Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 943

September 2008

Constructive Data Mining: Modeling Argentine Broad Money Demand

Neil R. Ericsson and Steven B. Kamin

Constructive Data Mining:
Modeling Argentine Broad Money Demand

Neil R. Ericsson and Steven B. Kamin*

Abstract: This paper assesses the empirical merits of PcGets and Autometrics—two recent algorithms for computer-automated model selection—using them to improve upon Kamin and Ericsson's (1993) model of Argentine broad money demand. The selected model is an economically sensible and statistically satisfactory error correction model, in which cointegration between money, inflation, the interest rate, and exchange rate depreciation depends on the inclusion of a "ratchet" variable that captures irreversible effects of inflation. Short-run dynamics differ markedly from the long run. Algorithmically based model selection complements opportunities for the researcher to contribute value added in the empirical analysis.

Keywords: Argentina, Autometrics, broad money, dynamic specification, cointegration, conditional models, currency substitution, dollarization, error correction, exogeneity, hyperinflation, irreversibility, model design, model selection, money demand, PcGets, ratchet effect.

JEL classifications: C52, E41.

*Forthcoming in Jennifer L. Castle and Neil Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, 2008. The authors are staff economists in the Division of International Finance, Board of Governors of the Federal Reserve System, Washington, D.C. 20551 U.S.A.; and they may be reached on the Internet at ericsson@frb.gov and kamins@frb.gov respectively. The authors are grateful to Julia Campos, Jennifer Castle, Dale Henderson, David Hendry, Katarina Juselius, Jaime Marquez, Bent Nielsen, Anders Rahbek, Neil Shephard, and two referees for helpful comments. All numerical results were obtained using PcGive Version 12.00, Autometrics Version 1.5, and PcGets Version 1.02: see Doornik and Hendry (2007), Doornik (2008), and Hendry and Krolzig (2001).

1 Introduction

We are delighted to contribute to this Festschrift in honor of David F. Hendry. As discussed in Ericsson (2004), David has contributed to numerous areas of econometrics and economics, including:

- money demand,
- error correction models and cointegration,
- exogeneity,
- model development and design,
- econometric software,
- economic policy,
- consumers' expenditure,
- Monte Carlo methodology,
- the history of econometrics, and
- the theory of economic forecasting.

We draw on David's contributions to the first six topics to assess and improve upon Kamin and Ericsson's (1993) model of Argentine broad money demand, focusing on model design and cointegration analysis. Recent developments by David and co-authors in computer-automated model selection help us obtain a more parsimonious, empirically constant, data-coherent, error correction model for broad money demand in Argentina. Cointegration between money, inflation, the interest rate, and exchange rate depreciation depends on the inclusion of a "ratchet" variable that captures irreversible effects of inflation.

To better understand money demand and currency substitution in a hyperinflationary economy, Kamin and Ericsson (1993) develop an empirical model of broad money (M3) in Argentina for monthly data over 1978–1993, a period including hyperinflation and a subsequent decline in inflation to a rate close to contemporary U.S. and European levels. Kamin and Ericsson's underlying economic theory is a standard money demand model, augmented by short-run nonlinear dynamics and a ratchet effect from inflation. Their empirical model clarifies the relative importance of factors determining money demand and currency holdings. Also, the structure of broad money demand in Argentina does not appear to have changed over the 1980s and 1990s. Rather, the fall in demand during the late 1980s and into the 1990s is explained by changes in the determinants of money demand itself.

That said, the analysis in Kamin and Ericsson (1993) has three notable shortcomings. First, their cointegration analysis excludes a trend, which may have affected inferences. Second, in their single equation modeling of Argentine money demand, Kamin and Ericsson augment the data from the cointegration analysis with an impulse dummy (for a known asset freeze from the Plan Bonex) and an asymmetric term in price acceleration. While both variables are stationary in principle, their exclusion from the cointegration analysis could have affected the results obtained. Third, an alternative single equation model might have been obtained if a different model search path had been followed.

Following the approach in Ericsson (2008, Chapters 9 and 10), the current paper addresses these issues, as follows. Cointegration is re-analyzed, including the impulse dummy, the asymmetric inflation term, and a trend. The cointegrating vector in this expanded framework is similar to the one obtained by Kamin and Ericsson (1993). Path dependence in model selection is examined by using two model selection algorithms: David Hendry and Hans-Martin Krolzig's (2001) PcGets, and Jurgen Doornik and David Hendry's (2007) Autometrics. Kamin and Ericsson's (1993) analysis is robust to multi-path searches by both algorithms; at the same time, Autometrics finds an even more parsimonious specification. The details of the model improvement highlight the strengths and the limitations of computer-automated model selection. Our approach thus illustrates new techniques, which shed light on existing results. And, re-examination of an existing dataset with new techniques is very much in the spirit of other work in this area, including Hendry and Mizon (1978), Engle and Hendry (1993), Doornik, Hendry, and Nielsen (1998), and Hendry (2006).

This paper is organized as follows. Section 2 briefly describes the economic theory and the data. Section 3 summarizes the cointegration analysis and error correction model for Argentine money demand in Kamin and Ericsson (1993). Section 4 re-analyzes the long-run properties of Argentine money demand on the expanded dataset. Section 5 then designs a single equation model of money demand, using the algorithms for computer-automated model selection in PcGets and Autometrics. Depending upon the modeling strategy, pre-search testing, choice of required regressors, and representation and choice of the initial general model, PcGets and Autometrics obtain several distinct—albeit similar—final models in their general-to-specific selection processes. Additional analysis of those models obtains a final specification that is similar to—but more parsimonious than—the one in Kamin and Ericsson (1993). That final specification appears well-specified with empirically constant coefficients; and its economic interpretation is straightforward. Section 6 concludes.

For expositional convenience, two conventions are adopted. First, "domestic" means Argentine. Second, Argentine currency is always denominated in pesos (the Argentine currency at the end of the sample) although historically other currencies were used.

2 Economic Theory and the Data

This section first discusses the theory of money demand (Section 2.1) and then considers the data themselves (Section 2.2).

2.1 Economic Theory

The standard theory of money demand posits:

$$M^d/P \quad = \quad q(Y, \mathbf{R}) \; , \tag{1}$$

where M^d is nominal money demanded, P is the price level, Y is a scale variable, and $\mathbf{R}$ (in bold) is a vector of returns on various assets. The function $q(\cdot,\cdot)$ is increasing in Y, decreasing in those elements of $\mathbf{R}$ associated with assets excluded from M, and increasing in those elements of $\mathbf{R}$ for assets included in M.

Three assets for Argentine residents are considered: broad money (M3), domestic goods, and dollars. Their nominal returns are denoted R, Δp, and Δe, where E is the exchange rate (domestic/foreign), variables in lowercase are in logarithms, and Δ is the difference operator. This choice of assets and returns seems reasonable. Relatively few peso instruments outside of M3 were held in significant quantities during most of the sample period. Also, the interest rate on dollar deposits was small and unvarying relative to Δe, so it was excluded in calculating the return on dollar-denominated assets.

Empirical models below employ (1) in its standard log-linear form, with two modifications. First, the scale variable is omitted, as in Cagan's (1956) money demand model for hyperinflationary economies.[1] Second, following Enzler, Johnson, and Paulus (1976), Simpson and Porter (1980), Piterman (1988), Melnick (1990), Ahumada (1992), and Uribe (1997) *inter alia*, the money demand equation includes a ratchet variable, which is the maximum inflation rate to date, denoted Δp^{max}. Higher inflation rates may induce innovations to economize on the use of domestic money balances. Once inflation subsides, these innovations are unlikely to disappear immediately (if at all), leading to a long-lived negative effect of inflation on money demand. Hence, Δp^{max} may proxy for financial innovation, be it a shift toward dollar usage or toward other forms of economizing on domestic money holdings.

With these two modifications, equation (1) has the following form:

$$m - p \quad = \quad \gamma_0 \, + \, \gamma_1 R \, + \, \gamma_2 \Delta p \, + \, \gamma_3 \Delta e \, + \, \gamma_4 \Delta p^{max} \; . \tag{2}$$

Anticipated signs of coefficients are $\gamma_1 > 0$, $\gamma_2 < 0$, $\gamma_3 < 0$, and $\gamma_4 \leq 0$. Broad money is composed primarily of interest-bearing deposits, so the interest rate R should exert a positive effect on money demand. The coefficients on Δp and Δe should be negative: goods and dollars are alternatives to holding money. Because Δp^{max} increases monotonically throughout the sample, a strictly negative γ_4 implies irreversible reductions in money demand due to historically higher rates of inflation.

[1] A preliminary investigation found little role for Y in the cointegration analysis or in error correction modeling. This is consistent with Ahumada's (1992) evidence on currency demand, and may be due to the relatively stationary nature of real GDP in Argentina over the sample period.

If R, Δp, and Δe enter equation (2) only as relative rates of return, then $\gamma_2 + \gamma_3 = -\gamma_1$, and equation (2) can be rewritten as:

$$m - p \quad = \quad \gamma_0 \; - \; \gamma_2(R - \Delta p) \; - \; \gamma_3(R - \Delta e) \; + \; \gamma_4 \Delta p^{max} \, . \qquad (3)$$

Equation (3) links real money demand to two opportunity costs and the ratchet effect. This representation is particularly useful when interpreting empirical error correction models in the context of multiple markets influencing money demand.

