ECONOMIC RESEARCH REPORTS

AN EXAMINATION OF ECONOMETRIC TESTS OF THE PROPOSITIONS CENTRAL TO THE NEW CLASSICAL MACROECONOMICS

by

Roman Frydman

and

Peter Rappoport

R.R. #85-04 January 1985

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

August 1984
Revised January 1985

Please do not quote
without permission.

Comments welcome.

AN EXAMINATION OF ECONOMETRIC TESTS OF THE PROPOSITIONS CENTRAL TO THE NEW CLASSICAL MACROECONOMICS

Roman Frydman and Peter Rappoport*

C. V. Starr Center for Applied Economics

and

Department of Economics
New York University
New York, NY 10003

*Roman Frydman gratefully acknowledges support from the National Science Foundation and the Hoover Institution at Stanford University. The authors would like to express their indebtedness to the participants of the NBER Time Series Seminar, especially Larry Christiano and Alfonso Novales, for helpful discussion and several constructive suggestions.

TABLE OF CONTENTS

Introduction

In much of contemporary macroeconomic theory, the influence of events on variables of interest differs according as the events are anticipated by economic agents or not. Faced with the problem of modelling these anticipations, economists have opted for the assumption of rational expectations. Following Muth's (1961) original definition this has been interpreted as mandating the economist to specify expectations as predictions based on the "objective probability distribution" of events, conditional on the information possessed by economic agents.

In theoretical applications the rational expectations hypothesis has been implemented by assuming that the "objective probability distribution" is the same as the probability distribution implied by the model being analyzed by the economist.[1] The distinguishing feature of empirical implementations of the rational expectations hypothesis is that the investigator has to specify a model generating the "objective probability distribution" of the variable(s) forecasted by economic agents. The adequacy of the investigator's model has to be judged from the empirical performance of his or her forecasting equation. Thus, while the rational expectations hypothesis is supposed to describe how rational agents form their forecasts,[2] it is silent on how the investigator can accurately replicate these expectations from empirical data.

The rational expectations approach has been used extensively in empirical studies of three central hypotheses of the New Classical Macroeconomics. One of these, the "Lucas proposition," postulates a relationship between the

---

[1]This procedure is discussed in detail in the Introduction to Frydman and Phelps (1983).

[2]Frydman (1983a) has argued that there is a distinction between the individual rationality postulate in economics and the rational expectations hypothesis.

variance of the unanticipated component of a nominal variable and the slope of the "output-inflation tradeoff."[3] Another asserts that anticipated money supply or nominal demand growth does not affect real output. The third is the rational expectations hypothesis itself.[4]

In testing these hypotheses the investigators have had to face the problem of specifying empirically the rational expectations of economic agents. The usual practice has been to model rational expectations as the least squares projection of the variable being forecast on a list of variables considered by the investigator to be available to agents.[5] Several investigators have recently recognized that this procedure may lead to mismeasurement of rational expectations of agents and considered effects of this mismeasurement on the tests of hypotheses of the New Classical Macroeconomics.[6]

The existing analyses of effects of mismeasurement of rational expectations have either assumed that the measurement error is a white noise process or relied on the fact that this error is a projection error and thus uncorrelated with the regressors of the investigator's forecasting equation. The current literature lacks a more complete characterization of the properties

---

[3]The output-inflation tradeoff has been examined empirically by Lucas (1973), Cukierman and Wachtel (1979), Froyen and Waud (1980), Attfield and Duck (1983) and Kormendi and Meguire (1984), among others.

[4]Recent studies of the policy-neutrality hypothesis include Barro (1977, 1978), Barro and Rush (1980), Gordon (1982), Leiderman (1980) and Mishkin (1983). The last two papers also test the rational expectations hypothesis.

[5]For an example of this approach see Barro (1977). The same procedure for empirical modelling of rational expectations has also been used in other contexts.

[6]For example, Frydman and Schankerman (1981), Startz (1983) and Abel and Mishkin (1983) examine the test of rationality. The latter two authors and Attfield (1983) also study the test of short-run policy neutrality. Kormendi and Meguire (1983) consider the effects of mismeasurement of rational expectations on tests of the Lucas proposition.

of the investigator's error in measuring rational expectations of agents. In addition, existing results have been derived only in special cases and cover only particular aspects of the problem. These drawbacks seriously limit the validity and applicability of these results.[7] Despite these limitations some authors have suggested that their results possess general validity. For example, Abel and Mishkin (1983, p. 3) summarize their analysis by stating that "the exact specification of the relevant information set used in rational forecasts is not necessary for the cross-equation tests of rationality . . . and short-run neutrality to have desirable asymptotic properties."[8,9] Two recent reviewers of Mishkin's (1983) book have stressed this conclusion.[10] Attfield and Duck (1983) and Kormendi and Meguire (1984) justify their empirical tests of the Lucas proposition by appealing to special "plausible assumptions" about the properties of their error in measuring rational expectations.[11]

This paper attempts a thorough examination of the implications of mismeasurement of rational expectations of economic agents. We examine its effects on the statistical properties of the estimators and test statistics employed in empirical studies of the three New Classical propositions, discussed above.

---

[7] This statement will be substantiated throughout this paper.

[8] More detailed comments on the Abel and Mishkin results are contained in section 3.2 of this paper.

[9] Abel and Mishkin also apply their framework of analysis to tests of the rational expectations-efficient capital markets hypothesis. This area is not dealt with in our paper.

[10] Kaufman (1984, p. 397) regards this finding as one of the most important results of Mishkin's book. The result is also singled out for mention by Sheffrin (1984).

[11] These assumptions will be discussed in section 6 of our paper.

We demonstrate that, for the short-run neutrality and rational expectations hypotheses, these statistical properties depend critically on the nature of the mismeasurement committed by the investigator. Mismeasurement can result because the investigator, while succeeding in modelling the forecast errors as a white noise process, uses a Wold representation of lower dimension than the agents'. In this case, we prove that, while the estimators are consistent, the test statistics used in the literature do not possess an asymptotic χ^2 distribution. We also consider the case where the forecast errors are not necessarily white noise, and would arise, for example, from structural misspecification. Here we prove the estimators to be inconsistent and the test statistics to be invalid.[12] Rappoport (1984) provides empirical evidence that the second of these two cases describes the principal models that have been used in the literature. In conjunction with our results, this evidence implies that the statistical inferences drawn in those studies on these two hypotheses are invalid. This implication applies equally to studies that support and to those that contradict the short-run neutrality hypothesis.

As for the Lucas proposition, we demonstrate that mismeasurement vitiates inference, irrespective of the type of mismeasurement that has occurred.

In view of these results, we propose alternative methods for empirical investigation of the questions posed by the New Classical Macroeconomics. Estimators of crucial parameters are consistent when the measured forecast errors are white noise, although the resulting test statistic does not have a

---

[12] We also note that the result that consistency of the estimators hinges on the properties of the forecast errors makes the problem considered here different from the standard errors-in-variables problem. In addition, the measurement error is not uncorrelated with the true value of the variable measured and it occurs in a non-linear and multi-equation model estimated by full-information methods. The techniques used to derive the results in this paper may thus be of interest in their own right.

χ^2 distribution. However, we suggest a GLS-type procedure that corrects for this problem. The appropriateness of this procedure hinges on the investigator's success in reducing the measured forecast errors to white noise. We briefly survey some recent empirical evidence and provide a suggestive example to indicate some potential difficulties in achieving this in practice. We have not been able to come up with an alternative procedure for testing for rationality in the presence of other types of mismeasurement. However, the short-run neutrality hypothesis may be tested using an instrumental variable estimator. This procedure does not require an explicit proxy for rational expectations, and so is not vulnerable to mismeasurement problems.

Finally, we mention an interesting hypothesis that is immune to mismeasurement, and that has not been tested in the literature. This hypothesis states that the short-run effect of nominal variables on aggregate real output is the same irrespective of whether their movements are anticipated or not. Accordingly we name this the hypothesis of "Irrelevance of the Anticipated-Unanticipated Distinction" (IAUD).

The structure of the paper is as follows. Section 1 sets out the framework for analyzing mismeasurement of economic agents' rational expectations. It discusses the implications of mismeasurement for the interpretation of the restrictions used in the literature to characterize rational expectations, and distinguishes between the two types of mismeasurement to which we alluded above. Section 2 describes explicitly the model of real output and money growth which is used to test the Lucas proposition and the hypotheses of short-run neutrality and rationality. The third section contains the principal results on estimation and inference in tests of the short-run neutrality and rationality hypotheses. The proofs of these results are contained in an

Appendix. The implications of the results in section 3 are summarized in section 4, and alternative methods of estimation and testing are discussed in section 5. In section 6 we turn to the Lucas proposition. Section 7 presents a schematic summary of our results.

1. Mismeasurement of Rational Expectations

The rational expectations hypothesis has been used widely as a recipe for modelling expectations empirically. It is interpreted as implying that agents will use all available information to make forecasts of the variable of interest, say x,[13] that are unbiased and have minimum error variance. This has led to the practice of modelling rational expectations as the least squares projection of x on the available information (Sargent, 1973, p. 167). Indeed, the accepted procedure has come to involve running a least squares regression of x_t on a list of variables, say z_t, considered to be available to agents for predicting x_t one period prior to its occurrence.[14] However, this procedure does not guarantee the accuracy of the econometrician's measurement of rational expectations. Several authors, notably Startz (1983), Abel and Mishkin (1983) and Attfield (1983), have considered the effects of omitting variables from the equation for x. This section extends the work of these authors and develops a framework that describes explicitly the difference between rational expectations of economic agents and the investigator's measurement of them.

---

[13] We shall introduce the following notational conventions at this stage. Lower case letters with a time subscript refer to individual time-series observations. Capital letters with a time subscript refer to lists of T time-series observations, ending at the date of the subscript. Thus, for example, if x_t is the observation on x at time t, then X_T is $(x_1',\ldots,x_T')'$. Unadorned lower case letters will be used to refer to variables generalically.

[14] The fact that z_t has the time subscript "t" should not necessarily be construed that agents use the realizations of time t to forecast x_t.

We model rational expectations, x_t^{re}, as the least squares projection of x_t on the information set ψ_t. We assume, for the sake of concreteness, that ψ_t exhausts the information available at t − 1 for forecasting x_t. It can contain, for example, "raw" variables, functions of variables and parameters, or products of dummy and stochastic variables. The set ψ_t is concocted in such a way that the time series x_t^{re} may be written as the least squares projection of x_t on ψ_t. Thus, the characterization of rational expectations used here does not rule out, for example, expectations mechanisms that are non-linear in variables, have time-varying coefficients, or that summarize switches among members of a set of expectations functions.[15] The definition of x_t^{re} implies that

$$x_t = x_t^{re} + e_t. \tag{1}$$

Here, $\{e_t\}$ is a sequence of i.i.d. (white noise) random variables, and

$$x_t^{re} = \psi_t \delta_0 \tag{2}$$

where

$$\delta_0 = \operatorname*{plim}_{T\to\infty}(\Psi_T'\Psi_T)^{-1}\Psi_T'X_T$$

$$\Psi_T = (\Psi_1',\ldots,\Psi_T')'$$

$$X_T = (x_1,\ldots,x_T)'$$

In this paper we assume that the probability limits (plim) in expressions like the one for δ_0 exist.

In contrast, we assume that the investigator projects x_t only on $z_t \subset \psi_t$, yielding a proxy for rational expectations, $\hat{x}_t^{re}$, defined by

$$\hat{x}_t^{re} = z_t \delta \tag{3}$$

---

[15] We do not pass judgment on whether these expectations are plausible or feasible as models of rational expectations of agents.

where[16]

$$\delta = \underset{T\to\infty}{\text{plim}}(Z_T'Z_T)^{-1}Z_T'X_T$$

$$Z_T = (z_1',\ldots,z_T')'$$

The difference between x_t^{re} and $\hat{x}_t^{re}$ may be expressed by projecting the former on z_t. As e_t is white noise, this results in the decomposition

$$x_t^{re} = z_t\delta + v_t = \hat{x}_t^{re} + v_t \tag{4}$$

where

$$v_t = [\psi_t - z_t \underset{T\to\infty}{\text{plim}}(Z_T'Z_T)^{-1}Z_T'\Psi_T]\delta_0.$$

Thus, v_t is the error in measurement of x_t^{re} committed by the investigator. It is orthogonal to z_t in the sense that $\underset{T\to\infty}{\text{plim}}\ Z_T'V_T = 0$. The combination of equations (1) and (4) exhibits the breakdown of x_t into the observable and unobservable components of rational expectations ($z_t\delta$ and v_t respectively) and the white noise rational forecast error (e_t):

$$x_t = z_t\delta + v_t + e_t. \tag{5}$$

1.1. Interpreting the Rational Expectations Hypothesis

In the framework of this paper, the investigator is able to study the formation of expectations and their effects only via the variables in z and their interaction with other variables. In order to formulate the restrictions implied by rational expectations that can be observed by the investigator, we need a representation, similar to (4), of an arbitrary expectations variable,

---

[16] In fact, the investigator has to estimate δ, and so (if he or she uses data from time 1 to T) the numerical value used is $z_t(Z_T'Z_T)^{-1}Z_T'X_T$. Since our analysis is conducted in terms of probability limits of estimators, we abstract from this additional problem in this paper.

x_t^e. It is obvious that just as x_t^{re} and x_t may be decomposed into components explained by a least squares projection on z_t and the associated residual, so may x_t^e. Formally this decomposition is given by

$$x_t^e = z_t\delta^* + v_t^* \tag{6}$$

where

$$v_t^* \equiv x_t^e - z_t\delta^*$$

$$\delta^* = \text{plim}(Z_T'Z_T)^{-1}Z_T'X_T^e ,$$

and v_t^* and z_t are orthogonal in the sense defined above for v_t and z_t.

