EMPIRICAL MODEL SELECTION: FRIEDMAN AND SCHWARTZ REVISITED

Neil R. Ericsson*

March 16, 2010

Abstract: Using annual data from Friedman and Schwartz (1982), Hendry and Ericsson (1991a) developed an empirical model of the demand for broad money in the United Kingdom over 1878–1970. Model selection was a central issue in assessing the merits of their model, so this paper re-evaluates that model with Autometrics, a recent third-generation algorithm for computer-automated model selection. Hendry and Ericsson's model is remarkably robust to the model selection path, as characterized through variations in the algorithm's settings for target size, pre-search testing, fixity of regressors, indicator saturation, representation of the general model, and choice of dependent variable. This paper also assesses the empirical merits of Autometrics, using it to improve upon Hendry and Ericsson's (1991a) model.

Keywords: Autometrics, broad money, dynamic specification, cointegration, conditional models, error correction, Friedman and Schwartz, model design, model selection, money demand, United Kingdom.

JEL classifications: C52, E41.

Neil R. Ericsson
Stop 15, Division of International Finance, Federal Reserve Board
2000 C Street, N.W., Washington, D.C. 20551 U.S.A.
(1) (202) 452–3709 (office)
(1) (202) 263–4850 (fax)
ericsson@frb.gov (email)
www.federalreserve.gov/research/staff/ericssonneilr.htm (webpage)

*The author is a staff economist in the Division of International Finance, Board of Governors of the Federal Reserve System, Washington, D.C. 20551 U.S.A., and is currently on leave to the Department of Economics, The George Washington University, 2115 G Street, N.W., Washington, D.C. 20052 U.S.A. He may be reached on the Internet at ericsson@frb.gov and ericsson@gwu.edu. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. The author is grateful to David Hendry for helpful discussions. All numerical results were obtained using PcGive Versions 13.0 and 13.1 and Autometrics Version 1.5e in OxMetrics Versions 6.0 and 6.10: see Doornik and Hendry (2007) and Doornik (2009).

1 Introduction

Friedman and Schwartz (1982) and Hendry and Ericsson (1991a) developed alternative empirical models of the demand for broad money in the United Kingdom on data spanning 1878–1970. Empirical model selection was a central issue in assessing the relative merits of those models, as reflected by the title to Friedman and Schwartz's (1991) response to Hendry and Ericsson (1991a): "Alternative Approaches to Analyzing Economic Data". To assess the role of model selection in Hendry and Ericsson (1991a), the current paper re-evaluates Hendry and Ericsson's model with Autometrics, Jurgen Doornik and David Hendry's (2007) third-generation algorithm for computer-automated model selection. Hendry and Ericsson's model is remarkably robust to the model selection path, as characterized through variations in the algorithm's settings for target size, pre-search testing, fixity of regressors, indicator saturation, representation of the general model, and choice of dependent variable.

This paper also assesses the empirical merits of Autometrics, using it to improve on Hendry and Ericsson's (1991a) model. The selected model is an economically sensible and statistically satisfactory error correction model, in which cointegration between money, prices, income, and the short-term interest rate depends on nonlinear dynamics, both through nonlinear transformation of the variables themselves and through a nonlinear error correction term. 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.

This paper is organized as follows. Section 2 briefly describes the economic theory and the data. Section 3 summarizes the model obtained by Hendry and Ericsson (1991a). Section 4 describes the model selection algorithm in Autometrics and the options considered in model selection. Section 5 then applies Autometrics to multiple representations of a general unrestricted model for U.K. broad money demand, verifies the robustness of Hendry and Ericsson's (1991a) model to the model selection process, and obtains a more parsimonious data-coherent representation; and it investigates the nonlinearities implied by the empirical model. Section 6 concludes.

2 Economic Theory and the Data

This section first summarizes 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:

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

where M^d is nominal money demanded, P is the price level, I 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 I, 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.

Empirical models below employ equation (1) in its standard log-linear form, with interest rates entering in levels and $q(\cdot, \cdot)$ homogeneous of degree one in income:

$$m - p - i \quad = \quad a_0 \; + \; a_1' R \; . \tag{2}$$

Capital letters denote both the generic name and the level; logs are in lowercase. The anticipated sign of a coefficient in the vector a_1 is positive if the associated asset is within M, and negative if the associated asset is outside M. See Friedman and Schwartz (1982), Hendry and Ericsson (1991a), and Ericsson, Hendry, and Prestwich (1998a) for further discussion in the context of Friedman and Schwartz's dataset.

2.2 The Data

This subsection describes the data available and considers some of their basic properties.

Table 1 lists primary series from Friedman and Schwartz (1982) for the United Kingdom: the broad measure of money M2 (denoted M, and defined as notes and coin plus checking and savings accounts), income (I, real net national income), prices (P, the deflator for net national income), a short-term interest rate (RS), and a long-term interest rate (RL). The variables D_1 and D_3 are dummies for World War I and World War II respectively. The variable D_2 is a data-based dummy for "an upward demand shift [for money during 1921–1955], produced by economic depression and war"; see Friedman and Schwartz (1982, pp. 228, 281). Under the quantity theory of money, the income elasticity is unity, so a key derived variable is velocity V, with the log of inverse velocity being $m - p - i$, which is the left-hand side variable in equation (2).

Friedman and Schwartz (1982, Table 4.9) give a full listing of the data, which are annual over 1871–1975. In order to match results in Hendry and Ericsson (1991a), estimation is over 1878–1970 ($T = 93$) unless otherwise noted. Details on the data's sources, construction, and caveats appear in Friedman and Schwartz (1982, Chapters 3–5) and Hendry and Ericsson (1991a, Section I and Data Appendix). Further analysis appears in Escribano (1985, 2004), Campos, Ericsson, and Hendry (1990),

Table 1: Friedman and Schwartz's (1982) U.K. money demand dataset.

Notation	Definition	Units
	Raw data	
M	Nominal broad money stock	£ million
I	Real net national income (NNI)	1929 £ million
P	NNI deflator	1929 = 100
RS	Short-term interest rate	percentage (annual rate)
RL	Long-term interest rate	percentage (annual rate)
	Transformed and constructed data	
D_1	Dummy variable for World War I (1914–1918)	0 or 1
D_2	Liquidity shift dummy (1921–1955)	0 or 1
D_3	Dummy variable for World War II (1939–1945)	0 or 1
D_5	$D_1 + D_3$	0 or 1
V	Velocity of money ($= (I \cdot P)/M$)	–

Hendry and Ericsson (1991a), Ericsson, Hendry, and Prestwich (1998a, 1998b), and Teräsvirta and Eliasson (2001).

Figure 1 shows the time series of (m, p), $(m - p, i)$, (v, RS), and $(\Delta m, \Delta p)$ as a 2×2 panel: p is adjusted for its mean in the first graph; i is adjusted for its mean in the second graph; RS is adjusted for its mean and range in the third graph; and uppercase delta Δ is the difference operator.[1] Figures typically appear as 2×2 panels of graphs, with each graph labeled by a suffix a, b, c, or d, as follows: $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$.

Figure 1*a* emphasizes the huge changes in money and prices over the century, noting that the data are in logs. From Figure 1*b*, real money and income increase roughly tenfold over the sample. Figure 1*c* shows marked variation in velocity, with some co-movement by RS. Notabaly, the short-term interest rate is less than 1% for over half the time during the 1930s and 1940s, presaging current discussion on a "zero lower bound"; cf. Bernanke (2009) and Kohn (2009). Figure 1*d* plots the

[1] The difference operator Δ is defined as $(1 - L)$, where the lag operator L shifts a variable one period into the past. Hence, for x_t (a variable x at time t), $Lx_t = x_{t-1}$ and so $\Delta x_t = x_t - x_{t-1}$. More generally, $\Delta_j^i x_t = (1 - L^j)^i x_t$ for positive integers i and j. If i or j is not explicit, it is taken to be unity.

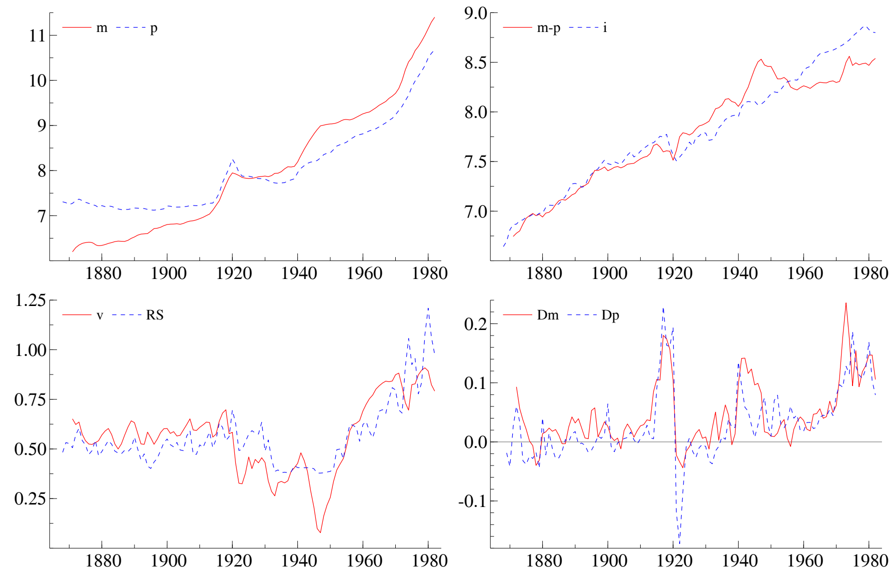

Figure 1: The logarithms of nominal money and prices (m and p), and of real money and income ($m - p$ and i); the logarithm of velocity and the short-run interest rate (v and RS); and nominal money growth and inflation (Δm and Δp).

growth rate of nominal money and the inflation rate, which move relatively closely overall. There is one notable departure between them: 1920–1922, when prices fall by 25% but nominal money by only 5%. Ericsson, Hendry, and Prestwich (1998a) further characterizes these series in terms of their univariate time series properties.

3 Previous Results

Friedman and Schwartz (1982) and Hendry and Ericsson (1991a) developed alternative empirical models of the demand for broad money in the United Kingdom on data spanning 1878–1970. As shown in Hendry and Ericsson (1991a), Friedman and Schwartz's model appears mis-specified on several accounts, revealing the potential for an improved model. Using recursive procedures on the annual data, Hendry and Ericsson (1991a) obtain a better-fitting, constant, dynamic error correction model. Results on exogeneity and encompassing imply that Hendry and Ericsson's money demand model is interpretable as a model of money but *not* of prices because its constancy holds only conditional on contemporaneous prices. This section summarizes that empirical model of U.K. broad money demand.

Hendry and Ericsson (1991a) test for and find cointegration between the log of inverse velocity ($m - p - i$) and the level of the short-run interest rate (RS). The estimated cointegrating relation is their equation (9):

$$\widehat{(m-p-i)}_t \quad = \quad -0.309 \quad - \quad 7.00 \ RS$$
$$T = 98 \ [1873\text{–}1970] \quad \mathrm{R}^2 = 0.56 \quad \hat{\sigma} = 10.86\% \ , \tag{3}$$

where a circumflex ˆ on the dependent variable denotes its fitted value, the subscript t is the time index, R^2 is the squared multiple correlation coefficient, and $\hat{\sigma}$ is the estimated residual standard error. Below, the residual from equation (3) is denoted $\hat{u}$.

