Nonlinear Error-Correction Money Demand: Transactions-Theory Based Models

Alvaro Escribano
Department of Economics
Universidad Carlos III de Madrid
alvaroe@eco.uc3m.es
(34-91-6249854)

and

Clinton A. Greene
Department of Economics
University of Missouri
St. Louis, MO 63121-4499
clinton_greene@umsl.edu
(314) 516-5565

October 31, 2002

Abstract:

Forecast tests are used to compare linear and non-linear error-correction models of US M1 1948-1994. Transactions theory implies aggregate money is best modeled by powers of a rational-polynomial or a simpler exponential approximation. These implied non-linear models are distinct from smooth-transition models.

Fourth-quarter-ahead forecasts from linear, smooth-transition, exponential and rational-polynomial forms are compared. Forecasts are rigorous in the sense that the forecast period is omitted when estimating the cointegrating relationship. We introduce a measure of forecast encompassing which relies upon ex-ante estimation of pooled forecasts, more closely matching the problem faced by forecasters.

JEL: C53, C52, C22, C15, E41, E47

1. Introduction

Optimal inventory management is often an (S,s) or trigger-target rule. Such rules create complex effects at the aggregate level as in Caplan (1985), Cabellero and Engle (1991) or Greene (2001). Consideration of money as an inventory with costly management can imply simple rules as in Miller and Orr (1966) or more elaborate variants as Vickson (1985). Greene (1999) shows optimal management at the individual level implies a non-linear error-correction model for aggregate money holdings. This non-linear form is distinct from smooth-transition models which have been suggested using heuristic arguments as in Sarno, Taylor and Peel (2002). Escribano (1997a,b and 1998) and Escribano and Mira (2002) have developed the statistical theory for non-linear cointegrated processes and non-linear error-correction.

A number of studies have found a reliable long-run money demand relationship such as Swanson (1998) and Cutler, Davies, Rhodd and Schwarm (1997). On the other hand attempts to exploit dynamic models of money for short-run policy purposes have been less successful. For instance, Coenen, Levin and Wieland (2001) attempt to use money as an "indicator variable" for GDP growth with some success, but find the magnitude of the information content is small. Since monetary theory implies short-run relationships are non-linear, it should not be surprising that linear models have been of limited usefulness in a policy context. Before applying short-run models in broader contexts it is essential to develop empirical non-linear models of aggregate money which perform better than linear models.

This paper presents empirical results for a number of non-linear error-correction models of quarterly US M1 1948.2-1995.1.[1] Non-linear forms considered include the rational-

[1] After this date the Board of Governors no longer measures M1 as held and managed by depositors in checkable accounts. The so-called "retail" or "shadow" sweep accounts introduced in 1994 are an accounting device which allows banks to report a portion of M1 deposits as MMDA's for the purpose of evading reserve requirements.

polynomial and exponential models derived from monetary theory in Greene (1999) and the smooth-transition model as in Teräsvirta (1994) and Teräsvirta and Eliasson (2001). All these are compared to a linear model.

We examine fourth-step-ahead forecasts. Following Bhansali (2002), the models used to forecast over this longer horizon are not specified as direct extensions of standard one-step-ahead models (which would employ first quarterly lags). Instead, the fourth-quarter-ahead models are specified to reflect available information when predicting four quarters ahead. The dependent variable is a four quarter change (y_t -y_{t-4}). This is modeled as a function of its own lags four or more quarters back, and of an error-correction term employing the fourth (not the first) lag of aggregate money.

Fourth-step-ahead forecasts and associated encompassing tests are of particular interest for two reasons. First, forecasters are often interested in time-horizons beyond one quarter ahead. Second, the monetary theory outlined below implies the non-linear characteristics of aggregate money operate through the error-correction mechanism. The weight or importance of error-correction is larger for longer horizon forecasts. If this weight is indeed a non-linear function, then linear approximations will be less accurate over longer time horizons. Any non-linear effects will be more pronounced in the longer time-horizon forecasting exercise.

The formal forecast encompassing test regressions employ the entire series of forecasts. This ex-post treatment of forecasts is standard practice, but neglects an issue relevant to forecasters. Even if both competing forecast series (say from a linear and non-linear model) are statistically significant in the pooled model (neither model encompasses the other), this does not

---

Depositors are usually unaware of the sweep account, do not receive MMDA interest and can access their funds as un-restricted checkables. As estimated in Anderson and Rasche (2001), in 1996 more than 5% of demand deposits are unreported, and by the year 2000 almost half of M1 deposits are unreported.

imply a better forecasting model will employ both forecasts. The pooled regression model may itself be unstable. Section 4 introduces an informal approach to measuring the empirical usefulness of pooled forecasts which also avoids the problems inherent in ex-post testing.

Sections 2 and 3 present the economic and statistical background needed to understand the empirical regression models. Section 4 presents the details of the approaches adopted in evaluating forecast performance. In Section 5 we specify and fit these models to the data. After applying a number of lag-truncation criteria this results in about a dozen empirical variants. Forecasts used in the encompassing tests are rigorous in the sense that all model coefficients including those for the cointegrating relationship come from the sub-sample used to project into the future.

2. Nonlinear Money Demand As Based Upon Transactions Theory

The economic theory which lies behind these non-linear models is very simple but requires some adjustment of perspective. Suppose most agents manage their funds using trigger-target rules and there is a random component to individual net receipts. Then for each individual desired or optimal money holdings is not a scalar quantity. At any point in time most agents will find their balance well within the acceptable interval and will passively accept the balance as driven by daily exigencies. This has three implications:

First, individual money demand as usually interpreted is actually the expected value of random holdings given the influence of income and interest rate on the triggers and targets. For any individual actual holdings wander within an interval. Only when the management interval is breached (the balance is "too low" or "too high") is active control applied. In the aggregate context, this interval-valued money demand then leads to a second implication.

This second implication is that aggregate lagged money is an alternative to “money demand” as a predictor of current holdings. Every receipt implies a debit. If agents never adjusted their portfolios then aggregate money holdings would be constant over time. The predictive usefulness of lagged money is diminished only to the extent some portion of the economy's agents have adjusted their balance by selling or buying non-money assets.

Third, the more often such portfolio adjustments occur, the more useful becomes " long run money demand" (as traditionally defined) as a predictor of current holdings. Long run demand and lagged money are complementary predictors with optimal weights which depend on the width of management intervals. Management interval width (and long-run or average demand over long time horizons) varies as a function of traditional variables such as income and interest rates. Wide management intervals imply more weight on lagged money and a non-linear function is needed to model the optimal weighting scheme.

Greene (1999) shows the optimal weights are as given below in Equation 3.1. The parameters have known analytic values if all agents have the same triggers and targets. If triggers and targets are heterogeneous then an exponential function is a good approximation for the aggregate economy. This is essentially a varying-coefficients model in which coefficient variation is not inherently random nor a function of time, but rather coefficients vary as a deterministic non-linear function of long run demand itself. In the next section we show how this leads to the non-linear error-correction model.

3. Nonlinear Error Correction Money Demand Models

The traditional economic justification for error-correction models is an extension of partial-adjustment derived from quadratic adjustment-costs, as in Kennan (1979), Hansen and

Sargent (1980), Nickel (1985), Dolado, Galbraith and Banerjee (1991), Gregory, Pagan and Smith (1990) and Escribano and Pfann (1998). However as discussed above, in this paper lagged money represents aggregation effects rather than smooth adjustment.