Combining equations (4) and (6), the condition that expectations are rational may be expressed as

$$z_t\delta + v_t = z_t\delta^* + v_t^* \tag{7}$$

By projecting both sides of (7) onto z_t, it can be seen that this condition is equivalent to the restrictions:

$$\delta = \delta^* \tag{8a}$$

and

$$v_t = v_t^* \quad \text{for all } t. \tag{8b}$$

The hypothesis that expectations are rational is expressed in the literature by the restriction on the parameters of (7)[17]

$$\delta = \delta^* \tag{9}$$

We shall refer to this restriction as the "RE" hypothesis. It is apparent

---

[17]An example of this restriction in the context of tests of short-run policy neutrality is presented in the next section.

that it falls short of the demands of rational expectations which are given by equations (8). This means that one has to exercise great care when doing such things as "treating rational expectations as a maintained hypothesis." In our analysis, this will always mean that equations (8) hold, rather than just equation (9) on its own. However, when discussing "tests of the rational expectations hypothesis" that have been carried out in the literature, we refer to the tests of (9).

Comparison of equations (8) and (9) makes it clear that the latter is only a necessary and not a sufficient condition for rational expectations. That is, the condition $\delta = \delta^*$ encompasses a much larger class of expectations than rational expectations given by the projection of x_t on ψ_t. The following conditions are jointly sufficient for x_t^e to be a member of this class:

(i) Expectations are equal to the projection of x_t on the information set $\psi_t^* \subseteq \psi_t$.

(ii) The information set used by the investigator, z_T, is a linear combination of the information set used by agents, ψ_T^*. That is, $Z_T = \Psi_T^* A$, where $Z_T = (z_1', \ldots, z_T')'$, $\Psi_T^* = (\psi_1^{*\prime}, \ldots, \psi_T^{*\prime})'$, and A is a matrix with rows and columns conformable in number to the columns of Ψ_T^* and Z_T, respectively.

Under these two conditions we have from (6) and (3):

$$\delta^* = \text{plim}(Z_T'Z_T)^{-1}Z_T'X_T^e$$

$$= \text{plim}(Z_T'Z_T)^{-1}Z_T'\Psi_T^*(\Psi_T^{*\prime}\Psi_T^*)^{-1}\Psi_T^{*\prime}X_T$$

$$= \text{plim}(Z_T'Z_T)^{-1}A'\Psi_T^{*\prime}X_T$$

$$= \text{plim}(Z_T'Z_T)^{-1}Z_T'X_T = \delta$$

We also have

$$v_t^* = \psi_t^* \ \text{plim}(\Psi_T^{*\prime}\Psi_T^*)^{-1}\Psi_T^{*\prime}X_T - Z_t\delta^*$$

$$= \psi_t^* \ \text{plim}(\Psi_T^{*\prime}\Psi_T^*)^{-1}\Psi_T^{*\prime}X_T - Z_t\delta$$

$$v_t = \psi_t \ \text{plim}(\Psi_T^{\prime}\Psi_T)^{-1}\Psi_T^{\prime}X_T - Z_t\delta$$

Thus, when $\psi_t^* \neq \psi_t$, conditions (i) and (ii) define in addition to rational forecasts a large class of non-rational forecasts for which $\delta = \delta^*$, but $v_t \neq v_t^*$.

A plausible example of the class of non-rational forecasts satisfying the conditions of (i) and (ii) occurs when agents form expectations by projecting x onto their information ψ_t^*. Then, our result implies that the restriction tested in the literature, $\delta = \delta^*$, is satisfied as long as the agents use a "better" information set than does the investigator, that is, $z_t \subseteq \psi_t^* \subset \psi_t$.[18]

1.2. Measured Forecast Errors

The class of proxies for rational expectations used by the investigator is characterized by equation (5), or equivalently by the variables in z, given ψ. We will distinguish between two subclasses. The first contains models which are Wold representations of x.[19] Since the set of variables used by the investigator (z_t) is assumed to be a proper subset of the exhaustive information set (ψ_t), the Wold representation obtained by the investigator is accordingly

---

[18]This result can be shown to hold in contexts other than the policy neutrality-rational expectations literature. For example, it holds in the large class of models analyzed by Startz (1983). Note that the result may not cover the more general class of models containing forecasts of endogenous variables.

[19]Since (5) is not generally in moving average form, this definition of potential models of x requires that an autoregressive representation exists. See Sargent (1979, p. 263) for further remarks on this point. We include in this subclass models of x which are implied by Wold representations of an appropriate transformation (like differencing) of x.

assumed to be of lower dimension (involving fewer variables) than the true representation. All models of the form (5) that are not Wold representations of x are contained in the second subclass. It comprises, for example, distributed lag models of x constructed by the investigator that are not Wold representations. It also comprises regression models combining time-series and structural variables (e.g., indicators of shifts in regimes, supply shocks and various nonstationarities).

We will now describe the properties of the measured forecast errors that arise from each of the subclasses above. The properties of the errors v and v + e in (5) depend on the true process generating x, embodied in the information set ψ, the projection of x on ψ, and the variables in z selected by the investigator in the construction of his or her proxy for rational expectations. If the model of x is a Wold representation of x, the forecast error, v + e, is, by construction, a sequence of i.i.d. random variables. In such a case we call v + e a "Wold measured forecast error" (Wold error, for short). However, the measurement error, v, is a sequence of stationary random variables which, in general, are correlated over time. This fact is a consequence of the Wold decomposition theorem and is illustrated by the following example.

<u>Example 1</u>: We assume that $\psi_t = (x_{t-1}, z_{t-1}, x_{t-2}, z_{t-2}, \ldots)$ and suppose that x_t and z_t have the following true bivariate Wold representation:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 z_{t-1} + e_t \qquad (10)$$

$$z_t = \zeta_t - \alpha_3 \zeta_{t-1}$$

The assumption that (9) - (10) is a Wold representation of x and z means that e_t and ζ_t are white noise processes uncorrelated with each other at all leads and lags. Furthermore, the assumption that (10) - (11) is the true representation means that e_t and ζ_t are uncorrelated with variables in ψ_t. Thus equation (10) is an example of equations (1) and (2).

We also suppose that the investigator attempts to construct the univariate Wold representation for x. Using (11) and (10) we obtain

$$x_t = \alpha_1 x_{t-1} + \alpha_2 \zeta_{t-1} - \alpha_2 \alpha_3 \zeta_{t-2} + e_t \tag{12}$$

Using Wold's theorem we can write the error term in (12) as

$$\alpha_2 \zeta_{t-1} - \alpha_2 \alpha_3 \zeta_{t-2} + e_t = \mu_t - \phi \mu_{t-1} \tag{13}$$

where μ_t is a white noise error process.[20]

Combining (13) and (12) we obtain the univariate Wold representation of x:

$$\frac{1 - \alpha_1 L}{1 - \phi L} x_t = \mu_t \tag{14}$$

Equation (14) is an example of equation (5) with $\mu_t = v_t + e_t$. We note that by construction μ_t is uncorrelated with lagged values of x_t. However, μ_t is correlated with lagged values of z_t. This can be seen by comparing equation (13), rewritten as,

$$\mu_t = v_t + e_t = \frac{1}{1 - \phi L} (\alpha_2 \zeta_{t-1} - \alpha_2 \alpha_3 \zeta_{t-2} + e_t) \tag{15}$$

with equation (11). Thus, the use of a lower dimensional Wold representation when the true representation is higher dimensional can be interpreted as an instance of omitting a variable (z_{t-1}) from the forecasting equation for x_t. The example presents a special case of the omitted variable problem, since the

[20]The parameter ϕ can be explicitly computed in terms of α_2, α_3, Var(ζ_t), and Var(e_t). Although the explicit expression for ϕ plays no role in our analysis, the fact that ϕ depends on these structural parameters is important for section 6.

the resulting forecast error, v + e, is still white noise, although it is correlated with lagged values of the omitted variable.

Equation (15) also allows us to derive the explicit expression for v_t:

$$v_t = \frac{1}{1 - \phi L} (\alpha_2 \zeta_{t-1} - \alpha_2 \alpha_3 \zeta_{t-2} + \phi e_{t-1}) \tag{16}$$

Since $|\phi| < 1$ this shows that v is an ARMA(1,2) process.[21]

The above discussion would seem to suggest that it is simple to formulate the forecasting model for x with white noise forecast errors using only a subset of the variables required for the true multivariate representation of x. In fact, empirical attempts to model such representations have not to date succeeded in this task.[22] Moreover, there are plausible a priori reasons for the difficulties encountered in applying the Wold theorem in practice. It has been noted by others (cf. Neftci and Sargent (1978)) that the rate of growth of the money supply may change with changes in policy rules, economic structure or institutions in financial markets. If x is not covariance stationary during the sample period due to any of these or other changes, the Wold theorem is not applicable. The recognition of covariance nonstationarity of x requires structural modelling of changes in the process generating x.

In view of the above considerations and in order to analyze the validity of inferences in studies that have failed to generate Wold representations or

---

[21]The fact that v_t, which is unobservable, will in general be correlated over time does not seem to have been recognized in the literature analyzed in this paper. Attfield (1983) implicitly assumed that v is a white noise process (see his equation (2) for y_t on p. 282). Abel and Mishkin (1983) also carry out their analysis under this implicit assumption (for example, see their Theorem on p. 9). Kormendi and Meguire (1984) explicitly assume that x and v are white noise processes in their discussion of the effects of mismeasurement of rational expectations.

[22]A discussion of these empirical studies can be found in section 4.

true structural models of x, we shall also consider "non-Wold measured forecast errors" (non-Wold errors, for short), v + e. These arise in the second subclass of models described above. Non-Wold errors are not white noise and/or are correlated with lagged values of explanatory variables in the investigator's model of x.[23] It is also important to note that non-Wold errors can be either stationary or non-stationary. Models that are not Wold representations of x and some properties of non-Wold errors can be illustrated by an example.

<u>Example 2</u>: We suppose that the true process generating x is given by an AR(1) process whose coefficient changes at $t = t_0$. Thus,

$$x_t = \delta_1 x_{t-1} + e_t \quad \text{for } t > t_0 \tag{17}$$

$$x_t = \delta_2 x_{t-1} + e_t \quad \text{for } t \leq t_0 \tag{18}$$

Combining (17) and (18) we have:

$$x_t = \delta_1 x_{t-1} + (\delta_2 - \delta_1) d_t x_{t-1} + e_t \tag{19}$$

where $d_t = 0$ for $t > t_0$, and $d_t = 1$ for $t \leq t_0$. Equation (19) is an example of equations (1) and (2) with $\psi_t = (x_{t-1}, d_t x_{t-1}, x_{t-2}, d_{t-1} x_{t-2}, \ldots,)$.

Furthermore, we suppose that the investigator does not take into account the fact that there has been a structural change at $t = t_0$ and he or she fits an AR(1) process to the data for the period $t = 1, \ldots, t_0, \ldots, T$. Thus, the investigator's equation for x_t omits $d_t x_{t-1}$. The resulting proxy for rational expectations (see (3)) is given by

$$\hat{x}_t^{re} = x_{t-1}\delta \tag{20}$$

[23] Of course, v + e is orthogonal by construction to the "contemporaneous" values (z_t) of the regressors in (5).

where[24]

$$\delta = \frac{\delta_1(1 - \delta_2^2) + \delta_2(1 - \delta_1^2)}{2 - \delta_1^2 - \delta_2^2}$$

Using (5) in (17) and (18) we have

$$v_t = (\delta_1 - \delta)x_{t-1} \quad \text{for } t > t_0 \tag{21}$$

$$v_t = (\delta_2 - \delta)x_{t-1} \quad \text{for } t \leq t_0 \tag{22}$$

Equations (21) and (22) imply that:

$$v_t = \delta_1 v_{t-1} - \delta e_{t-1} \quad \text{for } t > t_0 \tag{23}$$

$$v_t = \delta_2 v_{t-1} - \delta e_{t-1} \quad \text{for } t \leq t_0 \tag{24}$$

Thus, in this case both v + e and v are correlated over time and covariance nonstationary. Straightforward, but tedious, computations using (17) - (18) and (20) - (22) yield the following expression for the autocorrelation of v + e at lag j:

$$\rho_j^{v+e} = \frac{(\delta_2 - \delta_1)^2(\delta_1^j(1 - \delta_1^2) + \delta_2^j(1 - \delta_2^2)]}{(2 - \delta_1^2 - \delta_2^2)[2(2 - \delta_1^2 - \delta_2^2) + (\delta_1 - \delta_2)^2]} \tag{25}$$

We evaluated the expression for ρ_1^{v+e} on the grid of values $\delta_i = 0,.1,\ldots,.9$ for i = 1,2. The largest value of ρ_1^{v+e} among the one hundred computed is .06 (when $\delta_1 = .3$ and $\delta_2 = .9$, or vice versa). Furthermore, it is clear from (25) that higher order autocorrelations are much smaller than ρ_1^{v+e}. Abstracting

---

[24]Note that $\delta = \text{plim}(X'_{T-1}X_{T-1})^{-1}X'_{T-1}X_T$ where the plim is taken on both sides of the point $t = t_0$, that is, T ranges from $-\infty$ to $+\infty$. This modification of the expression for δ below (3) will be used, whenever applicable, throughout this paper.

from sampling error, the Box-Pierce test based only on ρ_1^{v+e} would require 266 years of quarterly observations to reject the hypothesis of white noise errors at the 5% level. While no problems with the model would thus be diagnosed the correlation between $v_t + e_t$ and lagged values of x_{t-1} remains substantial. In the above example, although $\rho_1 = .06$, it can be shown that

$$\frac{1}{Var(e_t^2)} \underset{T\to\infty}{plim} \sum_{t=t_0-T}^{t_0+T} (v_t + e_t)x_{t-2} = .32.$$

2. The Framework Used to Test Short-Run Policy Neutrality and Rational Expectations

The policy neutrality-rational expectations literature is couched in a model whose central equation is a generalization of the Lucas (1973) supply function. In our discussion we will use the following form of this function:[25]

$$y_t = \sum_{i=0}^{M-1} \beta_i (x_{t-i} - x_{t-i}^e) + \sum_{i=0}^{M-1} \theta_i x_{t-i}^e + u_t. \tag{26}$$

Here, y is the deviation of output from its natural rate; x is the rate of growth of the money supply; x_t^e is the forecast of x_t based on information available at time t - 1; u is a spherically distributed random variable uncorrelated with any variables on the right hand side of the equation; and the parameter vectors $\beta = (\beta_0,\ldots,\beta_{M-1})$ and $\theta = (\theta_0,\ldots,\theta_{M-1})$ reflect sensitivity of output to unanticipated and anticipated x, respectively.