Hendry and Ericsson (1991a, Table 3) then estimate a fifth-order ADL in m, p, i, rs, and rl, also including three dummy variables (D_{1t}, D_{2t}, D_{3t}) and the variables $\hat{u}_{t-1}$, $\hat{u}^2_{t-1}$, and $\hat{u}^3_{t-1}$, where the latter allow error correction to enter the ADL nonlinearly. That ADL is replicated in Appendix Appendix B. as Table B8. Hendry and Ericsson simplify that 36-coefficient ADL to a much more parsimonious 9-coefficient conditional error correction model (ECM), given by their equation (10):

$$\begin{aligned}
\Delta\widehat{(m-p)}_t \quad = \quad & \underset{(0.061)}{0.452} \ \Delta(m-p)_{t-1} - \underset{(0.041)}{0.101} \ \Delta^2(m-p)_{t-2} \\
& - \underset{(0.040)}{0.599} \ \Delta p_t + \underset{(0.046)}{0.394} \ \Delta p_{t-1} - \underset{(0.005)}{0.021} \ \Delta rs_t \\
& - \underset{(0.018)}{0.062} \ \Delta_2 rl_t - \underset{(0.550)}{2.545} \ (\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} \\
& + \underset{(0.0022)}{0.0054} + \underset{(0.545)}{3.736} \ (D_{1t} + D_{3t})
\end{aligned} \tag{4}$$

$T = 93$ [1878–1970)] $\quad \mathrm{R}^2 = 0.8704 \quad \hat{\sigma} = 1.4240\% \quad Inn : F(27, 57) = 0.50$
$AR : F(2, 82) = 1.39 \quad ARCH : F(1, 91) = 1.66 \quad Normality : \chi^2(2) = 1.93$
$RESET : F(2, 82) = 0.21 \quad Hetero : F(15, 77) = 0.98 \quad Form : F(36, 56) = 0.83.$

Hendry and Ericsson (1991a) show that equation (4) is statistically satisfactory and has a straightforward economic interpretation.

Statistically, 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 (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 the degrees of freedom fill the parentheses. Estimated standard errors are in parentheses $(\cdot)$, below coefficient estimates. See Doornik and Hendry (2007) for details and references.

Economically, the long-run coefficients in equation (3) satisfy sign restrictions that are consonant with a money demand function. The empirical parameterization in equation (4) exhibits multiple equilibria, with two corresponding to the long-run solution (3) and a third being the solution in (3) shifted by 20%. In certain sense, then, these results are consistent with Friedman and Schwartz's use of an adjustment factor of about that order. However, in equation (4), the dating of the disequilibria is determined by the values of $\hat{u}_{t-1}$ and operates over the entire sample, not just for the period 1921–1955, as in Friedman and Schwartz (1982).

The short-run variables and coefficients in equation (4) have a straightforward interpretation. The sizes of the short-run coefficients imply large immediate responses to changes in inflation and interest rates, but slow adjustment subsequently via the error correction term to remaining disequilibria. Inflation enters as $\Delta p_t + \Delta^2 p_t$ (approximately), which is a predictor of next period's inflation, optimal if prices vary quadratically. Thus, equation (4) has a forward-looking interpretation, albeit one based on *data functions* rather than on *models* of the right-hand side variables. That interpretation follows from the theoretical and empirical developments on such predictors in Flemming (1976), Hendry and Ericsson (1991b), and Campos and Ericsson (1999). The coefficients in (4) correspond to nearly orthogonal decision variables, consistent with equation (4) representing a contingent plan of agents who partition available information into conceptually separate entities.

In spite of the apparent robustness of equation (4), its design has three notable shortcomings. First, equation (4) may depend on the path taken in the limited (manual) model selection process undertaken by Hendry and Ericsson (1991a). Second, and equally, that model selection process may have missed better specifications. Third, equation (4) was estimated conditional on the value of the coefficients in equation (3), whereas joint estimation of the two equations may be preferable. The remainder of the current paper addresses these three issues.

4 Computer-automated Model Selection

This section describes the model selection algorithm in Autometrics (Section 4.1) and the options considered in model selection (Section 4.2).

4.1 The Algorithm in Autometrics

Hoover and Perez (1999) proposed an automated model-selection algorithm that incorporated many of the features of the "Hendry" or LSE methodology. Hendry and Krolzig (2001) developed a second-generation algorithm called PcGets, which extended and improved upon Hoover and Perez's 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 13. Autometrics utilizes one-step and multi-step simplifications along multiple paths following a tree search method. Diagnostic tests serve as additional checks on the simplified models, and encompassing tests 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, Autometrics functions 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 algorithm is eminently sensible, noting that the data generation process itself is congruent and is as parsimonious as feasible.

The current subsection summarizes Autometrics as an automated model-selection algorithm, thereby providing the necessary background for interpreting its application in Section 5. For ease of reference, the algorithm is divided into three "stages", denoted Stage 0, Stage 1, and Stage 2. For full details of Autometrics's algorithm, see Doornik and Hendry (2007) and Doornik (2009). 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, indicator saturation, and F pre-search tests. Stage 0 involves three parts: the estimation and evaluation of the general model, inclusion of impulse dummies for all observations, 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 to continue with the simplification procedure, perhaps ignoring the offending diagnostic statistic or statistics.

Second, and optionally, Autometrics performs block additions and searches of impulse dummies for all observations in a process known as indicator saturation (IS). Doing so generates a robust regression estimator, and it provides a check for parameter constancy. See Hendry, Johansen, and Santos (2008) and Johansen and Nielsen (2009) for recent developments.

Third, and also optionally, Autometrics attempts to drop various sets of potentially insignificant variables. Autometrics does so by dropping all variables at a given lag, starting with the longest lag. Autometrics 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, 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 Autometrics tries deleting those variables to obtain a simpler model. If a simplification is rejected, Autometrics 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 Autometrics obtains only one terminal model, then that model is the final model. However, because Autometrics pursues multiple simplification paths in Stage 1, Autometrics may obtain multiple terminal models. To resolve such a situation, Autometrics creates a union model from those terminal models and tests each terminal model against that union model. Autometrics 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 (possibly iteratively), by 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.

In short, Autometrics is general-to-specific, multi-path, iterative, and encompassing, with diagnostic tests providing additional assessments of statistical adequacy, and with options for pre-search simplification. Autometrics can be characterized as having 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) indicator saturation (Stage 0), and
   (c) multi-path (and possibly iterative) selection of the final model (Stages 1 and 2).

Table 2: Design options and choices in the model selection process.

Design options	**Choices**
1. Target size	5%
	1%
2. Pre-search testing (Stage 0)	Switched on ("Yes")
	Switched off ("No")
3. Fixity of regressors	All variables from (4) are fixed.
	Only the intercept is fixed.
	No variables are fixed.
4. Indicator Saturation (Stage 0)	Switched on ("Yes")
	Switched off ("No")
5. Nesting of equation (4)	Natural nesting
	Minimal nesting
	Implicit nesting
	Redundant nesting
6. Choice of the dependent variable	$\Delta(m-p)_t$
	Δm_t

Below, Section 5.1 summarizes the actual simplifications found by Autometrics in practice, thereby providing additional insight into Autometrics's algorithm.

4.2 Algorithm Options

By being multi-path, the searches in Autometrics allow the investigation of equation (4)'s robustness and the examination of the empirical properties of Autometrics's algorithm itself. In addition, choices for six "design" options within the model selection process permit further insights. This subsection discusses the nature of these choices.

Table 2 lists the six design options and, for each option, the choices considered. Some observations on these six options are helpful before proceeding to empirical model selection.

Target size. Autometrics requires the modeler to choose which tests are calculated and to specify the critical values for those tests. In principle, 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, Autometrics offers several options for the "target size", which incorporates pre-designated selections of test statistics and critical values.

The "target size" is meant to equal "the proportion of irrelevant variables that survives the [simplification] process" (Doornik, 2009, p. 100). In the analysis below, Autometrics's target size is either 5% or 1%, which are values that appear to approximate the liberal and conservative strategies in PcGets. 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, in part on the class of regressors examined (and, in particular, whether indicator saturation is considered), and in part on the objectives of the modeling exercise. Also, the two approaches *may* generate similar or identical results, as seen below.

Pre-search testing. For pre-search testing, the selected option is either pre-search for both variable reduction and lag reduction, or no pre-search for either.

Fixity of regressors. Autometrics offers the option of designating certain variables as "fixed". Fixed variables are forced to always be included in regression, whereas free variables (variables that are not fixed) may be deleted by the algorithm.

Indicator saturation. See Hendry, Johansen, and Santos (2008) and Johansen and Nielsen (2009) for discussion and recent developments.

The final two options (nesting, and the choice of the dependent variable) benefit from a digression on the nature of the general unrestricted model that serves as the starting point in the general-to-specific modeling by Autometrics. Each of the two options affects the representation of the initial general model and so can affect the final model selected. In simplifying the initial model, Autometrics imposes only "zero restrictions", i.e., the algorithm can set coefficients to be equal only to zero, and not to other values. 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.

Nesting of equation (4). The analysis in Section 5 below considers four possibilities for the nesting of equation (4) in the general unrestricted model. These possibilities are denoted "natural" nesting, "minimal" nesting, "implicit" nesting, and "redundant" nesting. As will be seen, the choice of nesting affects the representation of the general unrestricted model. By considering alternative representations, Section 5 expands the range of paths and the range of potential terminal models, which can help robustify results.

In *natural nesting*, the general unrestricted model is formulated as an ECM that nests equation (4) in a very straightforward and "natural" manner.

$$
\begin{aligned}
\Delta(m-p)_t = \; & \delta_1\Delta(m-p)_{t-1} + \delta_2\Delta^2(m-p)_{t-2} + \delta_3\Delta p_t + \delta_4\Delta p_{t-1} + \delta_5\Delta rs_t \\
& + \delta_6\Delta_2 rl_t + \delta_7(\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} + \delta_8 + \delta_9(D_{1t} + D_{3t}) \\
& + \textstyle\sum_{i=3}^{4}\gamma_{0i}\Delta(m-p)_{t-i} + \sum_{i=2}^{4}\gamma_{1i}\Delta p_{t-i} + \sum_{i=0}^{4}\gamma_{2i}\Delta i_{t-i} \\
& + \textstyle\sum_{i=1}^{4}\gamma_{3i}\Delta rs_{t-i} + \sum_{i=1}^{4}\gamma_{4i}\Delta rl_{t-i} \\
& + \gamma_{a1}m_{t-1} + \gamma_{a2}p_{t-1} + \gamma_{a3}i_{t-1} + \gamma_{a4}rs_{t-1} + \gamma_{a5}rl_{t-1} \\
& + \gamma_{b1}\hat{u}_{t-1} + \gamma_{b2}\hat{u}^2_{t-1} + \gamma_{c2}D_{2t} + \gamma_{c3}D_{3t} + v_t \qquad (5)
\end{aligned}
$$

Coefficients associated with the variables in equation (4) are denoted by δ_i or δ_{ij}; those not associated with the variables in equation (4) are denoted by γ_i or γ_{ij}; and v_t is the error term.