As briefly exposited above in section 2 we know that money demand is to be interpreted as the expected value of random holdings given the influence of income and interest rate on the triggers and targets. In particular from Greene (1999),

$$E(M_t) \cong b_t M_{t-1} + [1 - b_t]LRD_t \tag{3.1a}$$

$$b_t = \left[1 - \frac{\beta_0 + \beta_1 LRD_t}{\beta_2 + \beta_3 LRD_t + \beta_4 LRD_t^{\,2} + \beta_5 LRD_t^{\,3}}\right]^{\beta_6} \tag{3.1b}$$

where LRD_t is the antilog of $\log(M_t) = \beta_0 + \beta_1\log(Y_t) + \beta_2\log(R_t)$. (3.1c)

Equation (3.1c) indicates that if the three variables are I(1) they are cointegrated with a single cointegration vector which is linear in logs. The objective now is to derive an empirical money demand model from (3.1a). In order to do that we decompose M_t as,

$$M_t = E(M_t) + u_t . \tag{3.2}$$

The term $E(M_t)$ is the expected value of M_t, given valid initial conditions, and u_t is a stationary error term, with zero mean and constant variance. From (3.1a) and (3.2) we get that,

$$M_t = b_t M_{t-1} + [1 - b_t]LRD_t + u_t. \tag{3.3}$$

Under the assumption that LRD_t is I(1) we can write,

$$LRD_t = LRD_{t-1} + v_t \tag{3.4}$$

where v_t is also stationary with zero mean, constant variance and independent of u_t.

After some simple algebra we get from (3.4) that

$$(M_t - LRD_t) = b_t (M_{t-1} - LRD_{t-1}) – b_t v_t + u_t \tag{3.5}$$

which can be expressed as a Dickey-Fuller (DF), Dickey and Fuller(1979), type of unit root test equation with a time varying and stochastic-unit root coefficient (b_t) and with heteroscedastic errors (w_t),

$$\Delta(M_t - LRD_t) = (b_t -1)(M_{t-1} - LRD_{t-1}) + w_t \tag{3.6a}$$

$$w_t = – b_t v_t + u_t. \tag{3.6b}$$

However, it is more convenient to write the system of equations (3.6a, b) as a single equation *nonlinear error correction model* with a constant variance error term,

$$\Delta M_t = (b_t -1)(M_{t-1} - LRD_{t-1}) + (1– b_t)\Delta LRD_t + u_t. \tag{3.7}$$

From equation (3.1b) we know that b_t is a nonlinear function of LRD_t and in particular

$$(b_t –1) = \left[1 - \frac{\beta_0 + \beta_1 LRD_t}{\beta_2 + \beta_3 LRD_t + \beta_4 LRD_t^{\ 2} + \beta_5 LRD_t^{\ 3}}\right]^{\beta_6} -1 = F(LRD_t, \beta) \tag{3.8}$$

which is a power of a *rational polynomial* in LRD_t. Rational polynomials are very general nonlinear functional forms based on Padé´s approximant, see Escribano (1997).

Let the first term of the RHS of equation (3.7) be $(b_t -1)(M_{t-1} - LRD_{t-1})= F(LRD_t, \beta)(M_{t-1} - LRD_{t-1})$ and the second term be $(1– b_t)\Delta LRD_t = F(LRD_t, \beta)\Delta LRD_t \cong g(\Delta LRD_t,\alpha)$. With this notation, equation (3.7) can be written as the following nonlinear error correction (NEC) model.

$$\Delta M_t = F(LRD_t,\beta)(M_{t-1} - LRD_{t-1}) + g(\Delta LRD_t,\alpha) + u_t. \quad (3.9)$$

In general, the error term, u_t , will be serially correlated. Therefore, in order to do estimation and inference in those models we could follow the usual two alternatives: *first*, use the autocorrelation consistent variance-covariance matrix of the errors, see for example Newey and West (1987) and Andrews (1991) if the estimators are consistent , or *second*, include extra lags of the variables of the model, as in the usual error correction literature, see Engle and Granger (1987) and Hendry (1995). When the explanatory variables are in levels, like the error correction terms (M_{t-1} - LRD_{t-1}), the lagged variables are highly correlated and therefore, it is usually enough to include one or two lags of them. However, to determine the correct number of lags of the other stationary variables we could follow a standard general-to-specific rule, see for example the applications of NEC modeling applied to different macroeconomic and financial variables of Escribano (1986), Hendry and Ericsson (1991), Granger and Lee (1989), Escribano and Granger(1998) and Granger (1998), Escribano and Pfann (1998) and Teräsvirta and Eliasson (2001).

If we do this then equation (3.8) will have the following *nonlinear error correction* (NEC) representation with white noise errors (ε_t) independent of the explanatory variable (ΔLRD_t),

$$\varphi(L)\Delta M_t = F(LRD_t,\beta)(M_{t-1} - LRD_{t-1}) + \theta(L)g(\Delta LRD_t, \alpha) + \varepsilon_t \quad (3.10)$$

where $\varphi(L)$ and $\theta(L)$ are finite order polynomials in the lag operator, $L^k x_t = x_{t-k}$, with all the roots of $|\varphi(z)| = 0$ and $|\theta(z)| = 0$ outside the unit circle. Given that LRD_t is I(1) and generated by

(3.4) then the first term of the RHS of (3.9) is $F(LRD_t,\beta)$ $(M_{t-1} - LRD_{t-1}) = F(LRD_{t-1}+ \Delta LRD_t,\beta)$ $(M_{t-1} - LRD_{t-1})$.

General NEC approximations are available. Consider for example consider the following two cases:

(i) When $F(LRD_{t-1} +\Delta LRD_t, \beta)(M_{t-1} - LRD_{t-1}) \cong f(M_{t-1} - LRD_{t-1},\gamma)$, from (3.9) we get the following NEC model,

$$\varphi(L)\Delta M_t = f(M_{t-1} - LRD_{t-1}, \gamma) + \theta(L)g(\Delta LRD_t, \alpha) + \varepsilon_t. \qquad (3.11)$$

(ii) When $F(LRD_{t-1} +\Delta LRD_t, \beta)(M_{t-1} – LRD_{t-1}) \cong f_1(M_{t-1} – LRD_{t-1},\gamma_1)$ $+ f_2(\Delta LRD_t\,(M_{t-1} – LRD_{t-1}),\gamma_2)$ from (3.9) we get an alternative expression for the NEC model,

$$\varphi(L)\Delta M_t = f_1(M_{t-1} - LRD_{t-1},\gamma_1) + f_2(\Delta LRD_t(M_{t-1} - LRD_{t-1}),\gamma_2)+ \theta(L)g(\Delta LRD_t, \alpha) + \varepsilon_t. \qquad (3.12)$$

These are nested in the class of general NEC models analyzed by {} which have the form of *nonlinear error correction* (NEC) models introduced by Escribano (1986, 1997). However, in the empirical sections of this paper we restrict ourselves to nonlinear approximations of the types considered in Equations (3.8) and (3.11).

From Escribano and Mira (2002) a key condition for having a type of Granger´s Representation Theorem, see Engle and Granger (1987), in nonlinear models with linear cointegrated variables, (equation (3.1c)), is that the slope of the error correction term, $(M_{t-1}$ - $LRD_{t-})$, be bounded $-2 < d[F(LRD_t,\beta)(M_{t-1} - LRD_{t-1})] / d(M_{t-1} - LRD_{t-1}) < 0$, where d[.]/d(.) is the slope of the nonlinear error correction.

Cubic polynomials can be justified as simple parametric approximation to any unknown functions which are fourth order continuously differentiable by using Taylor series approximations. They can also help in selecting between alternative functional forms, see Teräsvirta (1994) and Escribano and Jorda (2001). However, cubic polynomials are unbounded functions and they do not satisfy the sufficient conditions introduced by Escribano and Mira (2002) for having a NEC representation theorem. This problem was solved by Escribano (1997) using certain types of rational polynomials bounded by increasing linear functions and therefore satisfying the sufficient conditions.