Employing the representation for x_t^e in (6), the output equation may be expressed as

$$y_t = \sum_{i=0}^{M-1} \beta_i (x_{t-i} - z_{t-i}\delta^*) + \sum_{i=0}^{M-1} \theta_i z_{t-i}\delta^* + \eta_t^* \tag{27}$$

---

[25] In empirical studies the output equation also contains other variables. The omission of those variables from (26) simplifies our presentation without materially altering our results.

where

$$\eta_t^* = \sum_{i=0}^{M-1} (\theta_i - \beta_i) v_{t-i}^* + u_t.$$

The equation for money growth is (5), repeated here for convenience:

$$x_t = z_t \delta + v_t + e_t \tag{5}$$

It will greatly facilitate the derivation of our results is we cast the system of equations (27) and (5) in matrix notation. To this end, we use upper case letters without a time subscript to represent collections of variables in the output equation as arrays. For example,

$$X = (X_T, \ldots, X_{T-(M-1)})$$

$$Z = (Z_T, \ldots, Z_{T-(M-1)}),$$

and so on. X is a $T \times M$ matrix, and Z is a $T \times KM$ matrix, since Z_T, Z_{T-1}, etc. are $T \times K$ matrices. This permits us to rewrite all of T observations of (27) as:

$$Y_T = (X - Z(I_M \otimes \delta^*))\beta + Z(I_M \otimes \delta^*)\theta + V^*(\theta - \beta) + U_T \tag{28}$$

Using the same notation, the corresponding version of the money growth equation is

$$X_T = Z_T \delta + V_T + E_T = Z_T \delta + \Lambda_T \tag{29}$$

The short-run neutrality of money, to which we shall refer as the hypothesis of the "ineffectiveness of anticipated policy" (IAP), is modelled by the restriction:

$$\theta \equiv 0 \tag{30}$$

The rational expectations (RE) hypothesis is expressed in the literature by the restriction on the parameters of (28) and (29)

$$\delta = \delta^{*} \tag{31}$$

The following combinations of hypotheses expressed in (30) and (31) have been considered in the literature:

<u>IAP-RE</u>: Both hypotheses are treated as testable. The constrained version of (28) and (29) is characterized by the set of restrictions

$$\theta = 0 \quad \text{and} \quad \delta = \delta^{*} \tag{32a}$$

while the unconstrained system is characterized by

$$\theta \neq 0 \quad \text{and} \quad \delta \neq \delta^{*} \tag{32b}$$

<u>IAP|RE</u>: IAP is treated as a testable hypothesis while RE is maintained. The restrictions corresponding to the constrained and unconstrained systems are given by

$$\theta = 0; \text{ and } \quad \delta = \delta^{*} \tag{33a}$$

and

$$\theta \neq 0; \text{ and } \quad \delta \neq \delta^{*} \tag{33b}$$

respectively.

<u>RE</u>: Some authors (Leiderman (1980), Mishkin (1983, Ch. 6)) have attempted to test the RE restriction (31) without maintaining any other restrictions. Abel and Mishkin (1983, p. 20) demonstrated that θ, $\delta(=\delta^{*})$, and β are just identified, when the RE restriction is imposed. However, θ and δ^{*} are not separately identified when the RE restriction is not imposed during estimation. Therefore, likelihood ratio tests that compare restricted and unrestricted

models yield on evidence on the validity of the RE hypothesis. Consequently, we do not consider this case further.

As Abel and Mishkin (p. 20) point out, the following combination of hypotheses does not suffer from this problem:

RE|IAP: RE is treated as testable, while IAP is maintained. Formally, we have

$$\delta = \delta^*; \text{ and } \theta = 0 \tag{34a}$$

$$\delta \neq \delta^*; \text{ and } \theta = 0 \tag{34b}$$

We should also add that following our analysis in the previous section we will always impose the condition that $v_t = v_t^*$ for all t in (28) whenever we consider the hypothesis that expectations are rational.[26]

2.1. Estimation Methods

Two methods of estimation have been employed in the literature. One of the procedures (cf. Mishkin (1983), pp. 17-18) involves minimization of the objective function

$$S(\beta, \theta, \delta, \delta^*) = S_X S_Y \tag{35}$$

where

$$S_X = (X_T - Z_T\delta)'(X_T - Z_T\delta)$$

$$S_Y = [Y_T - X\beta - Z(I_M \otimes \delta^*)(\theta - \beta)]'[Y_T - X\beta - Z(I_M \otimes \delta^*)(\theta - \beta)]$$

subject to the various constraints imposed in testing the RE and IAP hypotheses.

---

[26] Note that if only (31) is imposed, as is the practice in the literature, the error term in the output equation would contain $v_t^*, v_{t-1}^*, \ldots$. This would create additional difficulties in the interpretation of tests carried out in the literature. This point will be further discussed in section 3.1.3. and footnote 32.

We shall refer to this procedure as the Least Product Squares (LPS) method.[27]

We recall from the previous section that the constrained system is the same for all three hypotheses, that is, it imposes both IAP and RE restrictions in estimation. Each test involves a comparison of this constrained system with a particular unconstrained system. We use the subscripts "c" and "u" to refer to estimators and their associated probability limits in the constrained and unconstrained systems, respectively. We also use a tilde to signify a least product of squares (LPS) estimator and a symbol without a tilde to denote a probability limit of an estimator. Thus $\tilde{\delta}_c$ denotes the estimator of δ in the constrained system and δ_c denotes the probability limit of $\tilde{\delta}_c$. However, note that we will use $\tilde{\beta}_c$ to denote the estimator of β in the constrained system, although the parameter β itself is never constrained.

The second procedure used in the literature (cf. Barro (1977)) is the two-stage method. The first stage involves OLS estimation of the money growth equation (29) and yields a proxy $Z_T\hat{\delta}$. This proxy is used in the second stage to estimate β and θ in the output equation (28).

The approach to modelling of rational expectations outlined in the previous section tracks the unobservable expectations variable by following the behavior of a portion of this variable, $z_t\delta$. In addition to the problems that this introduces for interpreting the RE hypothesis, it creates problems for estimation and inference. We now turn to these issues.

3. Estimation and Inference in Tests of the RE and IAP Hypotheses

In this section, we analyze the effects of mismeasurement of rational expectations on the statistical properties of the estimators and test statistics used in empirical studies of the neutrality and rationality

---

[27] The usual motivation for this procedure is that it is equivalent to full-information maximum likelihood procedure. This equivalence needs to be qualified. It will be further discussed in section 5.

hypotheses. Broadly stated, we prove that, in the case of non-Wold errors, crucial parameter estimators are inconsistent and the test statistics derived from them are uninterpretable. In the case of Wold errors, consistency of the estimators would be preserved. However, the test statistics used in the literature would still not possess asymptotic χ^2 distributions.

In addition to the distinction between Wold and non-Wold errors, it is important to differentiate between two specifications of the output equation. One involves both contemporaneous and lagged values of anticipated and unanticipated money growth. The other contains only the contemporaneous values. We shall refer to these as the "lagged" and "no-lags" cases, respectively.

3.1. The "Lagged" Case

3.1.1. Intertemporal Correlations Among z, v, and e.

Our point of departure is the observation that in the model (28) - (29) intertemporal correlations among z, v, and e will play a role in the asymptotic properties of estimators and test statistics. The correlations have to be specified before these asymptotic properties can be derived. Some of the correlations are known a priori. As far as others are concerned, we will allow for the possibility of correlations which cannot be ruled out a priori. We will distinguish, whenever differences arise, between Wold and non-Wold errors.

(i) Correlations between e and z. Since e is regarded as the innovation in the true representation of x, it will not be correlated with lagged values of z. However, correlations of z with the lagged values of e cannot be ruled out a priori. For example, z will be correlated with lagged values of e if z contains lagged values of x. These observations are formalized in the following conditions:[28]

$$\text{plim}\ \frac{1}{T}\ Z_T'E_{T-i} = \begin{cases} 0_K & \text{if } i \leq 0 \\ \text{not necessarily zero} & \text{if } i > 0 \end{cases} \tag{36}$$

(ii) Correlations between e and v + e. The same argument as in (i) implies that e is not correlated with lagged values of v + e. If v + e is a Wold error it is correlated with lagged values of e by construction (cf. (15)). Alternatively, if v + e is a non-Wold error, correlations between v + e and lagged values of e cannot be ruled out a priori (cf. (23) - (24)). Thus we have:

$$\text{plim}\ \frac{1}{T}\ (V_T + E_T)'E_{T-i} = \begin{cases} 0 & \text{if } i \leq 0 \\ \text{not necessarily zero} & \text{if } i > 0 \end{cases} \tag{37}$$

(iii) Correlations between z and v + e. If v + e is a Wold error it is uncorrelated with lagged values of z. However, z will in general be correlated with lagged values of v + e (cf. (11) and (15)). These observations can be summarized in:

$$\text{plim}\ \frac{1}{T}\ (V_T + E_T)'Z_{T-i} = \begin{cases} 0_K & \text{if } i \leq 0 \\ \text{not necessarily zero} & \text{if } i > 0 \end{cases} \tag{38a}$$

In contrast, if v + e is a non-Wold error, correlations between v + e and lagged as well as future values of z cannot be ruled out (cf. (21) - (22)). Thus we have:

---

[28] As above, we shall use capital letters to refer to vectors or matrices of time series of T columns (observations), the date of the last observation being the same as the subscript of the array. Thus, for example, $E_{T-i} = (e_{-(i-1)},\dots,e_0,\dots,e_{T-i})'$. The symbol 0_K denotes the K-element zero column vector.

$$\text{plim}\ \frac{1}{T}\ (V_T + E_T)'Z_{T-i} = \begin{cases} 0_K & \text{if } i = 0 \\ \text{not necessarily zero} & \text{if } i \neq 0 \end{cases}$$

3.1.2. Assumptions

The following assumptions are common to all our results in this section:

(i) The true output equation is:

$$Y_T = (X - X^e)\beta + X^e\theta + U_T \text{ and } \beta \neq \theta.$$

(ii) The condition that expectations are rational is (correctly) represented by the equations (25): $\delta = \delta^*$ and $v_t = v_t^*$ for all t.

(iii) $\text{plim}\ \frac{1}{T}\ Z'U_T = \text{plim}\ \frac{1}{T}\ V'U_T = 0$

$\text{plim}\ \frac{1}{T}\ E'E = \sigma_\varepsilon^2 I_M$

$\text{plim}\ \frac{1}{T}\ Z'Z < \infty$ and nonsingular.

(iv) The probability limits of $\frac{1}{T}$ times the following cross-product matrices are nonsingular for all non-zero values of the parameter d:

$X'(X - Z(I_M \otimes d))$, $(X - Z(I_M \otimes d))'(X - Z(I_M \otimes d))$,

$(X,Z(I_M \otimes d))'(X,Z(I_M \otimes d))$.

We assume in (i) that $\beta \neq \theta$. It can be seen from equation (28) that if $\beta = \theta$ no mismeasurement problems arise in the output equation. However, the condition $\beta = \theta$ assuredly does not describe the maintained hypothesis in the tests of the IAP and RE hypotheses carried out in the literature. Under IAP this condition implies $\beta = 0$. Furthermore, if $\beta = \theta$, then the output equation can give no information concerning RE. This condition will be further discussed in section 5.

