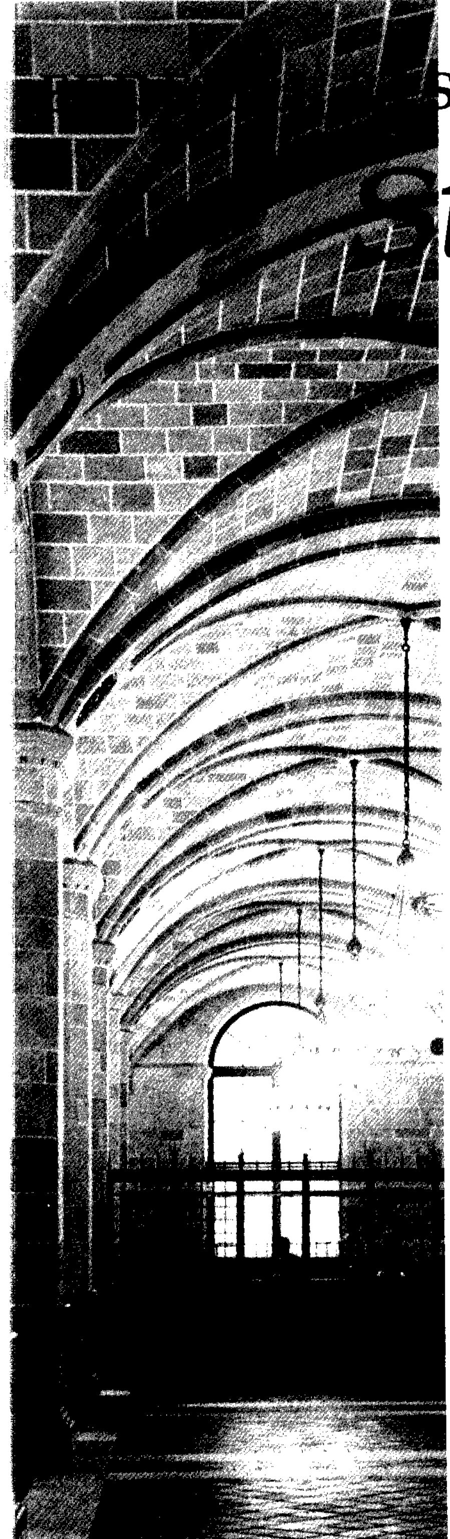

SERVE BANK OF NEW YORK

Staff Reports

An Efficient, Three-step Algorithm for Estimating Error-correction Models with an Application to the U.S. Macroeconomy

Michael D. Boldin

Number 6
October 1995

An Efficient, Three-step Algorithm for Estimating Error-correction Models with an Application to the U.S. Macroeconomy

Michael D. Boldin
Federal Reserve Bank of New York

Abstract:

This paper describes a three-step algorithm for estimating a system of error-correction equations that can be easily programmed using least-squares procedures. Nonetheless, the algorithm is both statistically and computationally efficient and when iterated gives maximum likelihood estimates of cointegration effects. Most important, the algorithm can handle different levels of cointegration, over-identified systems, breaks in trend, and complicated specifications for the short-run dynamics. The procedure is demonstrated with some small macroeconometric models, which suggest that breaks in the long-run trends for output and money are both statistically and economically significant in the 1961-94 period.

I. Introduction

Many important macroeconomic variables—such as (the log of) GDP, price indices, and monetary aggregates—seem to have a nonstationary trend that takes the form of a unit-root (random walk) process. These variables are therefore often labeled as being integrated of order one and are used in difference form in econometric models to ensure stationarity. However, a group of variables may share a common unit-root component and be cointegrated. Under cointegration, certain combinations of the variables are stationary or trend reverting; this fact can be used to retain statistically valid information from the levels of the variables. In fact, Engle and Granger (1987) showed a direct connection between cointegration and error-correction models (ECMs), which increased the usage of these types of econometric specifications that explicitly separate short-run dynamics from long-run trends.

This paper describes an efficient, three-step algorithm for estimating a system of error-correction equations (a vector-error correction model) that both fully accounts for cross-equation cointegration restrictions and can be programmed in almost any statistical package using least-squares procedures. The algorithm also helps shed light on the calculations that underlie Johansen's (1988) maximum likelihood estimation procedure for these models. In fact, when iterated, the three-step algorithm provides equivalent results. Although the Johansen procedure is computationally faster, the three-step algorithm does not require a software package that computes eigenvalues and eigenvectors. This feature makes it easier to program, especially when there are different levels of cointegration, overidentified systems, breaks in trend, or complicated specifications for the short-run equations.[1]

Some examples are presented at the end of the paper that demonstrate the flexibility of the three-step approach. Small macroeconometric models are estimated that allow for trend shifts without abandoning the error-correction approach. The results suggest that these breaks in the long-run relationships are both statistically and economically important.

II. The Three-step Algorithm for a Simple Example

In the simple case of two endogenous variables, y_1 and y_2, that are cointegrated over the sample period t=1,2,3,...T, a representative vector error-correction model is:

$$(1) \qquad v_t = y_{1,t} - \alpha\, y_{2,t}$$

$$(2) \qquad \Delta y_{1,t} = \gamma_1\, v_{t-1} + \beta_{1,1}\, \Delta y_{1,t-1} + \beta_{1,2}\, \Delta y_{2,t-1} + e_{1,t}$$

$$(3) \qquad \Delta y_{2,t} = \gamma_2\, v_{t-1} + \beta_{2,1}\, \Delta y_{1,t-1} + \beta_{2,2}\, \Delta y_{2,t-1} + e_{2,t}$$

Equation 1 defines the long-run relationship between y_1 and y_2, while equations 2 and 3 describe the short-run dynamics. In this system, $[1\ -\alpha]$ is the

---

[1]. GAUSS and EVIEWS (the Windows version of MicroTSP) programs for estimating the models that were used for this paper (with the data) are available upon request. I also have a programming outline to help those that are interested in writing algorithms for other statistical packages such as RATS or SAS. I will gladly review and keep an archive of programs that are written by others to implement the three-step estimation procedure.

cointegrating vector and the γ coefficients measure the speed-of-adjustment effects that move the variables toward the long-run equilibrium condition of v=0 and lead to trend-reversion.