2.2 The Data

This subsection describes the data available and considers some of their basic properties. The data are a broad measure of money (M3), as measured by all peso-denominated currency and domestic bank deposits (M, millions of pesos); the domestic consumer price index (P, 1968 = 1.00); the interest rate on domestic peso-denominated 30-day fixed-term bank deposits (R, fraction at a monthly rate); and the free-market exchange rate (E, in pesos per dollar). Also, p is transformed to the variable $\max(0, \Delta^2 p)$ [denoted $\Delta^2 p^{pos}$] to measure the differential effect of positive (rather than negative) accelerations in prices, as in Ahumada (1992). The variable $\Delta^2 p^{pos}$ is interpretable as allowing asymmetric short-run effects of inflation, similar to Δp^{max} allowing asymmetric long-run effects. All data are monthly and seasonally unadjusted, over January 1977–January 1993. Allowing for lags and transformations, estimation is over February 1978–January 1993 (T = 180) unless otherwise noted. Two dummy variables are also used: B, an impulse dummy for the beginning of the Plan Bonex (January 1990); and S, the seasonal dummy. Kamin and Ericsson (1993, Appendix) provide further details on the data.

Figure 1*a* plots the logarithms of nominal money and prices (m and p), which are notable by spanning orders of magnitude. Sharp increases in both series are visible around 1985 and 1989. While M is the variable of central interest in this study, its evolution is most easily understood in light of the various rates of return. Figure 1*b* plots the (monthly) inflation rate Δp, along with the generated ratchet variable Δp^{max}. Figure 1*c* plots Δp and the interest rate R, which move closely together, albeit with inflation being more volatile on a month-to-month basis. Figure 1*d* graphs R and the depreciation in the nominal exchange rate Δe, which also move closely together, with exchange rate depreciation being highly volatile. That said, real ex post *monthly* returns are commonly in excess of (plus-or-minus) two per cent, in large part owing to the high variability in the inflation rate.

The overall behavior of inflation (and so of R and Δe) can be characterized by periods of increasing inflation, followed by government "plans" to reign in inflation. The acceleration of prices during the early 1980s was sharply reversed in mid-1985 by

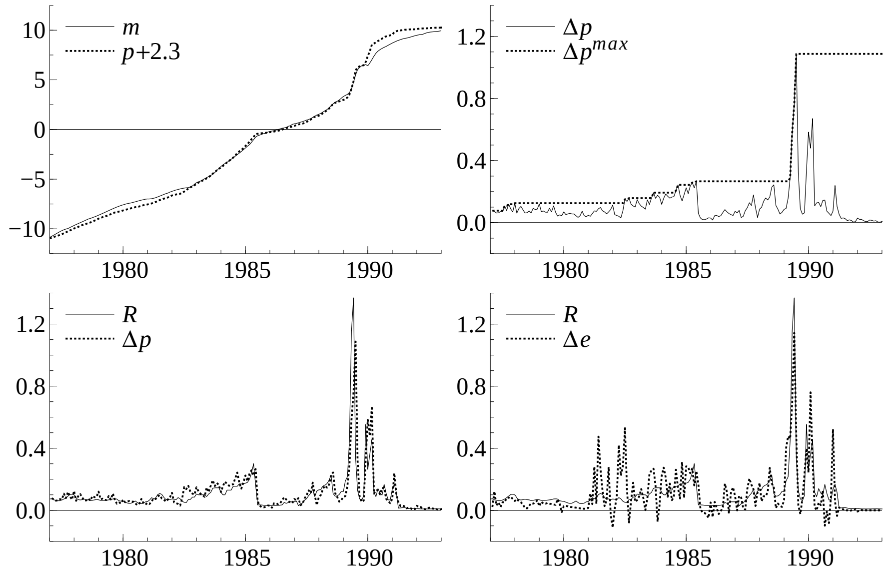

Figure 1: The logarithms of nominal money and prices (m and p), inflation Δp and maximal inflation Δp^{max}, R and Δp, and R and Δe.

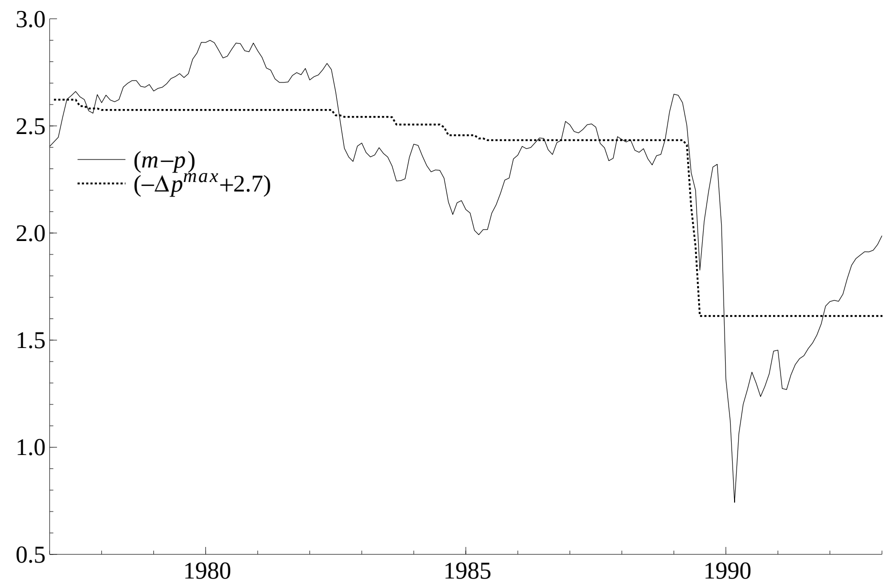

Figure 2: The logarithm of real money $(m-p)$ and the negative of the maximal inflation rate (Δp^{max}), adjusted for means.

the Plan Austral, which combined wage, price, and exchange rate freezes with some fiscal adjustment. Reductions in the fiscal deficit were not sufficient to eliminate inflationary pressures, which resumed in earnest by 1987. The August 1988 Plan Primavera ("Spring Plan") followed, and it aimed to limit the growth of prices and the official exchange rate to 4 percent per month. While inflation fell temporarily, the real exchange rate appreciated and the fiscal situation deteriorated. In February 1989, the Central Bank floated the exchange rate for financial transactions, which promptly depreciated sharply; and inflation rapidly increased to a record 197 percent per month in July 1989.

Under newly elected President Menem, the authorities announced a new program similar to the Plan Austral. Initially, inflation fell dramatically; but appreciation of the real exchange rate forced the Central Bank to float the commercial exchange rate, which quickly depreciated in value and spurred price inflation. In January 1990, the authorities attempted to restrain inflation by freezing most domestic peso-denominated bank time deposits and converting them to 10-year dollar-denominated bonds known as Bonex. The so-called Plan Bonex had little immediate effect upon inflation, but it did further reduce the Argentine public's faith in their financial system. By March 1990, when inflation reached 95.5 percent per month, broad money reached a record low of 3.1 percent of GDP.

Subsequently, inflation declined to single-digit levels due to a reduction in monetary emission made possible by concerted efforts to achieve fiscal adjustment. The fiscal deficit declined from over 20 percent of GDP in 1989 to about 3 percent in 1990 and 2 percent in 1991. In March 1991, the government announced the "Convertibility Program," which fixed the exchange rate against the dollar and required the Central Bank to hold international reserves equivalent to the monetary base. Subsequently, the inflation rate fell to under 1 percent per month.

Figure 2 graphs the log of real money $(m-p)$ and the negative of the ratchet variable Δp^{max}. Real money initially increases gradually, then falls abruptly by 20% in 1982. After continuing to fall through 1984, real money increases until the hyperinflation in 1989, when it plummets to approximately half its "pre-hyper" level. Even after very low inflation in subsequent years, real money did not return to its level of early 1989. Declines in real money are closely correlated with increases in the ratchet variable, although the stability of a relation between these variables may be an issue, noting the remaining large deviations between them.[2]

[2] For further analytical and empirical discussion of the Argentine economy, see Howard (1987), the World Bank (1990), Kamin (1991), Kiguel (1991), Manzetti (1991), Beckerman (1992), Kamin and Ericsson (1993), and Helkie and Howard (1994). See Dominguez and Tesar (2007) for a history of the post-1990 period.

3 Previous Results

This section summarizes the model of Argentine money demand developed by Kamin and Ericsson (1993).

Kamin and Ericsson (1993) test for and find cointegration between real money, the interest rate, the inflation rate, exchange rate depreciation, and the ratchet variable; and the ratchet variable is key to finding cointegration. While the interest rate and the exchange rate do not appear to be weakly exogenous, there are only minor differences between system estimates of the cointegrating vector and the solved long-run coefficients from a conditional single-equation autoregressive distributed lag (ADL) model. So, Kamin and Ericsson (1993) model broad money as a single-equation conditional error correction model (ECM).

In their single equation modeling, Kamin and Ericsson (1993) start with a seventh-order ADL that has 63 coefficients and simplify it to a more restricted "intermediate" ADL with only 30 coefficients. Kamin and Ericsson (1993) then further simplify to obtain the following 16-coefficient model, which is their equation (6).

$$
\begin{aligned}
\widehat{\Delta(m-p)}_t = \; & \underset{\substack{(0.028) \\ [0.035]}}{0.264}\,\Delta(m-p)_{t-1} + \underset{\substack{(0.031) \\ [0.032]}}{0.091}\,\Delta^2(m-p)_{t-5} \\
& - \underset{\substack{(0.040) \\ [0.041]}}{0.740}\,\Delta^2 p_t + \underset{\substack{(0.040) \\ [0.054]}}{0.101}\,\Delta^2 p_{t-5} - \underset{\substack{(0.078) \\ [0.089]}}{0.594}\,\Delta^2 p_t^{pos} \\
& + \underset{\substack{(0.018) \\ [0.021]}}{0.059}\,\Delta\Delta_6 p_t + \underset{\substack{(0.022) \\ [0.018]}}{0.182}\,\Delta^2 R_t + \underset{\substack{(0.044) \\ [0.045]}}{0.536}\,(R-\Delta p)_{t-1} \\
& + \underset{\substack{(0.022) \\ [0.022]}}{0.103} - \underset{\substack{(0.0078) \\ [0.0080]}}{0.0337}\,(m-p)_{t-1} - \underset{\substack{(0.017) \\ [0.019]}}{0.069}\,\Delta e_{t-1} \\
& - \underset{\substack{(0.010) \\ [0.010]}}{0.028}\,\Delta p_{t-1}^{max} - \underset{\substack{(0.038) \\ [0.039]}}{0.216}\,B_t + \underset{\substack{(0.032) \\ [0.020]}}{0.179}\,B_{t-3} \\
& + \underset{\substack{(0.64) \\ [0.43]}}{2.45}\,S_{t-6} + \underset{\substack{(0.62) \\ [0.74]}}{5.09}\,S_{t-12}
\end{aligned}
\qquad (4)
$$