In *minimal nesting*, the general unrestricted model is formulated explicitly as an ADL, except that the variables in equation (4) are included in the general model. Redundancies between those variables and the other variables in the ADL are eliminated by dropping certain variables in the ADL. The minimally nesting model is thus the unrestricted ADL, transformed minimally such that the resulting model explicitly nests equation (4). The minimally nesting model is as follows.

$$
\begin{aligned}
\Delta(m-p)_t = \; & \delta_1\Delta(m-p)_{t-1} + \delta_2\Delta^2(m-p)_{t-2} + \delta_3\Delta p_t + \delta_4\Delta p_{t-1} + \delta_5\Delta rs_t \\
& + \delta_6\Delta_2 rl_t + \delta_7(\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} + \delta_8 + \delta_9(D_{1t} + D_{3t}) \\
& + \gamma_{a1}m_{t-1} + \textstyle\sum_{i=4}^{5}\gamma_{0i}m_{t-i} + \gamma_{a2}p_{t-1} + \sum_{i=3}^{5}\gamma_{1i}p_{t-i} \\
& + \textstyle\sum_{i=0}^{5}\gamma_{2i}i_{t-i} + \sum_{i=1}^{5}\gamma_{3i}rs_{t-i} + \sum_{i=1}^{5}\gamma_{4i}rl_{t-i} \\
& + \gamma_{b1}\hat{u}_{t-1} + \gamma_{b2}\hat{u}^2_{t-1} + \gamma_{c2}D_{2t} + \gamma_{c3}D_{3t} + v_t \qquad (6)
\end{aligned}
$$

Equations (5) and (6) are identical, except in the way that they represent the variables that do *not* enter equation (4).

In *implicit nesting*, the general unrestricted model is formulated as an ADL. The variables in equation (4) are included in the general model, but they enter (for the most part) implicitly through the corresponding log-levels. The implicitly nesting model can be formulated as follows.

$$\begin{aligned}\Delta(m-p)_t &= \textstyle\sum_{i=1}^{4}\delta_{0i}m_{t-i} + \sum_{i=0}^{4}\delta_{1i}p_{t-i} + \sum_{i=0}^{1}\delta_{2i}rs_{t-i} \\ &\quad + \delta_{40}rl_t + \delta_{42}rl_{t-2} + \delta_{52}\hat{u}_{t-1}^2 + \delta_{53}\hat{u}_{t-1}^3 + \delta_6 + \delta_{71}D_{1t} + \delta_{73}D_{3t} \\ &\quad + \textstyle\sum_{i=5}^{5}\gamma_{0i}m_{t-i} + \sum_{i=5}^{5}\gamma_{1i}p_{t-i} \\ &\quad + \textstyle\sum_{i=0}^{5}\gamma_{2i}i_{t-i} + \sum_{i=2}^{5}\gamma_{3i}rs_{t-i} + \gamma_{41}rl_{t-1} + \sum_{i=3}^{5}\gamma_{4i}rl_{t-i} \\ &\quad + \gamma_{b1}\hat{u}_{t-1} + \gamma_{c2}D_{2t} + v_t \qquad (7)\end{aligned}$$

Hendry and Ericsson (1991a), in their Table 3, start with the implicitly nesting model (4) in their model simplification process. Also, below, when the implicitly nesting model has all variables from equation (4) fixed, that means that all variables in equation (7) that have a coefficient δ_i or δ_{ij} are fixed. There are 18 such variables in the implicitly nesting model (7), whereas there are only 9 such variables in the naturally nesting model (5) and in the minimally nesting model (6)

In *redundant nesting*, the general unrestricted model is the union of the general unrestricted models for natural nesting, minimal nesting, implicit nesting. The redundantly nesting model is as follows.

$$\begin{aligned}\Delta(m-p)_t &= \delta_1\Delta(m-p)_{t-1} + \delta_2\Delta^2(m-p)_{t-2} + \delta_3\Delta p_t + \delta_4\Delta p_{t-1} + \delta_5\Delta rs_t \\ &\quad + \delta_6\Delta_2 rl_t + \delta_7(\hat{u}_{t-1} - 0.2)\hat{u}_{t-1}^2 + \delta_8 + \delta_9(D_{1t} + D_{3t}) \\ &\quad + \textstyle\sum_{i=3}^{4}\gamma_{0i}\Delta(m-p)_{t-i} + \sum_{i=2}^{4}\gamma_{1i}\Delta p_{t-i} + \sum_{i=0}^{4}\gamma_{2i}\Delta i_{t-i} \\ &\quad + \textstyle\sum_{i=1}^{4}\gamma_{3i}\Delta rs_{t-i} + \sum_{i=1}^{4}\gamma_{4i}\Delta rl_{t-i} \\ &\quad + \textstyle\sum_{i=1}^{5}\gamma_{0i}m_{t-i} + \sum_{i=0}^{5}\gamma_{1i}p_{t-i} \\ &\quad + \textstyle\sum_{i=0}^{5}\gamma_{2i}i_{t-i} + \sum_{i=0}^{5}\gamma_{3i}rs_{t-i} + \sum_{i=0}^{5}\gamma_{4i}rl_{t-i} \\ &\quad + \gamma_{b1}\hat{u}_{t-1} + \gamma_{b2}\hat{u}_{t-1}^2 + \gamma_{b3}\hat{u}_{t-1}^3 \\ &\quad + \gamma_{c1}D_{1t} + \gamma_{c2}D_{2t} + \gamma_{c3}D_{3t} + v_t \qquad (8)\end{aligned}$$

While equation (8) includes several purely collinear variables, the algorithm in Autometrics is designed to account for those redundancies, searching along the multiple paths implied by eliminating statistically equivalent sets of redundant variables.

Choice of the dependent variable. Equations (5)–(8) are all written with $\Delta(m-p)_t$ as the dependent variable. However, each of those equations includes Δp_t, m_{t-1}, and p_{t-1} as regressors; hence each equation could be written equivalently with $(m-p)_t$ as the dependent variable, or Δm_t, or m_t. As with the choice of nesting, the choice among these four possible dependent variables affects the parameterization of the general unrestricted model, and hence the range of paths and the range of potential terminal models.

In particular, rewriting equations (5)–(8) with m_t as the left-hand side (LHS) variable may help disentangle the dynamic representation induced by working with differences and differentials of the underlying variables. Section 5 thus considers simplifications from equations (5)–(8) with m_t as the dependent variable. Those equations are as follows.

The naturally nesting model is:

$$\begin{aligned} m_t \;=\; & \delta_1 \Delta(m-p)_{t-1} + \delta_2 \Delta^2 (m-p)_{t-2} + \delta_3 \Delta p_t + \delta_4 \Delta p_{t-1} + \delta_5 \Delta rs_t \\ & + \delta_6 \Delta_2 rl_t + \delta_7 (\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} + \delta_8 + \delta_9 (D_{1t} + D_{3t}) \\ & + \textstyle\sum_{i=3}^{4} \gamma_{0i} \Delta(m-p)_{t-i} + \sum_{i=2}^{4} \gamma_{1i} \Delta p_{t-i} + \sum_{i=0}^{4} \gamma_{2i} \Delta i_{t-i} \\ & + \textstyle\sum_{i=1}^{4} \gamma_{3i} \Delta rs_{t-i} + \sum_{i=1}^{4} \gamma_{4i} \Delta rl_{t-i} \\ & + \gamma_{a1} m_{t-1} + \gamma_{a2} p_{t-1} + \gamma_{a3} i_{t-1} + \gamma_{a4} rs_{t-1} + \gamma_{a5} rl_{t-1} \\ & + \gamma_{b1} \hat{u}_{t-1} + \gamma_{b2} \hat{u}^2_{t-1} + \gamma_{c2} D_{2t} + \gamma_{c3} D_{3t} + v_t. \end{aligned} \tag{9}$$

The minimally nesting model is:

$$\begin{aligned} m_t \;=\; & \delta_1 \Delta(m-p)_{t-1} + \delta_2 \Delta^2 (m-p)_{t-2} + \delta_3 \Delta p_t + \delta_4 \Delta p_{t-1} + \delta_5 \Delta rs_t \\ & + \delta_6 \Delta_2 rl_t + \delta_7 (\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} + \delta_8 + \delta_9 (D_{1t} + D_{3t}) \\ & + \gamma_{a1} m_{t-1} + \textstyle\sum_{i=4}^{5} \gamma_{0i} m_{t-i} + \gamma_{a2} p_{t-1} + \sum_{i=3}^{5} \gamma_{1i} p_{t-i} \\ & + \textstyle\sum_{i=0}^{5} \gamma_{2i} i_{t-i} + \sum_{i=1}^{5} \gamma_{3i} rs_{t-i} + \sum_{i=1}^{5} \gamma_{4i} rl_{t-i} \\ & + \gamma_{b1} \hat{u}_{t-1} + \gamma_{b2} \hat{u}^2_{t-1} + \gamma_{c2} D_{2t} + \gamma_{c3} D_{3t} + v_t. \end{aligned} \tag{10}$$

The implicitly nesting model is:

$$
\begin{aligned}
m_t \;=\; & \sum_{i=1}^{4}\delta_{0i}m_{t-i} + \sum_{i=0}^{4}\delta_{1i}p_{t-i} + \sum_{i=0}^{1}\delta_{2i}rs_{t-i} \\
& + \delta_{40}rl_t + \delta_{42}rl_{t-2} + \delta_{52}\hat{u}_{t-1}^2 + \delta_{53}\hat{u}_{t-1}^3 + \delta_6 + \delta_{71}D_{1t} + \delta_{73}D_{3t} \\
& + \sum_{i=5}^{5}\gamma_{0i}m_{t-i} + \sum_{i=5}^{5}\gamma_{1i}p_{t-i} \\
& + \sum_{i=0}^{5}\gamma_{2i}i_{t-i} + \sum_{i=2}^{5}\gamma_{3i}rs_{t-i} + \gamma_{41}rl_{t-1} + \sum_{i=3}^{5}\gamma_{4i}rl_{t-i} \\
& + \gamma_{b1}\hat{u}_{t-1} + \gamma_{c2}D_{2t} + v_t. \qquad (11)
\end{aligned}
$$

The redundantly nesting model is:

$$
\begin{aligned}
m_t \;=\; & \delta_1\Delta(m-p)_{t-1} + \delta_2\Delta^2(m-p)_{t-2} + \delta_3\Delta p_t + \delta_4\Delta p_{t-1} + \delta_5\Delta rs_t \\
& + \delta_6\Delta_2 rl_t + \delta_7(\hat{u}_{t-1} - 0.2)\hat{u}_{t-1}^2 + \delta_8 + \delta_9(D_{1t}+D_{3t}) \\
& + \sum_{i=3}^{4}\gamma_{0i}\Delta(m-p)_{t-i} + \sum_{i=2}^{4}\gamma_{1i}\Delta p_{t-i} + \sum_{i=0}^{4}\gamma_{2i}\Delta i_{t-i} \\
& + \sum_{i=1}^{4}\gamma_{3i}\Delta rs_{t-i} + \sum_{i=1}^{4}\gamma_{4i}\Delta rl_{t-i} \\
& + \sum_{i=1}^{5}\gamma_{0i}m_{t-i} + \sum_{i=0}^{5}\gamma_{1i}p_{t-i} \\
& + \sum_{i=0}^{5}\gamma_{2i}i_{t-i} + \sum_{i=0}^{5}\gamma_{3i}rs_{t-i} + \sum_{i=0}^{5}\gamma_{4i}rl_{t-i} \\
& + \gamma_{b1}\hat{u}_{t-1} + \gamma_{b2}\hat{u}_{t-1}^2 + \gamma_{b3}\hat{u}_{t-1}^3 \\
& + \gamma_{c1}D_{1t} + \gamma_{c2}D_{2t} + \gamma_{c3}D_{3t} + v_t. \qquad (12)
\end{aligned}
$$

Equations (5)–(12) are all equivalent representations. In estimation, their residuals will be identical, and hence will be their residual sums of squares and their equation standard errors. The value of R^2, however, will not be identical for equations with different dependent variables. Also, because transformations of variables result in reparameterizations, the coefficients $(\delta_i, \gamma_i, \gamma_{ij})$ in a given nesting above are *not* necessarily the same as the coefficients $(\delta_i, \gamma_i, \gamma_{ij})$ in an alternative equivalent nesting. However, for each pair of equations drawn from (5)–(12), there is a linear nonsingular mapping between the set of coefficients in one equation and the set of coefficients in the other equation.