This model assumes that all variables have zero means (making constant terms unnecessary) and that one lag of each dependent variable is sufficient to remove serial correlation in equations 2 and 3. These assumptions keep the notation simple and are not crucial in deriving the estimation procedure. At this stage, we also assume that $E[e'e]=\Sigma$ is diagonal (that is, there is no concurrent or contemporaneous correlation among the error terms). The implications of relaxing this restriction are discussed in Section IV. In addition, requiring the y_1 variables to have a coefficient of one in equation 1 is a convenient normalization that does not affect fully iterated, three-step results.[2]

Engle and Granger's two-step method first regresses y_1 on y_2 to estimate the coefficient α and compute the long-run equations error term: v. The second step uses lags of v and Δy as explanatory variables in OLS estimation of equations 2 and 3, which capture short-run relationships. While the parameters are consistently estimated with this procedure (Stock 1987), they are not fully efficient since the cointegrating vector, α, is estimated independently of the other parameters.

Engle and Yoo (1991) show that α can be revised by running a third-step regression that uses $(e_{1,t}/\sigma_1)$ as the dependent variable. Here, adjusted error terms from the second step estimation of equation 2 are stacked over

[2] To understand the need for a normalization convention, note that by multiplying the right-hand side of equation 1 by an arbitrary constant will not affect the β values or the fit of 2 and 3, only the scale of the γs.

analogous terms from estimates of equation 3. The single explanatory variable in this case is $(\gamma_i y_{2,t}/\sigma_i)$ with analogous stacking. The adjustment by the reciprocal of the equation standard errors, which yields a weighted regression setup, accounts for likely differences in the standard errors for each short-run equation (heteroskedasticity). The regression of $(e_{i,t}/\sigma_i)$ on $(\gamma_i y_{2,t}/\sigma_i)$ yields the incremental adjustment to α that corresponds to a step in a Newton-Raphson optimization routine and iterating until convergence yields maximum likelihood estimates. Most important, Engle and Yoo show that the first iteration's revision to α is asymptotically efficient.

One shortcoming of the Engle and Yoo procedure is that all other coefficients are fixed in the third step. It is possible to relax this condition and improve on the estimates of the cointegrating vector (in the sense that they improve the fit of the short-run equations). For example, a better third step can be derived by substituting for equation 1 in equations 2 and 3, multiplying through the γ terms, moving $y_{1,t-1}$ to the left-hand side, and creating a single regression equation:

$$(4) \qquad (\Delta y_{i,t} - \gamma_i \, y_{1,t-1}) = \alpha \, (\gamma_i \, y_{2,t-1}) + \beta_{i,1} \, \Delta y_{1,t-1} + \beta_{i,2} \, \Delta y_{2,t-1} + e_{i,t}$$

where the T observations for i=1 are stacked above the T observations for i=2. With equation 4, it is made clear that cointegration places cross-equation restrictions on the effect of lagged-level variables, y_1 and y_2, while the coefficients (β_i) for the lagged differenced variables can differ in the two short-run equations. Therefore, it is possible to fix only the γs and σs and estimate equation 4 by defining separate lagged Δy variables for i=1 and i=2 to yield five truly independent variables: $\gamma_i y_{2,t-1}$ and two sets of $\Delta y_{1,t-1}$ and $\Delta y_{2,t-1}$.

To expand on this framework, we can estimate revisions to γ simultaneously with revisions to α. By ignoring the β effects for the time being (for reasons that will become clear below), we create a linearized version of the model:

$$(5) \qquad \Delta y_{i,t} - \gamma_i (y_{1,t-1} - \alpha\, y_{2,t-1}) = \tilde{\alpha}\, (\gamma_i\, y_{2,t-1}) + \tilde{\gamma}_i\, (y_{1,t-1} - \alpha\, y_{2,t-1})$$

where the γ and α terms on the left-hand side are taken as given and the incremental revisions to these parameters are denoted by $\tilde{\gamma}$ and $\tilde{\alpha}$.[3] Adding back the $\beta\ \Delta y$ terms, equation 5 could be estimated via a stacked regression equation that is conditional on initial estimates of γ and α. However, this setup may be excessively large because most "small" macroeconometric models have at least four endogenous variables and use four to twelve lags of each in the explanatory equations. In these instances, one can greatly reduce the size of the matrix that must be constructed and then inverted to compute the revised α and γ parameters.

