

The firm's budget constraint can be specified as:

$$d_t = f(k_t) - i_t\left[1 + h(m_t, k_t, \psi_i)\right] - w_t - m_t + \frac{m_{t-1}}{(1 + \pi_{t-1})} \tag{5}$$

The optimization problem for the firms, then, can be written as:

$$\max_{i_t, m_t} \quad V = \int_{t=0}^{\infty} \left[f(k_t) - i_t \left(1 + h\left(\frac{i_t}{k_t}, \frac{m_t}{k_t}, \psi_i\right) \right) + \right. \tag{6}$$

subject to :

$$\dot{k}_t = i_t - \delta k_t \qquad with \qquad k_0 > 0 \tag{7}$$

here the interest rate is assumed variable, so that firms face a path of interest rates $\{r_t\}_{t=0,\infty}$ instead of a constant rate. Hence, the discount rate has to be properly adjusted to $e^{-\int_t^s r_s ds}$. The Hamiltonian of the problem is:

$$H_t = e^{-\int_t^s r_s ds}\left[f(k_t) - i_t\left[1 + h(\frac{i_t}{k_t}, \frac{m_t}{i_t}, \psi_i)\right] - m_t + \frac{\quad}{(1}\right. \tag{8}$$

where q_t is the shadow price of capital. The optimality conditions of the problem are:

$$(a) \qquad \frac{\partial H_t}{\partial i_t} = e^{-\int_{s=t}^{\infty} r_s ds}\left[1 + h(\cdot) + \frac{i_t}{k_t} h'_{i/k}(\cdot) - \cdot\right.$$

$$(b) \qquad \frac{\partial H_t}{\partial m_t} = e^{-\int_{s=t}^{\infty} r_s ds}\left[h'_{m/i}(\cdot) - 1 - \frac{1 + r_t}{1 + \pi_t}\right] = \tag{9}$$

$$(c) \qquad -\frac{\partial H_t}{\partial k_t} = \frac{dq_t}{dt} = -f'(k_t) + \left[i_t/k_t\right]^2 h'(\cdot) -$$

$$(d) \qquad \lim_{t\to\infty} q_t k_t e^{-\int_t^s r_s ds} = 0$$

Equation (a) shows that investment will take place until the shadow value of installed capital equals the marginal cost of an extra unit of capital. Then, whenever q_t=1, investment will be zero, as the shadow value of capital equals its replacement cost. Equation (b), is the condition for the maintaining money balances, which yield a money demand function once function h(.) is appropriately specified. Equation (c) corresponds to equation of motion of real wealth, while (d) is the standard no-Ponzi game condition.

REFERENCES

Apt, J. and J. Quiroz (1992): "Una Demanda de Dinero Mensual para Chile, 1983:1-1992:8", ***Revista de Análisis Económico***, vol. 7, pp. 103-139.

Arrau, P. and J. De Gregorio (1994): "Financial Innovation and Money Demand: Application to Chile and Mexico", ***Review of Economics and Statistics,*** vol. 75, pp. 524-530.

Arrau, P.; J. De Gregorio, C. Reinhart and P. Wickham (993): "The Demand for Money in Development Countries: Assessing the Role of Financial Innovation," ***IMF***, (mimeo).

Bernanke, B. and M. Gertler (1987): "Banking and Macroeconomic Equilibrium," ***in New Approaches to Monetary Economics***, W.A. Barnett and K. Singleton (eds.); Cambridge University Press.

Bordo, M. and S. Jonung (1990): "The Long Run Behavior of Velocity of Circulation: the Institutional Approach Revisited," ***Journal of Policy Modeling***, vol. 12, pp. 165-197.

Boyd, J. and E. Prescott (1986): "Financial Intermediary Coalitions," ***Journal of Economic Theory***, vol. 38, pp. 211-232.

Caballero, R. (1992): "A Fallacy of Composition," ***American Economic Review***, vol. 82, pp. 1279-1292.

Caballero, R. (1994): "On the Dynamics of Aggregate Investment," in A. Solimano and L. Serven (eds..), ***Striving for Growth after Adjustment***, The World Bank.

Caballero, R. and E.M.R.A. Engel (1991): "Dynamic S,s Economies," ***Econometrica***, vol. 59. pp. 1659-1686.

Caplin A.S. and and D.F. Spulber (1987): "Menu Costs and Neutrality of Money," ***Quarterly Journal of Economics***, vol. 102, pp. 703-725.

Cochrane, J. (1988): "How Big Is the Random Walk in GNP?," ***Journal of Political Economy***, vol. 96, pp. 893-920.

Cooley, T. And E. Prescott (1976): "Estimation in the Presence of Stochastic Parameter Variation,"***Econometrica***, vol. 44, pp. 167-184.