$T = 180$ [1978(2)–1993(1)] $\quad \mathrm{R}^2 = 0.9489 \quad \hat{\sigma} = 2.192\% \quad dw = 2.08$
$Inn_1 : F(47, 117) = 1.47^{+} \quad Inn_2 : F(14, 150) = 1.46 \quad AR : F(7, 157) = 0.61$
$ARCH : F(7, 150) = 2.75^{*} \quad Normality : \chi^2(2) = 0.59 \quad RESET : F(1, 163) = 0.43$
$Hetero : F(26, 137) = 0.99 \quad Form : F(102, 61) = 0.71 \quad Chow : F(33, 131) = 0.76,$

where a circumflex ˆ on the dependent variable denotes its fitted value, the subscript t is the time index, $\Delta\Delta_6 p_t = \Delta(p_t - p_{t-6})$, R^2 is the squared multiple correlation coefficient, and $\hat{\sigma}$ is the estimated residual standard error. The long-run solution to equation (4) is:

$$m - p \quad = \quad 3.05 \quad + \quad 15.93 \ (R - \Delta p) \ - \quad 2.04 \ \Delta e \ - \quad 0.84 \ \Delta p^{max} \,. \qquad (5)$$

Kamin and Ericsson (1993) show that equation (4) has a straightforward economic interpretation and is statistically satisfactory. Economically, the long-run coefficients in (5) satisfy sign restrictions that are consonant with a money demand function. The short-run variables and coefficients in (4) are also easily understood. Each short-run variable enters as a second difference (an acceleration), which is a natural transformation to stationarity for a potentially I(2) variable. The coefficient on $\Delta^2 p_t$ is close to -1, implying that, in the short run, agents are in essence adjusting nominal (and not real) money.[3] The lag lengths on $\Delta^2(m-p)_{t-5}$, $\Delta^2 p_{t-5}$, and $\Delta\Delta_6 p_t$ are consistent with agents' adjustments for seasonality in the data. The variable $\Delta\Delta_6 p_t$ is also consistent with a natural data-based predictor of future (seasonal) inflation, extending the theoretical and empirical developments on such predictors in Flemming (1976), Hendry and Ericsson (1991), and Campos and Ericsson (1999). And, the coefficient on $\Delta^2 p_t^{pos}$ is very negative and statistically significant, implying stronger reactions to rising inflation than to falling inflation.

The estimated money demand function also shed lights on the dollarization of the Argentine economy. Kamin and Ericsson (2003) reinterpret the ratchet effect in light of data measuring the extent of dollarization. Specifically, the reduction in peso money demand attributable to the ratchet effect is comparable in magnitude to the estimated stock of total dollar assets held domestically by Argentine residents, where those assets are estimated from U.S. Treasury data. This suggests that secular reductions in the demand for pesos reflect substitution into dollars rather than mere economizing on peso balances (or other forms of financial innovation). Thus, the ratchet may proxy for dollar holdings, which relaxes the draconian assumption of true irreversibility.

Statistically, Kamin and Ericsson (1993) show that equation (4) is parsimonious and empirically constant and satisfies a variety of diagnostic tests. Equation (4) and the regressions below report diagnostic statistics for testing against various alternative hypotheses: residual autocorrelation (dw and AR), skewness and excess kurtosis ($Normality$), autoregressive conditional heteroscedasticity ($ARCH$), RESET ($RESET$), heteroscedasticity ($Hetero$ and $Form$), non-innovation errors relative to a more general model (Inn), and predictive failure ($Chow$, Chow's prediction interval statistic). The asymptotic null distribution is designated by $\chi^2(\cdot)$ or $F(\cdot,\cdot)$, where

[3] Hendry and Ericsson (1991, p. 853) and Baba, Hendry, and Starr (1992) find similar results for narrow money demand in the United Kingdom and the United States. Also, in keeping with this observation about $\Delta^2 p_t$, Kamin and Ericsson (1993) simplify the restricted intermediate ADL to obtain an alternative ECM where that ECM has Δm_t as the dependent variable.

the degrees of freedom fill the parentheses. Estimated standard errors are in parentheses $(\cdot)$, below coefficient estimates; heteroscedasticity-consistent standard errors are in brackets $[\cdot]$. See Doornik and Hendry (2007) for details and references.

In spite of the apparent robustness of equation (4), its design has shortcomings. The associated cointegration analysis excludes $\Delta^2 p^{pos}$, a linear trend, and an impulse dummy for the Plan Bonex. And, equation (4) may depend on the path taken for model selection. The remainder of the current paper addresses these issues.

4 Integration and Cointegration

This section presents unit root tests for the variables of interest (Section 4.1). Then, Johansen's maximum likelihood procedure is applied to test for cointegration among real money, inflation, the interest rate, exchange rate depreciation, $\Delta^2 p^{pos}$, the ratchet variable, and a linear trend (Section 4.2). Coefficient restrictions and the adjustment mechanism are examined in the Johansen framework.

4.1 Integration

Table 1 lists augmented Dickey–Fuller (ADF) statistics and related calculations for the data. In order to test whether a given series is I(0), I(1), I(2), or I(3), Table 1 calculates unit root tests for the original variables, for their changes, and for the changes of the changes. This permits testing the order of integration, albeit by testing adjacent orders of integration in a pairwise fashion. The largest estimated root ($\hat{\rho}$) is listed adjacent to each ADF statistic: this root should be approximately unity if the null hypothesis is correct. The lag length of the reported ADF regression is based on minimizing the AIC, starting with a maximum of twelve lags.

Nominal money, prices, and the exchange rate appear to be I(2). Real money, the nominal interest rate, inflation, and the inflation ratchet variable appear to be I(1). The ex post real interest rate and $R - \Delta e$ appear stationary.

4.2 Cointegration

Cointegration analysis helps clarify the long-run relationships between integrated variables. A brief review leads to the current analysis and places the latter in context.

Johansen's (1988, 1991) procedure is maximum likelihood for finite-order vector autoregressions (VARs) with variables that are integrated of order one [I(1)], and it is easily calculated for such systems. Various approaches exist for modeling possibly cointegrated I(2) variables. Johansen (1992b) proposes and implements a unified (vector autoregressive) system approach for the entire testing sequence going from

Table 1: ADF statistics for testing a unit root in various time series.

Variable[a,b]	lag ℓ	$t_{ADF(\ell)}$	$\hat{\rho}$	$\hat{\sigma}$ (%)	t-prob (%)	F-prob (%)	AIC
m	8	–2.81	0.988	5.318	2.1	52.1	–5.75
p	12	–2.95	0.984	7.741	17.8	—	–4.98
e	8	–3.07	0.970	12.63	0.1	43.7	–4.02
$m-p$	11	–3.08	0.934	6.912	2.9	47.1	–5.22
R	12	–2.06	0.821	9.208	8.2	—	–4.64
Δp^{max}	10	–1.74	0.983	2.631	5.7	62.9	–7.15
$R-\Delta p$	5	–4.64**	0.079	9.300	0.1	12.4	–4.65
$R-\Delta e$	1	–9.72**	0.069	10.65	1.4	24.2	–4.40
Δm	8	–2.28	0.849	5.431	15.8	48.6	–5.71
Δp	12	–2.45	0.810	7.905	8.1	—	–4.94
Δe	8	–2.76	0.703	12.96	13.7	84.6	–3.97
$\Delta(m-p)$	7	–3.95*	0.380	7.118	14.0	67.5	–5.18
ΔR	11	–5.46**	–1.952	9.334	2.4	43.5	–4.61
$\Delta(\Delta p^{max})$	9	–4.27**	0.487	2.657	3.0	76.4	–7.14
$\Delta(R-\Delta p)$	11	–6.16**	–5.268	9.670	1.1	81.7	–4.54
$\Delta(R-\Delta e)$	12	–7.49**	–7.758	10.71	9.1	—	–4.33
$\Delta^2 m$	6	–9.12**	–1.348	5.520	1.7	40.8	–5.69
$\Delta^2 p$	10	–4.96**	–1.517	8.072	11.7	43.9	–4.91
$\Delta^2 e$	6	–9.36**	–2.257	13.26	0.0	66.7	–3.94
$\Delta^2(m-p)$	9	–7.56**	–3.759	7.323	1.2	46.0	–5.11
$\Delta^2 R$	12	–7.87**	–11.81	9.888	4.1	—	–4.49
$\Delta^2(\Delta p^{max})$	10	–6.09**	–2.108	2.778	11.6	78.5	–7.04
$\Delta^2(R-\Delta p)$	11	–7.67**	–17.68	10.65	2.3	86.0	–4.35
$\Delta^2(R-\Delta e)$	12	–8.49**	–19.20	12.19	0.2	—	–4.07
$\Delta^2 p^{pos}$	5	–3.17$^+$	0.643	4.188	0.7	76.7	–6.24

Notes:

a. Twelfth-order ADF regressions were initially estimated, and the final lag length was selected to minimize the Akaike Information Criterion (AIC). The columns report the name of the variable examined, the selected lag length ℓ, the ADF statistic on the simplified regression ($t_{ADF(\ell)}$), the estimated coefficient on the lagged level that is being tested for a unit value ($\hat{\rho}$), the regression's residual standard error ($\hat{\sigma}$, in %), the tail probability of the t-statistic on the longest lag of the final regression (t-prob, in %), the tail probability of the F-statistic for the lags dropped (F-prob, in %), and the AIC.

b. All of the ADF regressions include an intercept, monthly dummies, and a linear trend. MacKinnon's (1991) approximate finite-sample critical values for the corresponding ADF statistics are –3.14 (10%), –3.44 (5%), and –4.01 (1%) for $T = 177$. In this table, and in the other results reported herein, rejection of the indicated null hypothesis is denoted by $^+$, *, and ** for the 10%, 5%, and 1% levels. Samples sizes are $T = 179$, $T = 178$, and $T = 177$ respectively for the three null hypotheses.