3.1.3. Behavior of Estimators and Test Statistics: The Case of Non-Wold Errors

In addition to assumptions (i) - (iv) we also assume:

(v) The intertemporal correlations among v, z and e are given by (36), (37) and (38b).[29]

We begin by considering the joint RE and IAP (IAP-RE) hypothesis. Using (32a) and (28) - (29) the constrained system can be written as:

$$Y_T = X\beta - Z(I_M \otimes \delta)\beta - V\beta + U_T \tag{39}$$

$$X_T = Z_T\delta + V_T + E_T \tag{29}$$

The unconstrained system comprises (29) and (28), repeated here for convenience:

$$Y_T = X\beta + Z(I_M \otimes \delta^*)(\theta - \beta) + V^*(\theta - \beta) + U_T. \tag{40}$$

As we noted in section 2, δ^* and θ in (40) are not separately identified. Only the product $\alpha = (I_M \otimes \delta^*)(\theta - \beta)$ can be identified. Thus, for any values of α, β and δ^*, a value of θ that satisfies the equation for α can be found. Without altering the substance of our argument, it is convenient to set $\delta^* = \delta$. Given the estimators $\tilde{\alpha}_u$ and $\tilde{\beta}_u$, the following expression formally defines the "estimator," $\tilde{\theta}_u$:

$$\tilde{\alpha}_u = (I \otimes \delta)(\tilde{\theta}_u - \tilde{\beta}_u). \tag{41}$$

In essence, this approach fixes δ^* by using the null hypothesis of RE.[30]

---

[29] Condition (39b) amounts to the assumption that $\text{plim} \frac{1}{T} (I \otimes \delta')Z'(V + E)$ is nonsingular. The results that follow can be proved even when this matrix is singular, as long as v + e is not white noise. Such a singularity can be interpreted as the result of Barro and Rush's (1980) procedure in which they include all significant lags of z in their final money growth equation. The proofs are omitted to save space and are available on request.

[30] This is convenient because we will evaluate all probability limits, denoted by "plim", under the assumption that null and maintained hypotheses H_0 are valid.

(Equivalently, we could have used the IAP restriction, $\theta = 0$, to derive an expression like (41) that would define the "estimator" $\tilde{\delta}_u^*$.)

The following theorem deals with the asymptotic behavior of the estimators involved in testing the IAP-RE hypothesis:

Theorem 1: Suppose that assumptions (i) - (v) hold. Let H_0 describe the joint IAP and RE hypothesis. That is:

$$H_0: \quad \delta = \delta^*, \; V = V^*, \text{ and } \theta = 0$$

$$H_A: \quad \delta \neq \delta^* \text{ or } V \neq V^* \text{ or } \theta \neq 0 \quad \text{(or both).}$$

Then,[31]

(A) $\operatorname*{plim}_{H_0} \tilde{\delta}_c = \delta_c \neq \delta$

(B) $\operatorname*{plim}_{H_0} \tilde{\beta}_c = \beta_c \neq \beta$ and $\operatorname*{plim}_{H_0} \tilde{\beta}_u = \beta_u \neq \beta$

(C) $\operatorname*{plim}_{H_0} \tilde{\theta}_u = \theta_u \neq 0$

Theorem 1 invalidates the test of the IAP-RE hypothesis employed in the literature in the case of non-Wold errors. It says that if δ^* is in fact equal to δ, the estimator of δ in the constrained system, $\tilde{\delta}_c$, does not converge to δ asymptotically. This result implies that the standard test statistic used in the literature does not have a χ^2 distribution asymptotically. This test statistic can be written as (cf. Mishkin (1983, p. 18)):

$$PLR = T \ln \left(\frac{\tilde{S}_X^c \tilde{S}_Y^c}{\tilde{S}_X^u \tilde{S}_Y^u} \right) \tag{42}$$

---

[31] The proof of this theorem and of other results of this paper are contained in the Appendix.

where $\tilde{S}_X^c, \tilde{S}_Y^c$ are S_X and S_Y in (35) evaluated at the estimators in the constrained system and $\tilde{S}_X^u$ and $\tilde{S}_Y^u$ are the same quantities for the unconstrained system.

The following corollary deals with the asymptotic behavior of the pseudo-likelihood ratio (PLR) statistic defined in (42):

<u>Corollary 1.1</u>: The PLR statistic used in testing the RE-IAP hypothesis is unbounded in probability limit. Thus, this test statistic does not have an asymptotic χ^2 distribution.

We now analyze the test of the RE hypothesis when the IAP hypothesis is maintained (RE|IAP). The constrained system is again given by (39) and (29). The unconstrained system comprises (29) and (from (34b)):

$$Y_T = X\beta - Z(I_M \otimes \delta^*)\beta - V^*\beta + U_T \tag{43}$$

The following theorem deals with the asymptotic behavior of the estimators involved in testing the RE|IAP hypothesis. Note that the subscript "u" now refers to a different unconstrained system than in Theorem 1.

<u>Theorem 2</u>: Suppose that assumptions (i) - (v) hold. Let H_0 describe the RE|IAP hypothesis. That is,

$$H_0: \ \delta = \delta^*, \ V = V^*; \text{ and } \theta = 0$$

$$H_A: \ \delta \neq \delta^* \text{ or } V \neq V^*; \text{ and } \theta = 0$$

Then,

(A) $\operatorname*{plim}_{H_0} \tilde{\delta}_c = \delta_c \neq \delta$

(B) $\operatorname*{plim}_{H_0} \tilde{\delta}_u^* = \delta_u^* \neq \delta$

(C) $\underset{H_0}{\text{plim}}\ \tilde{\beta}_c = \beta_c \neq \beta$ and $\underset{H_0}{\text{plim}}\ \tilde{\beta}_u = \beta_u \neq \beta$.

Theorem 2 implies a corollary similar to Corollary 1.1. with the PLR statistic appropriately defined for the RE|IAP hypothesis. We state this corollary without proof:

Corollary 2.1: The PLR statistic used in testing the RE|IAP hypothesis is unbounded in probability limit. Thus, this test statistic does not have an asymptotic χ^2 distribution.

Finally, we turn to the test of the IAP hypothesis when the RE hypothesis is maintained (IAP|RE). Once again, the constrained system is given by (39) and (29). From (33b), the unconstrained system comprises (29) and

$$Y_T = X\beta + Z(I_M \otimes \delta)(\theta - \beta) + V(\theta - \beta) + U_T \tag{43}$$

The following theorem summarizes the results for this hypothesis.

Theorem 3: Suppose that assumptions (i) - (v) hold. Let H_0 describe the IAP|RE hypothesis. That is,

H_0: $\theta = 0$; and $\delta = \delta^*$, $V = V^*$

H_A: $\theta \neq 0$; and $\delta = \delta^*$, $V = V^*$.

Then,

(A) $\underset{H_0}{\text{plim}}\ \tilde{\delta}_c = \delta_c \neq \delta$

(B) $\underset{H_0}{\text{plim}}\ \tilde{\delta}_u = \delta$

(C) $\underset{H_0}{\text{plim}}\ \tilde{\beta}_c = \beta_c \neq \beta$ and $\underset{H_0}{\text{plim}}\ \tilde{\beta}_u = \beta_u \neq \beta$

(D) $\underset{H_0}{\text{plim}}\ \tilde{\theta}_u = \theta_u \neq 0.$

There is a formal analogy between the proof of this Theorem and that of Theorem 1 which turns on $\tilde{\delta}_u$. In Theorem 1 we set $\delta^* = \delta$ in order to focus on the properties of the "estimator" $\tilde{\theta}_u$, in view of the lack of identification of δ^* and θ. Here, δ^* in the unconstrained system is genuinely identified by setting it equal to δ in the money growth equation. Consequently, θ is also just identified in the unconstrained system and $\tilde{\theta}_u$ is a bona fide estimator. It is also clear that Theorem 3 implies a corollary analogous to Corollary 1.1.

The consideration of the IAP|RE hypothesis naturally leads to the analysis of the asymptotic behavior of the two-stage estimators of β and θ. We recall from section 2.1 that the two-stage procedure involves estimating the money growth equation (29) by OLS and then using the resulting proxy $Z_T\hat{\delta}$ to estimate β and θ in the second stage. Theorem 3 immediately implies the following corollary:

Corollary 3.1: The two-stage estimators of β and θ are both inconsistent.

Corollary 3.1. also implies that the standard test statistic used to test $\theta = 0$ does not have the F-distribution attributed to it in the literature (cf. Barro (1977)). Thus, the two-stage procedure used in testing the IAP|RE hypothesis is also not valid in the case of non-Wold errors.

We should also note that rational expectations appears as either a maintained or null hypothesis in all of our results in this section. This approach corresponds to the standard approach in the literature. However, it is possible to prove similar results when expectations are not assumed to be rational as long as one is prepared to make assumptions analogous to (36) - (38) above, concerning the intertemporal correlations among z and $x - x^e$

and v^*.[32]

The implication of our results in this section is that statistical inferences based on the least product of squares and two-stage estimation are not valid, when the investigator's equation for money growth generates forecast errors which are not white noise and may be correlated with lagged values of the explanatory variables. All of the critical parameter estimators are inconsistent, and referring computed pseudo-likelihood ratios to the χ^2 distribution does not provide tests of the hypothesis in question at the stated significance levels.

3.1.4. Behavior of Estimators and Test Statistics: The Case of Wold Errors

In addition to assumptions (i) - (iv) we assume:

(vi) The intertemporal correlations among v, z and e are given by (36), (37) and (38a).[33]

The following theorem is analogous to Theorem 1 for the case of Wold forecast errors.

Theorem 4: Suppose that assumptions (i) - (iv) and (vi) hold. Let H_0 describe a joint IAP and RE hypothesis formally stated in Theorem 1. Then,

(A) $\underset{H_0}{\text{plim}}\ \tilde{\delta}_c = \delta$

---

[32] In this case, it would be necessary to specify the correlations in more detail than is exhibited in (36) - (38). since correlations between past x^e and future $x - x^e$ cannot be ruled out (cf. Rappoport (1984), for example). It is the fact that these correlations are zero while others are not that drives the proofs of our results.

[33] We note that in contrast to assumption (v) made in the previous section, assumption (vi), via condition (38a), implies that the matrix $\text{plim}\,\frac{1}{T}\,(I \otimes \delta')Z'$ is upper triangular with a zero diagonal and thus singular. Thus proofs of the results in this section cannot rely on the invertibility of this matrix as did those of the previous section.

(B) $\operatorname*{plim}_{H_0} \tilde{\beta}_c = \beta_c \neq \beta$ and $\operatorname*{plim}_{H_0} \tilde{\beta}_u = \beta_u \neq \beta$ and $\beta_c = \beta_u$

(C) $\operatorname*{plim}_{H_0} \tilde{\theta}_u = 0$

What this theorem amounts to is that $\operatorname*{plim}_{H_0} \tilde{\alpha}_u = (I_M \otimes \delta)(\tilde{\theta}_u - \tilde{\beta}_u) = - \operatorname*{plim}_{H_0}(I_M \otimes \tilde{\delta}_c)\tilde{\beta}_c$. Moreover, since $\operatorname*{plim}_{H_0} \tilde{\beta}_c = \operatorname*{plim}_{H_0} \tilde{\beta}_u$ and $\operatorname*{plim}_{H_0} \tilde{\delta}_c = \delta$, the cross-equations restrictions implied by the RE-IAP hypothesis will be satisfied in the probability limit in both the constrained and unconstrained systems. However, $\tilde{\beta}_c$ and $\tilde{\beta}_u$ are inconsistent, and this fact which stems from the presence of v in the output equation causes the PLR statistic in (42) not to have an asymptotic χ^2 distribution. This result is presented in the following corollary:

Corollary 4.1: The PLR statistic converges in distribution to the inner product with itself of a normally distributed variable with a nonscalar covariance matrix. Thus, this test statistic does not have an asymptotic χ^2 distribution.

This corollary suggests that the test of the RE-IAP hypothesis can be salvaged in the case of Wold errors, by treating β as a nuisance parameter and correcting the standard test statistic for the presence of v and the inconsistency of $\tilde{\beta}_u$. We take this up in section 5.

We note, without formally stating the results, that in tests of the RE|IAP and IAP|RE hypotheses the least product of squares and two-stage estimators of all parameters except β are consistent. Again, the presence of v and inconsistency of the estimators of β cause PLR statistics not to be distributed as χ^2 asymptotically.

3.2. The "No-Lags" Case

Abel and Mishkin (1983) in their theoretical analysis and Gordon (1982) in his empirical study use "no-lags" version of the output equation, that is, they assume $\beta_i = \theta_i = 0$ for $i > 0$. Abel and Mishkin (1983, p. 3) claim that the asymptotic properties of the statistics used in the tests of the RE and IAP hypotheses are not affected by the misspecification of rational forecasts.[34] The previous subsection showed this to be false in the "lagged" case. The purpose of this section is to examine estimation and inference in the "no-lags" case.[35]

We begin by considering the joint IAP and RE (IAP-RE) hypothesis. Using (32a) and (28) - (29) the constrained system can be written as

$$Y_T = (X_T - Z_T\delta)\beta_0 - V_T\beta_0 + U_T \tag{44}$$

$$X_T = Z_T\delta + V_T + E_T \tag{29}$$

The unconstrained system ((32b)) comprises (29) and

$$Y_T = X_T\,\beta_0 + (Z_T\delta^*)(\theta_0 - \beta_0) + V_T^*(\theta_0 - \beta_0) + U_T \tag{45}$$

The crucial point here is that Z_T is orthogonal to $V_T + E_T$ by construction and to E_T by definition. Thus, Z_T is orthogonal to V_T. Moreover,

---

[34]In fairness to Abel and Mishkin we should point out that they do not study directly the properties of the relevant test statistics. Instead they focus on the equivalence of the tests of the RE and the IAP hypotheses with tests of Granger causality of z on y. In another version of the paper, Mishkin (1983, Ch. 3) seems to be aware of some of the problems to be analyzed in this section (see Remark on p. 50). However, he still concludes the chapter (p. 56) with the same claim as in his paper with Abel.