Cooper R .and J. Ejarque (1994): "Financial Intermediation and Aggregate Fluctuations: a Quantitative Analysis, ***NBER Working Paper*** # 4819.

Constantinides, G. (1991): "Habit Formation: a Resolution of the Equity Premium Puzzle," ***Journal of Political Economy***, vol. 98, pp. 519-543.

Corbo, V. and B. Piedrabuena (1995): "Reduciendo una Inflacion Cronica: Un Modelo de la Inflacion Chilena," in ***Empirical Studies of Inflation in Chile***, F. Morande and F. Rosende (eds). ILADES-Georgetown University.

Easterly, W; P. Mauro, and K. Schmidt-Hebbel (1994): "Money Demand and Seigniorage-Maximizing Inflation," ***Journal of Money, Credit and Banking*** (forthcoming).

Elbadawi, I.; R. Gallant and J. Souza (1987): "An Elasticity Can Be Estimated Consistently Without A Priori Knowledge of Functional Form," ***Econometrica***, vol. 35, pp. 1731-1751.

Engle, R. and C. Granger (1987): "Co-Integration and Error-Correction: Representation, Estimation, and Testing," ***Econometrica***, vol. 35, pp. 251-276.

Esptein Z. and L. Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," ***Journal of Political Economy***, vol. 96, pp.893-920.

Fischer, S. (1979): "Anticipations and the Nonneutrality of Money," ***Journal of Political Economy***, vol. 87, pp. 225-252.

Gallant, R. and H. White (1992): ***A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models,*** Basil Blackwell, N.Y.

García, C; C. Johnson, and F. Morandé (1995): "La Inflación en Chile. Una Nota Estadística Preliminar," in ***Empirical Studies of Inflation in Chile***, F. Morande and F. Rosende (eds). ILADES/Georgetwon University.

Garrido, J.M. and M. Valenzuela (1991): "Dinero y Precios en el Corto y Largo Plazo: Evidencia Empírica para Chile," ***Cuadernos de Economía,*** vol. 28, pp. 141-168.

Gertler, M. (1988): "Financial Structure and Aggregate Economic Activity: An Overview," ***Journal of Money, Credit and Banking,*** vol. 20, pp. 559-596.

Goldfeld, S.M. (1976): "The Case of Missing Money," ***Brookings Papers on Economic Activities***, vol. 3, pp. 638-730.

Goldfeld, S.M. and D. Sichel (1990): "The Demand for Money," in B.M. Friedman and F.H. Hahn (eds.); ***Handbook of Monetary Economics***; Elsevier-Science Publishers,

Granger, C.W.J. (1988): "Some Recent Development in a Concept of Causality," ***Journal of Econometrics***, vol. 39, pp. 199-211.

Granger, C.W.J. and Newbold, P. (1974): "Spurious Regressions in Econometrics," ***Journal of Econometrics***, vol. 2, pp. 111-120.

Granger, C.W.J. and T. Teräsvirta (1993): ***Modelling Nonlinear Economic Relationships***; Oxford University Press.

Hahn, G. and R. Hendrikson (1971): "A Table of Percentage Points of the Distribution of the Largest Absolute Value of *k* Student *t* Variates and its Applications, ***Biometrika***, vol. 2, pp. 323-332.

Hansen, L. P. and K. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectation Models," ***Econometrica***, vol. 50, pp. 895-916.

Heaton, J. and D. Lucas (1993): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Prices," ***NBER Working Paper Series*** # 4249.

Hendry, D. and Neale, A. (1991): "A Montecarlo Study of the Effect of Structural Breaks on Tests for Unit Roots," in P. Hackl and A.H. Weslund (eds..); ***Economic Structural Change: Analysis and Forecast***, Berlin; Springer-Verlag.

Herrera, L.O. and R. Vergara (1992): "Estabilidad de la demanda de dinero, cointegración y política monetaria," ***Cuadernos de Economía,*** vol. 29, pp. 35-54.

King, M. and M. Robson (1993): "A Dynamic Model of Investment and Endogenous Growth," ***Scandinavian Journal of Economics***, vol. 95, pp. 445-466.

Kuan, C.M. and H. White (1994): "Artificial Neural Networks: an Econometric Perspective," ***Econometric Reviews***, vol. 13, pp. 1-91.

Labán, R. (1991): "La Hipótesis de Cointegración y la Demanda por Dinero en Chile: 1974:1988," ***Cuadernos de Economía,*** vol. 28, pp. 169-187.

Lagos, L.F. (1991): "Un Comentario sobre las Reglas Monetarias," ***Cuadernos de Economía,*** vol. 28, pp. 21-28.