The trick to reducing computational burdens while efficiently estimating the stacked regression is to recognize that the OLS estimate of b_1 in:

$$y_t = b_1\, x_{1,t} + b_2\, x_{2,t}$$

can be computed by first regressing y on x_2, saving the resulting residuals

[3] I am grateful to Professor Christopher Sims of Yale University for pointing out that a revision to $\tilde{\gamma}$ can be made simultaneously with the revision to $\tilde{\alpha}$. My initial draft of this paper only revised $\tilde{\alpha}$ in step 3. Also, note that attempts to improve the revisions by including variables that capture the multiplicative (cross-term) effects of $\tilde{\alpha}$ and $\tilde{\gamma}$ will create a singularity problem in the regression equation.

as r_y, regressing x_1 on x_2, saving the resulting residuals as r_{x_1}, and finally, regressing r_y on r_{x_1}. The result is equivalent to the b_1 estimate from regressing y on both x_1 and x_2 simultaneously. This property conveniently carries over to the case where $x_{1,t}$ and $x_{2,t}$ are vectors of more than one explanatory variable. Also, the standard error term for the second step's b_1 coefficient is the one in the original model (after a degrees of freedom correction, which is discussed below).

To apply this two-stage regression concept to equation 5, we first separately regress the difference and levels variables, Δy_1, Δy_2, y_1, and y_2, on the lagged difference variables, $\Delta y_{1,t-1}$ and $\Delta y_{2,t-1}$, and save the resulting residuals as $R_{\Delta y_1}$, $R_{\Delta y_2}$, R_{y_1}, and R_{y_2}, respectively. Next, we construct the third-step regression's dependent variable as $\{(R_{\Delta y_i,t}-\gamma_i R_{y_1,t-1}+\gamma_i \alpha R_{y_2,t-1})/\sigma_i\}$, with the i=1 variables stacked above the i=2 variables. The two explanatory variables are $\{(\gamma_i R_{y_i,t-1}/\sigma_i)\}$ and $\{(R_{y_1,t-1}-\alpha R_{y_2,t-1})/\sigma_i\}$, with analogous stacking.[4] Regression estimates of this equation yield incremental revisions to α and γ that simultaneously account for revisions to the β's.

III. Efficiency and a Comparison to the Johansen Method

This algorithm is efficient on many levels. First, because the estimates from the first two steps are consistent, one iteration of three-step procedure yields asymptotically efficient estimates. Second, my third step only takes the σs as given. A more efficient step for iterating would require a nonlin-

[4] Note that a separate right-hand side term for $\gamma_i R_{y_1}$ is not needed because of the normalization assumption. Also, since v is a linear function of y_1 and y_2, the second right-hand variable is equivalent to R_v.

ear optimization procedure. Third, if the procedure is iterated, the construction of the R variables decreases the computation time by reducing the effective size of the regression problem.[5] Fourth, and probably most important, iterating until convergence yields exact maximum likelihood estimates that fully account for cross-equation, cointegration restrictions and that have all the desirable properties of correctly specified maximim likelihood estimates (MLEs).[6]

While many other methods for estimating cointegration parameters are also asymptotically efficient, including Engle and Yoo's, most are single-equation based (for example, Saikkonen 1991 and Phillips and Loretan 1991). In contrast, the three-step procedure uses information from the entire system so that fully iterated results will be equivalent to results from Johansen's procedure. In fact, the construction of the R variables should be

---

[5] For various models I have confirmed that my algorithm converges noticeably faster than both the Engle and Yoo algorithm and a procedure that ignores simultaneous revisions to γ, and instead reestimates the γs by repeating the second step. Engle and Yoo also claim that the effect of iterating is likely to be small since their procedure regresses differences variables on levels variables. While each iteration tends to produce relatively small changes, an example in Section VI shows that the final or "converged to" estimates can be quite different from the results of the first iteration.

[6] Convergence means that the first-order conditions are satisfied. However, since the α and γ parameters enter the model in a multiplicative fashion and cross-term effects are ignored in the third step, convergence cannot be guaranteed for all starting values. In fact, some restricted models that were considered for Section VI did not converge, a result that proved useful in determining the robustness of a particular specification. However, for a correctly specified model, the likelihood function is concave over the admissible parameter set, guaranteeing convergence to the MLEs when the starting values are not too far from the true values.

familiar to econometricians who are conversant with the Johansen procedure. To be explicit, the MLE formula minimizes:

$$(5) \qquad \Sigma_t \, (R_{\Delta y,t} - (\gamma\alpha')\, R_{y,t})\, \Sigma^{-1}\, (R_{\Delta y,t} - (\gamma\alpha')\, R_{y,t})'$$

for a given Σ, where γ and α are 2x1 in this example. In equation 5, the construction and use of the R variables implicitly controls for optimal values of β for any estimate of $(\gamma\alpha')$. An explicit normalization scheme for α is not used since only the space spanned by $\gamma\alpha'$ can be identified (in other words, specific values for α and γ cannot be determined). Johansen (1988) shows how to form the concentrated likelihood function by substituting the combined $(\gamma\alpha')$ effect into the formula for Σ and by using eigenvalue calculations to solve the corresponding first-order (optimization) conditions. The three-step procedure relies on an iterated technique to reach the same set of parameter values.

The following section describe how models that are more complicated than equations 1 to 3 can be specified and programmed in a relatively efficient manner. For example, adding exogenous trends or placing restrictions on the long-run relationships is straightforward. Also, the three-step algorithm allows the short-run equations to be handled more flexibly. In much of the empirical work with ECMs, the short-run equations are given the same lagged differenced variables. This trait, which is likely to lead to compromises that overparameterize some equations and underparameterize others, results in a loss in efficiency. Obviously, the benefits of parsimony should not be abandoned, and diagnostic tests that can detect overfitting problems should be considered.