[35]This model is of interest on its own due to the "observational equivalence" problems discussed by McCallum (1979). Our analysis also has implications for the "equivalence" of neutrality and Granger causality tests discussed by Sargent (1973), McCallum (1979) and Nelson (1979). To save space this implication will not be pursued here.

intertemporal correlations among z, v and e do not play any role in the asymptotic behavior of the estimators. These facts imply the following theorem, irrespective of whether the errors are Wold or non-Wold:

Theorem 5: Suppose that assumptions (i) - (iv) hold (appropriately specialized to the model with "no lags"). Let H_0 denote the null RE-IAP hypothesis described in Theorem 1. Then, for both the case of Wold and non-Wold forecast errors,

(A) $\operatorname{plim}_{H_0} \tilde{\delta}_c = \delta$

(B) $\operatorname{plim}_{H_0} \tilde{\beta}_c = \operatorname{plim}_{H_0} \tilde{\beta}_u \neq \beta_0$

(C) $\operatorname{plim}_{H_0} \tilde{\theta}_u = 0$

where $\tilde{\beta}_c$, $\tilde{\beta}_u$, and $\tilde{\theta}_u$ are scalars.[36]

While the probability limits of the estimators are thus unaffected by the nature of the measured forecast errors, the distinction between Wold and non-Wold errors is important for the properties of the test statistic PLR. In the case of Wold errors Corollary 4.1 applies. However, the non-Wold class encompasses cases when v is nonstationary. Then the possibility of deriving a limiting distribution and its form, should one exist, depend on the specific type of nonstationarity of v. We note that in general this problem appears intractable.

Analogous results can be shown to hold here for the least product of squares estimators and the pseudo-likelihood ratio statistics involved in the

---

[36]Note that part A demonstrates that the assertion by Abel and Mishkin (1983, p. 11) that when rational expectations are mismeasured $\tilde{\delta}_c$ is inconsistent is not correct.

testing of the IAP|RE and RE|IAP hypotheses and the two-stage estimators involved in testing of the IAP|RE hypothesis.

4. Implications for the Validity of Inferences in Previous Empirical Studies

4.1. The "Lagged" Case

The most prominent of the studies employing the model with lags have been undertaken by Barro (1977, 1978), Barro and Rush (1980), Leiderman (1980) and Mishkin (1983). Our results in sections 3.1.3. and 3.1.4. imply that the behavior of the estimators of parameters in those models crucially depends on the properties of the forecast errors of the investigator's money growth equation. Rappoport (1984) examined the prediction performance of the money growth models of Barro and Rush (1980) and Mishkin (1983), among others. He found that the forecast errors generated by those models were autocorrelated and correlated with explanatory variables. This empirical evidence in conjunction with our results in section 3.1.3. implies that the statistical inferences that have been made in those studies are not valid. All of the critical parameter estimators are inconsistent. Consequently, referring the resulting pseudo-likelihood ratio statistics to a χ^2 distribution does not produce a test whose significance level is interpretable.

4.2. "No-Lags" Case

Gordon (1982) used a model with no lags in his examination of the IAP hypothesis.[37] He employed a two-stage procedure and used nominal demand growth, money growth and velocity as alternative nominal variables. Gordon

[37] In our discussion we mention only the issues related to our analysis in section 3.2. However, we note that in his estimation of the standard errors Gordon does not seem to take into account the two-stage nature of his procedure (Murphy and Topel (1983)).

rejected the IAP hypothesis. According to the results in section 3.2. the estimator of θ_0 in the Gordon study is consistent. Since Gordon used ordinary least squares in his estimation of the output equation, our results also imply that the validity of his tests crucially depends on whether his equations predicting nominal GNP (pp. 1099-1101) are the true equations in the sense that the mismeasurement error v_t equals zero for all t. This requires further study.[38]

5. Alternative Methods

In this section we will outline alternative procedures that attempt to alleviate the difficulties related to the mismeasurement of rational expectations. The order in which we present our approach reflects the order in which we feel it is natural to proceed in the investigation of the questions posed by the New Classical Macroeconomics.

5.1. An Alternative Hypothesis

The hallmark of the theory of output espoused by the New Classical Macroeconomics is that changes in the money stock, or some other nominal variable, have different impacts, depending on whether they are anticipated or not. The Lucas (1973) supply function represents an explicit version of this difference in which unanticipated money growth affects output ($\beta \neq 0$), and the IAP hypothesis holds ($\theta = 0$). Perhaps surprisingly, none of the studies in this literature tested the following null hypothesis:

$$H_0: \quad \theta = \beta \tag{46a}$$

against

$$H_A: \quad \theta \neq \beta \tag{46b}$$

[38] We note that Gordon (p. 1101) reports two "marginally significant" changes in the structure of his nominal demand equation. This would seem to indicate that $v \neq 0$.

We note that if $\theta = \beta$ the effects of mismeasurement disappear from the output equation (28). Under this condition it is thus possible to estimate β and θ consistently from a regression of y_t on distributed lags of $x_t - \hat{x}_t^{re}$ and $\hat{x}_t^{re}$. Tests of (46a) remain valid for any arbitrary choice of regressors in the money growth equation. In particular they can be run for the forecast functions proposed by Barro and Rush (1980), Mishkin (1983), and Sheffrin (1979), although none of these functions can be considered to be adequate representations of rational expectations (cf. Rappoport (1984)).[39]

In addition to its econometric features the null hypothesis $\theta = \beta$ has an important economic meaning. It states that irrespective of whether money growth is anticipated or unanticipated, its effect on the behavior of deviations of aggregate real output from the natural level is the same. Accordingly, we shall call (46a) the hypothesis of the "Irrelevance of the Anticipated Unanticipated Distinction" (IAUD).[40]

Since the tests of IAUD can be carried out even if expectations are misspecified we propose that IAUD be tested first. If IAUD is not rejected, then the next stage should involve the test of the hypothesis that "money matters" in the equation for output.[41] More formally, if $\theta = \beta$ we can write

$$Y_T = X\beta + U_T \tag{47}$$

and test

---

[39]The empirical performance of H_0 in (46a) will be examined in the sequel to this paper.

[40]It is apparent that the cross-equation restrictions imposed by rationality cannot be tested as long as IAUD holds. In fact, it is only when the power of tests of (46a) is considered (and thus $\theta \neq \beta$) that the method of expectations formation and the possibility of their being misspecified play a role.

[41]Alternatively, one could combine (46a) and (48a) into one joint hypothesis that $\theta = \beta$ and both are equal to zero.

$$H_0: \quad \beta = 0 \tag{48a}$$

against

$$H_A: \quad \beta \neq 0 \tag{48b}$$

It is worth noting that the test of (48a) is similar to the tests of the "St. Louis equation" (Anderson and Jordan (1968)). The crucial difference is that we suggest that this test be carried out after the test of IAUD ($\theta = \beta$). Otherwise, if in fact IAUD is rejected the equation for output given in (47) is misspecified since it does not allow for the distinction between unanticipated and anticipated components of money growth.

If the equality of θ and β is rejected we can proceed to test the IAP and RE hypothesis.

5.2. An Instrumental Variable Procedure for Testing IAP|RE[42]

Consider the following (basic) version of the output equation, when the RE hypothesis is maintained:

$$Y_T = (X - X^{re})\beta + X^{re}\theta + U_T \tag{49}$$

Since the RE hypothesis is maintained we can utilize (1) to transform (49) into

$$Y_T = X\theta + E(\theta - \beta) + U_T \tag{50}$$

Since e is an innovation in the money growth process, it is uncorrelated with the lagged values of x. Thus, we can proceed as follows:

1. Using the set of instruments $P = (X_{T-M},\ldots,X_{T-(2M-1)})$ estimate θ in (50) consistently.

2. Using the residuals from the previous step fit a moving average

---

[42]We are indected to Larry Christiano for suggesting this procedure.

representation of the error process $E(\theta - \beta) + U_T$ in (50).

3. Using the standard χ^2-statistic, test the IAP|RE hypothesis that $\theta = 0$.

An important property of this procedure is that it does not require the money growth process to be modelled explicitly. It does not, however, permit a test of the RE hypothesis,[43] for which we have to reintroduce into the output equation the distinction between anticipated and unanticipated money growth. Thus, we are again faced with the problems in testing the IAP and RE hypotheses discussed in the previous sections.

5.3. Tests of the IAP and RE Hypotheses Using the Least Product of Squares Estimators

Our results imply that the least product of square and two-stage estimators of all parameters, except β, are consistent if either a model excludes lags of anticipated and unanticipated components from the output equation or the equation for money growth generates forecast errors which are white noise and uncorrelated with the lagged values of explanatory variables. Postponing until later a discussion of the "no-lags" case, we first take up the question of obtaining a money growth equation with Wold errors.

It is apparent that the crucial step is to generate a forecasting equation for x with stationary forecast errors that are uncorrelated with the lagged values of explanatory variables. Subsequently, the filter implied by a Wold representation of the errors can be applied to the equation for x to obtain a money growth equation where forecast errors are white noise. Unfortunately, recent empirical evidence suggests that it may not be easy to obtain stationary errors. Friedman (1984) assembled evidence on the structural

[43]It is clearly important to be able to test the RE hypothesis given its widespread use in the literature examined in this paper and in other areas of macroeconomics.

instability of the money growth and other macroeconomic variables. Similarly, Blanchard and Watson (1984) suggested that those macroeconomic variables behave differently over different NBER reference cycles. Earlier summary statistics presented by Sims (1982) also point toward covariance nonstationarity of the money supply. In any case, a successful attempt at modelling structural changes in the money supply process (resulting in stationary errors) has to be made before the RE hypothesis can be tested.[44]

After obtaining a forecasting equation for money growth with Wold errors, we can proceed to estimation and inference related to the IAP and RE hypotheses. In order to deal with these issues we repeat (29) for convenience and rewrite it as follows:

$$Y_T = [X - Z(I_M \otimes \delta)\beta_c + E\beta - (V + E)(\beta - \beta_c) + U_T \tag{51}$$

$$X_T = Z_T\delta + V_T + E_T \tag{29}$$

where $\beta_c = \operatorname*{plim}_{H_0} \tilde{\beta}_c = \operatorname*{plim}_{H_0} \tilde{\beta}_u$ and for example we can take H_0 to be the IAP-RE hypothesis.

It would seem from (51) and (29) that in addition to the error term in (51) being correlated over time, the errors of the two equations are correlated with each other. However, expression (A39) in the Appendix implies that

$$\text{plim}\ \frac{1}{T}\ [E\beta - (V + E)(\beta - \beta_c) + U_T]'[V_T + E_T] = 0 \tag{52}$$

Thus, the errors in (51) and (52) are uncorrelated asymptotically.[45] Therefore,

---

[44] Epps and Singleton (1983) proposed a test of stationarity of time-series.

[45] The usual assumption made in the literature is that those errors are uncorrelated in finite sample, that is, the system (40) and (35) is recursive in finite sample. This is clearly not valid here.

if we treat β as a nuisance parameter and write the objective function in terms of $\beta_c = \underset{H_0}{\text{plim}}\ \tilde{\beta}_c = \underset{H_0}{\text{plim}}\ \tilde{\beta}_u$, the resulting system of equations is recursive asymptotically. This discussion suggests the following procedure:

1. Estimate the system (39) and (29) by the least product of squares method, to yield a consistent estimator of the error term $\eta_T = E\beta - (V + E)(\beta - \beta_c) + U_T$ in (51).

2. Since the money growth equation is supposed at this stage to contain Wold errors, $\eta_T = E\beta - (V + E)(\beta - \beta_c)$ is a covariance stationary process. Thus, fit a Wold representation to η.

3. Using the representation of η_T obtained in the previous step, estimate the parameters of (51) and (29) by minimizing

$$S^c_{GLS}(\beta_c,\delta) = S^c_X S^c_{Y,GLS}$$

where S^c_X is defined in (35) and

$$S^c_{Y,GLS} = \{Y_T - [X - Z(I_M \otimes \delta)]\beta_c\}'V_c^{-1}\{Y_T - [X - Z(I_M \otimes \delta)]\beta_c\}$$

where $V_c^{-1} = \text{Cov}(\eta)$.

4. Perform similar estimation for the unconstrained system and form the likelihood-ratio statistic[46]

$$LR = 2T \ln \left(\frac{\tilde{S}^c_{GLS}}{\tilde{S}^u_{GLS}} \right)$$

The final point concerns estimation and inference in the "no lags" case. Despite the fact that consistency of the estimators of parameters (including

[46] We recall from Corollary 4.1. that the pseudo-likelihood ratio statistic defined in (42) does not have an asymptotic χ^2 distribution. It is clear that LR defined above does have an asymptotic χ^2 distribution.

β_c) does not require Wold errors, stationarity of those errors is still necessary for correct inference. This is apparent from our discussion in section 3.2. and the second step of the procedure outlined above.

The main conclusion that emerges from this discussion is that obtaining a forecasting equation for money growth, or some other nominal variable, with stationary forecast errors seems necessary for correct inferences concerning the RE hypothesis.