5 Model Selection Revisited

This section applies Autometrics to multiple representations of a general unrestricted ADL for U.K. broad money demand, verifies the robustness of equation (4) to the model selection process, and obtains a more parsimonious data-coherent representation (Section 5.1); and it investigates the nonlinearities implied by equations (3)–(4) (Section 5.2). The details of the model improvement highlight the strengths and the limitations of computer-automated model selection.

5.1 Empirical Model Selection

Following the approach in Ericsson (2008, Chapters 9 and 10) and Ericsson and Kamin (2009), the current current subsection uses Autometrics to assess the possible path dependence of equation (4). The initial general model is estimated; and the algorithm simplifies that general model under each of the 192 permutations implied by the choices listed in Table 2. While the algorithm does obtain multiple distinct final models, equation (4)—or simple variants of it—appears statistically sensible. Additional analysis of those models results in a final specification that is similar to but more parsimonious than the one in equation (4). That final specification appears well-specified with empirically constant coefficients; and its economic interpretation is straightforward. These results bolster the model design in Hendry and Ericsson (1991a) and offer an improvement on it.

The choice of the general unrestricted model is fundamental in general-to-specific modeling. However, the general model for this subsection's model searches has already been evaluated extensively by Hendry and Ericsson (1991a), so details of the general model in its eight representations are relegated to Appendix B. Table B2 lists the estimates and standard errors for equation (5), i.e., the naturally nesting ECM representation of the unrestricted fifth-order ADL model, with $\Delta(m-p)_t$ as the dependent variable. The standard diagnostic statistics do not reject, except first-order ARCH (marginally so at the 5% level). Tables B3–B5 list the estimates and standard errors for the representations with minimal nesting, implicit nesting, and redundant nesting that have $\Delta(m-p)_t$ as the dependent variable. Tables B6–B9 likewise list the estimates and standard errors for the representations with natural nesting, minimal nesting, implicit nesting, and redundant nesting that have m_t as the dependent variable.

Table 3 summarizes Autometrics's model simplifications with no pre-search testing and no indicator saturation under variations in target size, fixity of regressors, and nesting of equation (4): 24 different scenarios in total. For comparison, Table 4 summarizes Autometrics's model simplifications with no pre-search testing but *with* indicator saturation for the same variations in target size, fixity of regressors, and nesting of equation (4). In these tables and the ones that follow, 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 models estimated" is the total number of distinct models estimated in the 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.

Several features of the simplifications in Table 3 are notable. First, comparing the upper and lower halves of the table, the choice of target size has little effect on the selected model: coefficients tend to be either highly significant or not significant at all. Two variables in equation (4) may be of only marginal importance statistically: $\Delta^2(m-p)_{t-2}$ and $\Delta_2 rl_t$. Their role in equation (4) is re-examined below.

Second, the choice of fixity of regressors also has little effect on the selected model, except when no regressor is fixed. In that situation, the intercept is often deleted, forcing the regression through the origin, while a variable in log-levels is retained as a proxy for the intercept. As a matter of practice, it seems sensible to retain the intercept throughout the model selection process. Little is gained by dropping it if it is statistically insignificant; and dropping the intercept can be both economically and statistically devastating to a model if the intercept actually is important. This contrasts with the potential roles of the intercept in models for forecasting; see Clements and Hendry (1998, 1999).

Third, the choice of nesting has little effect on the selected model, except for implicit nesting, when the number of variables associated with equation (4) is 18, rather than only 9 for natural, minimal, and redundant nesting. Model selection under redundant nesting leads to a similar or identical model as under natural nesting, indicative of the ability of Autometrics to successfully handle exactly collinear regressors. Occasionally, though, when Autometrics is applied to a natural nesting, it obtains a more parsimonious model with a better fit (in terms of $\hat{\sigma}$) than it obtains for a redundant nesting. For example, see Table 3 for a target size of 1% with the intercept fixed: natural nesting results in ($k_f = 8,\ \hat{\sigma} = 1.465$), whereas redundant nesting results in ($k_f = 9,\ \hat{\sigma} = 1.471$). While the differences in k_f and in $\hat{\sigma}$ are small, Autometrics applied to this redundant nesting would have selected the model with a smaller k_f and $\hat{\sigma}$ if it had found it, but it didn't. Autometrics does allow different degrees of "search effort", so these discrepancies suggest an area for further research. Along a similar vein, Ericsson (2009, Chapter 10) and Ericsson and Kamin (2009) note that, for a *given* representation of the general unrestricted model, Autometrics occasionally 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.

Table 3: Statistics on computer-automated model selection from an ECM of U.K. money demand with $\Delta(m-p)_t$ as the dependent variable, with no pre-search and no impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	No	HE (10)	Natural	no IS	Δmp	36	9	2	0	1.424
5%	No	Intercept	Natural	no IS	Δmp	36	9	256	4	1.424
5%	No	None	Natural	no IS	Δmp	36	9	250	7	1.416
5%	No	HE (10)	Minimal	no IS	Δmp	36	9	2	0	1.424
5%	No	Intercept	Minimal	no IS	Δmp	36	9	248	6	1.424
5%	No	None	Minimal	no IS	Δmp	36	9	247	14	1.419
5%	No	HE (10)	Implicit	no IS	Δmp	36	18	2	0	1.457
5%	No	Intercept	Implicit	no IS	Δmp	36	11	188	8	1.494
5%	No	None	Implicit	no IS	Δmp	36	11	137	7	1.494
5%	No	HE (10)	Redundant	no IS	Δmp	62	9	2	0	1.424
5%	No	Intercept	Redundant	no IS	Δmp	62	8	451	14	1.465
5%	No	None	Redundant	no IS	Δmp	62	9	269	12	1.422
1%	No	HE (10)	Natural	no IS	Δmp	36	9	2	0	1.424
1%	No	Intercept	Natural	no IS	Δmp	36	8	195	4	1.465
1%	No	None	Natural	no IS	Δmp	36	8	347	9	1.461
1%	No	HE (10)	Minimal	no IS	Δmp	36	9	2	0	1.424
1%	No	Intercept	Minimal	no IS	Δmp	36	8	230	7	1.465
1%	No	None	Minimal	no IS	Δmp	36	8	348	13	1.461
1%	No	HE (10)	Implicit	no IS	Δmp	36	18	2	0	1.457
1%	No	Intercept	Implicit	no IS	Δmp	36	11	348	12	1.494
1%	No	None	Implicit	no IS	Δmp	36	9	482	13	1.580
1%	No	HE (10)	Redundant	no IS	Δmp	62	9	2	0	1.424
1%	No	Intercept	Redundant	no IS	Δmp	62	9	310	18	1.471
1%	No	None	Redundant	no IS	Δmp	62	8	744	16	1.455

Table 4: Statistics on computer-automated model selection from an ECM of U.K. money demand with $\Delta(m-p)_t$ as the dependent variable, with no pre-search but with impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	No	HE (10)	Natural	IS	Δmp	129, 36	35	31	3	0.729
5%	No	Intercept	Natural	IS	Δmp	129, 24	24	10	0	0.939
5%	No	None	Natural	IS	Δmp	129, 32	32	14	0	0.768
5%	No	HE (10)	Minimal	IS	Δmp	129, 58	56	26	2	0.269
5%	No	Intercept	Minimal	IS	Δmp	129, 49	47	32	2	0.432
5%	No	None	Minimal	IS	Δmp	129, 25	25	5	0	0.919
5%	No	HE (10)	Implicit	IS	Δmp	129, 75	75	4	0	0.062
5%	No	Intercept	Implicit	IS	Δmp	129, 27	24	36	2	0.989
5%	No	None	Implicit	IS	Δmp	129, 66	66	7	0	0.141
5%	No	HE (10)	Redundant	IS	Δmp	155, 34	34	6	0	0.775
5%	No	Intercept	Redundant	IS	Δmp	155, 50	50	3	0	0.395
5%	No	None	Redundant	IS	Δmp	155, 36	33	26	2	0.763
1%	No	HE (10)	Natural	IS	Δmp	129, 9	9	1	0	1.424
1%	No	Intercept	Natural	IS	Δmp	129, 9	9	4	0	1.409
1%	No	None	Natural	IS	Δmp	129, 9	9	2	0	1.377
1%	No	HE (10)	Minimal	IS	Δmp	129, 9	9	1	0	1.424
1%	No	Intercept	Minimal	IS	Δmp	129, 12	9	37	4	1.409
1%	No	None	Minimal	IS	Δmp	129, 7	7	3	0	1.492
1%	No	HE (10)	Implicit	IS	Δmp	129, 18	18	1	0	1.457
1%	No	Intercept	Implicit	IS	Δmp	129, 25	19	61	5	1.122
1%	No	None	Implicit	IS	Δmp	129, 18	12	74	5	1.401
1%	No	HE (10)	Redundant	IS	Δmp	155, 9	9	1	0	1.424
1%	No	Intercept	Redundant	IS	Δmp	155, 12	12	4	0	1.419
1%	No	None	Redundant	IS	Δmp	155, 15	14	24	2	1.314

Table 4 summarizes results for the same model selection process as in Table 3, but with indicator saturation. For a target size of 5%, final models have many more regressors than the corresponding models in Table 3, with the additional regressors being dozens of impulse dummies. Hendry (2010) notes the need to control size when the initial number of regressors is very large, e.g., $k_1 > T$, as it is in Table 4. For a smaller target size of 1%, the final models in Table 4 are very similar to (and sometimes identical to) those in Table 3. However, the impulse dummy for 1936 often appears significant, even with a target size of 1%.

Results with pre-search testing and with m_t as the dependent variable are very similar to the results in Tables 3 and 4, so the corresponding tables are reported in Appendix A. In summary, pre-search testing often (but not always) reduces the number of models estimated; and it often reduces the number of paths that need to be searched in Stage 1, sometimes markedly so. As a consequence, pre-search testing frequently reduces the number of multiple terminal models and, in some instances, obtains the final model. Pre-search testing, however, has little or no effect on the specification of the finally selected model. Ericsson and Kamin (2009) find similar effects from pre-search testing when modeling Argentine money demand on a dataset with a similar number of observations and starting from a general model with a similar number of regressors. That said, pre-search testing may still have a more substantive role when working with datasets with many more observations and general models with many more regressors.

Changing the dependent variable from $\Delta(m-p)_t$ to m_t identifies two potential reparameterizations of interest. With m_t as the dependent variable, the coefficient on m_{t-1} is approximately unity, so the dependent variable could be transformed to Δm_t. Also, the coefficients on Δp_{t-1} and $\Delta(m-p)_{t-1}$ in equation (4) are both approximately +0.4, so those two variables could be replaced by the single variable Δm_{t-1}.Data transformations permit a final representation that is more highly parsimonious than previously obtained.