Escribano and Mira (1997b) gave sufficient conditions for consistency and asymptotic normality of the nonlinear least squares (NLS) estimator of the parameters of models (3.10) and (3.11). Both NEC models are based on the linearity (log-linear) of the cointegration relationship with a nonlinear (or time varying) error correction adjustment (NEC). It is possible to extend the analysis to nonlinear cointegration relationships, along the lines of Escribano and Mira (1997a) and Saikkonen and Choi (2000), using NLS estimators of the unknown cointegrating parameters but this is out of the scope of this paper.

4. Forecast Encompassing Tests and Pooled Model Forecasts

We will be using coefficients as estimated through time t to generate a forecast for time t+4. The usual encompassing test takes a series of such forecasts from two models $\hat{y}_1, \hat{y}_2$ and estimates a regression equation of the form

$$y_t = b_0 + b_1\hat{y}_{1t} + b_2\hat{y}_{2t} + \varepsilon_t. \tag{4.1}$$

An encompassing test of $b_1 = 0$ may or may not impose $b_0 = 0$ and/or $b_2 = 1$. But if one accepts $b_1 = 0$ and rejects $b_2 = 0$ then model two encompasses model one. In the tests applied below we follow Andrews, Minford and Riley (1996) in imposing $b_2 = 1$ but leaving b_0 unrestricted in testing the null $b_1 = 0$.[2]

There are two problems with such an encompassing test. First, if the sample is large enough then at conventional test size model two could encompass model one even if the difference in forecast performance is small. If both competing models are linear then this is not a nuisance. But suppose model one is linear and model two is non-linear, includes more coefficients to estimate and requires a grid search to ensure estimated coefficients minimize the sum of squared residuals. Then the fact that model two statistically encompasses model one is not enough to convince us of the practical superiority of model two.

Second, the joint model may not be stable. This is of particular importance if the model forecasts are complements, i.e. suppose one can reject $b_1 = 0$ (given $b_2 = 1$) and can also reject $b_2 = 0$ (given $b_1 = 1$). In such a case it is tempting to conclude the joint model of Equation 4.1 with unrestricted coefficients is the better forecasting model. But in formal encompassing tests Equation 4.1 is used to estimate coefficients ex post, the regression is not used to forecast ex ante. Results can differ if one uses the joint model to forecast. Most importantly, the rmse of one-step-ahead forecasts from 4.1 can be larger than those of the individual model forecasts.

We propose a check on standard ex-post encompassing tests which also provides a measure of the size of forecast improvement. In this check there are three quantities to be compared. The first two quantities are the root-mean-squared forecast errors of the individual

[2] For an introduction to forecast encompassing tests see Charemza and Deadman (1997). A more current review can be found in Newbold and Harvey (2002).

models. The third uses Equation 4.1 to generate true forecasts. In particular, obtain coefficient estimates ($\hat{b}_{0,T}, \hat{b}_{1,T}, \hat{b}_{2,T}$) using the forecast data through time T. Then use these estimates and the pooled model to forecast one-step-ahead:

$$\hat{y}_{T+1} = \hat{b}_{0,T} + \hat{b}_{1,T}\hat{y}_{1,T+1} + \hat{b}_{2,T}\hat{y}_{2,T+1} \tag{4.2}$$

The rmse of the pooled-model forecasts (from Equation 4.2) is then compared to that of the individual models. Constructing these ex ante forecasts provides a check on the performance of the pooled model.[3] It also can help the reader to consider the "economic" significance of the statistical tests. If one individual model statistically encompasses others but this model is particularly costly to estimate or use then the difference in performance may not be large enough in magnitude to justify using the apparently superior model. This is particularly the case when comparing linear and non-linear models or when comparing forecasts which must be commercially purchased.

5. Empirical Money Demand Application of the Nonlinear Error Correction Model

Several flexible parametric functions can be used in practice. Escribano (1986) estimated the first NEC model using of cubic polynomials in (M_{t-1} - LRD_{t-1}) while Escribano (1997) considered also rational polynomials. Greene (1999) considered the following exponential approximation of Equation (3.1b), b_t = $[\exp(b_0/LRD_t^2)]$, which corresponds to a particular parametric selection of the nonlinear functions F(.) and g(.) in (3.9). Other alternative parametric functions are piece-wise linear adjustments, see Granger and Lee (1989), threshold adjustments, see Balke and Fomby (1997), bilinear adjustments as in Peel and Davidson (1998) or smooth transition regressions as in van Dijk and Franses (2000). Non parametric approaches, like kernel or smoothing splines are also possible to implement with NEC models, see Escribano (1986, 1997), but are out of the scope of this paper.

Empirical Model Specification for Fourth-Step-Ahead Forecasts

The cointegrating relationship is estimated via a simple (log-linear) levels regression of real US M1 per household on a constant, real GDP per household and a long-term interest rate. Previous researchers such as Cutler, Davies, Rhodd and Schwarm (1997) and Swanson (1998) have found cointegration for similar variables.[4] Denoting the fitted values from the levels regression as $\overline{M}_t$ then $\overline{M}_t$ is our estimate of LRD_t in the above equations.

Results will be presented for models designed to generate fourth-quarter-ahead forecasts. The dependent variable is the change over four quarters, $\Delta_4 m_t = m_t - m_{t-4}$. Likewise the first lag of the dependent variable included in the regression is $\Delta_4 m_{t-4} = m_{t-4} - m_{t-8}$. Aggregate money from m_t back through m_{t-3} is taken to be unknown (future) in the model specification. We suppose that the path of non-dependent regressors is known, so $\Delta_4 \overline{M}_t = \overline{M}_t - \overline{M}_{t-4}$ is also included. And the error-correction mechanism is $m_{t-4} - \overline{M}_{t-4}$.

Alternatively, one could use the usual one-step-ahead model (with first lags of all variables) to recursively forecast the dependent t+1, t+2, t+3 and finally t+4 periods ahead. Bhansali (2002) shows there is no difference between the two approaches *if* the lag-truncation

---

[3] Models are always simplifications of the true data generation process. Thus they should be regarded as miss-specified and unstable a priori. The question is whether such miss-specification is severe enough to matter, in this case, is ex post estimation a good guide to ex ante performance?
[4] For our data the Phillips-Perron test (with four lags) for a unit root in the error-correction term is rejected at a p-value of less than one-percent.

and non-dependent regressors included precisely correspond to the actual process generating the data. But if we must regard models as simplifications of the actual process (most obviously by omitting infinite lags or other relevant variables) then using the one-step model for longer time-horizon forecasts will result in larger expected prediction errors. Unless we have absolute confidence in the specification, it is better to specify the fourth-step-ahead problem as a separate modeling exercise, and this practice is followed here. To summarize the variables employed:

$\Delta_4 m_t$ =	4-quarter changes of real money per household.[5]
$(m_{t-4} - \overline{M}_{t-4})$	Error-correction term given information available at time t-4.
$\Delta_4 \overline{M}_t$	4-quarter changes of independent variable (long-run demand).

For each model type we begin with four lags of all variables (except the error-correction term) and then simplify via a general-to-specific approach, dropping first those variables which contribute least to sample fit.[6] Final lag truncation is determined by four criteria, potentially generating four variants. First, we drop variables to minimize the SER (standard error of regression). Second, we minimize the AIC (Akaike information criterion). Third, we continue to drop variables until all are individually significant for a nominal size of five-percent (via an analysis-of-variance F-test, not local estimates of t-statistics). Fourth, we continue to drop variables until the implied joint restriction is rejected for a nominal test size of ten-percent. These criteria are listed in order of complexity, there at least as many lags retained in the specifications which minimize the SER as there are lags retained in those which minimize the AIC.