6. The Implications of Mismeasurement of Rational Expectations for Tests of the Lucas Proposition

Lucas (1973) provided a formal model in which variations of real output around its natural rate are caused by individuals' inability to distinguish between relative price changes and general inflation. This model implies an output equation of the form:

$$y_t = \beta_0(x_t - x_t^{re}) + u_t \tag{53}$$

(where y_t is again the deviation of output from its natural rate).[47] The coefficient β_0 is linked to expectation formation by the following expression:

$$\beta_0 = \frac{\alpha\sigma_D^2}{\alpha\sigma_D^2 + \sigma_e^2} \tag{54}$$

where α is the slope of the individual producer's supply curve, σ_D^2 is the variance of relative demand shocks, and σ_e^2 is the variance of the rational expectations forecast error, as before. It is apparent that the slope of the "output-inflation tradeoff," β_0, varies negatively with the variance of the

[47] Note that the IAP hypothesis is treated as maintained in (53). This has been the standard assumption in studies of the Lucas proposition. The failure of the IAP hypothesis to hold would introduce serious problems in addition to those considered here.

unanticipated component of the nominal variable, σ_e^2. That is,

$$\frac{\partial \beta_0}{\partial \sigma_e^2} < 0 \tag{55}$$

If the investigator knew the true model for x given in (1) - (2) the relationship in (55) could be tested by comparing (β_0, σ_e^2) pairs across countries.[48]

When rational expectations are mismeasured, two problems stand in the way of this procedure. If the investigator does not know (2) and so attempts to measure x^{re} by regressing x and z, this yields a residual with variance $\sigma_v^2 + \sigma_e^2$ instead of σ_e^2. The estimator of β_0 is then construed by regressing y_t on the residual from the projection of x on z. The resulting estimator, $\hat{\beta}_0$, has the following property (see (A48)):

$$b_0 \equiv \text{plim } \hat{\beta}_0 = \frac{\sigma_e^2}{\sigma_v^2 + \sigma_e^2} \beta_0 \tag{56}$$

In order to test (55) the investigator has at his or her disposal only the estimates of b_0 and $\sigma_v^2 + \sigma_e^2$, since β_0 and σ_e^2 are unobservable. An examination of (56) demonstrates that $\frac{\partial b_0}{\partial(\sigma_v^2 + \sigma_e^2)}$ bears no necessary relationship to $\frac{\partial \beta_0}{\partial \sigma_e^2}$. For example, even if (55) is valid, b_0 could be positively related to $\sigma_v^2 + \sigma_e^2$, if countries with high values of $\sigma_v^2 + \sigma_e^2$ also have high values of σ_e^2. More damaging to this test of the Lucas proposition is the case where β_0 is positively related to σ_e^2 (and $\sigma_v^2 + \sigma_e^2$), and yet

[48] In this paper we omit consideration of the difficulties posed for estimation by the possibility that β varies over time, evidence for which has been presented by Froyen and Waud (1980). This problem is analyzed in detail by Frydman (1983b).

$\frac{b_0}{\beta_0}$ is sufficiently negatively related to $\sigma_u^2 + \sigma_e^2$, so that $\frac{\partial b_0}{\partial(\sigma_v^2 + \sigma_e^2)} < 0$ when $\frac{\partial \beta_0}{\partial \sigma_e^2} > 0$.

Evidently, which of these cases occurs depends on the way in which movements in $\sigma_v^2 + \sigma_e^2$ are shared among σ_v^2 and σ_e^2. Since σ_v^2 and σ_e^2 are unobservable any assumptions concerning their behavior across countries seem to be untestable. This implies that when rational expectations are mismeasured the results of the tests of the Lucas proposition based on the examination of pairs $(b_0, \sigma_v^2 + \sigma_e^2)$ across countries seem impossible to interpret.[49]

The difficulties of interpretation of the tests of the Lucas proposition are even more serious in the model in which the output equation contains lags of unanticipated money growth. The version of the output equation is thus

$$Y_T = (X - X^e)\beta + U_T \tag{57}$$

where all variables are defined as before. Assume that x^e is proxied by the least squares projection of x on z.

Before we proceed with a general analysis we consider a recent study of the Lucas proposition by Kormendi and Meguire (1984). They use a model of the form of (57). From the expression (A11) in the Appendix, it follows that

$$\text{plim } \hat{\beta} = \text{plim}[(V + E)'(V + E)]^{-1}(V + E)'E\beta. \tag{58}$$

Although Kormendi and Meguire recognize the problem of mismeasurementof of rational expectations, they study the effects of this mismeasurement under

[49] Attfield and Duck (1983, p. 448) are aware of the problem discussed here. They suggest two arbitrary assumptions on the behavior of σ_v^2 and σ_e^2 across countries which allow them to interpret the sign of $\frac{\partial b_0}{\partial(\sigma_v^2 + \sigma_e^2)}$ to be the same as the sign of $\frac{\partial \beta_0}{\partial \sigma_e^2}$.

the assumption that v is a white noise processes, uncorrelated at all leads and lags with e. x is also assumed to be white noise. In such a case (58) becomes

$$\text{plim } \hat{\beta}_i = \frac{\sigma_e^2}{\sigma_v^2 + \sigma_e^2} \beta_i \qquad i = 0,\ldots,M-1 \tag{59}$$

Thus the assumption that v is white noise and uncorrelated with lagged values of e reduces the problem to the one discussed above in the model with no lags of unanticipated money growth in the output equation (see (54)). However, as we pointed out in section 1.1., v is, in general, correlated with lagged values of e and these correlations cannot be observed. Thus, we have to examine the implications of (58) for tests of the Lucas proposition without the assumption that v is a white noise process orthogonal at all leads and lags to e.

Equation (58) demonstrates that each element of plim $\hat{\beta}$ is a linear combination of the components of β, β_i, $i = 0,\ldots,M-1$, with weights given by the intertemporal correlations among v + e and e. Our analysis in section 1.1. implies that v + e is in general correlated with the lagged values of e. This is true even if v + e is a matrix of white noise Wold error. Moreover, correlations between v + e and lagged values of e are unobservable, since they depend on the parameters of the true process generating x.[50] In general, therefore, nothing can be said about the Lucas proposition by examining the relationship between $\hat{\beta}$ and the estimate of $\sigma_v^2 + \sigma_e^2$.

---

[50]For an example of this fact in the case of Wold forecast errors see the proof of Theorem 3. For the case of non-Wold forecasts errors see Example 2 in section 1.1.

7. Summary of Results

This paper has focused on the econometric implications of the investigator's mismeasurement of rational expectations. We have considered two different assumptions concerning the structure of the measured forecast errors and two specifications of the standard model of aggregate real output and money growth employed in the New Classical Macroeconomics. In each case we have the implications of mismeasurement for inference concerning four widely tested hypotheses. It may, perhaps, be convenient to summarize our results in the following table:

<table>
<tr><th></th><th></th><th colspan="2">RE-IAP, RE|IAP, IAP|RE</th><th>Lucas Proposition</th></tr>
<tr><th></th><th></th><th>Estimators</th><th>Test Statistics</th><th>Relation between β and σ_e^2</th></tr>
<tr><td rowspan="2">No Lags Case</td><td>Non-Wold Errors</td><td rowspan="3">Consistent except for β</td><td>Not χ^2 asymptotically uncorrectable in general</td><td rowspan="2">Interpretable only under special assumptions concerning cross-country and intertemporal variations in σ_v^2 and σ_e^2</td></tr>
<tr><td>Wold Errors</td><td rowspan="2">Not χ^2 asymptotically correctable using procedure in section 5</td></tr>
<tr><td rowspan="2">Lagged Case</td><td>Wold Errors</td><td>Uninterpretable unless v is white noise; then reduces to No Lags case</td></tr>
<tr><td>Non-Wold Errors</td><td>Inconsistent except for unconstrained δ in IAP|RE</td><td>Asymptotic distribution not computable in general</td><td>Uninterpretable</td></tr>
</table>

APPENDIX

In order to simplify the computation of matrix partial derivatives of (35) it will be convenient to write equation (28) in the following equivalent ways:

$$\begin{aligned} Y_T &= X\beta + L\delta^* + V^*\gamma + U_T \\ &= X\beta + J^*\gamma + V^*\gamma + U_T \end{aligned} \tag{A1}$$

where we have used the following abbreviations:

$$\gamma = \theta - \beta; \quad J^* = J(\delta^*) = Z(I_M \otimes \delta^*); \quad L = Z\Phi(I_K \otimes \gamma)$$

Here Φ is a KM × KM matrix which permutes appropriately the columns of Z. Using this notation we can compute the following first derivatives of the objective function S defined in (35):

$$\frac{\partial S}{\partial \beta} = - S_X(Y_T - X\beta - J^*\gamma)'X \tag{A2}$$

$$\frac{\partial S}{\partial \gamma} = - S_X(Y_T - X\beta - J^*\gamma)'J^* \tag{A3}$$

$$\frac{\partial S}{\partial \delta^*} = - S_X(Y_T - X\beta - L\delta^*)'L \tag{A4}$$

$$\frac{\partial S}{\partial \delta} = - S_Y(X_T - Z_T\delta)'Z_T \tag{A5}$$

To simplify the proofs of our results and shorten our presentation we will first prove three lemmas.

The estimator $\tilde{\delta}_c$ obtained in the constrained estimation utilizes information from both output and money growth equations. However, the following lemma will allow us to consider the "implied" estimator of δ^* obtained from the output equation "alone":

Lemma 1-A: Let $\tilde{\delta}_Y^*$ and $\tilde{\delta}_X$ be the solutions to $\frac{\partial S}{\partial \delta^*} = 0$ and $\frac{\partial S}{\partial \delta} = 0$, where the other parameters are evaluated at values satisfying the first order conditions for a minimum of S. Then,

$$\text{plim } \tilde{\delta}_c = \delta \quad \text{iff} \quad \text{plim } \tilde{\delta}_Y^* = \delta$$

Proof: First note that under the constraint $\delta = \delta^*$ the first order conditions can be solved for the Lagrange multiplier and written as:

$$\frac{\partial S}{\partial \delta^*} + \frac{\partial S}{\partial \delta} = 0 \tag{A6}$$

Equation (A6) can be combined with (A4) - (A5) to yield:

$$S_X(Y_T - X\beta - L\delta^*)'L + S_Y(X_T - Z_T\delta)'Z_T = 0 \tag{A7}$$

Let $\tilde{\delta}_Y^*$ and $\tilde{\delta}_X$ be the solutions to $\frac{\partial S}{\partial \delta^*} = 0$ and $\frac{\partial S}{\partial \delta} = 0$, where the other parameters are evaluated at values satisfying $\frac{\partial S}{\partial \beta} = \frac{\partial S}{\partial \gamma} = 0$. Then the estimator of δ in the constrained system, $\tilde{\delta}_c$, can be obtained from (A4) - (A5) and (A7), and it is given by

$$\tilde{\delta}_c - \delta = [(\tilde{L}'\tilde{L})\tilde{S}_X + (Z_T'Z_T)\tilde{S}_Y]^{-1}[(\tilde{L}'\tilde{L})\tilde{S}_X(\tilde{\delta}_Y^* - \delta) + (Z_T'Z_T)\tilde{S}_Y(\tilde{\delta}_X - \delta)] \tag{A8}$$

Since plim $\tilde{\delta}_X = \delta$ by construction, it follows that plim $\tilde{\delta}_c = \delta$ iff plim $\tilde{\delta}_Y^* = \delta$.

Q.E.D.

The next lemma derives the necessary and sufficient condition for the consistency of $\tilde{\delta}_c$.

Lemma 2-A: In the constrained system (39) and (29)

$$\underset{H_0}{\text{plim}}\ \tilde{\delta}_c = \quad \text{iff} \quad \text{plim } \frac{1}{T} J'M_{X-J}E = 0$$

where $M_{X-J} = I - (X - J)[(X - J)'(X - J)]^{-1}(X - J)'$, and H_0 refers to the RE and IAP conditions.

Proof: From Lemma 1-A the consistency of $\tilde{\delta}_c$ can be examined by considering the behavior of $\tilde{\delta}^*_Y$. Setting $\frac{\partial S}{\partial \delta^*} = 0$ in (A4) yields:

$$\tilde{\delta}^*_Y = (\tilde{L}'\tilde{L})^{-1}\tilde{L}'(Y_T - X\tilde{\beta}_c) \tag{A9}$$

Using the expression for Y_T ((39)), under H_0, expression (29), and the identities $J\gamma = L\delta$ and $J\tilde{\gamma} = \tilde{L}\delta$ equation (A9) can be transformed to yield:

$$\tilde{\delta}^*_Y - \delta = (\tilde{L}'\tilde{L})^{-1}\tilde{L}'[-(X - J)\tilde{\beta}_c + E\beta + U_T] \tag{A10}$$

The estimator $\tilde{\beta}_c$ of β in the constrained system can be obtained from (A2) and (A3):

$$\tilde{\beta}_c = [(X - \tilde{J}^c)'(X - \tilde{J}^c)]^{-1}(X - \tilde{J}^c)'E\beta + o_p(1) \tag{A11}$$

where $\tilde{J}^c = Z(I_M \otimes \tilde{\delta}_c)$. Plugging (A11) into (A10) yields:

$$\tilde{\delta}^*_Y - \delta = (\tilde{L}'\tilde{L})^{-1}\tilde{L}'\{I - (X - J)[(X - \tilde{J}^c)'(X - \tilde{J}^c)]^{-1}(X - \tilde{J}^c)\}E\beta + o_p(1) \tag{A12}$$

Equation (A12) and assumption (iv) imply that the necessary and sufficient condition for the consistency of $\tilde{\delta}^*_Y$ (and thus $\tilde{\delta}_c$) is

$$\text{plim}\,\frac{1}{T}\,\tilde{L}'\{I - (X - (X - J)[(X - \tilde{J}^c)'(X - \tilde{J}^c)]^{-1}(X - \tilde{J}^c)\}E\beta = 0 \tag{A13}$$

Multiplying (A13) by δ' on the left hand side and using $\delta'\tilde{L}' = -\tilde{\beta}_c'J'$ we can write (A13) equivalently as

$$\text{plim}\,\frac{1}{T}\,\tilde{\beta}_c'J'\{I - (X - J)[(X - \tilde{J}^c)'(X - \tilde{J}^c)]^{-1}(X - \tilde{J}^c)\}E\beta = 0 \tag{A14}$$

Given assumption (iv) and noting that $\tilde{\beta}_c$ is a product of a nonsingular matrix and β and that (A14) should hold for arbitrary values of β gives the following condition equivalent to (A14):

$$\operatorname{plim} \frac{1}{T} J'\{I - (X - \tilde{J})[(X - \tilde{J}^c)'(X - \tilde{J}^c)]^{-1}(X - J^c)'\}E = 0 \tag{A15}$$

If δ_c is consistent, condition (A15) reduces to

$$\operatorname{plim} \frac{1}{T} J'\{I - (X - J)[(X - J)'(X - J)'\}E = 0$$

i.e.,

$$\operatorname{plim} \frac{1}{T} J'M_{X-J}E = 0 \tag{A16}$$

Since (A15) is a sufficient condition for the consistency of $\tilde{\delta}_c$, if (A15) holds $\tilde{\delta}_c$ is consistent. In turn, for a consistent $\tilde{\delta}_c$ condition (A15) takes the form of (A16). Q.E.D.