I(2) to I(1) to I(0). His empirical application uses data on U.K. narrow money demand, which appear to have the same orders of integration as the Argentine series above. For the U.K. data, Johansen (1992b) tests for and finds that nominal money and prices (which are I(2)) cointegrate with a $(+1 : -1)$ cointegrating vector to give real money, which is I(1). He then tests for and finds that real money, inflation, real income, and interest rates (all of which are I(1)) cointegrate. Because the I(2) Argentine variables m and p appear to cointegrate as the I(1) variable $m - p$, the cointegration analysis here *begins* with the variables $m-p$, Δp, R, Δp^{max}, Δe, $\Delta^2 p^{pos}$, and a linear trend.

Empirically, the lag order of the VAR is not known *a priori*, so some testing of lag order may be fruitful in order to ensure reasonable power of the Johansen procedure. Given the number of variables, the number of observations, and the data's periodicity, the largest system considered is a seventh-order VAR of $m-p$, Δp, R, Δp^{max}, Δe, and $\Delta^2 p^{pos}$. In that VAR, the linear trend is restricted to lie in the cointegration space; and an intercept, seasonal dummies, and the Plan Bonex dummy B (and three of its lags) enter freely. Empirically, the seventh lag may be statistically insignificant, but no further lag restrictions appear feasible, so inferences below are for the seventh-order VAR.

Table 2 reports the standard statistics, 95% critical values (c.v.'s), and estimates for Johansen's procedure applied to this seventh-order VAR. The maximal eigenvalue and trace eigenvalue statistics (λ_{max} and λ_{trace}) strongly reject the null of no cointegration in favor of at least one cointegrating relationship, and likely in favor of two cointegrating relationships. However, parallel statistics with a degrees-of-freedom adjustment (λ^a_{max} and λ^a_{trace}) suggest only one cointegrating relationship. Because the VAR for Table 2 uses a large number of degrees of freedom in estimation, inferences are based on the adjusted eigenvalue statistics.

Table 2 also reports the standardized eigenvectors and adjustment coefficients, denoted β' and α in a common notation. The first row of β' is the estimated cointegrating vector, which can be written in the form of (2):

$$\begin{aligned} m - p \quad = \quad & \text{intercept} \; - \underset{(3.72)}{10.89}\, \Delta p \; + \underset{(4.48)}{17.53}\, R \; - \underset{(0.32)}{1.20}\, \Delta p^{max} \\ & - \underset{(1.55)}{6.17}\, \Delta e \; - \underset{(10.68)}{62.69}\, \Delta^2 p^{pos} \; + \underset{(0.0020)}{0.0028}\, t. \end{aligned} \tag{6}$$

All coefficients have their anticipated signs. Also, the trend t appears to be statistically insignificant: $\chi^2(1) = 1.31$ [0.252], where the asymptotic p-value is in square brackets. And, the hypothesis of "relative rates of return" in (3) appears acceptable. Numerically, the sum of the coefficients on Δp and Δe (-17.06) is approximately equal to minus the coefficient on R (17.53). Statistically, that restriction cannot be rejected: $\chi^2(1) = 0.04$ [0.850]. Jointly, the restrictions on the trend and rates of

Table 2: A cointegration analysis of the Argentine money demand data.

Rank r	$r=0$	$r\le 1$	$r\le 2$	$r\le 3$	$r\le 4$	$r\le 5$	$r\le 6$
Log-likelihood	2497.21	2528.69	2551.50	2563.38	2571.45	2576.82	2578.52
Eigenvalue λ_r	–	0.295	0.224	0.124	0.086	0.058	0.019
				Null hypothesis			
	$r=0$	$r\le 1$	$r\le 2$	$r\le 3$	$r\le 4$	$r\le 5$	
$\lambda_{\max}$	62.98**	45.61**	23.77	16.13	10.75	3.38	
$\lambda^a_{\max}$	48.28*	34.97	18.22	12.37	8.24	2.59	
95% c.v.	43.97	37.52	31.46	25.54	18.96	12.25	
λ_{trace}	162.6**	99.64**	54.03	30.26	14.13	3.38	
λ^a_{trace}	124.7*	76.39	41.42	23.20	10.83	2.59	
95% c.v.	114.9	87.31	62.99	42.44	25.32	12.25	
				Eigenvectors β'			
Variable	$m-p$	Δp	R	Δp^{max}	Δe	$\Delta^2 p^{pos}$	*trend*
	1	10.89	–17.53	1.20	6.17	62.69	–0.0028
	0.08	1	–0.78	0.04	–0.30	0.50	0.0003
	–0.25	–2.40	1	–0.29	–0.49	4.66	–0.0002
	–0.61	–8.49	47.66	1	–17.90	–5.39	–0.0062
	–1.43	15.50	–18.97	–0.38	1	7.94	–0.0092
	–0.63	0.34	–2.35	–1.27	0.35	1	0.0058
Variable				Adjustment coefficients α			
$m-p$	–0.020	0.265	0.083	0.002	0.010	0.010	
Δp	0.015	–0.365	0.015	–0.002	–0.013	–0.002	
R	0.034	0.085	–0.067	0.002	–0.020	0.015	
Δp^{max}	0.016	–0.137	–0.005	0.002	–0.004	0.002	
Δe	–0.048	0.877	0.093	0.011	–0.034	–0.012	
$\Delta^2 p^{pos}$	0.005	–0.344	–0.052	0.000	–0.009	0.001	
				Weak exogeneity test statistics			
	$m-p$	Δp	R	Δp^{max}	Δe	$\Delta^2 p^{pos}$	
$\chi^2(1)$	6.57*	3.70$^+$	9.89**	12.6**	4.58*	0.59	
				Statistics for testing the significance of a given variable in $\beta' x$			
	$m-p$	Δp	R	Δp^{max}	Δe	$\Delta^2 p^{pos}$	*trend*
$\chi^2(1)$	2.71$^+$	2.79$^+$	7.65**	3.89*	7.44**	16.7**	1.31
				Multivariate statistics for testing trend stationarity			
	$m-p$	Δp	R	Δp^{max}	Δe	$\Delta^2 p^{pos}$	
$\chi^2(5)$	48.4**	45.7**	43.7**	56.6**	30.6**	24.2**	

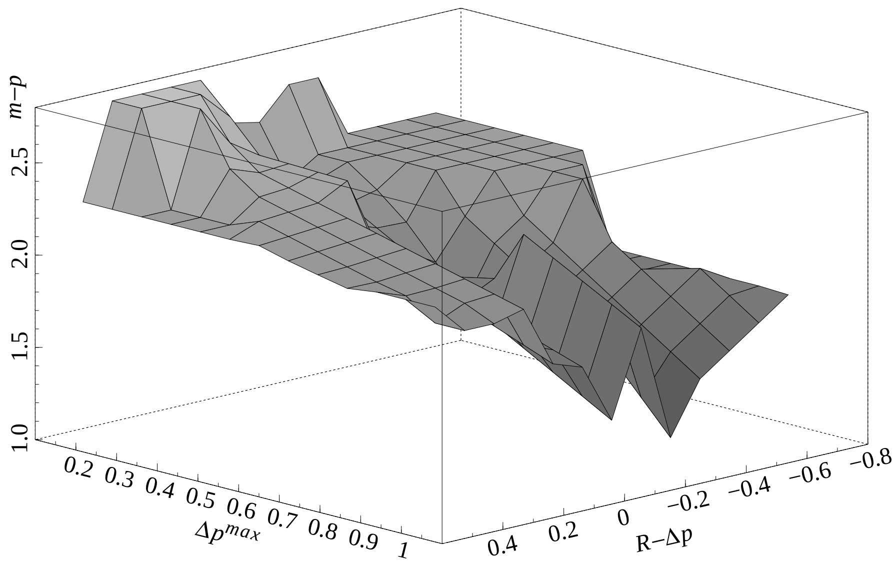

Figure 3: The logarithm of real money $(m-p)$, plotted against the maximal inflation rate (Δp^{max}) and the real interest rate $(R-\Delta p)$.

return also appear acceptable: $\chi^2(2) = 1.39$ [0.498].

Table 3 reports the estimated values of α and β when estimated unrestrictedly, and when estimated with a zero coefficient on the trend imposed, with the hypothesis of "relative rates of return" imposed, and with both of those restrictions imposed. The similarity of coefficient estimates across the various restrictions points to the robustness of the results and is partial evidence in favor of those restrictions.

Thus, the nominal interest rate and inflation enter the long-run money demand function as the ex post real rate, with a semi-elasticity of about eleven, which is about unity at annual rates. The nominal interest rate relative to the exchange-rate depreciation has about half that effect on money demand. Money demand is highly sensitive to the movement of inflation, both through $\Delta^2 p^{pos}$ and through the ratchet variable Δp^{max}. In particular, for each additional percent in the maximal monthly inflation rate over the relative past, the coefficient on Δp^{max} implies approximately one percent lower money holdings.

Figure 3 plots key aspects of equation (6)—namely, the relationship between the variables $(m-p)$, Δp^{max}, and $(R-\Delta p)$. Real money holdings fall as Δp^{max} increases and as the return on money relative to goods $(R-\Delta p)$ declines.

Returning to Table 2, the coefficients in the first column of α measure the feedback effects of the (lagged) disequilibrium in the cointegrating relation on the variables in the vector autoregression. Specifically, -0.020 is the estimated feedback coefficient

Table 3: Just-identified and over-identified estimates of β and α, with corresponding estimated standard errors, from a cointegration analysis of Argentine money demand.