[5] As detailed in the Data and Estimation Appendix, we impose the nominal adjustment mechanism.
[6] Since the first known lag for the change in money is $m_{t-4} - m_{t-8}$, the fourth lag is $m_{t-8} - m_{t-12}$.

Model Types

It is helpful to think of the non-linear models as divided into two broad classes. The first class models the variation in the weight on the error-correction term as in Equations 3.1b and 3.8. Here we have two contenders, powers of a rational-polynomial and the exponential approximation previously employed in Greene (1999). These models are non-linear in LRD or $\overline{M}_t$ (but not in the error-correction term) and correspond most faithfully to the monetary theory outlined above. The exponential model is simply an approximation of the more heavily parameterized rational-polynomial. These models will be denoted "Exp" or "Rat" respectively for exponential and rational-polynomial.

The smooth-transition model is non-linear in the error-correction term and thus constitutes a second class of model. Although not rigorously derived from transactions theory, heuristic arguments have been used by Sarno (1999) and Sarno, Taylor and Peel (2002) to tie it to the complex dynamics induced by trigger-target rules as modeled in Greene (2001). This class of models is increasingly being used in the applied literature and represents an important competitor.

Finally we should note that the smooth-transition model includes the linear model (with a constant weight on the error-correction term) as a sub-case. So within sample these must fit better than linear models. But the exponential models (as in Equation 3.1) do not include the linear sub-case.[7]

[7] In principle the rational-polynomial model also admits the linear sub-case. But in order to achieve convergence of non-linear least-squares we were forced to simplify and restrict the form of Equation 3.8. More detail is in the Data and Estimation Appendix.

6. Empirical Results

Table 1 shows estimation results for examples (marginal 5% lag selection) of these fourth-step-ahead models. Note the weight on the error-correction term for the linear model is more than 0.25 where 0.05 is typical for quarterly one-step-ahead models. The p-value of the reset statistic is significant for the smooth-transition model (0.012) and (depending upon tastes in test size) possibly for the rational-polynomial model (0.077). The smooth-transition model appears to improve on the SER of the linear model by about 5%, but as we will see this advantage does not carry-over as strongly into the ex-ante forecasts. Likewise the ex-post fit of the linear model (SER 200) is tighter than that of the rational-polynomial model (SER =211), but this will also not carry over to ex-ante forecasting. This implies there is some instability in these models of practical importance.

Table 2 shows the formal (ex post) encompassing tests. The first rows compare the linear and non-linear models. Here the linear model is usually encompassed by the non-linear models (accept zero weight on linear forecasts, reject zero weight on alternative model). For instance, the first entries show the tests comparing the Min SER version of the linear and smooth-transition model. A null of no weight on the linear model forecasts can be accepted with a p-value of 0.734. But a null of no weight on the smooth-transition model can be rejected (p-value of 0.047). This continues across the rows. Where the linear model is not encompassed, neither is the alternative non-linear encompassed so the models are complements. The only exception is for the smooth-transition model. For the min AIC and marginal 5% lag criteria the linear and smooth models are close substitutes (p-values are greater than 10%).

The second and third sets of rows of Table 2 compare the non-linear models. For a test size of 10% then in most cases neither model dominates and both contribute information not in the other model's forecasts, i.e. they are complements as judged by these ex post tests. This might imply that a joint forecasting model employing pairs of non-linear model forecasts will improve upon the individual forecasts But recall these tests are ex post. Results of actual ex ante joint model forecasts are presented in the other tables.

Table 3 presents the first set of ex ante results. The second column compares individual models to the linear variants. Although the smooth model ex post SER was about 5% less than that of the linear model (in Table 1), here it forecasts only 1.5-2.4% more tightly than the linear model. In contrast, the exponential model forecasts improve on the linear (with the exception of the joint 10% version) by a more substantial 7-9%.[8] The rational-polynomial models also improve upon their linear counterparts by 3.8-7.1%.

The third column of Table 3 shows the results of using the linear and non-linear forecasts in a joint model to forecast one-step-ahead (as in Equation 4.2). In most cases the joint models forecast only slightly more accurately (1.7% for the min SER smooth-transition and linear forecasts pooled in a joint model) or forecast less accurately (-1.9% for the min AIC versions of the same models). The only exceptions are for the joint 10% version of the exponential model (17.6% reduction in rmse) and the min SER version of the rational-polynomial model (13.9% reduction). These exceptions are intriguing since these joint models forecast with an rmse smaller (206 and 204) than that for any other single- or joint-model based forecasts. But the joint linear/nonlinear models do not consistently perform better than the individual model forecasts, implying the evidence for a joint model is weak.

[8] The exponential joint 10% model drops only one variable from the marginal 5% model. The implied marginal restriction is rejected with a p-value of 0.1%. Nonetheless we held to the mechanical exercise and dropped the variable since the joint restriction implied in dropping variables up until that point was not rejected.

In Table 4 the rmse of the smooth-transition models are compared to the other non-linear models. The second column shows that with the exception of the joint 10% version of the exponential model, the rmse of the other individual models is smaller than that of the smooth-transition models. As in Table 3, the third column of Table 4 shows the joint ex ante forecasting models do not consistently perform better than the individual models. Hence there is little reason to consider the smooth-transition forecasts as complements to the other non-linear forecasts.

Finally, Table 5 compares the exponential and rational-polynomial models. Again with the exception of the joint 10% version, the exponential model forecasts more tightly than the rational-polynomial model (most of the entries in the second column are negative). But the differences are not large. Likewise the joint model consistently improves upon the individual models, but only by 1.8-5.9%. We suspect for most purposes this improvement is not large enough to justify estimating both non-linear models.[9]

Ex Post Tests vs. Joint Forecasts

In Table 2 there are a number of cases in which both model forecasts are significant (statistically) in the formal tests. Consider first the middle columns which compare the exponential model to the linear and smooth-transition. There are five cases in which the p-values for both the first and alternative models are 5% or less. In order down the columns, they are the linear min AIC and joint 10% variants, and the smooth-transition min SER, min AIC and joint 10% variants. The third column of Table 3 shows that the ex ante forecasts are not superior for the corresponding model which uses the min AIC linear and exponential forecasts in a joint model. Table 4 shows that corresponding joint exponential and smooth-transition model also

[9] And recall that the exponential form is an approximation of the more general rational-polynomial form.

does not improve on the individual models ex ante. Turning back to the third set of columns in Table 2, there are six cases in which the formal tests imply the models are complements. Only for half of these do we find a smaller joint model rmse in the ex ante forecasts (of Tables 6 and 7). So the formal ex post tests are not a reliable predictor of the usefulness of the ex ante joint model forecasts.

Comparing the Coefficient Variation of the Nonlinear Models

All these non-linear models posit variation in the weight placed upon the error-correction term. The smooth-transition model allows this weight to have a linear component plus a component which is a function of the error-correction term itself, i.e. the difference between actual money and the static prediction/long-run demand. The rational-polynomial and exponential models posit this weight as a function of long-run demand. There are two useful ways to consider the variation in this weight. First consider Figure 1, which shows the variation in this weight over time as implied by the marginal 5% versions of the 4th-step-ahead non-linear models (as estimated in Table 1). The linear weight is constant at -0.28. The smooth-transition model implies the most variation in the coefficient or weight. On average over time this weight is farther from zero than for the other models.

The exponential and rational-polynomial models represent the same underlying theory and in some respects behave in a similar manner. The weight on the error-correction term trends away from zero until about 1981.4 (observation 135), at which time growing long-run demand implies a trend back towards zero. In contrast, the weight assigned by the smooth-transition model has less of a trend but often moves on average counter to the other non-linear model weights. For instance for observations 95-130 (1972-80) the smooth-transition weight moves towards zero while the exponential model weight is moving away from zero.