The final of our auxiliary lemmas describes the probability limit of the "estimator" $\tilde{\theta}_u$ defined by (41).

Lemma 3-A: Consider the unconstrained system (40) and (29) of tha IAP-RE hypothesis described in Theorem 1.

$$\operatorname*{plim}_{H_0} \tilde{\theta}_u = 0 \quad \text{if} \quad \operatorname{plim} \frac{1}{T} J'M_{X-J}E = 0$$

Proof: Utilizing the first order conditions $\frac{\partial S}{\partial \beta} = \frac{\partial S}{\partial \gamma} = 0$ and (41) we obtain

$$X'X\tilde{\beta}_u + X'J\tilde{\gamma}_u = X'E\beta + o_p(1) \tag{A17}$$

$$J'X\tilde{\beta}_u + J'J\tilde{\gamma}_u = J'E\beta + o_p(1) \tag{A18}$$

Subtracting (A18) from (A17) and using $\tilde{\gamma}_u = \tilde{\theta}_u - \tilde{\beta}_u$ yields:

$$\tilde{\beta}_u = [(X - J)'(X - J)]^{-1}[(X - J)'E\beta - (X - J)']\tilde{\theta}_u] + o_p(1) \tag{A19}$$

Substituting (A19) into (A17) and rearranging gives:

$$X'M_{X-J}J\tilde{\theta}_u = X'M_{X-J}E\beta + o_p(1) \tag{A20}$$

Since $(X - J)'M_{X-J} = 0$ equation (A20) implies

$$\text{plim}\ \frac{1}{T}\ J'M_{X-J}J\tilde{\theta}_u = \text{plim}\ \frac{1}{T}\ J'M_{X-J}E\beta \tag{A21}$$

Using assumption (iv) and noting that (A21) should hold for an arbitrary β we get

$$\text{plim}\ \tilde{\theta}_u = 0 \quad \text{if} \quad \text{plim}\ \frac{1}{T}\ J'M_{X-J}E = 0$$

Q.E.D.

We will now utilize Lemmas 2-A and 3-A in the proof of Theorem 1.

Proof of Theorem 1:

Parts A and C: From Lemmas 2-A and 3-A we have to examine the condition:

$$\text{plim}\ \frac{1}{T}\ J'M_{X-J}E = 0 \tag{A16}$$

Assumption (iv) permits $J'(X - J)$ to be inverted. The necessary condition for $\tilde{\delta}_c$ to be consistent and for $\underset{H_0}{\text{plim}}\ \tilde{\theta}_u = 0$ becomes

$$\text{plim}\ \frac{1}{T}\ [(X - J)'(X - J)][J'(X - J)]^{-1}J'E = \text{plim}\ \frac{1}{T}\ (X - J)'E \tag{A22}$$

Note that the (1,1)th element of the matrix on the right hand side of (A22), $\text{plim}\ \frac{1}{T}\ (X - J)'E$, is σ_ε^2 by assumption (ii). However, (36) implies that the first column of $\text{plim}\ \frac{1}{T}\ J'E$ is composed entirely of zeroes. Hence, the (1,1)th element of the matrix on the left hand side of (A22) is zero. Thus the necessary condition for $\tilde{\delta}_c$ to be consistent and for $\underset{H_0}{\text{plim}}\ \tilde{\theta}_u = 0$ is violated.

Q.E.D.

Part B: From Lemma 2-A and (A11) we have

$$\text{plim } \tilde{\beta}_c = \text{plim } [(X - J^c)'(X - J^c)]^{-1}(X - J^c)'E\beta \tag{A23}$$

where $J^c = Z(I_M \otimes \delta_c)$ and $\delta_c \neq \delta$. Assumptions (iv) and (v) imply that

$$\text{plim } \tilde{\beta}_c = \beta_c \neq \beta.$$

From (A21) and assumption (iv) we get

$$\text{plim } \tilde{\theta}_u = \text{plim } [J'M_{X-J}J]^{-1}J'M_{X-J}E\beta \tag{A24}$$

Substituting (A24) into (A19) and rearranging yields:

$$\text{plim } \tilde{\beta}_u = \text{plim } [(X-J)'(X-J)]^{-1}(X-J)'\{I - J[J'M_{X-J}J]^{-1}J'M_{X-J}\}E\beta \tag{A25}$$

Assumptions (iv) and (v) imply that

$$\text{plim } \tilde{\beta}_u = \beta_u \neq \beta \qquad \text{Q.E.D.}$$

Proof of Corollary 1.1: For PLR to converge asymptotically, in particular to a χ^2 distribution, it is necessary that

$$\underset{H_0}{\text{plim}} \frac{\tilde{S}^c_X \tilde{S}^c_Y}{\tilde{S}^u_X \tilde{S}^u_Y} = 1$$

From Theorem 1 plim $\tilde{\delta}_c \neq \delta$. However, plim $\tilde{\delta}_u = \delta$ by construction. Using the expression for S_X in (35) we immediately conclude that

$$\underset{H_0}{\text{plim}} \frac{\tilde{S}^c_X}{\tilde{S}^u_X} > 1$$

Since $\tilde{S}^c_Y$ is derived by imposing more constrains than $\tilde{S}^u_Y$, $\tilde{S}^c_Y \geq \tilde{S}^u_Y$. Hence,

$$\underset{H_0}{\text{plim}} \frac{\tilde{S}_X^c \tilde{S}_Y^c}{\tilde{S}_X^u \tilde{S}_Y^u} > 1.$$

Therefore, the pseudo-likelihood ratio statistic is unbounded asymptotically. In particular, PLR does not converge to a χ^2 distribution. Q.E.D.

Proof of Theorem 2: Part A is the same as part A of Theorem 1.

Part B: Analogously to (A9) and (A11) we have

$$\tilde{\delta}_u^* = (\tilde{L}'\tilde{L})^{-1}\tilde{L}'(Y_T - X\tilde{\beta}_u) \quad \text{(A26)}$$

and

$$\tilde{\beta}_u = [(X - \tilde{J}^u)'(X - \tilde{J}^u)]^{-1}(X - \tilde{J}^u)E\beta + o_p(1) \quad \text{(A27)}$$

where $\tilde{J}^u = Z(I_M \otimes \tilde{\delta}_u^*)$. Following the rest of the proof of Lemma 2-A with $\tilde{J}^u$ substituted for $\tilde{J}^c$ implies that the necessary and sufficient condition for the consistency of $\tilde{\delta}_u^*$ is plim $\frac{1}{T} J'M_{X-J}E = 0$. Then, from the proof of Part A of Theorem 1 this condition is violated and thus plim $\tilde{\delta}_u^* \neq \delta$. Q.E.D.

Part C: Since the constrained system is the same as for Theorem 1,

$$\underset{H_0}{\text{plim}}\ \tilde{\beta}_c = \beta_c \neq \beta.$$

Assumptions (iv) and (v) and expression (A27) imply that

$$\underset{H_0}{\text{plim}}\ \tilde{\beta}_u = \beta_u \neq \beta. \quad \text{Q.E.D.}$$

Proof of Theorem 3: Parts A, C and D are the same as in Theorem 1.

Part B: Using Lemma 1-A we consider the implied "estimator" $\tilde{\delta}_Y^*$. Analogously to (A26) we have

$$\tilde{\delta}_Y^* = (\tilde{L}'\tilde{L})^{-1}\tilde{L}'(Y_T - X\tilde{\beta}_u) \quad \text{(A28)}$$

Notice that in contrast to (A26) the unconstrained system contains the restriction $\delta = \delta^*$, but we do not impose the condition $\tilde{\beta} = -\tilde{\gamma}$ on the estimators. In this case (A28), (29) and (43) imply that

$$(\tilde{L}'\tilde{L})(\tilde{\delta}_Y^* - \delta) = \tilde{L}'[-X\tilde{\beta}_u - J\tilde{\gamma}_u + E\beta + U_T] \tag{A29}$$

or equivalently

$$\delta'(\tilde{L}'\tilde{L})(\tilde{\delta}_Y^* - \delta) = \tilde{\gamma}_u' J'[-X\tilde{\beta}_u - J\tilde{\gamma}_u + E\beta + U_T] \tag{A30}$$

The first order conditions $\frac{\partial S}{\partial \beta} = \frac{\partial S}{\partial \gamma} = 0$ yield the following equations:

$$\begin{bmatrix} \tilde{\beta}_u \\ \tilde{\gamma}_u \end{bmatrix} = \begin{bmatrix} X'X & X'J^u \\ \tilde{J}^{u\prime}X & \tilde{J}^{u\prime}\tilde{J}^u \end{bmatrix}^{-1} \begin{bmatrix} X'E \\ \tilde{J}^{u\prime}E \end{bmatrix} + o_p(1) \tag{A31}$$

where $\tilde{J}^u = Z(I_M \otimes \tilde{\delta}_u)$. We again note that $\tilde{\delta}_u$ is the estimator of δ in the unconstrained system (under $\theta \neq 0$ and $\delta = \delta^*$ in (43)), and is distinct from $\tilde{\delta}_Y^*$ (see (A8)). Plugging (A31) into (A30) results in

$$\delta'(\tilde{L}'\tilde{L})(\tilde{\delta}_Y^* - \delta) = \tilde{\gamma}_u'\{-[J'X, J'J] \begin{bmatrix} X'X & X'J^u \\ \tilde{J}^{u\prime}X & \tilde{J}^{u\prime}\tilde{J}^u \end{bmatrix}^{-1} \begin{bmatrix} X'E \\ \tilde{J}^{u\prime}E \end{bmatrix} \tag{A32}$$

We now show that $\underset{H_0}{\text{plim}}\, \tilde{\delta}_u = \delta$ is a solution to (A32). By Lemma 1-A the probability limit of the left hand side of (A32) is equal to zero and the term in curly brackets is $o_p(1)$, since $\tilde{J}^u = J + o_p(1)$. Thus, the consistent estimator $\tilde{\delta}_u$ solves the first order conditions. Q.E.D.

<u>Proof of Theorem 4</u>: When the investigator succeeds in finding a Wold representation for x, explicit expressions can be obtained for the elements of plim $\frac{1}{T} J'M_{X-J}E$. The first part of the proof derives these explicit expressions. Let equation (29) be of the form

$$x_t = a(L)x_t + b(L)w_t + v_t + e_t \tag{A33}$$

where w_t is some set of variables, a(L) and b(L) are respectively scalar and vector polynomials in the lag operator. We assume that $a_0 = 0$ and $1 - a(L)$ is invertible. This is one equation in a multivariate Wold representation of (x,w). The other equations may be written in moving average form as

$$w_t = c(L)\zeta_t + d(L)(v_t + e_t) \tag{A34}$$

where c(L) is a matrix conformable to w and ζ, ζ is a vector white noise error process uncorrelated at all leads and lags with v + e.

Substituting (A34) into (A33) and expressing x in moving average form results in

$$x_t = h(L)(v_t + e_t) + g(L)\zeta_t \tag{A35}$$

where $h(L) = (1 - a(L))^{-1}(b(L)d(L) + 1)$ and $g(L) = (1 - a(L))^{-1}b(L)c(L)$. Denoting by h_i the coefficient of h(L) corresponding to L^i we have

$$E[x_t(v_{t-i} + e_{t-i})] = h_i\sigma^2_{v+e} \quad \text{for all } i > 0 \tag{A36}$$

where $\sigma^2_{v+e} = Var(v_t + e_t)$. We also note that (for an example see (14) and (15)):

$$E[x_t e_{t-i}] = h_0 {}^{i}_{v+e,e} + h_1\gamma^{i-i}_{v+e,e} + \dots h_i\sigma^2_e$$

where $\gamma^i_{v+e,e} = E[(v_t + e_t)e_{t-i}]$.