Variable corresponding to an element of β' or α'						
$m-p$	Δp	R	Δp^{max}	Δe	$\Delta^2 p^{pos}$	*trend*
Estimate of β' (just-identified)						
1	10.89	−17.53	1.20	6.17	62.69	−0.0028
	(3.72)	(4.48)	(0.32)	(1.55)	(10.68)	(0.0020)
Estimate of β' (zero coefficient on trend imposed)						
1	10.45	−14.49	0.92	3.51	45.28	0
	(2.71)	(3.27)	(0.14)	(1.07)	(7.75)	
Estimate of β' (rates-of-return restriction imposed)						
1	10.64	−16.60	1.21	5.96	58.67	−0.0027
	(3.55)	(3.87)	(0.27)	(1.46)	(8.14)	(0.0019)
Estimate of β' (trend and rates-of-return restrictions imposed)						
1	10.19	−13.58	0.94	3.38	41.48	0
	(2.56)	(2.76)	(0.12)	(1.01)	(5.51)	
Estimate of α' (just-identified)						
−0.020	0.015	0.034	0.016	−0.048	0.005	
(0.007)	(0.007)	(0.012)	(0.004)	(0.021)	(0.005)	
Estimate of α' (zero coefficient on trend imposed)						
−0.024	0.016	0.050	0.020	−0.051	0.003	
(0.010)	(0.009)	(0.016)	(0.005)	(0.029)	(0.007)	
Estimate of α' (rates-of-return restriction imposed)						
−0.021	0.015	0.036	0.017	−0.050	0.005	
(0.008)	(0.007)	(0.012)	(0.004)	(0.022)	(0.005)	
Estimate of α' (trend and rates-of-return restrictions imposed)						
−0.026	0.017	0.054	0.022	−0.054	0.004	
(0.011)	(0.010)	(0.017)	(0.006)	(0.031)	(0.007)	

for the money equation. The negative coefficient implies that lagged excess money induces smaller holdings of current money. The coefficient's numerical value implies slow adjustment to remaining disequilibrium. The estimated coefficient is numerically smaller than those for quarterly broad money demand (e.g., -0.26, -0.15, and -0.20 in Taylor (1986)) and monthly currency demand (e.g., -0.14 for Argentina in Ahumada (1992)). However, smaller adjustment coefficients are plausible with high-frequency data for a broad aggregate.

The third block from the bottom of Table 2 reports values of the statistic for testing weak exogeneity of a given variable for the cointegrating vector. Equivalently, the statistic tests whether or not the corresponding row of α is zero; see Johansen (1992a, 1992b). If the row of α is zero, disequilibrium in the cointegrating relationship does not feed back onto that variable. Surprisingly, inflation (including in its form $\Delta^2 p^{pos}$) may be weakly exogenous. However, the interest rate, the exchange rate, and the ratchet variable are not weakly exogenous, justifying a systems approach to analyzing cointegration.

The penultimate block in Table 2 reports statistics for testing the significance of individual variables in the cointegrating vector. Each variable is significant, except the linear trend.

The final block in Table 2 reports values of a multivariate statistic for testing the trend stationarity of a given variable. Specifically, this statistic tests the restriction that the cointegrating vector contains all zeros except for a unity corresponding to the designated variable and an unrestricted coefficient on the trend, with the test being conditional on the presence of exactly one cointegrating vector; see Johansen (1995, p. 74). For instance, the null hypothesis of trend-stationary real money implies that the cointegrating vector is $(1\ 0\ 0\ 0\ 0\ 0\ *)'$, where "$*$" represents an unrestricted coefficient on the linear trend. Empirically, all of the stationarity tests reject with p-values less than 0.1%. By being multivariate, these statistics may have higher power than their univariate counterparts. Also, the null hypothesis is the stationarity of a given variable rather than the nonstationarity thereof, and stationarity may be a more appealing null hypothesis. That said, these rejections of stationarity are in line with the *inability* in Table 1 to reject the null hypothesis of a unit root in each of $m - p$, Δp, R, Δp^{max}, and Δe.

Because R, Δp^{max}, and Δe are not weakly exogenous for the cointegrating vector, inferences in a single equation for broad money could be hazardous if the cointegrating vector is estimated jointly with the equation's dynamics; see Hendry (1995). One solution is to model a subsystem. A second solution is to construct an error correction term from the system estimates and then develop a single equation ECM that uses that system-based error correction term. A third solution—adopted below—is to develop a single equation ECM from the single equation ADL, noting that the system estimate of the cointegrating relationship is numerically close to the ADL's long-run solution. See Hendry and Doornik (1994) and Juselius (1992) for paradigms of the first two approaches.

5 Computer-automated Model Selection

This section first describes the model selection algorithms in PcGets and Autometrics (Section 5.1) and then applies these algorithms to an ECM representation of an ADL for Argentine money demand (Section 5.2).

5.1 The Algorithms in PcGets and Autometrics

Hendry and Krolzig (2001) develop a computer program PcGets, which extends and improves upon Hoover and Perez's (1999) automated model-selection algorithm; see also Hendry and Krolzig (1999, 2003, 2005) and Krolzig and Hendry (2001). Doornik and Hendry (2007) implement a third-generation algorithm called Autometrics, which is part of PcGive version 12. PcGets and Autometrics utilize one-step and multi-step simplifications along multiple paths, diagnostic tests as additional checks on the simplified models, and encompassing tests to resolve multiple terminal models. Both analytical and Monte Carlo evidence show that the resulting model selection is relatively non-distortionary for Type I errors. At an intuitive level, PcGets and Autometrics function as a series of sieves that aim to retain parsimonious congruent models while discarding both noncongruent models and over-parameterized congruent models. This feature of the algorithms is eminently sensible, noting that the data generation process itself is congruent and is as parsimonious as feasible.

The remainder of the current subsection summarizes PcGets and Autometrics as automated model-selection algorithms, thereby providing the necessary background for interpreting their application in Section 5.2. For ease of reference, the algorithm in PcGets is divided into four "stages", denoted Stage 0, Stage 1, Stage 2, and Stage 3. For full details of PcGets's algorithm, see Hendry and Krolzig (2001, Appendix A1). Hendry and Krolzig (2003) describe the relationship of the general-to-specific approach to other modeling approaches in the literature, and Hoover and Perez (2004) extend the general-to-specific approach to cross-section regressions.

Stage 0: the general model and F pre-search tests. Stage 0 involves two parts: the estimation and evaluation of the general model, and some pre-search tests aimed at simplifying the general model before instigating formal multi-path searches.

First, the general model is estimated, and diagnostic statistics are calculated for it. If any of those diagnostic statistics is unsatisfactory, the modeler must decide what to do next—whether to "go back to the drawing board" and develop another general model, or whether to continue with the simplification procedure, perhaps ignoring the offending diagnostic statistic or statistics.

Second, PcGets attempts to drop various sets of potentially insignificant variables. PcGets does so by dropping all variables at a given lag, starting with the longest lag. PcGets also does so by ordering the variables by the magnitude of their t-ratios and either dropping a group of individually insignificant variables or (alternatively) retaining a group of individually statistically significant variables. In effect, an F pre-search test for a group of variables is a single test for multiple simplification paths, a characteristic that helps control the costs of search. If these tests result in a statistically satisfactory reduction of the general model, then that new model is the starting point for Stage 1. Otherwise, the general model itself is the starting point for Stage 1.

Stage 1: a multi-path encompassing search. Stage 1 tries to simplify the model from Stage 0 by searching along multiple paths, all the while ensuring that the diagnostic tests are not rejected. If all variables are individually statistically significant, then the initial model in Stage 1 is the final model. If some variables are statistically insignificant, then PcGets tries deleting those variables to obtain a simpler model. PcGets proceeds down a given simplification path only if the models along that path have satisfactory diagnostic statistics. If a simplification is rejected or if a diagnostic statistic fails, PcGets backtracks along that simplification path to the most recent previous acceptable model and then tries a different simplification path. A terminal model results if the model's diagnostic statistics are satisfactory and if no remaining regressors can be deleted.

If PcGets obtains only one terminal model, then that model is the final model, and PcGets proceeds to Stage 3. However, because PcGets pursues multiple simplification paths in Stage 1, PcGets may obtain multiple terminal models. To resolve such a situation, PcGets creates a union model from those terminal models and tests each terminal model against that union model. PcGets then creates a new union model, which nests all of the surviving terminal models; and that union model is passed on to Stage 2.

Stage 2: another multi-path encompassing search. Stage 2 in effect repeats Stage 1, applying the simplification procedures from Stage 1 to the union model obtained at the end of Stage 1. The resulting model is the final model. If Stage 2 obtains more than one terminal model after applying encompassing tests, then the final model is selected by using the Akaike, Schwarz, and Hannan–Quinn information criteria. See Akaike (1973, 1981), Schwarz (1978), and Hannan and Quinn (1979) for the design of these information criteria, and Atkinson (1981) for the relationships between them.

Stage 3: subsample evaluation. Stage 3 re-estimates the final model over two subsamples and reports the results. If a variable is statistically significant in the full sample and in both subsamples, then the inclusion of that variable in the final model is regarded as "100% reliable". If a variable is statistically insignificant in one or both subsamples or in the full sample, then its measure of reliability is reduced. A variable that is statistically insignificant in both subsamples and in the full sample is regarded as being "0% reliable". The modeler is left to decide what action, if any, to take in light of the degree of reliability assigned to each of the regressors.

PcGets thus has two components:

1. Estimation and diagnostic testing of the general unrestricted model (Stage 0); and
2. Selection of the final model by
   (a) pre-search simplification of the general unrestricted model (Stage 0),
   (b) multi-path (and possibly iterative) selection of the final model (Stages 1 and 2), and
   (c) post-search evaluation of the final model (Stage 3).