Finally, note that after 1969 (approximately observation 85) the rational-polynomial weight settles at close to the fixed weight of the linear model. In general the variation in weight of the exponential model is greater than that of the rational-polynomial. Part of the explanation for the better performance of the exponential model lies in this greater differentiation from the constant linear weight.

Figures 2 and 3 show the functional variation in weights. The smooth-transition model allows for variation as a function of the error-correction term, while the driving variable for the other models is long-run demand ($\overline{M}_t$). Given the sample variation in the driving variables, the smooth-transition and exponential models imply weights that vary (respectively) from -0.2 to -0.6 and from -0.2 to -0.48. While the variation for the rational-polynomial model is from about -0.2 to -0.3.

Returning to the second column of Table 3, though the smooth-transition model implies the most extreme variation in the error-correction weight it improves upon the linear forecast rmse by only 1.5-2.4%. In contrast, even the milder coefficient variation (-0.2 to -0.3) of the rational-polynomial models improve upon the linear forecast rmse by 3.8-7.1%. Since the weight changes more slowly for the rational-polynomial and exponential models, it is not as crucial to get the specification and timing correct. But if the wide and sudden swings in weight implied by the smooth-transition model (as in Figure 1) are correct in principle, then it becomes very important to get the timing correct. The smooth-transition model fits more tightly than the other models when estimated over the whole sample (SER = 191 in Table 1) but improves upon the linear model only slightly when forecasting ahead and does not improve upon the other non-linear models. It appears it is not possible to correctly model the extreme changes in optimal smooth-transition coefficient weights with the correct timing and magnitude except ex-post.

7. Conclusions

Economic theory implies non-linear models, and the statistical theory for empirical estimation allows for applications using non-linear error-correction models. As a group the non-linear models examined here forecast with smaller (root-mean-squared) errors than linear models. In the statistical tests the non-linear models often encompassed the linear, but the linear never encompassed the non-linear models. Although the smooth-transition model fits better than other models when estimated ex post, it does not forecast nearly as well. This is likely due to the wild coefficient swings implied by the model, which must be forecast precisely to be useful.

Whether the gains of non-linear models are worth abandoning the convenience of linear models will depend on the consumer's purposes and constraints. But the exponential model is not difficult to estimate. Since it does not admit the linear sub-case, its superior performance was not a foregone conclusion in this study. And it is formally tied to monetary theory.

In the formal encompassing test it often appears that forecasts are complements, implying the superior performance of a joint model. But such ex post results have not carried over into ex ante forecast performance. It may be that conventional test sizes are not rigorous enough. But a useful check is to be had in actually constructing the ex ante joint forecasts.

Literature Cited

Anderson, R.G. and Rasche, R.H. (2001), "Retail Sweep Programs and Bank Reserves, 1994-1999", *Federal Reserve Bank of St. Louis Review*, 83(1),51-72

Andrews, M. J., Minford, A.P.L. and Riley, J. (1996), "On Comparing Macroeconomic Models Using Forecast Encompassing Tests", *Oxford Bulletin of Economics and Statistics*, 58, 279-305

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica*, 59, 817-58

Balke N.S. and Fomby T.B. (1997) "Threshold Cointegration". *International Economic Review*. Vol. 38, 3.

Bhansali, R. J. (2002), "Multi-Step Forecasting", in *A Companion to Economic Forecasting*, Clements, M. P. and Hendry, D. F. (Eds), Oxford: Blackwell Publishers, pp206-221

Baumol, W.J. (1952), "The Transactions Demand for Cash -An Inventory Theoretic Approach", *Quarterly Journal of Economics*, 545-556

Cabellero, R. and Engle, E. (1991), "Dynamic (S,s) Economies", *Econometrica*, 59(6), 1659-86

Caplin, A. S. (1985), "The Variability of Aggregate Demand with (S, s) Inventory Policies", *Econometrica*, 53, 1395-1410

Charemza, W.W. and Deadman, D.F. (1997), *New Directions in Econometric Practice*, Second Edition, (Aldershot, Edward Elgar Publishing Ltd.)

Coenen, G., Levin, A. and Wieland, V. (2001), "Data Uncertainty and the Role of Money as an Information Variable for Monetary Policy", European Central Bank working paper #84

Cutler, Davies, Rhodd and Schwarm (1997), "The Demand for M1 in a Large Macroeconomic System: Evidence from Cointegration Analysis", *Journal of Macroeconomics*, 19(1), 53-78

Dolado, J.J., J. W. Galbraith and A. Banerjee(1991) "Estimating Intertemporal Quadratic Adjustment Cost Models with Integrated Series" International Economic Review, vol. 32, pp. 919-936.

Dickey, D.F. and Fuller, W. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74(366), 427-31

Engle, R. F. and C.W.J. Granger (1987) "Cointegration and Error Correction: Representation, Estimation and Testing". *Econometrica*, 55 (2). 251-276.

Escribano A. and O. Jorda (2001) "Testing Nonlinearity: Decisión Rules for Selecting Between

Logistic and Exponential STAR Models" *Spanish Economic Review*, 3, 193-209.

Escribano, A. (1986), "Nonlinear Error Correction: The Case of Money Demand in the U.K. (1878-1970)". Chapter IV, Ph.D. Dissertation, University of California San Diego.

Escribano, A. (1997), "Nonlinear Error Correction: The Case of Money Demand in the U.K. (1878-1970)". W.P. series , Universidad Carlos III de Madrid.

Escribano, A. and C.W.J. Granger (1998), "Investigating the Relationship Between Gold and Silver Prices". *Journal of Forecasting*, 17, 81-107.

Escribano, A. and Mira, S. (2002), "Nonlinear Error Correction Models". *Journal of Time Series Analysis*, 23, 409-522.

Escribano, A. and Mira, S. (1997a), "Nonlinear Cointegration with Mixing Errors". Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, Working Paper 97-25.

Escribano, A. and Mira, S. (1997b), "Nonlinear Error Correction Models". Departamento de Estadística y Econometría, *Universidad Carlos III de Madrid, Working Paper* 97-26.

Escribano, A. and Pfann, G. (1998), "Non-linear Error Correction, Asymmetric Adjustment and Cointegration". *Economic Modelling*, 15(2), 197-216

Greene, C. A. (1999), "Non-Linear Money Demand: Transactions-Theory Based Results with an Empirical Demonstration", mimeo, Department of Economics, University of Missouri.

Greene, C. A. (2001), "Trigger-Target Rules and the Dynamics of Aggregate Money Holdings", *Journal of Economic Dynamics and Control*, 25, 1193-1219

Gregory A.W., A. Pagan and G. Smith (1990) "Estimating Linear Quadratic Models with Integrated Processes". *Rochester Center for Research, working paper* 247, 1-34.

Hansen, B.E. (2000), "Testing for Structural Change in Conditional Models", *Journal of Econometrics*, 97, 93-115

Hansen, L.P. and T.J. Sargent(1980) "Formulating and Estimating Dynamic Linear Rational Expectations Models", *Journal of Economics Dynamics and Control*, 2, 7-46.

Harvey, A. C. (1990), *The Econometric Analysis of Time Series*, 2nd ed., (Cambridge Massachusetts, MIT Press)

Kennan, J. (1979) "The Estimation of Partial Adjustment Models with Rational Expectations", *Econometrica*, 47, 1441-1455.

Lucas, R. E. (1988), “Money Demand in the United States: A Quantitative Review”, *Carnegie-Rochester Conference Series on Public Policy*, 29(0), 137-67

Miller, M.H. and Orr, D. (1966), “A Model of the Demand for Money by Firms”, *Quarterly Journal of Economics*, 413-435

Newbold, P. and Harvey, D. I. (2002), “Forecast Combination and Encompassing”, in *A Companion to Economic Forecasting*, Clements, M.P. and Hendry, D.F., (eds), (Oxford, Blackwell Publishers Ltd)

Newey, W.K. and West, K.D. (1987), “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”. Econometrica 55, 703-8.