IV. Generalizing the Three-step Algorithm

A point to note about the three-step procedure described above is that the variance-covariance matrix for the equation error terms, Σ, was assumed to be diagonal (no contemporaneous correlation). Typically, users of Johansen's method do not restrict Σ, and the three-step algorithm must be amended to replicate the corresponding MLE cointegration vectors. The easiest way to make the change for the model above is to add contemporaneous values of Δy_1 to equation 3. The necessary adjustments for computing the R variables are described below. Note that because of identification problems, a similar adjustment cannot simultaneously be made to equation 2. However, adding $\Delta y_{2,t}$ to the right-hand side of equation 2, instead of adding $\Delta y_{1,t}$ to the right-hand side of equation 3, will not change the final cointegrating vector estimate because it does not change the value of the likelihood function.[7]

Other possible changes are the addition of trends, dummy variables, and other exogenous variables to equations 1, 2, and 3. For instance, it is not necessary to include all error-correction terms or to have the same lag length in every short-run equation. Also, the addition of other variables to the right-hand side of the short-run equations is an easy task with the three-step procedure; as long as these exogenous variables are stationary, addition will not violate the basic assumptions of the error-correction framework. In fact, other relevant variables can improve the efficiency of estimates of the long-run relationships by making the short-run equations

---

[7] The α estimates can change greatly, however, when contemporaneous effects are excluded from the short-run equations.

more precise. In comparison, the statistical software packages that have built-in Johansen procedures are currently quite limited in their ability to experiment with and test different specifications for the short-run equations.

The following steps present the generalized three-step algorithm for k=1,2,..K cointegrating relationships, i=1,2,..N endogenous variables and the N short-run equations. In these directions, X_i represents the entire set of right-hand side variables, save the cointegration terms, in the equation that explains Δy_i. Similarly, Z_k represents the entire set of right-hand side variables in the k^{th} cointegrating (levels) equation. Identification issues concerning the appropriate sets of variables for the long- and short-run equations are discussed in the next section.

(1) Specify and estimate each of the K cointegrating relationships by OLS methods using one of the endogenous variables (y_k) as a left-hand side variable. The right-hand side candidates (Z_k) are the other endogenous (some y variables may be excluded) and possibly exogenous variables such as a constant and a time trend. Save the residuals as the initial error correction (v_k) terms for step 2. (Since the error-correction term is entered with a lag in the step 2 regressions, the sample period should be t=0 to t=T-1 for this step when the short-run equations are estimated for the period t=1 to t=T.)

(2) Again using OLS procedures, estimate the N short-run equations for each endogenous, differenced variable (Δy_i) including both short-run effects (X_i) and the lagged error-correction terms (v_k) on the right-hand side. Save each equation's standard error (σ_i) and error correction coefficients ($\gamma_{k,i}$) for the third step.

(3) A. Regress each of the differenced variables (Δy_i) on the corresponding set of short-run variables (X_i) and save the residuals as a vector: ($R_{\Delta y_i}$). Regress each of the left-hand side levels variables (y_k) in step 1 on each set of short-run variables (X_i) and save the residual as a vector: ($R_{y_{k,i}}$). Next, regress each of the right-hand variables in the step one regressions on each set of short-run variables (X_i) and save these residuals as a matrix: ($R_{z_{k,i}}$).

B. Set the third step's left-hand side variable as $\{(R_{\Delta y_1}-\Sigma_k\gamma_{k,1}R_{v_{k,1}})/\sigma_1\}$ stacked over $\{(R_{\Delta y_2}-\Sigma_k\gamma_{k,2}R_{v_{k,2}})/\sigma_2$ and so on, where $R_{v_{k,i}}=(R_{y_{k,i}}-\alpha_k R_{z_{k,i}})$, forming NT rows. Create the right-hand side variables for $\tilde{\alpha}$ as $\{(\gamma_{1,1}R_{z_{1,1}})/\sigma_1\ (\gamma_{2,1}R_{z_{2,1}})/\sigma_1\ \ldots\ (\gamma_{K,1}R_{z_{K,1}})/\sigma_1\}$ stacked over $\{(\gamma_{1,2}R_{z_{1,2}})/\sigma_2\ (\gamma_{2,2}R_{z_{2,2}})/\sigma_2\ \ldots\ (\gamma_{K,2}R_{z_{K,2}})/\sigma_2\}$ and so on. The right-hand side variables for $\tilde{\gamma}$ start with $\{(\alpha_1 R_{z_{1,1}})/\sigma_1\ (\alpha_2 R_{z_{2,1}})/\sigma_1\ \ldots\ (\alpha_K R_{z_{K,1}})/\sigma_1\}$ with zeros in the remaining rows. Each set of remaining $\tilde{\gamma}$ variables (i=2,3,...n) have zeros in the first T(i-1) rows, followed by $\{(\alpha_1 R_{z_{1,i}})/\sigma_i\ (\alpha_2 R_{z_{2,i}})/\sigma_i\ \ldots\ (\alpha_K R_{z_{K,i}})/\sigma_i\}$ and zeros in any remaining rows.

C. Regress the TNx1 left-hand side variable in step 3B on the two sets of right-hand side variables (for $\tilde{\alpha}$ and $\tilde{\gamma}$) and use the resulting coefficients to compute incremental revisions to α and γ. Finally, the most direct way to estimate σ_i for step 3B is to compute the standard error for $(R_{\Delta y_1}-\Sigma_k\gamma_{k,1}R_{v_{k,1}})$.

As explained above, this algorithm will converge to the MLE solution by iterating between 3B and 3C. Note that step 3A does not need to be repeated, since the R variables do not depend on any estimated coefficients.