This subsection's description of these four stages summarizes the algorithm in PcGets. Below, Section 5.2 summarizes the actual simplifications found by PcGets in practice, thereby providing additional insight into PcGets's algorithm.

PcGets requires the modeler to choose which tests are calculated and to specify the critical values for those tests. In PcGets, the modeler can choose the test statistics and their critical values directly, although doing so is tedious because of the number of statistics involved. To simplify matters, PcGets offers two options with pre-designated selections of test statistics and critical values. These two options are called "liberal" and "conservative" model selection strategies. The liberal strategy errs on the side of keeping some variables, even although they may not actually matter. The conservative strategy keeps only variables that are clearly significant statistically, erring in the direction of excluding some variables, even although those variables may matter. Which strategy is preferable depends in part on the data themselves and in part on the objectives of the modeling exercise, although (as below) the two approaches may generate similar or identical results.

The algorithm in PcGets is general-to-specifc, multi-path, iterative, and encompassing, with diagnostic tests providing additional assessments of statistical adequacy, and with options for pre-search simplification. The algorithm in Autometrics shares these characteristics with the algorithm in PcGets; hence, many of the remarks above about PcGets apply directly to Autometrics. However, Autometrics (unlike PcGets) uses a tree search method, with refinements on pre-search simplification and on the objective function. See Doornik and Hendry (2007) and Doornik (2008) for details.

5.2 Modeling of Argentine Money Demand Revisited

Using PcGets and Autometrics, the current subsection assesses the possible path dependence of equation (4). The initial general model is estimated; and the algorithms simplify that general model under each of the 24 permutations implied by the list of choices below. While the algorithms do obtain multiple distinct final models, equation (4)—or simple variants of it—appears statistically sensible; and one variant obtained by Autometrics is even more parsimonious than (4). These results bolster the model design in Kamin and Ericsson (1993) and offer an improvement on it.

The multi-path searches in PcGets and Autometrics allow investigation of equation (4)'s robustness and examination of the empirical properties of the two algorithms themselves. In addition, four choices within the model selection process permit further insights. In PcGive, these choices concern the following.

1. Model strategy: either liberal ("L") or conservative ("C").
2. Pre-search testing (Stage 0): either switched on ("Yes") or off ("No").
3. The representation of the initial general ECM (three options): either the representation as tabulated, or either of two representations that explicitly nest equation (4). The latter two representations are distinguished by whether the variables from (4) are "free" or "fixed". Fixed variables are forced to always be included in regression, whereas free variables may be deleted by the algorithm.
4. The choice of the general model: either the unrestricted ADL, or the *intermediate* ADL described in Section 3.

For model strategy (choice #1), the options in Autometrics do not correspond precisely to PcGets's liberal and conservative strategies. Instead, Autometrics allows the user to select a "target size", which is meant to equal "the proportion of irrelevant variables that survives the [simplification] process" (Doornik, 2008). In the analysis below, Autometrics's target size is either 5% or 1%, which appear to approximate liberal and conservative strategies in PcGets. For pre-search testing (choice #2), the selected option in Autometrics is either pre-search for both variable reduction and lag reduction, or no pre-search for either—in order to match PcGets as closely as possible. The third choice above is identical for PcGets and Autometrics, as is the fourth choice.

For both PcGets and Autometrics, the third choice (the representation of the initial general ECM) can affect the final model selected. In simplifying the initial model, PcGets and Autometrics impose only "zero restrictions", i.e., the algorithms can set coefficients to be equal only to zero. Although a linear model is invariant to nonsingular linear transformations of its data, the coefficients of that model are *not* invariant to such transformations. For example, a model with regressors x_t and x_{t-1} is invariant to including the regressors Δx_t and x_{t-1} instead; but the deletion of x_{t-1}

results in two different simplifications, depending on the representation. See Campos and Ericsson (1999) for additional discussion.

Table 4 lists the estimates and standard errors for the ECM representation of the unrestricted seventh-order ADL model of $m-p$, Δp, R, Δp^{max}, Δe, and $\Delta^2 p^{pos}$. The standard diagnostic statistics do not reject. The implied coefficient on the error correction term appears to be highly significant statistically, with a t-ratio of -3.53. The intermediate ADL (in ECM representation) is Table 4, but re-estimated with the "boxed-in" coefficients in Table 4 set to zero. Kamin and Ericsson (1993) show that the estimated coefficients in this intermediate model are close to those in the unrestricted ECM in Table 4; and the intermediate ADL is a statistically acceptable reduction of Table 4, with $Inn : F(33, 117) = 1.42$ [0.091]. For ease of reference, the intermediate ADL is denoted Table 4*.

Table 5 summarizes PcGets's model simplifications under the 24 different scenarios described above; Table 6 does likewise for Autometrics. In these tables, k_1 is the number of regressors in the general model for multi-path searches, k_f is the number of coefficients in the final specific model for multi-path searches, the "number of paths" is the number of different simplification paths considered in a multi-path search, the "number of terminal models" is the number of distinct terminal specifications after a multi-path search, and $\hat{\sigma}$ is the residual standard error of the final specific model. If multi-path searches are iterated, the table lists values for each iteration, where appropriate. The "number of models estimated" is the total number of distinct models estimated in the multi-path search.

Several features of the simplifications in Tables 5 and 6 are notable. First, pre-search testing typically reduces the number of paths that need to be searched in Stage 1, and often markedly so. As a consequence, pre-search testing frequently reduces the number of multiple terminal models and, in some instances, obtains the final model. Second, if the initial general model is the intermediate ECM (Table 4*, rather than the general ECM in Table 4), that choice is in effect a pre-search, albeit an informal one. That choice also typically obtains a single terminal model on the initial multi-path search. Third, a conservative strategy generally obtains a more parsimonious model than a liberal strategy, as expected. Fourth, Kamin and Ericsson's (1993) model results from a conservative-like strategy, as is apparent from examining the specifications of the final models in Tables 5 and 6. Fifth, the 1% and 5% target sizes in Autometrics appear closely comparable to the liberal and conservative strategies in PcGets. That said, in several instances, Autometrics dominates PcGets by obtaining a more parsimonious model with a better fit (in terms of $\hat{\sigma}$), whereas PcGets never dominates Autometrics in that sense. This outcome reflects differences in the algorithms' details. Finally, data transformations through the "nesting" approach permit a final representation that is more highly parsimonious than previously

Table 4: An unrestricted error correction representation for real money conditional on inflation, the interest rate, and the change in the exchange rate.

Variable[a,b,c]	Lag j							
	0	1	2	3	4	5	6	7
$\Delta(m-p)_{t-j}$	-1 $(-)$	0.270 (0.088)	0.154 (0.093)	-0.019 (0.094)	0.029 (0.079)	0.204 (0.079)	-0.125 (0.072)	
Δp_{t-j}	-0.768 (0.101)	0.177 (0.149)	-0.041 (0.140)	0.144 (0.136)	0.121 (0.117)	0.168 (0.119)	-0.275 (0.122)	-0.053 (0.079)
ΔR_{t-j}	0.222 (0.053)	0.031 (0.159)	0.172 (0.149)	0.140 (0.143)	-0.039 (0.128)	0.002 (0.096)	-0.131 (0.090)	
$\Delta(\Delta p^{max}_{t-j})$	-0.223 (0.155)	-0.169 (0.184)	0.327 (0.281)	0.599 (0.297)	-0.104 (0.268)	-0.211 (0.264)	0.132 (0.317)	
$\Delta^2 p^{pos}_{t-j}$	-0.406 (0.162)	-0.049 (0.154)	0.180 (0.161)	-0.114 (0.156)	-0.168 (0.133)	0.003 (0.137)	0.181 (0.130)	-0.017 (0.112)
Δe_{t-j}	-0.004 (0.020)	-0.040 (0.022)	-0.014 (0.023)	0.025 (0.022)	-0.040 (0.022)	0.011 (0.023)	0.032 (0.023)	-0.032 (0.022)
$(m-p)_{t-j}$		-0.053 (0.015)						
R_{t-j}		0.441 (0.158)						
Δp^{max}_{t-j}		-0.055 (0.019)						
B_{t-j}	-0.353 (0.139)	0.178 (0.078)	0.058 (0.073)	0.286 (0.084)				
S_{t-j}	0.160 (0.047)		-1.62 (1.20)	-0.08 (1.21)	0.41 (1.10)	-0.21 (1.16)	2.81 (1.02)	
S_{t-j-6}		0.06 (0.97)	0.10 (1.03)	-0.55 (1.01)	0.86 (1.03)	0.20 (1.02)	4.35 (1.02)	

$T = 180$ [1978(2)–1993(1)] $\quad R^2 = 0.968 \quad \hat{\sigma} = 2.058\%$

$dw = 2.02 \quad AR : F(7,110) = 0.67$ [0.695] $\quad LM_p : F(1,116) = 0.08$ [0.773][d]

$ARCH : F(7,103) = 0.92$ [0.491] $\quad Normality : \chi^2(2) = 0.77$ [0.682]

$Hetero : F(109,7) = 0.08$ [1.000] $\quad RESET : F(1,116) = 1.89$ [0.171]

Notes:

a. The dependent variable is $\Delta(m-p)_t$. Even so, the equation is in levels, not in differences, noting the inclusion of the regressor $(m-p)_{t-1}$.

b. The variables $\{S_{t-i}\}$ are the seasonal dummies, except that S_0 is the intercept. February is S_{t-2}, March is S_{t-3}, etc. For readability, the coefficients and estimated standard errors for the seasonal dummies have been multiplied by 100.

c. The 33 coefficients that are "boxed in" are set equal to zero in the partially restricted intermediate error correction representation denoted Table 4*.

d. The statistic LM_p is the Lagrange multiplier statistic for testing the imposed restriction of long-run price homogeneity.