Nickell, S. (1985) “Error Correction, Partial Adjustment and All That: An Expository Note”. *Oxford Bulletin of Economics and Statistics*, 47, 119-129.

Salmon, M. (1982), “Error Correction Mechanisms”, *Economic Journal*, 92, 615-629

Sarno, L. (1999), "Adjustment Costs and Nonlinear Dynamics in the Demand for Money: Italy, 1861-1991", *International Journal of Finance and Economics*, 4, 155-177

Sarno, L., Taylor, M.P, and Peel, D. (2002), "Non-Linear Equilibrium Correction in US Real Money Balances, 1869-1997", *Journal of Money, Credit and Banking*, (Forthcoming)

Swanson, N. R. (1998), “Money and Output Viewed through a Rolling Window”, *Journal of Monetary Economics*,

Teräsvirta, T. (1994) “Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models” *Journal of the American Statistical Association*, 89, 208-218.

Teräsvirta, T. and A. Eliasson (2001) “Nonlinear Error Correction and the U.K. Money Demand for Broad Money, 1878-1993”. *Journal of Applied Econometrics*, 16, 277-288.

Van Dijk, D., and Ph. H. Franses (2000) “Nonlinear Error-correction Models for Interest Rates in the Netherlands”, *in Nonlinear Econometric Modelling in Time Series*, eds. W. Barnett et. al., Cambridge University Press, Chapter 8, 203-227.

Vickson, R. G. (1985), "Simple Optimal Policy for Cash Management: The Average Balance Requirement Case", *Journal of Financial & Quantitative Analysis*, 20(3), 353-369

Data and Estimation Appendix

Data begins 1947.1. The period after the introduction of interest-bearing checking accounts involved serious transitional effects. We don't want to be accused of designing models specialized to this period. So we excluded estimation of the dependent variable 1982.1-1987.3. Allowing for the lags of the unrestricted 4-lag versions of the models, this means we begin to employ data with a date of 1985.1, so only data 1982.1-1984.4 is entirely excluded from the estimation. All data is seasonally adjusted Money and income are measured as real per household. This removes the common but economically meaningless trend created by population growth (replication). The cointegrating levels regression is estimated via a standard log-linear static levels (Engle-Granger) regression. Linear and non-linear models are estimated in non-logarithmic levels. For the linear model White's heteroskedasticity test (1-4 lags) does not reject homoskedastic errors. Phillips-Perron test rejects unit root for the dependent variable (1% critical value -2.578, empirical value is -3.946).

Original data for M1 is monthly and in billions.

For the years 1947.1-60.4 M1 is from Table 2, Section 1; *Supplement to Banking & Monetary Statistics*, Board of Governors of the Federal Reserve System (1962). Denote this series here as M1A.

For the years 1961.1-1996.1 from Federal Reserve Bank of Saint Louis, Historical data files, in "FRED" (http://www.stls.frb.org/) as of June 1997, denoted there as "M1 Money Stock". Denote this series here as "M1B". These monthly figures for M1A and M1B are converted to quarterly averages. Then let M1(1947.1-1960.4) = M1A -0.166667 and
M1(1961.1-1996.1) = M1B. Note M1B(1960.4) = M1A(1960.4) -0.166667

HOUSEHOLDS:

Original data is annual in thousands, an estimate for March of each year so the figures are entered for the first quarter of the corresponding year. Other quarters are calculated as a simple linear interpolation. Thus the calculation for the second quarter of each year is
$H = 0.75H_{\text{year } t} + 0.25H_{\text{year } t+1}$.
Household data is from Table A-2; *Household and Family Characteristics, March 1994, Population Characteristics, Current Population Reports, P20-483* (Issued Sept. 1995), U.S. Department of Commerce.

GDP Deflator:

From Summary National Income and Product Series, 1929-96, Table 3; *Survey of Current Business*, U.S. Dept. of Commerce, May 1997.

Dependent Variable for Engle-Granger Levels Regression:

M = (nominal M1 per household)/(GDP deflator)

Dependent Variable for One-Step-Ahead Error-Correction Models:

Δm_t = Δ(nominal M1 per household)/(GDP deflator at time t)

This (along with the form of the error-correction term discussed below) imposes the nominal adjustment mechanism. Previous experience shows this results in simpler dynamics and fewer lagged variables needed to ensure uncorrelated residuals.

Independent Variables:

Y = REAL GDP per household:

Figures for real GDP are from Summary National Income and Product Series, 1929-96, Table 2A; *Survey of Current Business*, U.S. Dept. of Commerce, May 1997.

R = Long-Term Bond Rate:

From Federal Reserve Bank of Saint Louis, Historical data files, in FRED (http://www.stls.frb.org/) as of June 1997, denoted as "LTGOVBD, Long-Term U.S. Government Bond Yield (10 years or more), Including Flower Bonds, Average of Daily Figures". Monthly figures converted to quarterly averages.

Long-run Demand and Error-Correction:

Let $\overline{M}$ = the fitted values from the least-squares regression $M = b_0 Y^{b1} R^{b2} + e$. Then to impose the nominal-adjustment mechanism let the error-correction term

ECM_t = [(nominal M1 per household at time t-1)/(GDP deflator at time t) - $\overline{M}_{t-1}$].

The set of regressors employed are Δm, $\Delta\overline{M}$, and the ECM.

Fourth-Step-Ahead Models:

The dependent variable is nominal money per household minus the same lagged four quarters, deflated by the current GDP deflator. Changes in long-run demand are measured as $\overline{M}_t - \overline{M}_{t-4}$. The error-correction term is [(nominal M1 per household at time t-4)/(GDP deflator at time t) - $\overline{M}_{t-4}$]. In conjunction with the definition of the dependent variable this again imposes the nominal-adjustment mechanism.

Comments on Estimation:

Even taking b_t as fixed, the model of Equation 3.7 is not a standard error-correction model. As exposited in Hendry and Richard (1983) this is the error-correction model with dynamics restricted to partial-adjustment. To make 3.7 directly comparable to a linear error-correction model we must allow the second term take a different value: $\Delta M_t = (b_{1t} - 1)(M_{t-1} - LRD_{t-1}) + (1 - b_{2t})\Delta LRD_t + u_t$. We began estimation of the exponential version with separate b_{1t} and b_{2t} and tested the restriction $b_{1t} = b_{2t}$ as part of the general-to-specific lag selection. For one-step-ahead models this was accepted but this was rejected for the 4th-step-ahead models, hence the two terms in Table 1 for the exponential model.

Estimating the rational-polynomial model with b_t as in Equation 3.8 proved to be very difficult. Even with grid search restrictions were required. We began with all the possible permutations of restrictions ($B_2 = B_1 = 0$, $B_2 = B_1 = B_5 = 0$ etc) and picked the most complex which could be reliably estimated. These considerations forced us to impose $b_{1t} = b_{2t}$. The coefficients were restricted to be non-negative, except for B_6 of Equation 3.8 which was restricted to be ≥ 1.