V. Identification and Testing

A. Long-run Identification

If more than one cointegrating relationship is estimated, each must be linearly independent. In the absence of a theoretically justified structure, one way to achieve independence is to exclude any variable used as a left-hand side variable in step 1 from the right-hand side of the other first-step regressions. This particular design is not restrictive, however. A linear combination of two cointegrating vectors is also a valid cointegrating vector, and other designs can be constructed from the results. Another possibility with the three-step algorithm is to entirely exclude some variables from the cointegration scheme (an example is given in Section VI) to yield an overidentified system. The conventional Johansen procedure only handles exactly identified systems.[8]

B. Short-run Identification

For contemporaneous correlation effects, a "triangular structure" is most convenient. Also, it is consistent with the Cholesky decomposition schemes that are popular in VAR analysis. When this structure is set up to identify contemporaneous causation, the first equation excludes all contemporaneous variables, the second equation adds only the first differenced endogenous variable (on a contemporaneous basis), the third equation adds

---

[8] Johansen (1991) shows the formulas for applying linear restrictions, but this procedure also requires eigenvalue calculations and is not easily implemented within most statistical software packages. Pesaran and Shin (1994) discuss the identification assumptions of the Johansen procedure and derive two alternative MLE procedures for overidentified (constrained) cointegration schemes that share some features with the three-step approach. Their approach does not rely solely on OLS estimation procedures, however.

the first and second differenced endogenous variables, and so on.

Different contemporaneous schemes, including structural decompositions, could be designed and implemented by adding contemporaneous terms to the short-run equations in a nontriangular manner.[9] However, care must be taken to ensure that the system remains identified. Also, researchers should be aware that structural interpretations of the second step β and γ estimates will depend on the placement of contemporaneous variables in the right-hand side of the short-run equations, as a causal ordering among the variables would be assumed.

Most interesting is a finding below that the use of an over-identified system (shown in a case where zero contemporaneous correlation is assumed) can greatly affect estimates of long-run relationships. This should not be too surprising to users of unrestricted VARs who find impulse response function to be sensitive to the error decomposition scheme. In addition, each error correction effect need not be present in every short-run equation. In other words, some $\gamma_{k,i}$ may be set to zero by excluding the corresponding error-correction terms from the right-hand side of a short-run equation. This option can greatly affect both the short- and long-run properties of the model.

---

[9] Using u to represent the vector of correlated error terms of the short-run equations, the generalized, "structural" model has u=Ae with Σ_e being a diagonal matrix and $E[uu'] = A\Sigma_e A^1$. Since the u's are linear functions of Δy and the explanatory variables, the model can be reformulated to yield independent errors with contemporaneous Δy's in the right-hand side of some equations according to $A^{-1}\,\Delta y = A^{-1}\,X\,\beta + e$. See Hamilton (1994) for a comparison of recursive (triangular) and nonrecursive (structural) VARs. Little work has been done on combining cointegration restrictions and structural identification, however.

C. Constructing Parameter Standard Errors

Conventional formulas, such as $\sigma^2(X'X)^{-1}$ with $X=\{R_z\ R_v\}$, for the coefficient standard errors in the step 3 regression can be used to compute standard errors for both α and γ. It must be recognized that most statistical packages make a degrees-of-freedom adjustment of (1/(T-P)) to compute σ, where T is the sample size and P is the number of estimated coefficients. This formula is inappropriate here, however, because the number of right-hand variables in the third step is less than the number of coefficients in the entire system. To get an unbiased estimate of parameter uncertainty, OLS standard errors can be simply multiplied by $((T-P)/(T-M))^{1/2}$, where M is the actual number of coefficients in the model (including the implicit β terms). The alternative adjustment that is used below sets M=0, which yields maximum-likelihood (but not unbiased) estimates of the parameter standard errors.[10]

Since cointegration forces the correlation between R_z and R_v to go to zero as the sample size increases, asymptotically valid standard errors can be constructed by ignoring the correlation between the two sets of variables during the sample period. In other words, the standard errors for α and can be computed by inverting $(R_z'R_z)$, and the standard errors for γ can be computed by separately inverting $(R_v'R_v)$. This property is a direct consequence of the fact that the errors in the long-run equations are stationary, while the individual levels variables are not. (See Johansen 1991 for a detailed proof.) Usually only these asymptotics-motivated standard errors are reported, although they may not be appropriate for the sample size that is

[10] Note that the division by σ_i in step 3 makes the stacked equation's standard error equal to one and $(X'X)^{-1}$ can be used without any adjustment.

used in a particular study. Below I report both sets of parameter standard errors and discuss significant differences.

D. Testing

Various hypothesis tests can easily be accommodated. First, t-statistics are valid for testing point estimates of individual cointegration parameters as long as the null hypothesis does not imply that a lower order of cointegration is present. For example, a t-test of α=0 in equation 1 would not have a conventional distribution since it assumes that y_1 and y_2 are not related in the long run.[11]

While Wald or F-tests for multiple restrictions can be applied in a straightforward manner, likelihood ratio (LR) tests can be also computed by estimating both the null (H_o, restricted) and alternative (H_a, unrestricted) models. The test statistic is:

$$LR = T * \sum_i (\log \sigma_{ii|H_o} - \log \sigma_{ii|H_a})$$

and would be compared to the χ^2 distribution for q degrees of freedom, which corresponds to the additional number of parameters in H_a. This framework is particularly useful when the null model's specification includes restrictions on both the short-run equations and the cointegration

[11] Dickey and Rosanna (1994) provide a good overview of the problem.

parameters.[12] Again, it must be realized that when the null specification has one less cointegrating relationship than the alternative model, conventional test distributions (chi-square in this case) are not appropriate.