In light of these model searches, equation (4) is rewritten with Δm_t as the dependent variable; the regressors Δp_{t-1} and $\Delta(m-p)_{t-1}$ are transformed to Δp_{t-1} and Δm_{t-1}; and the regressors Δp_{t-1} and $\Delta^2(m-p)_{t-2}$ are dropped. The corresponding model, which improves on equation (4), is as follows.

$$\begin{aligned}\widehat{\Delta m_t} = {} & \underset{(0.044)}{0.381}\ \Delta m_{t-1} + \underset{(0.038)}{0.396}\ \Delta p_t - \underset{(0.005)}{0.025}\ \Delta rs_t \\ & - \underset{(0.017)}{0.059}\ \Delta_2 rl_t - \underset{(0.542)}{2.778}\ (\hat{u}_{t-1} - 0.2)\hat{u}^2_{t-1} \\ & + \underset{(0.0021)}{0.0056} + \underset{(0.532)}{4.010}\ (D_{1t} + D_{3t}) + \underset{(1.421)}{4.028}\ D_{1936t} \end{aligned} \tag{13}$$

$T = 93\ [1878\text{–}1970)]\quad \mathrm{R}^2 = 0.9102\quad \hat{\sigma} = 1.4011\%\quad Inn : F(28,58) = 0.5$
$AR : F(2,83) = 0.49\quad ARCH : F(1,91) = 0.24\quad Normality : \chi^2(2) = 5.58$
$RESET : F(2,83) = 0.19\quad Hetero : F(11,80) = 0.84\quad Form : F(21,70) = 0.19,$

where D_{1936t} is an impulse dummy for 1936. The coefficient on D_{1936t} is virtually identical to the coefficient on the war dummy $(D_{1t} + D_{3t})$, although a more complete economic explanation for D_{1936t} requires further investigation.

The coefficients in equation (13) are otherwise little changed from the corresponding ones in equation (4), except that the coefficients on $\Delta(m-p)_{t-1}$ and Δp_{t-1} are restricted to be equal, and the coefficient on $\Delta^2(m-p)_{t-2}$ is restricted to be zero. No tests reject at the 5% level. Transformation of the dependent variable from $\Delta(m-p)_t$ to Δm_t is consistent with Ss-type adjustment behavior for money, where short-run factors determine nominal money movements given the desired bands, and longer-run factors influence the levels of the bands; see Miller and Orr (1966), Milbourne (1983), and Smith (1986). Because equation (13) is invariant to whether Δm_t or $\Delta(m-p)_t$ is the dependent variable, equation (13) has virtually the same economic interpretation as equation (4), and it is more parsimonious than (4). Autometrics thus verifies the robustness of equation (4)'s specification; and it improves upon that specification, both through greater parsimony and through robust estimation by indicator saturation.

5.2 Nonlinearities in the Model

Equations (3) and (4) combined imply nonlinearity in the parameters. In practice, Hendry and Ericsson (1991a) dealt with this nonlinearity through a two-step procedure: the long-run relationship in equation (3) was estimated first; then the dynamic relationship in equation (4) was estimated, conditional on the estimates from equation (3). Equation (13) likewise conditions on the estimates from equation (3). In the spirit of Escribano (2004), the current subsection re-estimates the parameters in equations (3) and (13) jointly by nonlinear least squares and evaluates that formulation with diagnostic tools available in PcGive.

Estimating equation (3) and equation (13) jointly by nonlinear least squares obtains the following.

$$\begin{aligned}\widehat{\Delta m_t} = &\ \underset{(0.045)}{0.366}\,\Delta m_{t-1} + \underset{(0.038)}{0.397}\,\Delta p_t - \underset{(0.005)}{0.023}\,\Delta rs_t - \underset{(0.018)}{0.063}\,\Delta_2 rl_t \\ &- \underset{(0.805)}{1.687}\,([(m-p-i)_{t-1} + \underset{(0.038)}{0.299} + \underset{(0.698)}{5.997}\,RS_{t-1}] - \underset{(0.093)}{0.232})\,\tilde{u}^2_{t-1} \\ &+ \underset{(0.0025)}{0.0035} + \underset{(0.561)}{4.262}\,(D_{1t} + D_{3t}) + \underset{(0.014)}{4.270}\,D_{1936t}\end{aligned} \tag{14}$$

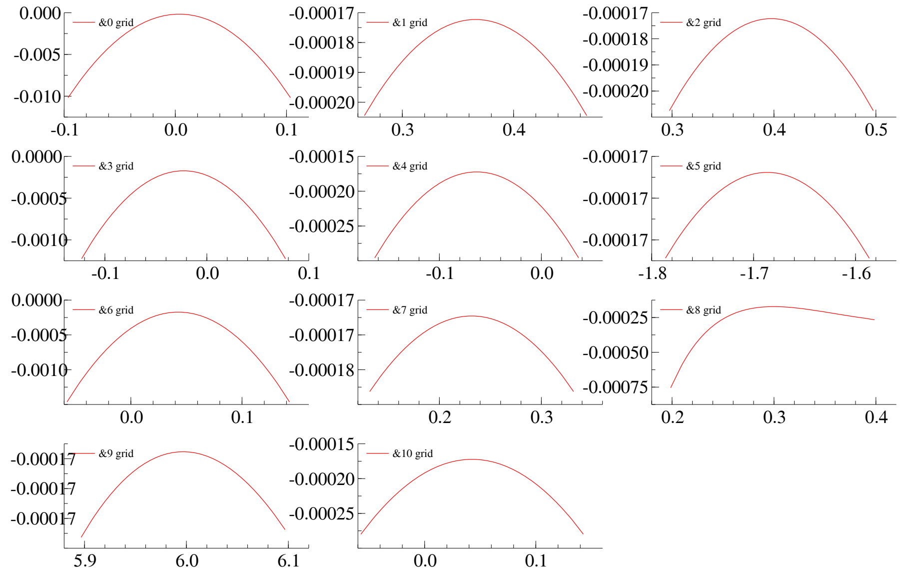

Figure 2: Grid plots for individual coefficients in equation (14).

$$T = 93 \; [1878\text{–}1970)] \qquad \mathrm{R}^2 = 0.9138 \qquad \hat{\sigma} = 1.3978\%$$
$$AR : F(2,80) = 0.65 \qquad ARCH : F(1,91) = 0.38 \qquad Normality : \chi^2(2) = 7.84^*$$
$$Hetero : F(19,73) = 0.10 \qquad Form : F(51,41) = 0.44,$$

where, in the interest of brevity, $\tilde{u}_{t-1}$ is precisely the term in square brackets $[\cdot]$. The coefficients in equation (14) are little changed from the corresponding ones in equations (3) and (13). The diagnostic statistics are not rejected, except that the normality test rejects at approximately the 2% level.

The remainder of this subsection further evaluates the numerical and statistical properties of equation (14) graphically. Numerical convergence appears good, as indicated by the grid plots for individual coefficients in Figure 2.

Figure 3 plots the actual and fitted values for equation (14) as time series, the actual and fitted values in a cross-plot, the corresponding residuals, and the histogram and estimated density of the residuals. Visually, the residuals exhibit no unusual properties, other than a minor skewness to the right.

Figure 4 plots the one-step residuals and the one-step, breakpoint, and forecast Chow statistics. The Chow statistics confirm the empirical constancy of equation (14). Only a single one-step Chow statistic is statistically significant at the 1% level, and that statistic is only barely significant at that level. None of the breakpoint Chow statistics is significant at the 1% level. That is, no split of the sample obtains a rejection of constancy. None of the forecast Chow statistics is significant at the 1% level, either. The diagnostic statistics and the statistics for testing parameter constancy all

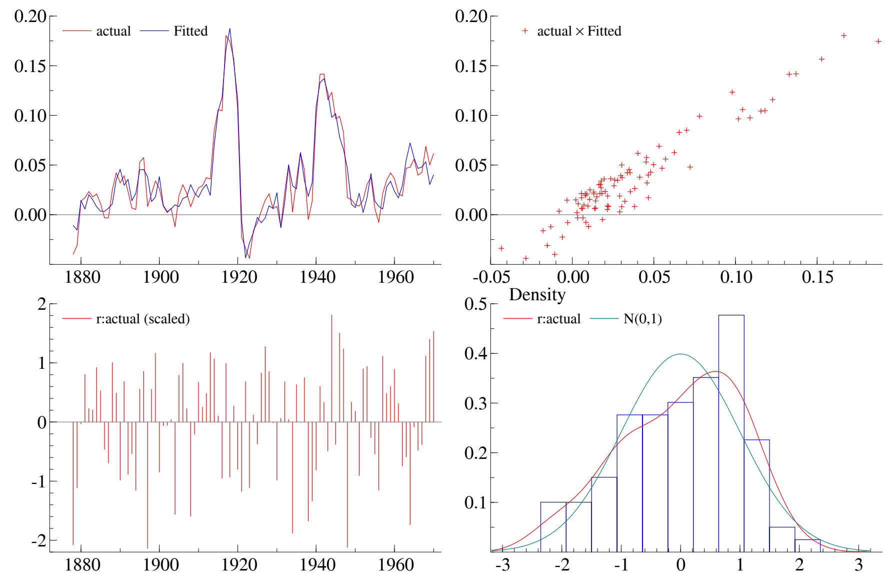

Figure 3: Graphic analysis for equation (14).

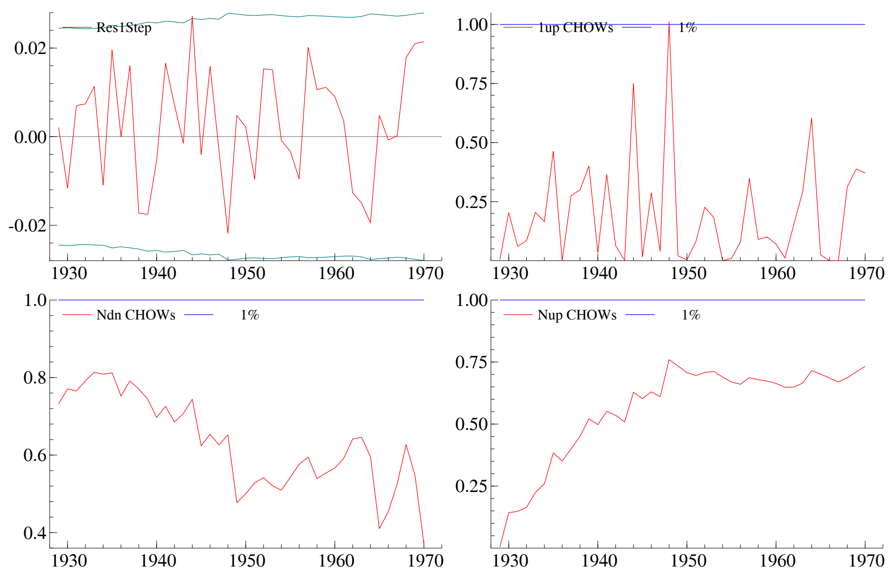

Figure 4: One-step residuals and the one-step, breakpoint, and forecast Chow statistics for equation (14).

point to equation (14) being a well-specified, empirically constant, congruent model. Nonlinear estimation of equation (14) offers some advantages over the two-step estimation of equations (3) and (13); and existing econometric software offers extensive analysis of such a nonlinear model.