<u>Parts A and C</u>: From Lemmas 2-A and 3-A we have to verify the condition $\text{plim}\,\frac{1}{T}\,J'M_{X-J}E = 0$, which it will be more convenient to express in the following equivalent way:

$$\operatorname{plim} \frac{1}{T} X'M_{X-J}E = 0 \tag{A38}$$

Expressions (A36) - (A37) and the assumption that (A33) is a Wold representation of x imply that

(a) $\operatorname{plim} \frac{1}{T} X'E$ is an upper triangular matrix with every diagonal element equal to $h_0\sigma_e^2$ and the typical (i,j)th element equal to $h_0\sigma_{v+e,e}^{(j-i)} + h_1\sigma_{v+e,e}^{(j-i-1)} + + \ldots + h_{(j-i)}\sigma_\varepsilon^2$, for $j > i$.

(b) $\operatorname{plim} \frac{1}{T} X'(V + E)$ is an upper triangular matrix with the typical (i,j)th element equal to $h_{(j-i)}\sigma_{v+e}^2$, for $j \geq i$.

(c) $\operatorname{plim} \frac{1}{T} [(V + E)'(V + E)]^{-1} = \frac{1}{\sigma_{v+e}^2} I$.

(d) $\operatorname{plim} \frac{1}{T} (V + E)'E$ is an upper triangular matrix with every diagonal element equal to σ_e^2 and the typical (i,j)th element equal to $\gamma_{v+e,e}^{(j-i)}$, for $j > i$.

Using (a) - (d) it can be tediously verified that

$$\operatorname{plim} \frac{1}{T} X'E - \operatorname{plim} \frac{1}{T} X'(V + E)[(V + E)'(V + E)]^{-1}(V + E)'E = 0$$

Q.E.D.

Part B: Using (A19), (A23) and the results $\underset{H_0}{\operatorname{plim}}\ \tilde{\delta}_c = \delta$ and $\underset{H_0}{\operatorname{plim}}\ \tilde{\theta}_u = 0$ proved above we immediately conclude that

$$\underset{H_0}{\operatorname{plim}}\ \tilde{\beta}_c = \frac{1}{\sigma_{v+e}^2} \operatorname{plim} \frac{1}{T} (V + E)'E = \underset{H_0}{\operatorname{plim}}\ \tilde{\beta}_u \neq \beta \tag{A39}$$

Q.E.D.

Proof of Corollary 4.1: We adapt the standard approach to the problem of convergence of likelihood ratio statistics to the problem of convergence of the pseudo-likelihood ratio statistic PLR defined in (42). PLR can be written as

$$PLR = T(\ln \tilde{S}_c - \ln \tilde{S}_u) \tag{A40}$$

It will be convenient here to express S_c and S_u in (35) as

$$S_c(\beta,\alpha,\delta,\gamma) \equiv S_c(\phi_c) = S_X S_Y + \lambda(\alpha - (I_M \otimes \delta)\beta)$$

$$S_u(\beta,\alpha,\delta) \equiv S_u(\phi_u) = S_X S_Y$$

where S_X is given in (35) and $S_Y = (Y_T - X\beta + Z\alpha)'(Y_T - X\beta + Z\alpha)$. Here, Z and α are understood to be redefined in such a way as to eliminate repetitions among the columns of Z. We also define $\bar{\phi}_u = \operatorname{plim}_{H_0} \tilde{\phi}_u$ and $\bar{\phi}_c = \operatorname{plim}_{H_0} \tilde{\phi}_c$.

An argument analogous to the one in Roy (1957, pp. 75-76) implies that the asymptotic distribution of PLR is the same as the asymptotic distribution of the following quantity:

$$PLR' = \sqrt{T}\,(\bar{\phi}_u - \tilde{\phi}_u)\left[\left.\frac{\partial^2 \ln S_u}{\partial^2 \phi_u}\right|_{\bar{\phi}_u}\right]\sqrt{T}\,(\bar{\phi}_u - \tilde{\phi}_u)' \tag{A41}$$

Note that PLR' is the same as PLR except that $\bar{\phi}_u$ is substituted for $\tilde{\phi}_c$, and thus only the unconstrained estimators have to be dealt with. Also, $\frac{\partial^2 \ln S_u}{\partial^2 \phi_u}$ is a matrix.

Using the Taylor expansion of $\left.\frac{\partial \ln S_u}{\partial \phi_u}\right|_{\tilde{\phi}_u} = 0$ around $\bar{\phi}_u$ and the fact $\bar{\phi}_u = \operatorname{plim} \tilde{\phi}_u$ yields

$$\sqrt{T}\ (\bar{\phi}_u - \tilde{\phi}_u) = \left[\frac{\partial^2 \ln S_u}{\partial^2 \phi_u} \Bigg|_{\bar{\phi}_u} \right] \sqrt{T}\ \frac{\partial \ln S_u}{\partial \phi_u} \Bigg|_{\bar{\phi}_u} + o_p(1) \tag{A42}$$

The Central Limit Theorem for multilinear processes proved by Parzen (1957) and Theorem 4 imply that $\sqrt{T}\ (\bar{\phi}_u - \tilde{\phi}_u)$ converges to a normal distribution with mean zero and covariance matrix

$$\textstyle\sum = \underset{H_0}{\text{plim}} \left[\frac{\partial^2 \ln S_u}{\partial^2 \phi_u} \Bigg|_{\bar{\phi}_u} \right]^{-1} \left[\sqrt{T}\ \frac{\partial \ln S_u}{\partial \phi_u} \Bigg|_{\bar{\phi}_u} \right] \left[\sqrt{T}\ \frac{\partial \ln S_u}{\partial \phi_u} \right]' \left[\frac{\partial^2 \ln S_u}{\partial^2 \phi_u} \Bigg|_{\bar{\phi}_u} \right]^{-1} \tag{A43}$$

We write $\sum$ as $A^{-1}BA^{-1}$, and note that for PLR' (and thus PLR) to have a χ^2 distribution asymptotically it is necessary that $A^{-1}B$ be idempotent. It can be shown that B is nonsingular and thus the necessary condition becomes

$$A^{-1}B = I \quad \text{or} \quad A = B \tag{A44}$$

We will now show that if $V \neq 0$ condition (A44) is violated in the model (28) - (29), under H_0 of Theorem 4. For convenience we rewrite S_u as

$$S_u = S_X(Y_T - Wb)'(Y_T - Wb) \tag{A45}$$

where $Wb = (X,Z)(\beta', -\alpha')'$. It is sufficient to prove that (A44) is violated for one of the components of ϕ_u. It can be shown that the submatrix of B,

$$\underset{H_0}{\text{plim}} \left[\sqrt{T}\ \frac{\partial \ln S_u}{\partial b} \Bigg|_{\bar{\phi}_u} \right] \left[\sqrt{T}\ \frac{\partial \ln S_u}{\partial b} \Bigg|_{\bar{\phi}_u} \right]'$$

is equal to

$$\underset{H_0}{\text{plim}} \frac{1}{(\frac{S_Y}{T})^2} \left[\frac{W'W}{T} \sigma_u^2 + \frac{W'\Xi\Xi'W}{T} \right] \tag{A46}$$

where $\Xi = [V - W \text{ plim} (\frac{W'W}{T})^{-1} \frac{W'V}{T}]\beta$. Note that Ξ is a component of the error in the output equation evaluated at $\bar{\Phi}_u$ and is a consequence of the inconsistency of $\tilde{\beta}_u$. In contrast, the corresponding submatrix of A,

$$\left. \underset{H_0}{\text{plim}} \frac{\partial^2 \ln S_u}{\partial b^2} \right|_{\bar{\Phi}_u}$$

is equal to

$$\underset{H_0}{\text{plim}} \frac{1}{(\frac{S_Y}{T})^2} [\frac{W'W}{T} (\sigma_u^2 + \text{plim} \frac{\Xi'\Xi}{T})]. \tag{A47}$$

Since Ξ is not i.i.d., (A46) is equal to (A47) iff $V = 0$. Thus, if $V \neq 0$ condition (A44) is violated. Q.E.D.

Proof of Theorem 5:

Parts A and C: The condition $\text{plim} \frac{1}{T} J'M_{X-J}E = 0$ holds here due to the orthogonality of Z_T, E_T and $V_T + E_T$. Q.E.D.

Part B: Using (A39) and the orthogonality of V_T and E_T, we have

$$\underset{H_0}{\text{plim}} \tilde{\beta}_c = \underset{H_0}{\text{plim}} \tilde{\beta}_u = \frac{\sigma_e^2}{\sigma_v^2 + \sigma_e^2} \beta \neq \beta \ . \tag{A48}$$

REFERENCES

Abel, Andrew, and Frederic S. Mishkin. "An Integrated View of Tests of Rationality, Market Efficiency and the Short-Run Neutrality of Monetary Policy." Journal of Monetary Economics, 11 (January 1983), 3-24.

Anderson, Leonall C., and Jerry L. Jordan. "A Monetarist Model for Economic Stabilization." Federal Reserve Bank of St. Louis, Review, 50, 1-24.

Attfield, Clifford L. F. "An Analysis of the Implications of Omitting Variables from the Monetary Growth Equation in a Model of Real Output and Unanticipated Money Growth." European Economic Review, 23 (December 1983), 281-290.

————, and Nigel W. Duck. "The Influence of Unanticipated Money Growth on Real Output." Journal of Money, Credit and Banking, 15 (November 1983), 442-454.

Barro, Robert J. "Unanticipated Money Growth and Unemployment in the United States." American Economic Review, 67 (March 1977), 101-115.

————. "Unanticipated Money, Output, and the Price Level in the United States." Journal of Political Economy, 86 (August 1978), 549-580.

————, and Mark Rush. "Unanticipated Money and Economic Activity." In Rational Expectations and Economic Policy, edited by Stanley Fischer. Chicago: University of Chicago Press, 1980.

Blanchard, Olivier, and Mark W. Watson. "Are Business Cycles All Alike?" NBER Working Paper No. 1392, June 1984.

Cukierman, Alex, and Paul Wachtel. "Differential Inflationary Expectations and the Variability of the Rate of Inflation: Theory and Evidence." American Economic Review, 69 (September 1979), 595-609.

Epps, T. W., and Kenneth J. Singleton. "A Test for Changes in the Distribution of a Time Series." Mimeo, April 1983.

Friedman, Benjamin M. "Money, Credit and Interest Rates in the Business Cycle." NBER Working Paper No. 1482, October 1984.

Froyen, Richard T., and Roger N. Waud. "Further International Evidence on Output-Inflation Tradeoffs." American Economic Review, 70 (June 1980), 409-421.

Frydman, Roman. "Individual Rationality, Decentralization, and the Rational Expectations Hypothesis." In Individual Forecasting and Aggregate Outcomes: "Rational Expectations" Examined, edited by Roman Frydman and Edmund S. Phelps. New York: Cambridge University Press, 1983a.

———. "Subjective Forecasts and the Output-Inflation Relationship." Mimeo, 1983b.

———, and Edmund S. Phelps. "Introduction." In Frydman and Phelps, eds.

———, and Mark Schankerman. "An Alternative Perspective on Econometric Tests of the Rational Expectations Hypothesis." C. V. Starr Research Report No. 81-14, April 1981.

Gordon, Robert J. "Price Inertia and Policy Ineffectiveness in the United States, 1890-1980." Journal of Political Economy, 90 (December 1982), 1087-1117.

Kaufman, Roger. Review of Frederic Mishkin's A Rational Expectations Approach to Macroeconometrics. Journal of Money, Credit, and Banking, 1984.

Kormendi, Roger C., and Philip G. Meguire. "Cross-Regime Evidence of Macroeconomic Rationality." Journal of Political Economy, 92 (October 1984), 875-908.

Leiderman, Leonardo. "Macroeconometric Testing of the Rational Expectations and Structural Neutrality Hypothesis for the United States." Journal of Monetary Economics, 6 (January 1980), 69-82.

Lucas, Robert E., Jr. "Some International Evidence on Output-Inflation Tradeoffs." American Economic Review, 63 (June 1973), 326-334.

McCallum, Bennett T. "On the Observational Inequivalence of Classical and Keynesian Models." Journal of Political Economy, 87 (April 1979), 394-402.

Mishkin, Frederic S. A Rational Expectations Approach to Macroeconometrics. Chicago: The University of Chicago Press, 1983.

Murphy, Kevin M., and Robert H. Topel. "Estimation and Inference in Two-step Econometric Models." Mimeo, October 1983.

Muth, John F. "Rational Expectations and the Theory of Price Movements." Econometrica, 29 (July 1961), 315-335.

Neftci, Salih, and Thomas J. Sargent. "A Little Bit of Evidence on the Natural Rate Hypothesis from the U.S." Journal of Monetary Economics, 4 (April 1978), 315-319.

Nelson, Charles R. "Granger Causality and the Natural Role Hypothesis." Journal of Political Economy, 87 (April 1979), 390-394.

Parzan, Emanuel. "A Central Limit Theorem for Multilinear Stochastic Processes." The Annals of Mathematical Statistics, 28 (1957), 252-256.

Rappoport, Peter. "Unfalsified Expectations: A New Approach to Modelling Expectations in Empirical Macroeconomics." Mimeo, December 1984.

Roy, K. P. "A Note on the Asymptotic Distribution of Likelihood Ratio." Calcutta Statistical Association Bulletin, 7 (March 1957), 73-77.

Sargent, Thomas J. "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment." Brookings Papers on Economic Activity, 2 (1973), 429-472.

———. Macroeconomic Theory. New York: Academic Press, 1979.

Sims, Christopher A. "Is There a Monetary Business Cycle?" American Economic Review, Papers and Proceedings, 73 (May 1973), 228-233.

Startz, Richard. "Testing Rational Expectations by the Use of Overidentifying Restrictions." Journal of Econometrics, 23 (December 1983), 343-352.