VI. Applications to U.S. Macroeconometric Data

The ability of the three-step algorithm to handle nonstandard cointegrating relationships is demonstrated in this section with some small macro-models that use quarterly data over the 1961-1994 period. This exploration, although tentative, shows that researchers should not ignore the possibility of breaks in both the economy's trend and its underlying structure. In essence, this analysis is motivated by Campbell and Perron's (1991) discussion of the problems that can result from using conventional unit-root tests to make specification decisions and from ignoring the possibility of breaks in trend. To save space, I have deleted tables with unit-root tests that would justify the primary cointegration schemes. The results are available on request and are consistent with the majority of the other studies on the subject: the log of nominal GDP, real GDP, nominal M2, real M2, and nominal interest rates are only stationary after differencing (at least over the 1960-94 period). The only controversial decision is whether the log of the GDP deflator must be double differenced to achieve stationarity. Therefore, cases where inflation does and does not have a unit root are con-

[12] However, as noted above, asymptotic independence between α and γ may be assumed. Also, the LR calculation assumes that both models were formulated to have diagonal error covariance matrices (Σ), which can always be achieved by judiciously choosing the placement of contemporaneous Δy variables. See footnote 9.

sidered below.

The first two models, in Tables 1a and 1b, estimate error-correction equations for the log of nominal output (py) and M2 (m). Four lags of the differenced variables are used in the short-run equations in this example and all that follow. The cointegrating coefficients are normalized to equal one for nominal GDP. In addition, following the format of equation 1, the M2 coefficient is given a sign that is consistent with it being a left-hand side variable in an OLS equation. Therefore, the sign is opposite what would be reported for cointegration equations that put all variables on the same side of the long-run equation. The important feature for this normalization scheme is that a positive sign signifies that nominal GDP and M2 tend to move in the same direction.

In the top panel of Table 1a, no additional time trend variables are used, and contemporaneous short-run effects are ignored. The first row reports the OLS first step results. A coefficient of almost exactly one on M2 suggests that M2 velocity (the ratio of Nominal GDP-to-M2) is stationary over the 1961-89 period. The first iteration's three-step coefficient for M2 is again very close to one, as are the fully iterated or MLE results. The second panel in Table 1a allows contemporaneous interaction between Δpy and Δm, and the results are quite similar. An M2 coefficient that is close to one seems to be very robust in this simple two-variable system.

The bottom panel of Table 1a shows the MLE error-correction coefficients for both models. The ordering is Δpy then Δm2 in the model that includes contemporaneous effects in the short-run equations. However, since the cross correlation is low (less than 0.05), this assumption has little impact on the results. The signs are as expected, and nominal output and

M2 growth are affected almost equally by deviations of M2 velocity from its mean. The individual t-statistics are not exceptionally high, however, even though formal cointegration tests easily reject the null hypothesis that M2 velocity had a unit-root component in the 1961-1989 period. In computing t-statistics, it is interesting that the two sets of parameter standard errors agree at the first four decimal places. (In other models, there are significant differences.)

Table 1b shows results for another nominal GDP-M2 model, with the sample expanded to 1994 and time-trend effects added. Again, the top panel results ignore contemporaneous short-run correlations, while the middle panel includes these effects. Experimenting with a few options, including a simple linear term, a quadratic trend, and piece-wise trend shifts in 1974 and 1991, I found plausible estimates and significant trends when a 1991 trend shift was added to an equation with a 1974 trend shift. (In this analysis, significance is defined by a t-statistic above two.) The OLS estimates and first-iteration three-step results show only an economically meaningful trend after 1990. However, in both panels, MLEs of both the 1974 and 1991 trend shift are significant, and an LR test of joint time trend effects gives a highly significant value of more than 16. Most important, the trend shift is found to be quite large—the coefficients suggest that nominal GDP will tend to increase at a rate that is almost 6.0 percent per year above the growth rate of M2.

By comparing Tables 1a and 1b, we can see that the addition of a time trend lowers the M2 coefficient significantly below one when the asymptotics-motivated standard errors are used in a t-test. Results for models that are not reported in the tables confirmed that this effect was not caused by expanding the sample period. However, the table shows substantially higher standard errors when the cross-correlation between α and γ is not set at zero. Also in contrast to the simpler model, error-correction effects are quite low for nominal GDP growth, while the statistical significance of $\gamma_{\Delta m2}$ remains high. Further exploration (not reported in the tables) shows that this condition is not a consequence of the ordering (Δpy then Δm2 in the table). Rather, it occurs when the sample period is increased to 1994.

Table 2 shows estimates for a model that splits Nominal GDP into real (y) and price (p) terms and adds T-bill (r) rates along with M2. Only a single cointegrating relationship with a simple time trend is allowed over the 1961-94 period. The most interesting finding is that parameter estimates are greatly affected by the inclusion of contemporaneous effects. For example, the role of interest rates in the cointegrating or equilibrium relationship overwhelms that of M2 in the model without contemporaneous effects. In the model with contemporaneous effects, the coefficients for M2 and interest rates are similar in magnitude. Also, the two sets of parameter standard errors are much closer with this specification. The results from this model should not be taken too seriously, however, since formal LR tests suggest that a second cointegrating relationship exists.

Table 3 presents a more interesting vector error-correction model for real GDP (y), real M2 (m2-p), T-Bill rates (r), and the inflation rate from the price deflator (Δp). The use of the inflation rate instead of the price level is in recognition of the fact that inflation rates are often found to be nonstationary. Two cointegrating equations are allowed in the model, and 1961-94 is used for the sample period. Formal LR tests pointed to at least two cointegrating relationships, with more than 95 percent confidence.

The possibility of a third cointegration equation cannot be discounted, however, and is discussed below.