Table 1
Fourth-Step-Ahead Models (Marginal 5% Lag Selection) Estimated 1949.4-1995.1

N = 159 $\Delta_4 m_t$: (mean, σ) = (255, 299)

Linear:

$$\Delta_4\hat{m}_t = -0.282(m_{t-4} - \overline{M}_{t-4}) + 1.29\Delta_4 m_{t-4} - 1.11\Delta_4 m_{t-5} + 1.48\Delta_4 m_{t-8} + 0.295\Delta_4 \overline{M}_{t-1}$$

R^2 =0.564, SER =200, AR(5th-8th) p-value =0.978, ARCH(5th-8th) p-value =1.00
Reset test (terms, p-value) =($\Delta\hat{m}^2$ $\Delta\hat{m}^3$ $\Delta\hat{m}^4$, 0.116), (max-F, 10% critical value) =(12.1, 16.2)

Smooth Transition:

$$\Delta_4\hat{m}_t = b_t(m_{t-4} - \overline{M}_{t-4}) + 1.15\Delta_4 m_{t-4} - 0.883\Delta_4 m_{t-5} + 1.29\Delta_4 m_{t-8} + 0.292\Delta_4 \overline{M}_{t-1} - 0.074\Delta_4 \overline{M}_{t-7}$$

$$b_t = -0.623 + 0.423(1 - \mathrm{EXP}[-0.263(10^{-5})(m_{t-4} - \overline{M}_{t-4})^2])$$

R^2 =0.611, SER =191 AR(5th-8th) p-value =0.888, ARCH(5th-8th) p-value =1.00
Reset test (terms, p-value) =($\Delta\hat{m}^2$ $\Delta\hat{m}^3$ $\Delta\hat{m}^4$, 0.012), (max-F, 10% critical value) = (6.34, 18.1)

Exponential:

$$\Delta_4\hat{m}_t = (b_{1t} - 1)(m_{t-4} - \overline{M}_{t-4}) + (1 - b_{2t})\Delta_4 \overline{M}_t + 1.33\Delta_4 m_{t-4} - 1.15\Delta_4 m_{t-5} + 1.57\Delta_4 m_{t-8}$$

$$-0.705\Delta_4 \overline{M}_t + 0.271\Delta_4 \overline{M}_{t-1}$$

$$b_{1t} = \mathrm{EXP}(-34.6(10^6)/\overline{M}^2_t) \quad b_{2t} = \mathrm{EXP}(-157(10^6)/\overline{M}^2_t)$$

R^2 =0.575, SER =199 AR(5th-8th) p-value =0.960, ARCH(5th-8th) p-value =1.00
Reset test (terms, p-value) =($\Delta\hat{m}^2$ $\Delta\hat{m}^3$ $\Delta\hat{m}^4$, 0.138), (max-F, 10% critical value) = (8.54, 19.7)

Rational Polynomial:

$$\Delta_4\hat{m}_t = (b_t - 1)(m_{t-4} - \overline{M}_{t-4}) + (1 - b_t)\Delta_4 \overline{M}_t$$

$$+1.32\Delta_4 m_{t-4} - 1.10\Delta_4 m_{t-5} + 1.50\Delta_4 m_{t-8} + 0.0787\Delta_4 \overline{M}_{t-1} - 0.0636\Delta_4 \overline{M}_{t-7}$$

$$b_t = [1 - (\overline{M}_t /(1.02*10^9 + 0.00116\overline{M}^3_t))]^{7.03*10^4}$$

R^2 =0.525, SER =211 AR(5th-8th) p-value =,0.790 ARCH(5th-8th) p-value =1.00
Reset test (terms, p-value) =($\Delta\hat{m}^2$ $\Delta\hat{m}^3$ $\Delta\hat{m}^4$, 0.077), (max-F, 10% critical value) = (6.37, 19.7)

Notes: Statistics include the Breusch-Godfrey AR test, the Engle LM ARCH test (the obs*R^2 versions asymptotically distributed $\chi^2(.)$), and the Ramsey Reset test. Critical values for the sup-F test are from Andrews (1993). If correctly specified then fourth-step-ahead models will have AR(4) errors, but not AR(4+k) for integer k ≥1. So we exclude lags one to four in the AR and ARCH tests.

Table 2
Fourth-Step-Ahead Encompassing Tests 1960.4-1995.1

First Model	Smooth4 First Encompassed	Smooth4 Alternative Encompassed	Exp4 First Encompassed	Exp4 Alternative Encompassed	Rat4 First Encompassed	Rat4 Alternative Encompassed
Linear4	(P-Val)	(P-Val)	(P-Val)	(P-Val)	(P-Val)	(P-Val)
Min SER	0.734	0.047	0.178	0.010	0.000	0.000
Min AIC	0.496	0.331	0.050	0.016	0.118	0.015
Marginal 5%	0.626	0.107	0.967	0.003	0.048	0.002
Joint 10%	0.685	0.031	0.000	0.000	0.922	0.000
Smooth4						
Min SER			0.041	0.024	0.031	0.007
Min AIC			0.034	0.015	0.077	0.013
Marginal 5%			0.244	0.005	0.077	0.013
Joint 10%			0.000	0.000	0.511	0.000
Exp4						
Min SER					0.794	0.000
Min AIC					0.003	0.001
Marginal 5%					0.001	0.004
Joint 10%					0.056	0.000

Notes: Estimation is from 1949.3. One-step-ahead forecasts are generated 1960.4-1995.1. The encompassing test allows for biased forecasts. In particular the reported p-values are for a null of $b_1 = 0$ from the regression equation $y_t = \hat{y}_{t,2} + b_0 + b_1\hat{y}_{t,1}$ as in Andrews et al (1996).

Table 3
Fourth-Step Ahead Linear vs. Non-Linear Models
Forecast Improvement of Individual and Joint Models 1970.1-1995.1

Model	Individual Model Forecast rmse	Non-Linear Forecast Improvement Over Linear rmse	Joint Model Forecast Improvement Over (Min) Individual rmse
Linear4			
Min SER	254.629		
Min AIC	“		
Marginal 5%	249.906		
Joint 10%	“		
Smooth4			
Min SER	250.460	1.6%	1.7%
Min AIC	245.623	1.7%	-1.9%
Marginal 5%	“	2.4%	-0.1%
Joint 10%	246.007	1.5%	2.9%
Exp4			
Min SER	230.918	9.3%	-3.3%
Min AIC	“	7.6%	-2.4%
Marginal 5%	228.434	9.3%	-5.4%
Joint 10%	276.155	-9.7%	17.6%
Rat4			
Min SER	237.073	6.9%	13.9%
Min AIC	240.489	3.8%	-1.3%
Marginal 5%	“	4.5%	-0.7%
Joint 10%	233.829	7.1%	2.9%

Notes: The forecast series for individual models begins with 1960.4, joint model forecasts begin with 1970.1. Joint model forecasts are calculated as in Equation 4.2.

Table 4
Smooth-Adjustment vs. Alternative Non-Linear Models
4th-Step Ahead Forecast Improvement of Individual and Joint Models 1970.1-1995.1:

Model	Individual Model Forecast rmse	Alternative Model Forecast Improvement Over Smooth rmse	Joint Model Forecast Improvement Over (Min) Individual rmse
Smooth4			
Min SER	250.460		
Min AIC	245.623		
Marginal 5%	“		
Joint 10%	246.007		
Exp4			
Min SER	230.918	7.8%	-2.2%
Min AIC	“	6.0%	-2.9%
Marginal 5%	228.434	7.0%	-5.1%
Joint 10%	276.155	-11.3%	20.3%
Rat4			
Min SER	237.073	5.3%	7.7%
Min AIC	240.489	2.1%	-1.0%
Marginal 5%	“	2.1%	-1.0%
Joint 10%	233.829	5.7%	2.7%