6 Conclusions

Computer-automated model selection with the software package Autometrics demonstrates the robustness of Hendry and Ericsson's (1991a) final error correction model and improves on that model by using multi-path searches that would be tedious and prohibitively time-consuming with standard econometric packages. Long-run money demand is driven by a negative effect from the short-run interest rate. 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, additional improvements to model selection algorithms may obtain further improvements on model specification. Computer-automated model selection algorithms are still under development; 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 (2009), Hendry, Johansen, and Santos (2008), Hoover, Demiralp, and Perez (2009), Hoover, Johansen, and Juselius (2008), and Johansen and Nielsen (2009).

Second, a computer-automated model selection algorithm is a tool. It is incredibly powerful for what it does, but it also has limitations. In particular, the empirical modeler still faces two important choices: the initial general model, and the particular parameterization of that model. As demonstrated above, a computer-automated model selection algorithm can help sort through parameterizations of the general model by evaluating those parameterizations one at a time, or jointly (as with "redundant nesting"). Still, insights by the empirical modeler may be of value in guiding the search algorithm, as in the choice of parameterization in Section 5.1. And, computer-automated model selection algorithms are not yet capable of sorting out the choice of dependent variable, even when the choice of conditioning set is not at issue.

Third, indicator saturation provides a potentially powerful tool for assessing the role of constructed dummy variables in existing empirical models. If those dummy variables are excluded from estimation, indicator saturation should detect their omission, with certain coefficients of the retained indicator variables being equal.

Fourth, analysis of historical datasets such as the one provided by Friedman and Schwartz (1982) may offer insights on current monetary policy. To wit, the short-term interest rate RS is less than 1% for over half the time during the 1930s and 1940s, presaging current discussions about a "zero lower bound"; cf. Bernanke (2009) and Kohn (2009).

Fifth, Hendry and Ericsson (1991a), Attfield, Demery, and Duck (1995), Ericsson, Hendry, and Prestwich (1998a, 1998b), and Escribano (2004) have extended Friedman and Schwartz's dataset through 2000; and they have developed "extended" models on that extended dataset using limited (manual) model selection processes. Re-examination of the extended dataset with computer-automated model selection is in order, as is further extension of the dataset itself. 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 (1998a).

Appendix A. Additional Results

This appendix tabulates additional results on computer-automated model selection from a general unrestricted model of U.K. money demand, with each table's results categorized by target size, fixity of regressors, and the nature of nesting in the general model. Table A1 serves as a guidepost for the tables in this appendix, along with Tables 3 and 4 in the text.

Table A1: A categorization of results on model specification searches.

Table	LHS variable	Pre-search?	Indicator saturation?
Table 3	$\Delta(m-p)_t$	No	No
Table 4	$\Delta(m-p)_t$	No	Yes
Table A2	$\Delta(m-p)_t$	Yes	No
Table A3	$\Delta(m-p)_t$	Yes	Yes
Table A4	m_t	No	No
Table A5	m_t	No	Yes
Table A6	m_t	Yes	No
Table A7	m_t	Yes	Yes

See the text for the definition of notation and abbreviations used in the tables below.

Table A2: Statistics on computer-automated model selection from an ECM of U.K. money demand with $\Delta(m-p)_t$ as the dependent variable, with pre-search but no impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	Yes	HE (10)	Natural	no IS	Δmp	14	9	37	1	1.424
5%	Yes	Intercept	Natural	no IS	Δmp	19	9	102	1	1.424
5%	Yes	None	Natural	no IS	Δmp	19	9	100	3	1.416
5%	Yes	HE (10)	Minimal	no IS	Δmp	13	9	33	1	1.424
5%	Yes	Intercept	Minimal	no IS	Δmp	16	9	62	1	1.424
5%	Yes	None	Minimal	no IS	Δmp	16	9	75	2	1.416
5%	Yes	HE (10)	Implicit	no IS	Δmp	20	19	31	1	1.424
5%	Yes	Intercept	Implicit	no IS	Δmp	29	11	328	14	1.494
5%	Yes	None	Implicit	no IS	Δmp	29	12	229	6	1.460
5%	Yes	HE (10)	Redundant	no IS	Δmp	24	9	54	0	1.424
5%	Yes	Intercept	Redundant	no IS	Δmp	24	12	172	9	1.388
5%	Yes	None	Redundant	no IS	Δmp	24	10	157	6	1.417
1%	Yes	HE (10)	Natural	no IS	Δmp	9	9	33	0	1.424
1%	Yes	Intercept	Natural	no IS	Δmp	18	8	97	2	1.465
1%	Yes	None	Natural	no IS	Δmp	18	8	153	4	1.458
1%	Yes	HE (10)	Minimal	no IS	Δmp	10	9	27	1	1.424
1%	Yes	Intercept	Minimal	no IS	Δmp	12	8	41	1	1.465
1%	Yes	None	Minimal	no IS	Δmp	12	8	46	2	1.465
1%	Yes	HE (10)	Implicit	no IS	Δmp	19	18	22	1	1.457
1%	Yes	Intercept	Implicit	no IS	Δmp	16	13	67	3	1.480
1%	Yes	None	Implicit	no IS	Δmp	16	12	65	2	1.477
1%	Yes	HE (10)	Redundant	no IS	Δmp	17	9	55	0	1.424
1%	Yes	Intercept	Redundant	no IS	Δmp	30	9	251	13	1.478
1%	Yes	None	Redundant	no IS	Δmp	29	11	301	14	1.411

Table A3: Statistics on computer-automated model selection from an ECM of U.K. money demand with $\Delta(m-p)_t$ as the dependent variable, with pre-search and impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	Yes	HE (10)	Natural	IS	Δmp	129, 36, 36	35	31	3	0.729
5%	Yes	Intercept	Natural	IS	Δmp	129, 24, 24	24	10	0	0.939
5%	Yes	None	Natural	IS	Δmp	129, 32, 32	32	14	0	0.768
5%	Yes	HE (10)	Minimal	IS	Δmp	129, 58, 56	56	12	0	0.269
5%	Yes	Intercept	Minimal	IS	Δmp	129, 49, 49	47	34	2	0.432
5%	Yes	None	Minimal	IS	Δmp	129, 25, 25	25	5	0	0.919
5%	Yes	HE (10)	Implicit	IS	Δmp	129, 75, 75	75	4	0	0.062
5%	Yes	Intercept	Implicit	IS	Δmp	129, 27, 27	24	36	2	0.989
5%	Yes	None	Implicit	IS	Δmp	129, 66, 66	66	7	0	0.141
5%	Yes	HE (10)	Redundant	IS	Δmp	155, 34, 34	34	6	0	0.775
5%	Yes	Intercept	Redundant	IS	Δmp	155, 50, 50	50	3	0	0.395
5%	Yes	None	Redundant	IS	Δmp	155, 36, 36	33	30	2	0.763
1%	Yes	HE (10)	Natural	IS	Δmp	129, 9, 9	9	1	0	1.424
1%	Yes	Intercept	Natural	IS	Δmp	129, 9, 9	9	4	0	1.409
1%	Yes	None	Natural	IS	Δmp	129, 9, 9	9	2	0	1.377
1%	Yes	HE (10)	Minimal	IS	Δmp	129, 9, 9	9	1	0	1.424
1%	Yes	Intercept	Minimal	IS	Δmp	129, 12, 10	9	20	2	1.409
1%	Yes	None	Minimal	IS	Δmp	129, 7, 7	7	3	0	1.492
1%	Yes	HE (10)	Implicit	IS	Δmp	129, 18, 18	18	1	0	1.457
1%	Yes	Intercept	Implicit	IS	Δmp	129, 25, 25	19	71	5	1.122
1%	Yes	None	Implicit	IS	Δmp	129, 18, 16	14	26	2	1.328
1%	Yes	HE (10)	Redundant	IS	Δmp	155, 9, 9	9	1	0	1.424
1%	Yes	Intercept	Redundant	IS	Δmp	155, 12, 12	12	4	0	1.419
1%	Yes	None	Redundant	IS	Δmp	155, 15, 15	14	24	2	1.314

Table A4: Statistics on computer-automated model selection from an ECM of U.K. money demand with m_t as the dependent variable, with no pre-search and no impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	No	HE (10)	Natural	no IS	m	36	10	2	0	1.421
5%	No	Intercept	Natural	no IS	m	36	11	246	6	1.394
5%	No	None	Natural	no IS	m	36	9	263	6	1.416
5%	No	HE (10)	Minimal	no IS	m	36	10	2	0	1.421
5%	No	Intercept	Minimal	no IS	m	36	11	216	7	1.385
5%	No	None	Minimal	no IS	m	36	9	278	7	1.416
5%	No	HE (10)	Implicit	no IS	m	36	18	2	0	1.457
5%	No	Intercept	Implicit	no IS	m	36	11	279	9	1.494
5%	No	None	Implicit	no IS	m	36	11	137	7	1.494
5%	No	HE (10)	Redundant	no IS	m	62	10	2	0	1.421
5%	No	Intercept	Redundant	no IS	m	62	12	373	13	1.380
5%	No	None	Redundant	no IS	m	62	8	358	15	1.456
1%	No	HE (10)	Natural	no IS	m	36	10	2	0	1.421
1%	No	Intercept	Natural	no IS	m	36	9	202	4	1.465
1%	No	None	Natural	no IS	m	36	8	223	4	1.458
1%	No	HE (10)	Minimal	no IS	m	36	10	2	0	1.421
1%	No	Intercept	Minimal	no IS	m	36	9	215	6	1.465
1%	No	None	Minimal	no IS	m	36	8	345	9	1.458
1%	No	HE (10)	Implicit	no IS	m	36	18	2	0	1.457
1%	No	Intercept	Implicit	no IS	m	36	11	289	8	1.494
1%	No	None	Implicit	no IS	m	36	12	291	10	1.473
1%	No	HE (10)	Redundant	no IS	m	62	10	2	0	1.421
1%	No	Intercept	Redundant	no IS	m	62	9	552	12	1.526
1%	No	None	Redundant	no IS	m	62	9	492	15	1.463

Table A5: Statistics on computer-automated model selection from an ECM of U.K. money demand with m_t as the dependent variable, with no pre-search but with impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	No	HE (10)	Natural	IS	m	129, 21	21	9	0	1.075
5%	No	Intercept	Natural	IS	m	129, 37	37	6	0	0.690
5%	No	None	Natural	IS	m	129, 45	45	6	0	0.408
5%	No	HE (10)	Minimal	IS	m	129, 26	23	44	3	1.023
5%	No	Intercept	Minimal	IS	m	129, 23	22	14	2	1.016
5%	No	None	Minimal	IS	m	129, 25	25	11	0	0.941
5%	No	HE (10)	Implicit	IS	m	129, 75	75	4	0	0.062
5%	No	Intercept	Implicit	IS	m	129, 27	24	36	2	0.989
5%	No	None	Implicit	IS	m	129, 45	41	107	5	0.655
5%	No	HE (10)	Redundant	IS	m	155, 53	51	62	4	0.426
5%	No	Intercept	Redundant	IS	m	155, 38	34	52	5	0.666
5%	No	None	Redundant	IS	m	155, 55	55	6	0	0.298
1%	No	HE (10)	Natural	IS	m	129, 10	10	1	0	1.421
1%	No	Intercept	Natural	IS	m	129, 11	10	11	2	1.409
1%	No	None	Natural	IS	m	129, 10	9	11	2	1.403
1%	No	HE (10)	Minimal	IS	m	129, 10	10	1	0	1.421
1%	No	Intercept	Minimal	IS	m	129, 9	9	3	0	1.486
1%	No	None	Minimal	IS	m	129, 11	8	21	2	1.458
1%	No	HE (10)	Implicit	IS	m	129, 18	18	1	0	1.457
1%	No	Intercept	Implicit	IS	m	129, 20	16	54	4	1.246
1%	No	None	Implicit	IS	m	129, 16	12	21	2	1.401
1%	No	HE (10)	Redundant	IS	m	155, 10	10	1	0	1.421
1%	No	Intercept	Redundant	IS	m	155, 21	19	48	4	1.217
1%	No	None	Redundant	IS	m	155, 9	9	3	0	1.399