The first long-run equation is specified to correspond to an IS (goods demand) equation with a break in 1974 (both a one-time level shift and a trend change). The second long-run equation is an LM (money demand) specification. A trend shift in 1974 is allowed in the IS equation, and a trend shift in 1991 is allowed in the LM equation.

I attempted to include both inflation and nominal interest rate terms in the IS equation and thereby estimate real rate effects, but the inflation coefficient was implausibly large. Therefore, only nominal interest rate effects are captured, and an unexpected positive, although insignificant, relationship with output is found.[13] With this model, the trend in output is significantly lower after 1974. Also, the point estimate of the level shift is large, but imprecisely estimated, a result that is consistent with Perron's (1989) conclusion about oil price shocks in the 1970s. For the LM equation, the elasticity between real M2 and real output is slightly greater than one, interest rate effects are insignificant, and a large trend shift occurs in 1991.

The bottom panel of Table 3 shows the error-correction terms (γ_{IS} and γ_{LM}) for each short-run equation. The most prominent findings are that the only statistically significant γ_{IS} coefficient is for real M2, while the γ_{LM} coefficients for output, real M2, and interest rates are significant. In fact, direct trend reversion for real output is quite weak ($\gamma_{IS\Delta y}$= -0.0308), even after controlling for a break in 1974. However, deviations of real M2 from trend

[13] These effects are not necessarily due to cointegration restrictions since unrestricted VARs often show puzzling interest rate effects (i.e., positive or expansionary monetary shocks raise interest rates and lower inflation).

have a strong positive impact on real output ($\gamma_{LM,\Delta y} = 0.1430$). Also, the sign of $\gamma_{IS,\Delta^2 p}$ is surprisingly negative, which implies that inflation falls when output is above trend. While the effect is insignificant, it also does not support the view that there is a strong short-run Phillips curve effect from changes in aggregate demand. Note that since inflation is not present in either long-run equation, a vertical long-run Phillips curve is assumed, a condition that presumably makes short-run inflation effects stronger.

Although a more complete analysis of other plausible macro-economic relationships and break dates are necessary before we can make strong conclusions, the IS and LM equations in Table 3 demonstrate how the vector-error correction framework can show breaks in the data. Because the IS effects are quite weak, future research on the effect of inflation in the long-run equations seems warranted. Both sets of standard errors show that only the two trend coefficients are clearly significant in this long-run equation. Also, t-statics for the γ_{IS} terms are relatively low. One way to try and solve this problem is to add a one-period ahead forecast of inflation (from the short-run inflation equation) to the IS relationship. Another option is to consider a cointegrating relationship that would capture stationarity in real interest rates. In fact, I attempted to add a third cointegrating equation for nominal interest rates and realized inflation to the model in Table 3, but the estimated real rates turned out to be both implausible and overly sensitive to seemingly minor alterations in the specifications. A possible solution to this negative finding would be to use predicted inflation in the cointegrating relationship for nominal rates. Finally, it would be interesting to add long-term rates and an equilibrium yield-curve relationship to the model. The three-step algorithm provides a way to explore these effects as well as breaks in trends.

VII. Conclusion

This paper introduces a three-step algorithm to compute cointegration effects in a system of equations. This procedure relies solely on least-squares calculations and is therefore more intuitive than the Johansen method, which requires eigenvalue and eigenvector calculations. Most important, the flexibility of the approach allows for testing more complicated cointegration schemes. This feature is successfully demonstrated through estimates of small IS/LM based macro-econometric models with breaks in trends.

References

Campbell, J., and P. Perron. 1991 "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots." *NBER Macroeconomics Annual 1991*.

Dickey, D., and R. Rosanna. 1994 "Cointegrated Time Series: A Guide to Estimation and Hypothesis Testing." *Oxford Bulletin of Economics and Statistics* 56: 325-53.

Engle, R. F., and C. W. Granger. 1987. "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55: 251-76.

Engle, R. F., and A. Yoo. 1991. "Cointegrated Economic Time Series: An Overview with New Results." In R. F. Engle and C.W. Granger, eds., *Long-Run Economic Relationships: Readings in Cointegration*." Oxford University Press.

Hamilton, J. 1994. *Time Series Analysis*. Princeton University Press, NJ.

Hendry, D., and A. Banerjee, eds., 1992. "Special Issue: Testing Integration and Cointegration." *Oxford Bulletin of Economics and Statistics* 54.

Johansen, S. 1988. "Statistical Analysis of Cointegrating Vectors." *Journal of Economic Dynamics and Control* 12: 231-54.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59: 1551-80.

Perron, P. 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57: 1361-1401.

Pesaran, H., and Y. Shin. 1994. "Long Run Structural Modelling." Department of Applied Economics Working Paper no. 9419, University of Cambridge.

Phillips, P.C.B., and M. Loretan. 1991. "Estimating Long-Run Economic Equilibria." *Review of Economic Studies* 58: 407-36.

Saikkonen, P. 1991. "Asymptotically Efficient Estimation of Cointegrating Regressions." *Econometric Theory* 7: 1-21.

Stock, James. 1987. "Asymptotic Properties of Least-Squares Estimators of Cointegrating Vectors." *Econometrica* 55: 1035-56.