Table 5: Statistics on computer-automated model selection by PcGets of models for Argentine money demand, categorized according to model strategy, pre-search testing, representation of the general model, and choice of general model.

Model strategy	Pre-search?	Representation?	k_1	k_f	Number of paths	Number of terminal models	$\hat{\sigma}$ (%)
		The general model is Table 4 or equivalent.					
L	No	Table 4	63, 31	19	55, 7	13, 1	2.132
L	No	Nested	63, 31, 26	24	56, 18, 9	9, 3, 5	1.952
L	No	Fixed	63, 25	22	53, 10	6, 3	1.954
L	Yes	Table 4	26	21	10	3	2.078
L	Yes	Nested	28	23	11	3	1.989
L	Yes	Fixed	23	22	6	2	1.986
C	No	Table 4	63, 32, 31	23	61, 21, 20	10, 5, 4	2.015
C	No	Nested	63, 31	21	61, 22	9, 5	2.007
C	No	Fixed	63, 24	21	55, 9	6, 3	1.988
C	Yes	Table 4	21	21	1	1	2.139
C	Yes	Nested	22, 21	21	10, 8	2, 1	2.073
C	Yes	Fixed	24, 18	18	11, 3	1, 1	2.137
		The general model is Table 4* or equivalent.					
L	No	Table 4*	30	20	18	1	2.149
L	No	Nested	30	18	20	1	2.137
L	No	Fixed	30	18	20	1	2.137
L	Yes	Table 4*	20	20	1	1	2.149
L	Yes	Nested	18	18	1	1	2.137
L	Yes	Fixed	18	18	1	1	2.137
C	No	Table 4*	30, 20	20	19, 3	1, 1	2.149
C	No	Nested	30, 18	18	22, 3	2, 1	2.137
C	No	Fixed	30, 18	18	22, 3	2, 1	2.137
C	Yes	Table 4*	20, 20	20	3, 3	1, 1	2.149
C	Yes	Nested	18, 18	18	3, 3	1, 1	2.137
C	Yes	Fixed	18, 18	18	3, 3	1, 1	2.137

Table 6: Statistics on computer-automated model selection by Autometrics of models for Argentine money demand, categorized according to target size, pre-search testing, representation of the general model, and choice of general model.

Target size	Pre-search?	Rep-resen-tation?	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
			The general model is Table 4 or equivalent.				
5%	No	Table 4	63, 41	23	706	10, 17	1.997
5%	No	Nested	63, 36	21	378	8, 13	1.978
5%	No	Fixed	63, 31	20	306	6, 7	2.003
5%	Yes	Table 4	57, 37	22	470	6, 12	2.008
5%	Yes	Nested	50, 32	22	371	6, 7	1.972
5%	Yes	Fixed	45, 30	22	255	8, 8	1.986
1%	No	Table 4	63, 37	19	751	10, 20	2.095
1%	No	Nested	63, 29	21	501	8, 10	1.978
1%	No	Fixed	63, 22	18	497	2, 2	2.096
1%	Yes	Table 4	41, 35	20	677	10, 20	2.072
1%	Yes	Nested	39, 26	20	394	5, 5	2.014
1%	Yes	Fixed	30, 23	19	168	4, 4	2.078
			The general model is Table 4* or equivalent.				
5%	No	Table 4*	30, 20	20	46	1, 1	2.149
5%	No	Nested	30, 18	18	65	1, 1	2.137
5%	No	Fixed	30, 18	18	65	1, 1	2.137
5%	Yes	Table 4*	25, 20	20	36	1, 1	2.149
5%	Yes	Nested	23, 18	18	40	1, 1	2.137
5%	Yes	Fixed	23, 18	18	43	1, 1	2.137
1%	No	Table 4*	30, 18	18	77	1, 1	2.211
1%	No	Nested	30, 14	14	139	1, 1	2.236
1%	No	Fixed	30, 16	16	2	1	2.192
1%	Yes	Table 4*	21, 20	20	30	1, 1	2.149
1%	Yes	Nested	19, 17	17	36	1, 1	2.168
1%	Yes	Fixed	19, 17	17	38	1, 1	2.168

obtained; see the boxed-in result for Autometrics on Table 6.

The corresponding model, which improves on equation (4), is as follows.

$$
\begin{aligned}
\widehat{\Delta(m-p)_t} = \; & \underset{\substack{(0.025)\\ [0.024]}}{0.281} \Delta(m-p)_{t-1} - \underset{\substack{(0.041)\\ [0.040]}}{0.759} \Delta^2 p_t \\
& - \underset{\substack{(0.078)\\ [0.090]}}{0.564} \Delta^2 p_t^{pos} + \underset{\substack{(0.017)\\ [0.022]}}{0.040} \Delta\Delta_6 p_t \\
& + \underset{\substack{(0.022)\\ [0.019]}}{0.180} \Delta^2 R_t + \underset{\substack{(0.044)\\ [0.041]}}{0.543} (R-\Delta p)_{t-1} \\
& + \underset{\substack{(0.022)\\ [0.025]}}{0.093} - \underset{\substack{(0.0078)\\ [0.0088]}}{0.0300} (m-p)_{t-1} - \underset{\substack{(0.018)\\ [0.019]}}{0.060} \Delta e_{t-1} \\
& - \underset{\substack{(0.010)\\ [0.011]}}{0.025} \Delta p_{t-1}^{max} - \underset{\substack{(0.034)\\ [0.030]}}{0.253} B_t + \underset{\substack{(0.032)\\ [0.023]}}{0.170} B_{t-3} \\
& + \underset{\substack{(0.62)\\ [0.37]}}{1.97} S_{t-6} + \underset{\substack{(0.62)\\ [0.74]}}{4.78} S_{t-12} \qquad (7)
\end{aligned}
$$

$T = 180$ [1978(2)–1993(1)] $\quad \mathrm{R}^2 = 0.9462 \quad \hat{\sigma} = 2.236\% \quad dw = 2.10$
$Inn_3 : F(49, 117) = 1.61^* \quad Inn_4 : F(16, 150) = 1.85^* \quad AR : F(7, 159) = 1.65$
$ARCH : F(7, 152) = 2.73^* \quad Normality : \chi^2(2) = 0.44 \quad RESET : F(1, 165) = 1.60$
$Hetero : F(22, 143) = 1.08 \quad Form : F(75, 90) = 0.92 \quad Chow : F(33, 133) = 0.86$

The coefficients in equation (7) are little changed from the corresponding ones in equation (4), except that the coefficients for $\Delta^2(m-p)_{t-5}$ and $\Delta^2 p_{t-5}$ are restricted to be zero. No tests reject at the 1% level (an implication of choices made in the algorithm's parameters), although some do at the 5% level. Equation (7) has virtually the same economic interpretation as equation (4), and it is more parsimonious than (4). PcGets and Autometrics thus verify the robustness of equation (4)'s specification, and Autometrics improves upon that specification.

6 Conclusions

Computer-automated model selection with the software packages PcGets and Autometrics demonstrates the robustness of Kamin and Ericsson's (1993) final error correction model and improves on it by using multi-path searches that would be tedious and prohibitively time-consuming with standard econometrics packages. Long-run money demand is driven by a negative ratchet effect from inflation, and by the opportunity cost of holding peso-denominated financial assets rather than Argentine goods or U.S. dollars. Short-run dynamics are consistent with an Ss-type inventory model that is interpretable as having either real or nominal short-run bounds.

Several general remarks are germane, and each suggests extensions to the current analysis. First, improvements to the model selection algorithms may and do obtain an improved model specification. Computer-automated model-selection algorithms are still in their youth—if not in their infancy—and considerable analytical, Monte Carlo, and empirical research is ongoing; see Hendry and Krolzig (1999, 2003, 2005), Krolzig and Hendry (2001), Hoover and Perez (2004), Doornik (2008), Hendry, Johansen, and Santos (2008), Hoover, Demiralp, and Perez (2008), Hoover, Johansen, and Juselius (2008), and Johansen and Nielsen (2008).

Second, insights by other researchers may improve the current model in a progressive research strategy. For example, Nielsen (2004), building on Hendry and von Ungern-Sternberg (1981), proposes an alternative measure of the opportunity cost of holding money that may better capture agents' behavior in a hyperinflationary environment. Preliminary tests for that alternative measure as an omitted variable in Table 4 do not reveal an improved specification, however. For instance, for a variable X in levels, define ∇X_t as $(X_t - X_{t-1})/(X_{t-1})$, which is X's percentage change, measured as a fraction. Omitted variables tests include $F(8, 109) = 1.23$ [0.290] for $\{\nabla P_{t-i};\ i = 0, \ldots, 7\}$, $F(24, 93) = 1.10$ [0.363] for $\{\nabla P_{t-i},\ \nabla P^{max}_{t-i},\ \Delta\nabla P^{pos}_{t-i};\ i = 0, \ldots, 7\}$, and $F(32, 85) = 1.01$ [0.474] for $\{\nabla P_{t-i}$, ∇P^{max}_{t-i}, $\Delta\nabla P^{pos}_{t-i}$, $\nabla E_{t-i};\ i = 0, \ldots, 7\}$. None of these tests reject at standard levels. Still, Table 4 is a relatively unrestricted model, so further investigation is merited, particularly because Δp_t differs substantially from ∇P_t at high inflation rates and hence the interpretation of Δp_t may be affected.

Third, Kongsted (2005) develops a procedure for testing the nominal-to-real transformation, which is only informally investigated herein for money by using the ADF statistics. Fourth, in the VAR, the variables Δp^{max} and $\Delta^2 p^{pos}$ are transformations of Δp, so further consideration of their joint distributional properties is desirable. Fifth, data observations after 1993 may be informative. Even so, mechanistic extensions of the existing data may not be sufficient, as when data definitions change, the array of available assets alters, and underlying economic conditions shift; see Ericsson, Hendry, and Prestwich (1998).

References

Ahumada, H. (1992) "A Dynamic Model of the Demand for Currency: Argentina 1977–1988", *Journal of Policy Modeling*, 14, 3, 335–361.