Table A6: Statistics on computer-automated model selection from an ECM of U.K. money demand with m_t as the dependent variable, with pre-search but no impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	Yes	HE (10)	Natural	no IS	m	15	11	42	1	1.394
5%	Yes	Intercept	Natural	no IS	m	19	11	83	1	1.394
5%	Yes	None	Natural	no IS	m	19	9	94	1	1.416
5%	Yes	HE (10)	Minimal	no IS	m	12	11	36	1	1.385
5%	Yes	Intercept	Minimal	no IS	m	16	11	65	2	1.385
5%	Yes	None	Minimal	no IS	m	16	9	64	1	1.416
5%	Yes	HE (10)	Implicit	no IS	m	20	19	31	1	1.424
5%	Yes	Intercept	Implicit	no IS	m	29	11	330	14	1.494
5%	Yes	None	Implicit	no IS	m	29	12	232	6	1.460
5%	Yes	HE (10)	Redundant	no IS	m	16	13	62	2	1.383
5%	Yes	Intercept	Redundant	no IS	m	21	13	116	4	1.384
5%	Yes	None	Redundant	no IS	m	21	11	121	4	1.408
1%	Yes	HE (10)	Natural	no IS	m	10	10	33	0	1.421
1%	Yes	Intercept	Natural	no IS	m	17	9	95	2	1.465
1%	Yes	None	Natural	no IS	m	17	8	112	3	1.458
1%	Yes	HE (10)	Minimal	no IS	m	11	10	27	1	1.421
1%	Yes	Intercept	Minimal	no IS	m	12	10	50	2	1.421
1%	Yes	None	Minimal	no IS	m	12	8	42	1	1.458
1%	Yes	HE (10)	Implicit	no IS	m	19	18	22	1	1.457
1%	Yes	Intercept	Implicit	no IS	m	16	13	71	3	1.480
1%	Yes	None	Implicit	no IS	m	16	12	70	2	1.477
1%	Yes	HE (10)	Redundant	no IS	m	16	10	54	0	1.421
1%	Yes	Intercept	Redundant	no IS	m	19	11	145	6	1.425
1%	Yes	None	Redundant	no IS	m	18	10	114	4	1.420

Table A7: Statistics on computer-automated model selection from an ECM of U.K. money demand with m_t as the dependent variable, with pre-search and impulse saturation, categorized by target size, fixity of regressors, and the nature of nesting in the general model.

Target size	Pre-search?	Fixity of regressors	Nesting	IS	LHS	k_1	k_f	Number of models estimated	Number of terminal models	$\hat{\sigma}$ (%)
5%	Yes	HE (10)	Natural	IS	m	129, 21, 21	21	9	0	1.075
5%	Yes	Intercept	Natural	IS	m	129, 37, 37	37	6	0	0.690
5%	Yes	None	Natural	IS	m	129, 45, 45	45	6	0	0.408
5%	Yes	HE (10)	Minimal	IS	m	129, 26, 26	23	44	3	1.023
5%	Yes	Intercept	Minimal	IS	m	129, 23, 23	22	14	2	1.016
5%	Yes	None	Minimal	IS	m	129, 25, 25	25	11	0	0.941
5%	Yes	HE (10)	Implicit	IS	m	129, 75, 75	75	4	0	0.062
5%	Yes	Intercept	Implicit	IS	m	129, 27, 27	24	36	2	0.989
5%	Yes	None	Implicit	IS	m	129, 45, 45	41	109	5	0.655
5%	Yes	HE (10)	Redundant	IS	m	155, 53, 53	51	62	4	0.426
5%	Yes	Intercept	Redundant	IS	m	155, 38, 37	34	26	2	0.666
5%	Yes	None	Redundant	IS	m	155, 55, 55	55	6	0	0.298
1%	Yes	HE (10)	Natural	IS	m	129, 10, 10	10	1	0	1.421
1%	Yes	Intercept	Natural	IS	m	129, 11, 10	10	6	0	1.409
1%	Yes	None	Natural	IS	m	129, 10, 9	9	6	0	1.403
1%	Yes	HE (10)	Minimal	IS	m	129, 10, 10	10	1	0	1.421
1%	Yes	Intercept	Minimal	IS	m	129, 9, 9	9	3	0	1.486
1%	Yes	None	Minimal	IS	m	129, 11, 9	8	11	1	1.458
1%	Yes	HE (10)	Implicit	IS	m	129, 18, 18	18	1	0	1.457
1%	Yes	Intercept	Implicit	IS	m	129, 20, 20	16	60	4	1.246
1%	Yes	None	Implicit	IS	m	129, 16, 15	14	29	2	1.328
1%	Yes	HE (10)	Redundant	IS	m	155, 10, 10	10	1	0	1.421
1%	Yes	Intercept	Redundant	IS	m	155, 21, 21	19	50	4	1.217
1%	Yes	None	Redundant	IS	m	155, 9, 9	9	3	0	1.399

Appendix B. The General Unrestricted Model

In Tables B2–B9 below, this appendix documents the estimated general unrestricted model in its various representations given by equations (5)–(12). Table B1 categorizes those representations by the choice of dependent (LHS) variable and the choice of nesting.

Table B1: A categorization of representations of the general unrestricted model.

Equation	Table	LHS variable	Nesting
(5)	Table B2	$\Delta(m-p)_t$	Natural
(6)	Table B3	$\Delta(m-p)_t$	Minimal
(7)	Table B4	$\Delta(m-p)_t$	Implicit
(8)	Table B5	$\Delta(m-p)_t$	Redundant
(9)	Table B6	m_t	Natural
(10)	Table B7	m_t	Minimal
(11)	Table B8	m_t	Implicit
(12)	Table B9	m_t	Redundant

Hendry and Ericsson (1991a, Table 3) corresponds to Table B8; and the variable ecm_t is defined as $(\hat{u}_t - 0.2)\hat{u}_t^2$.

Table B2: A general unrestricted model for UK broad money, with natural nesting, and with $\Delta(m-p)_t$ as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$	−1 (−)	0.416 (0.126)		−0.114 (0.136)	−0.000 (0.111)	
Δp_{t-j}	−0.553 (0.077)	0.359 (0.108)	−0.061 (0.123)	0.058 (0.113)	0.050 (0.097)	
Δi_{t-j}	0.087 (0.072)	0.009 (0.083)	0.056 (0.073)	0.031 (0.072)	0.065 (0.071)	
Δrs_{t-j}	−0.019 (0.009)	−0.001 (0.012)	−0.005 (0.013)	0.000 (0.010)	−0.004 (0.009)	
Δrl_{t-j}		−0.059 (0.075)	0.019 (0.054)	−0.064 (0.050)	0.007 (0.052)	
m_{t-j}		−0.100 (0.084)				
p_{t-j}		0.095 (0.093)				
i_{t-j}		0.110 (0.078)				
rs_{t-j}		−0.004 (0.013)				
rl_{t-j}		−0.016 (0.041)				
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.121 (0.223)			
$D_{j,t}$	−0.201 (0.137)		0.607 (1.751)	−0.369 (1.442)		3.993 (1.220)
$\Delta^2(m-p)_{t-j}$			−0.205 (0.138)			
$\Delta_2 rl_{t-j}$	−0.069 (0.045)					
ecm_{t-j}		−1.821 (1.285)				

$T = 93$ [1878–1970] $R^2 = 0.895$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B3: A general unrestricted model for UK broad money, with minimal nesting, and with $\Delta(m-p)_t$ as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$	−1 (−)	0.473 (0.151)				
Δp_{t-j}	−0.553 (0.077)	0.420 (0.210)				
Δi_{t-j}						
Δrs_{t-j}	−0.019 (0.009)					
Δrl_{t-j}						
m_{t-j}		−0.157 (0.109)			0.057 (0.147)	0.000 (0.111)
p_{t-j}		0.090 (0.150)		0.120 (0.213)	−0.065 (0.122)	−0.050 (0.074)
i_{t-j}	0.087 (0.072)	0.031 (0.121)	0.047 (0.090)	−0.024 (0.088)	0.034 (0.091)	−0.065 (0.071)
rs_{t-j}		−0.005 (0.012)	−0.004 (0.010)	0.005 (0.011)	−0.004 (0.011)	0.004 (0.009)
rl_{t-j}		−0.075 (0.070)	0.079 (0.095)	−0.084 (0.075)	0.071 (0.073)	−0.007 (0.052)
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.121 (0.223)			
$D_{j,t}$	−0.201 (0.137)		0.607 (1.751)	−0.369 (1.442)		3.993 (1.220)
$\Delta^2(m-p)_{t-j}$			−0.148 (0.121)			
$\Delta_2 rl_{t-j}$	−0.069 (0.045)					
ecm_{t-j}		−1.821 (1.285)				
$T = 93$ [1878–1970] $R^2 = 0.895$ $\hat{\sigma} = 1.5535\%$						

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B4: A general unrestricted model for UK broad money, with implicit nesting, and with $\Delta(m-p)_t$ as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$	−1 (−)					
Δp_{t-j}						
Δi_{t-j}						
Δrs_{t-j}						
Δrl_{t-j}						
m_{t-j}		0.316 (0.157)	−0.621 (0.226)	0.295 (0.242)	−0.091 (0.220)	0.000 (0.111)
p_{t-j}	−0.553 (0.077)	0.590 (0.120)	0.201 (0.119)	−0.176 (0.116)	0.083 (0.110)	−0.050 (0.074)
i_{t-j}	0.087 (0.072)	0.031 (0.121)	0.047 (0.090)	−0.024 (0.088)	0.034 (0.091)	−0.065 (0.071)
rs_{t-j}	−0.019 (0.009)	0.014 (0.016)	−0.004 (0.010)	0.005 (0.011)	−0.004 (0.011)	0.004 (0.009)
rl_{t-j}	−0.069 (0.045)	−0.075 (0.070)	0.147 (0.072)	−0.084 (0.075)	0.071 (0.073)	−0.007 (0.052)
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.485 (0.204)	−1.821 (1.285)		
$D_{j,t}$	−0.201 (0.137)	3.993 (1.220)	0.607 (1.751)	3.624 (1.024)		
$\Delta^2(m-p)_{t-j}$						
$\Delta_2 rl_{t-j}$						
ecm_{t-j}						