Table 1A. Nominal GDP and M2 ECM, 1961-1989

Long-run Coefficients

	py	m2
OLS	LHS	1.0074
Without contemporaneous effects		
3rd step	LHS	1.0071
MLE	LHS	1.0071
(se)		(0.0116)
(ase)		(0.0116)
With contemporaneous effects		
3rd step	LHS	1.0070
MLE	LHS	1.0070
(se)		(0.0113)
(ase)		(0.0113)

Error-Correction Coefficients

	Δpy	Δm2
Without contemporaneous effects		
γ	-0.0672	0.0490
(se)	(0.0323)	(0.0220)
(ase)	(0.0323)	(0.0220)
With contemporaneous effects		
γ	-0.0672	0.0507
(se)	(0.0323)	(0.0224)
(ase)	(0.0323)	(0.0224)

Notes: Each short run equation includes 4 lags of each differenced, endogenous variable. py: log of Nominal GDP. m2: log of M2. LHS: treated as a right-hand side variable with a coefficient of one (and the other variables are treated as left-hand side variables). se: parameter standard errors from full system. ase: asymptotocs-motivated parameter standard errors.

Table 1B. Nominal GDP and M2 ECM, 1961-1994

Long-run Coefficients

	py	m2	tt74+	tt91+
OLS	LHS	1.0185	-0.0347	0.9356
Without contemporaneous effects				
3rd step	LHS	1.0118	-0.0005	1.4064
MLE	LHS	0.8635	0.5713	1.1798
(se)	(0.1048)	(0.3920)	(0.6691)	
(ase)	(0.0625)	(0.2082)	(0.6325)	
With contemporaneous effects				
3rd step	LHS	1.0108	-0.0030	1.4041
MLE	LHS	0.8873	0.4794	1.2269
(se)	(0.0894)	(0.3306)	(0.6147)	
(ase)	(0.0572)	(0.1908)	(0.5804)	

Error-correction Coefficients

	Δpy	Δm2
Without contemporaneous effects		
γ	-0.0058	0.0398
(se)	(0.0131)	(0.0206)
(ase)	(0.0128)	(0.0085)
With contemporaneous effects		
γ	-0.0092	0.0438
(se)	(0.0147)	(0.0211)
(ase)	(0.0141)	(0.0094)

Notes: py: log of nominal GDP. m2: log of M2. tt74+: time trend starting in 1974:1. tt91+: time trend starting in 1991:2. See Table 1A for further explanations.

Table 2. GDP, Deflator, M2, and T-Bill Rate ECM, 1961-1994

Long-run Coefficients

	y	p	m2	r	trend
OLS	LHS	-0.6789	0.4009	0.7426	0.7933
Without contemporaneous effects					
3rd step	LHS	-0.6359	0.3509	0.9185	0.8328
MLE	LHS	-0.5061	0.1115	2.4702	1.0887
(se)	(0.1668)	(0.2505)	(1.4175)	(0.3548)	
(ase)	(0.1518)	(0.1611)	(0.5017)	(0.2476)	
With contemporaneous effects					
3rd step	LHS	-0.6372	0.3718	0.6294	0.8055
MLE	LHS	-0.6482	0.3817	0.4744	0.8130
(se)	(0.0566)	(0.0582)	(0.1847)	(0.0915)	
(ase)	(0.0554)	(0.0582)	(0.1845)	(0.0883)	

Error-correction Coefficients

	Δy	Δp	Δm2	Δr
Without contemporaneous effects				
γ	0.0546	0.0208	-0.0085	0.0667
(se)	(0.0417)	(0.0167)	(0.0165)	(0.0481)
(ase)	(0.0216)	(0.0097)	(0.0155)	(0.0202)
With contemporaneous effects				
γ	-0.1020	-0.0072	0.1048	0.1847
(se)	(0.0594)	(0.0262)	(0.0426)	(0.0550)
(ase)	(0.0576)	(0.0262)	(0.0397)	(0.0490)

Notes: y: log of real GDP. p: log of GDP deflator. m2: log of M2. r: 3 month T-Bill rate
See Table 1A for further explanations.

Table 3. Small (IS/LM) Macroeconometric Model, 1961-1994

Long-run Coefficients-- IS Equation

	y	r	trend	d74+	tt74+
OLS	LHS	0.1526	0.9445	-4.8778	-0.3152
3rd step	LHS	0.2084	0.8374	-3.0580	-0.2234
MLE	LHS	0.2319	0.8335	-3.1749	-0.2764
(se)	(0.5442)	(0.0810)	(2.6875)	(0.0812)	
(ase)	(0.3742)	(0.0689)	(2.6473)	(0.0795)	

Long-run Coefficients-- LM Equation

	rm2	y	r	trend	tt91+
OLS	LHS	1.0778	-0.9197	-0.0327	-1.3479
3rd step	LHS	1.2528	-0.2669	-0.1854	-1.2659
MLE	LHS	1.1800	-0.0682	-0.1318	-1.0825
(se)	(0.1453)	(0.2974)	(0.0891)	(0.1978)	
(ase)	(0.1077)	(0.1939)	(0.0767)	(0.1866)	

Error-correction Coefficients

	Δy	$\Delta rm2$	Δr	$\Delta^2 p$
γ_{IS}	-0.0308	-0.0684	0.0252	-0.0099
(se)	(0.0408)	(0.0309)	(0.0282)	(0.0107)
(ase)	(0.0211)	(0.0180)	(0.0208)	(0.0091)
γ_{LM}	0.1430	-0.0613	-0.0750	0.0205
(se)	(0.0429)	(0.0420)	(0.0360)	(0.0151)
(ase)	(0.0312)	(0.0285)	(0.0318)	(0.0141)

Notes: y: log of real GDP. p: log of GDP deflator. rm2: real M2 (log of M2 minus log of GDP deflator). r: 3 month T-Bill rate. d74+: dummy (0/1) for period starting 1974:1. tt74+: time trend starting in 1974:1. tt91+: time trend starting in 1991:2. See Table 1A for further explanations.