Akaike, H. (1973) "Information Theory and an Extension of the Maximum Likelihood Principle", in B. N. Petrov and F. Csáki (eds.) *Second International Symposium on Information Theory*, Akadémiai Kiadó, Budapest, 267–281.

Akaike, H. (1981) "Likelihood of a Model and Information Criteria", *Journal of Econometrics*, 16, 1, 3–14.

Atkinson, A. C. (1981) "Likelihood Ratios, Posterior Odds and Information Criteria", *Journal of Econometrics*, 16, 1, 15–20.

Baba, Y., D. F. Hendry, and R. M. Starr (1992) "The Demand for M1 in the U.S.A., 1960–1988", *Review of Economic Studies*, 59, 1, 25–61.

Beckerman, P. (1992) *The Economics of High Inflation*, St. Martin's Press, New York.

Cagan, P. (1956) "The Monetary Dynamics of Hyperinflation", Chapter 2 in M. Friedman (ed.) *Studies in the Quantity Theory of Money*, University of Chicago Press, Chicago, 23–117.

Campos, J., and N. R. Ericsson (1999) "Constructive Data Mining: Modeling Consumers' Expenditure in Venezuela", *Econometrics Journal*, 2, 2, 226–240.

Dominguez, K. M. E., and L. L. Tesar (2007) "International Borrowing and Macroeconomic Performance in Argentina", Chapter 7 in S. Edwards (ed.) *Capital Controls and Capital Flows in Emerging Economies: Policies, Practices, and Consequences*, University of Chicago Press, Chicago, 297–342 (with discussion).

Doornik, J. A. (2008) "Autometrics", in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, this volume.

Doornik, J. A., and D. F. Hendry (2007) *PcGive 12*, Timberlake Consultants Ltd, London (4 volumes).

Doornik, J. A., D. F. Hendry, and B. Nielsen (1998) "Inference in Cointegrating Models: UK M1 Revisited", *Journal of Economic Surveys*, 12, 5, 533–572.

Engle, R. F., and D. F. Hendry (1993) "Testing Super Exogeneity and Invariance in Regression Models", *Journal of Econometrics*, 56, 1/2, 119–139.

Enzler, J., L. Johnson, and J. Paulus (1976) "Some Problems of Money Demand", *Brookings Papers on Economic Activity*, 1976, 1, 261–280 (with discussion).

Ericsson, N. R. (2004) "The *ET* Interview: Professor David F. Hendry", *Econometric Theory*, 20, 4, 743–804.

Ericsson, N. R. (2008) *Empirical Modeling of Economic Time Series*, in preparation.

Ericsson, N. R., D. F. Hendry, and K. M. Prestwich (1998) "The Demand for Broad Money in the United Kingdom, 1878–1993", *Scandinavian Journal of Economics*, 100, 1, 289–324 (with discussion).

Flemming, J. S. (1976) *Inflation*, Oxford University Press, Oxford.

Hannan, E. J., and B. G. Quinn (1979) "The Determination of the Order of an Autoregression", *Journal of the Royal Statistical Society, Series B*, 41, 2, 190–195.

Helkie, W. L., and D. H. Howard (1994) "External Adjustment in Selected Developing Countries in the 1990s", *Journal of Policy Modeling*, 16, 4, 353–393.

Hendry, D. F. (1995) "On the Interactions of Unit Roots and Exogeneity", *Econometric Reviews*, 14, 4, 383–419.

Hendry, D. F. (2006) "Robustifying Forecasts from Equilibrium-correction Systems", *Journal of Econometrics*, 135, 1–2, 399–426.

Hendry, D. F., and J. A. Doornik (1994) "Modelling Linear Dynamic Econometric Systems", *Scottish Journal of Political Economy*, 41, 1, 1–33.

Hendry, D. F., and N. R. Ericsson (1991) "Modeling the Demand for Narrow Money in the United Kingdom and the United States", *European Economic Review*, 35, 4, 833–881 (with discussion).

Hendry, D. F., S. Johansen, and C. Santos (2008) "Automatic Selection of Indicators in a Fully Saturated Regression", *Computational Statistics*, 23, 2, 317–335, 337–339.

Hendry, D. F., and H.-M. Krolzig (1999) "Improving on 'Data Mining Reconsidered' by K. D. Hoover and S. J. Perez", *Econometrics Journal*, 2, 2, 202–219.

Hendry, D. F., and H.-M. Krolzig (2001) *Automatic Econometric Model Selection Using PcGets 1.0*, Timberlake Consultants Press, London.

Hendry, D. F., and H.-M. Krolzig (2003) "New Developments in Automatic General-to-Specific Modeling", Chapter 16 in B. P. Stigum (ed.) *Econometrics and the Philosophy of Economics: Theory-Data Confrontations in Economics*, Princeton University Press, Princeton, 379–419.

Hendry, D. F., and H.-M. Krolzig (2005) "The Properties of Automatic *Gets* Modelling", *Economic Journal*, 115, 502, C32–C61.

Hendry, D. F., and G. E. Mizon (1978) "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England", *Economic Journal*, 88, 351, 549–563.

Hendry, D. F., and T. von Ungern-Sternberg (1981) "Liquidity and Inflation Effects on Consumers' Expenditure", Chapter 9 in A. S. Deaton (ed.) *Essays in the Theory and Measurement of Consumer Behaviour: In Honour of Sir Richard Stone*, Cambridge University Press, Cambridge, 237–260.

Hoover, K. D., S. Demiralp, and S. J. Perez (2008) "Empirical Identification of the Vector Autoregression: The Causes and Effects of U.S. M2", in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, this volume.

Hoover, K. D., S. Johansen, and K. Juselius (2008) "Allowing the Data to Speak Freely: The Macroeconometrics of the Cointegrated Vector Autoregression", *American Economic Review*, 98, 2, 251–255.

Hoover, K. D., and S. J. Perez (1999) "Data Mining Reconsidered: Encompassing and the General-to-specific Approach to Specification Search", *Econometrics Journal*, 2, 2, 167–191 (with discussion).

Hoover, K. D., and S. J. Perez (2004) "Truth and Robustness in Cross-country Growth Regressions", *Oxford Bulletin of Economics and Statistics*, 66, 5, 765–798.

Howard, D. H. (1987) "Exchange Rate Regimes and Macroeconomic Stabilization in a Developing Country", International Finance Discussion Paper No. 314, Board of Governors of the Federal Reserve System, Washington, D.C., November.

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 2/3, 231–254.

Johansen, S. (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 59, 6, 1551–1580.

Johansen, S. (1992a) "Cointegration in Partial Systems and the Efficiency of Single-equation Analysis", *Journal of Econometrics*, 52, 3, 389–402.

Johansen, S. (1992b) "Testing Weak Exogeneity and the Order of Cointegration in UK Money Demand Data", *Journal of Policy Modeling*, 14, 3, 313–334.

Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S., and B. Nielsen (2008) "An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator", in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, this volume.

Juselius, K. (1992) "Domestic and Foreign Effects on Prices in an Open Economy: The Case of Denmark", *Journal of Policy Modeling*, 14, 4, 401–428.

Kamin, S. B. (1991) "Argentina's Experience with Parallel Exchange Markets: 1981–1990", International Finance Discussion Paper No. 407, Board of Governors of the Federal Reserve System, Washington, D.C., August.

Kamin, S. B., and N. R. Ericsson (1993) "Dollarization in Argentina", International Finance Discussion Paper No. 460, Board of Governors of the Federal Reserve System, Washington, D.C., November.

Kamin, S. B., and N. R. Ericsson (2003) "Dollarization in Post-hyperinflationary Argentina", *Journal of International Money and Finance*, 22, 2, 185–211.

Kiguel, M. A. (1991) "Inflation in Argentina: Stop and Go Since the Austral Plan", *World Development*, 19, 8, 969–986.

Kongsted, H. C. (2005) "Testing the Nominal-to-real Transformation", *Journal of Econometrics*, 124, 2, 205–225.

Krolzig, H.-M., and D. F. Hendry (2001) "Computer Automation of General-to-specific Model Selection Procedures", *Journal of Economic Dynamics and Control*, 25, 6–7, 831–866.

MacKinnon, J. G. (1991) "Critical Values for Cointegration Tests", Chapter 13 in R. F. Engle and C. W. J. Granger (eds.) *Long-run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford, 267–276.

Manzetti, L. (1991) *The International Monetary Fund and Economic Stabilization: The Argentine Case*, Praeger, New York.

Melnick, R. (1990) "The Demand for Money in Argentina 1978–1987: Before and After the Austral Program", *Journal of Business and Economic Statistics*, 8, 4, 427–434.

Nielsen, B. (2004) "Money Demand in the Yugoslavian Hyperinflation 1991–1994", Economics Working Paper No. 2004-W31, Nuffield College, University of Oxford, Oxford, December.

Piterman, S. (1988) "The Irreversibility of the Relationship Between Inflation and Real Balances", *Bank of Israel Economic Review*, 60, January, 72–83.

Schwarz, G. (1978) "Estimating the Dimension of a Model", *Annals of Statistics*, 6, 2, 461–464.

Simpson, T. D., and R. D. Porter (1980) "Some Issues Involving the Definition and Interpretation of the Monetary Aggregates", in Federal Reserve Bank of Boston (ed.) *Controlling Monetary Aggregates III*, Federal Reserve Bank of Boston, Boston, *Conference Series No. 23*, 161–234 (with discussion).

Taylor, M. P. (1986) "From the General to the Specific: The Demand for M2 in Three European Countries", *Empirical Economics*, 11, 4, 243–261.

Uribe, M. (1997) "Hysteresis in a Simple Model of Currency Substitution", *Journal of Monetary Economics*, 40, 1, 185–202.

World Bank (1990) *Argentina: Reforms for Price Stability and Growth*, World Bank, Washington, D.C.