Lee, T.H.; H. White; and C.W.J. Granger (1993): "Testing for Nonlinearities in Time Series Models. A Comparison of Neural Network Methods and Alternative Methods, ***Journal of Econometrics***, vol.56, pp. 269-290.

León, J. and R. Soto (1994): "Testing for Unit Roots under Structural Breaks: a Monte Carlo Study, ***Working Paper Series*** #1406, The World Bank.

Lieberman, C (1977): "The Transaction Demand for Money and Technological Change," ***Review of Economics and Statistics***, vol. 59, pp. 307-317.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," ***Econometrica***, vol. 46, pp.1429-1945.

Lucas, R. (1984): "Money in a Theory of Finance," in K. Brunner and A. Metzler (eds.), ***Essays on Macroeconomic Implications of Financial and Labour Markets and Political Processes***, Carnegie-Rochester Conference Series, vol.46, pp.9-41.

Lütkepol, H.(1993): "The Sources of the US Money Demand Instability," ***Empirical Economics***, vol.18, pp.720-743.

Mehra, R and E. Prescott (1985): "The Equity Premium: A Puzzle," ***Journal of Monetary Economics***, vol. 15, pp. 145-162.

Martner, R; D. Titelmann, and A. Uthoff (199): "Componentes Internos y Externos de la Inflación en Chile: Un Enfoque de Cointegración," in ***Empirical Studies of Inflation in Chile***, F. Morande and F. Rosende (eds). ILADES-Georgetown University.

Matte, R. and P. Rojas (1989): "Evolución Reciente del Mercado Monetario y una Estimación de la Demanda por Dinero en Chile," ***Cuadernos de Economía,*** vol. 26, pp. 21-28.

Moore, G; R. Porter and D. Small (1990): "Modeling the Disaggregated Demands for M2 and M1 in the 1980s: the US experience," in the Board of Governors of the Federal Reserve System, ***Financial Sectors in Open Economies. Empirical Analysis and Policy Issues***, Board of Governors, Washington D.C..

Phillips, P.C.B. and Loretan, M. (1991): "Estimating Long-Run Economic Equilibria," ***Review of Economic Studies***, vol. 55, pp. 407-436.

Poterba, J and L. Summers (1988): "Mean Reversion in Stock Prices, Evidence and Implications, ***Journal of Financial Economics***, vol. 22, pp. 27-60.

Rosende, F. and L.O. Herrera (1991): "Teoria y Politica Monetaria: Elementos para el Analisis," ***Cuadernos de Economía,*** vol. 28, pp. 55-94.

Samuelson, P. (1958): "An Exact Consumption-Loan Model of Interest with or Without the Social Contrivance of Money," ***Journal of Political Economy***, vol. 66, pp. 467-482.

Sidrauski, M. (1967): "Inflation and Economic Growth," ***Journal of Political Economy***, vol. 75, pp. 534-544.

Soklos, P.L (1993): "Income Velocity and Institutional Change: Some New Time Series Evidence, 1980-1986," ***Journal of Money, Credit and Banking,*** vol. 26, .

Soto, R. (1995): "Money Demand in a Model of Endogenous Financial Innovation," mimeo, Georgetown University.

Stiglitz, J. and A. Weiss (1981): "Credit Rationing and Markets with Imperfect Information," ***American Economic Review***, vol. 71, pp. 393-410.

Stinchcombe M. and H. White (1991): "Using Feedforward Network to Distinguish Multivariate Populations," Discussion Paper, University of California, San Diego.

Stock, J. and M. Watson (1989): "A Simple Estimator of Cointegrating Vectors in Higher-order Integrated Systems," ***Econometrica***, vol. 61, pp. 783-820.

Svensson, L .P. (1985): "Money and Asset Prices in a Cash-In-Advance Economy," ***Journal of Political Economy***, vol. 93, pp. 919-944.

Teräsvirta, T; C.F. Lin and C.W.J. Granger (1993): "Power of the Neural Network Linearity Test," ***Journal of Time Series Analysis,*** vol. 14, pp. 209-220.

Virén, M. (1992): "Financial Innovations and Currency Demand: New Evidence," ***Empirical Economics***, vol. 17, pp. 451-461.

Wallace, N. (1983): "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy," ***Federal Reserve Bank of Minneapolis Quarterly Review***, Winter, 1-7.

White, H. (1989): "Some Asymptotic Results for Learning in Single Hidden-layer Feed-foward Network Models," ***Journal of the American Statistical Association***, vol. 84, pp. 1003-1013.

White, H. (1989): "Some Asymptotic Results for Learning in Single Hidden-Layer Neural Networks," mimeo.

Wilson, C. (1989): "An Infinite Horizon Model with Money," in J. Green and J.A. Scheinkman (eds.), ***General Equilibrium, Growth, and Trade***, New York, Academic Press.