$T = 93$ [1878–1970] $R^2 = 0.895$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B5: A general unrestricted model for UK broad money, with redundant nesting, and with $\Delta(m-p)_t$ as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$	−1 (−)	0.053 (0.204)		0 (−)	0.057 (0.199)	
Δp_{t-j}	−0.643 (0.190)	0 (−)	0 (−)	0 (−)	−0.009 (0.248)	
Δi_{t-j}	0 (−)	0 (−)	0 (−)	0 (−)	0 (−)	
Δrs_{t-j}	0 (−)	0 (−)	0 (−)	0 (−)	0 (−)	
Δrl_{t-j}		0 (−)	0 (−)	0 (−)	0 (−)	
m_{t-j}		0.263 (0.311)	−0.420 (0.286)	0 (−)	0 (−)	0.057 (0.092)
p_{t-j}	0.090 (0.203)	0 (−)	0 (−)	0.120 (0.289)	0 (−)	−0.115 (0.152)
i_{t-j}	0.087 (0.098)	0.031 (0.164)	0.047 (0.122)	−0.024 (0.120)	0.034 (0.123)	−0.065 (0.097)
rs_{t-j}	−0.019 (0.012)	0.014 (0.022)	−0.004 (0.014)	0.005 (0.015)	−0.004 (0.015)	0.004 (0.012)
rl_{t-j}	−0.069 (0.061)	−0.075 (0.095)	0.147 (0.097)	−0.084 (0.102)	0.071 (0.100)	−0.007 (0.070)
$\hat{u}^j_{t-1}$		0.077 (0.125)	0.121 (0.302)	0 (−)		
$D_{j,t}$	−0.201 (0.185)	0.369 (1.955)	0.607 (2.374)	0 (−)		3.624 (1.389)
$\Delta^2(m-p)_{t-j}$			−0.148 (0.164)			
$\Delta_2 rl_{t-j}$	0 (−)					
ecm_{t-j}		−1.821 (1.743)				

$T = 93$ [1878–1970] $R^2 = 0.895$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B6: A general unrestricted model for UK broad money, with natural nesting, and with m_t as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$		0.416 (0.126)		−0.114 (0.136)	−0.000 (0.111)	
Δp_{t-j}	0.447 (0.077)	0.359 (0.108)	−0.061 (0.123)	0.058 (0.113)	0.050 (0.097)	
Δi_{t-j}	0.087 (0.072)	0.009 (0.083)	0.056 (0.073)	0.031 (0.072)	0.065 (0.071)	
Δrs_{t-j}	−0.019 (0.009)	−0.001 (0.012)	−0.005 (0.013)	0.000 (0.010)	−0.004 (0.009)	
Δrl_{t-j}		−0.059 (0.075)	0.019 (0.054)	−0.064 (0.050)	0.007 (0.052)	
m_{t-j}	−1 (−)	0.900 (0.084)				
p_{t-j}		0.095 (0.093)				
i_{t-j}		0.110 (0.078)				
rs_{t-j}		−0.004 (0.013)				
rl_{t-j}		−0.016 (0.041)				
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.121 (0.223)			
$D_{j,t}$	−0.201 (0.137)		0.607 (1.751)	−0.369 (1.442)		3.993 (1.220)
$\Delta^2(m-p)_{t-j}$			−0.205 (0.138)			
$\Delta_2 rl_{t-j}$	−0.069 (0.045)					
ecm_{t-j}		−1.821 (1.285)				
$T = 93$ [1878–1970] $R^2 = 0.99987$ $\hat{\sigma} = 1.5535\%$						

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B7: A general unrestricted model for UK broad money, with minimal nesting, and with m_t as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$		0.473 (0.151)				
Δp_{t-j}	0.447 (0.077)	0.420 (0.210)				
Δi_{t-j}						
Δrs_{t-j}	−0.019 (0.009)					
Δrl_{t-j}						
m_{t-j}	−1 (−)	0.843 (0.109)			0.057 (0.147)	0.000 (0.111)
p_{t-j}		0.090 (0.150)		0.120 (0.213)	−0.065 (0.122)	−0.050 (0.074)
i_{t-j}	0.087 (0.072)	0.031 (0.121)	0.047 (0.090)	−0.024 (0.088)	0.034 (0.091)	−0.065 (0.071)
rs_{t-j}		−0.005 (0.012)	−0.004 (0.010)	0.005 (0.011)	−0.004 (0.011)	0.004 (0.009)
rl_{t-j}		−0.075 (0.070)	0.079 (0.095)	−0.084 (0.075)	0.071 (0.073)	−0.007 (0.052)
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.121 (0.223)			
$D_{j,t}$	−0.201 (0.137)		0.607 (1.751)	−0.369 (1.442)		3.993 (1.220)
$\Delta^2(m-p)_{t-j}$			−0.148 (0.121)			
$\Delta_2 rl_{t-j}$	−0.069 (0.045)					
ecm_{t-j}		−1.821 (1.285)				

$T = 93$ [1878–1970] $R^2 = 0.99987$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B8: A general unrestricted model for UK broad money, with implicit nesting, and with m_t as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$						
Δp_{t-j}						
Δi_{t-j}						
Δrs_{t-j}						
Δrl_{t-j}						
m_{t-j}	−1 (−)	1.316 (0.157)	−0.621 (0.226)	0.295 (0.242)	−0.091 (0.220)	0.000 (0.111)
p_{t-j}	0.447 (0.077)	−0.410 (0.120)	0.201 (0.119)	−0.176 (0.116)	0.083 (0.110)	−0.050 (0.074)
i_{t-j}	0.087 (0.072)	0.031 (0.121)	0.047 (0.090)	−0.024 (0.088)	0.034 (0.091)	−0.065 (0.071)
rs_{t-j}	−0.019 (0.009)	0.014 (0.016)	−0.004 (0.010)	0.005 (0.011)	−0.004 (0.011)	0.004 (0.009)
rl_{t-j}	−0.069 (0.045)	−0.075 (0.070)	0.147 (0.072)	−0.084 (0.075)	0.071 (0.073)	−0.007 (0.052)
$\hat{u}^j_{t-1}$		0.077 (0.092)	0.485 (0.204)	−1.821 (1.285)		
$D_{j,t}$	−0.201 (0.137)	3.993 (1.220)	0.607 (1.751)	3.624 (1.024)		
$\Delta^2(m-p)_{t-j}$						
$\Delta_2 rl_{t-j}$						
ecm_{t-j}						

$T = 93$ [1878–1970] $R^2 = 0.99987$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

Table B9: A general unrestricted model for UK broad money, with redundant nesting, and with m_t as the dependent variable.

Variable[a,b]	Lag j (or index j)					
	0	1	2	3	4	5
$\Delta(m-p)_{t-j}$		0.053 (0.204)		0 (−)	0.057 (0.199)	
Δp_{t-j}	0.357 (0.190)	0 (−)	0 (−)	0 (−)	−0.009 (0.248)	
Δi_{t-j}	0 (−)	0 (−)	0 (−)	0 (−)	0 (−)	
Δrs_{t-j}	0 (−)	0 (−)	0 (−)	0 (−)	0 (−)	
Δrl_{t-j}		0 (−)	0 (−)	0 (−)	0 (−)	
m_{t-j}	−1 (−)	1.263 (0.311)	−0.420 (0.286)	0 (−)	0 (−)	0.057 (0.092)
p_{t-j}	0.090 (0.203)	0 (−)	0 (−)	0.120 (0.289)	0 (−)	−0.115 (0.152)
i_{t-j}	0.087 (0.098)	0.031 (0.164)	0.047 (0.122)	−0.024 (0.120)	0.034 (0.123)	−0.065 (0.097)
rs_{t-j}	−0.019 (0.012)	0.014 (0.022)	−0.004 (0.014)	0.005 (0.015)	−0.004 (0.015)	0.004 (0.012)
rl_{t-j}	−0.069 (0.061)	−0.075 (0.095)	0.147 (0.097)	−0.084 (0.102)	0.071 (0.100)	−0.007 (0.070)
$\hat{u}^j_{t-1}$		0.077 (0.125)	0.121 (0.302)	0 (−)		
$D_{j,t}$	−0.201 (0.185)	0.369 (1.955)	0.607 (2.374)	0 (−)		3.624 (1.389)
$\Delta^2(m-p)_{t-j}$			−0.148 (0.164)			
$\Delta_2 rl_{t-j}$	0 (−)					
ecm_{t-j}		−1.821 (1.743)				

$T = 93$ [1878–1970] $R^2 = 0.99987$ $\hat{\sigma} = 1.5535\%$

Notes:

a. For readability, the coefficients and estimated standard errors for the dummies $D_{1,t}$, $D_{2,t}$, $D_{3,t}$, and $D_{5,t}$ have been multiplied by 100.

b. The coefficients that are "boxed in" correspond to those in the restricted model, equation (**).

References

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.

Attfield, C. L. F., D. Demery, and N. W. Duck (1995) "Estimating the UK Demand for Money Function: A Test of Two Approaches", mimeo, Department of Economics, University of Bristol, Bristol, November.

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

Campos, J., N. R. Ericsson, and D. F. Hendry (1990) "An Analogue Model of Phase-averaging Procedures", *Journal of Econometrics*, 43, 3, 275–292.

Clements, M. P., and D. F. Hendry (1998) *Forecasting Economic Time Series*, Cambridge University Press, Cambridge.

Clements, M. P., and D. F. Hendry (1999) *Forecasting Non-stationary Economic Time Series*, MIT Press, Cambridge.

Doornik, J. A. (2009) "Autometrics", Chapter 4 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, 88–121.

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

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

Ericsson, N. R. (2009) *Empirical Modeling of Economic Time Series*, in preparation.

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

Ericsson, N. R., D. F. Hendry, and K. M. Prestwich (1998b) "Friedman and Schwartz (1982) Revisited: Assessing Annual and Phase-average Models of Money Demand in the United Kingdom", *Empirical Economics*, 23, 3, 401–415.

Ericsson, N. R., and S. B. Kamin (2009) "Constructive Data Mining: Modelling Argentine Broad Money Demand", Chapter 17 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, 412–439.

Escribano, A. (1985) "Non-linear Error-correction: The Case of Money Demand in the U.K. (1878–1970)", mimeo, University of California at San Diego, La Jolla, California, December.

Escribano, A. (2004) "Nonlinear Error Correction: The Case of Money Demand in the United Kingdom (1878–2000)", *Macroeconomic Dynamics*, 8, 1, 76–116.

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

Friedman, M., and A. J. Schwartz (1982) *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867–1975*, University of Chicago Press, Chicago.

Friedman, M., and A. J. Schwartz (1991) "Alternative Approaches to Analyzing Economic Data", *American Economic Review*, 81, 1, 39–49.

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.

Hendry, D. F., and N. R. Ericsson (1991a) "An Econometric Analysis of U.K. Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz", *American Economic Review*, 81, 1, 8–38.

Hendry, D. F., and N. R. Ericsson (1991b) "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.

Hoover, K. D., S. Demiralp, and S. J. Perez (2009) "Empirical Identification of the Vector Autoregression: The Causes and Effects of US M2", Chapter 2 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, 37–58.

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.

Johansen, S., and B. Nielsen (2009) "An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator", Chapter 1 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, 1–36.

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.

Milbourne, R. (1983) "Optimal Money Holding Under Uncertainty", *International Economic Review*, 24, 3, 685–698.

Miller, M. H., and D. Orr (1966) "A Model of the Demand for Money by Firms", *Quarterly Journal of Economics*, 80, 3, 413–435.

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

Smith, G. W. (1986) "A Dynamic Baumol-Tobin Model of Money Demand", *Review of Economic Studies*, 53, 3, 465–469.

Teräsvirta, T., and A.-C. Eliasson (2001) "Non-linear Error Correction and the UK Demand for Broad Money, 1878–1993", *Journal of Applied Econometrics*, 16, 3, 277–288.