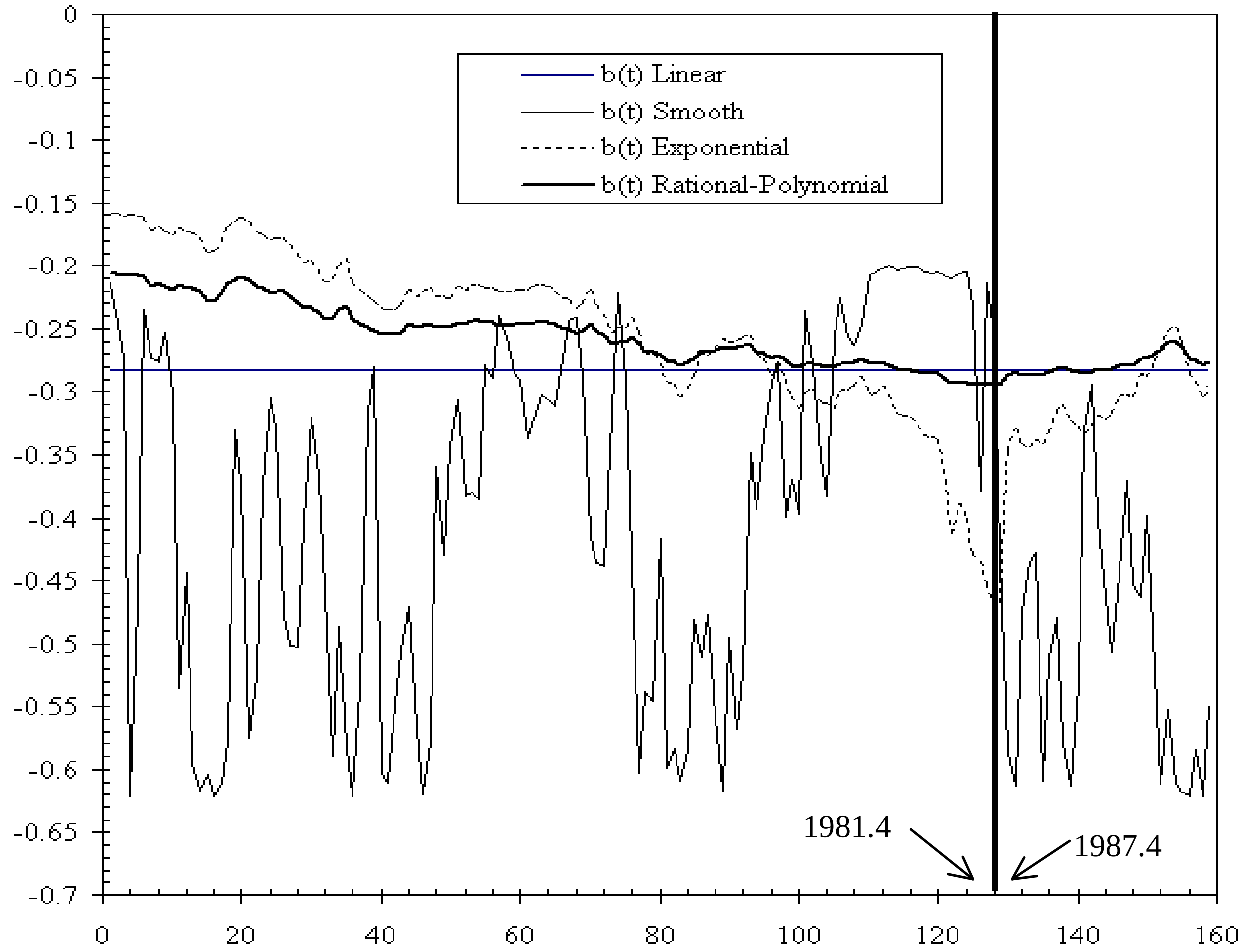
Notes: The forecast series for individual models begins with 1960.4, joint model forecasts begin with 1970.1. Joint model forecasts are calculated as in Equation 4.2.

Table 5
4th-Step Ahead Forecast Improvement of Individual and Joint Models 1970.1-1995.1: Exponential vs. Rational-Polynomial Model

Model	Individual Model Forecast rmse	Alternative Model Forecast Improvement Over Exponential rmse	Joint Model Forecast Improvement Over (Min) Individual rmse
Exp4			
Min SER	230.918		
Min AIC	“		
Marginal 5%	228.434		
Joint 10%	276.155		
Rat4			
Min SER	237.073	-2.7%	5.9%
Min AIC	240.489	-4.1%	1.8%
Marginal 5%	“	-5.3%	3.2%
Joint 10%	233.829	15.3%	4.5%

Notes: The forecast series for individual models begins with 1960.4, joint model forecasts begin with 1970.1. Joint model forecasts are calculated as in Equation 4.2.

Figure 1
Time Variation in Implied Weight on Error-Correction Term for Linear and Nonlinear Models



Notes: Estimation is 1949.4-95.1. For a discussion of the omitted period see Data and Estimation Appendix

Figure 2
Smooth-Transition Model
Implied Weight on Error-Correction Term as a Function of Same

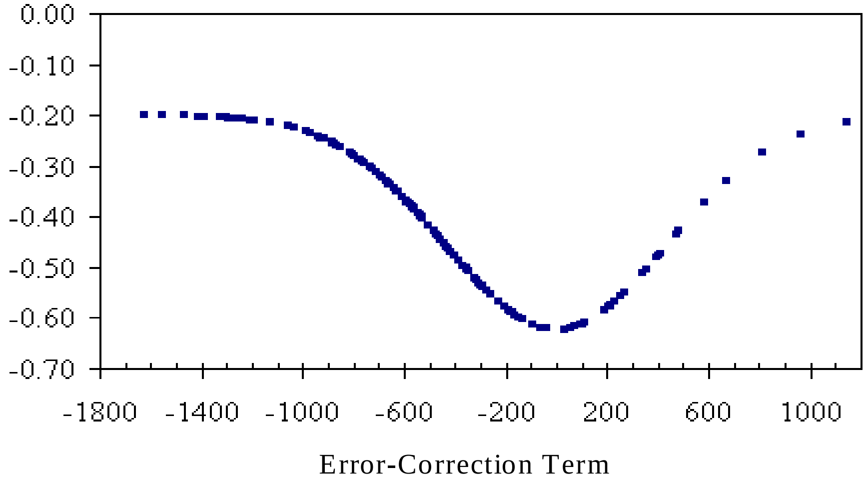
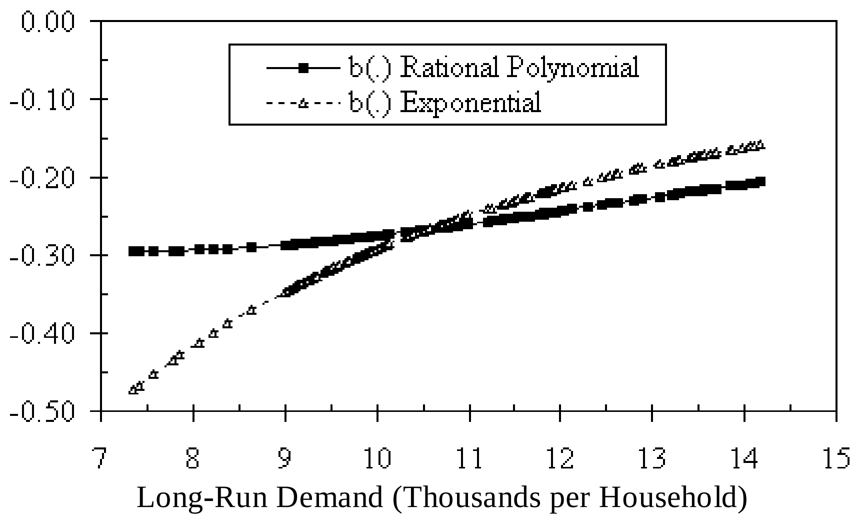


Figure 3
Rational-Polynomial and Exponential Models:
Implied Weight on Error-Correction Term as Function of Long-Run Demand



Notes: Estimates are from the models of Table 1. In the above figures "Long run demand" and "b(.)" denote (respectively) $\overline{M}_t$ and b_{1t} of Table 1 and the text.