SHOULD FIXED COEFFICIENTS BE REESTIMATED EVERY PERIOD FOR EXTRAPOLATION?

by

P.A.V.B. Swamy and Garry J. Schinasi

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.
Abstract

This paper demonstrates that forecast accuracy is not necessarily improved when fixed coefficient models are sequentially reestimated, and used for prediction, after updating the database with the latest observation(s). This is at variance with the now popular method (see Meese and Rogoff (1983, 1985)) of sequentially reestimating fixed coefficient models for prediction as new data "rolls" in. It is argued that although "rolling" may minimize the variance of predictions for some classes of estimators, "rolling" does not necessarily yield accurate predictions (i.e., predictions that are close to actual data). Minimizing the mean squared prediction errors is a necessary condition for maximizing the probability that a given predictor is more accurate than other predictors. This minimization need not require, and may even exclude, the most recent data. A by-product of the demonstration is that for predictors based on the same sample size, a predictor with smaller variance need not be more accurate than another predictor with a larger variance.
Should Fixed Coefficients be Reestimated Every Period for Extrapolation?

P.A.V.B. Swamy and Garry J. Schinasi*

1. Introduction

In making forecasts of future variables, some econometricians reestimate their models using all past data prior to each forecast period. The process involves fixing the starting date and the initial size of the sample and enlarging the sample by adding successive observations for reestimation and prediction as new data become available (see, e.g., Fromm and Klein (1976, p. 9) and Meese and Rogoff (1983, 1985)). It has been suggested that such a procedure improves forecast accuracy for two reasons: first, because a larger sample reduces the variances of fixed coefficient estimators; and second, because sequential estimation or "rolling" captures any variation in coefficients.

The theory underlying the common impression that one should always use all the available observations in estimation and prediction is more ambiguous than may be generally realized. Without proper theoretical justification, like an explicit risk minimizing motivation (favoring good predictions), a procedure that sequentially updates estimates of coefficients and predictions in a model assumed to have constant coefficients is meaningless.

The primary purpose of the paper is to demonstrate that any sequential method

---

* Views expressed in this paper are those of the authors and do not reflect the views of the Board of Governors or the staff of the Federal Reserve System. The comments by N. Ericsson, J. Marquez, P. von zur Muehlen and P. Tinsley are greatly appreciated.

1 This impression is also not fully supported by the asymptotic theory, some simple normal cases apart. For example, Lehmann (1983, pp. 352-388) analyzes various nonnormal situations showing that a parameter is more efficiently estimated even in large samples by discarding some sample observations than by using all the observations.
of estimating constant coefficients does not necessarily yield accurate forecasts. A by-product of the demonstration is that a predictor with a smaller variance need not be better than another predictor with a larger variance, even though both predictors are based on the same sample.²

The logic of the basic argument of this paper is as follows. If the objective of estimation is forecast accuracy, then one should prefer predictions that are close to actual realizations to predictions that are close to some other quantities (such as the means of predictors). Consequently, one should select the predictor that has the highest probability of being close to actual realizations. Although it is difficult to derive estimators based on this general criterion, a necessary condition for this probability to be a maximum is for the mean square error of a predictor (i.e., the expected squared deviation of a predictor from realization) to be a minimum. "Rolling" (i.e., sequential estimation) may minimize the forecast variance for broad classes of estimators, but minimizing variance does not necessarily minimize mean square error.³ Hence, for any given estimator, "rolling" may reduce the variance of a predictor, but it does not necessarily improve, and may

² The constant coefficient approach is inherently incorrect if some or all of the regression slopes change over time, sample enlargement notwithstanding.

³ In practice, many of these optimal predictors involve unknown parameters and their operational versions based on some sample estimators of the parameters may not satisfy all the conditions for achieving minimum variance. This makes it difficult to recognize an operational predictor with minimum variance in small samples. Because of this difficulty, attention has been shifted to the asymptotically optimal predictors which have the smallest asymptotic variance within the class of asymptotically normal predictors. This does not eliminate the difficulty because some asymptotically optimal predictors do not possess finite variances in small samples and hence violate a necessary condition for maximizing the probability of obtaining forecasts within an interval around actual realizations.
even reduce, forecast accuracy. Note that this is a theoretical result about
the forecasting properties of "rolling" regressions and it simply states that
there is no reason to believe that "rolling" improves forecast accuracy. The
logic of this argument extends to more general cases, where two or more
predictors are compared for forecast accuracy using the same sample.

Section 2 introduces and briefly discusses two forecast criteria
that compare forecasts with realizations rather than the means of predictive
distributions. We then discuss conventional estimators satisfying certain
sampling properties in Section 3. Operationally, some versions of these
estimators may not have finite variances. Using the realization-based crite-
ron, Section 4 shows that the best predictor must have minimum mean square
error, and that there is no reason why the method of sequential estimation
achieves this. Conclusions are presented in Section 5.

2. **Forecasting Criteria Based on Realizations**

Suppose that we are interested in predicting the value that would
be taken by the random variable \( y_t^* \) in a future period \( T+s \), where \( T \) is the
terminal period of the currently available sample observations on \( y_t^* \).\(^4\)
We now define two criteria for comparing predictors.

**Criterion of Highest Concentration**

The criterion of highest concentration compares predictions with actual
realizations and is defined as follows: an operational predictor \( \hat{y}_{T+S} \) of the

---

\(^4\) We distinguish a random variable from its value by an asterisk. For example, \( y_t \) is the value taken by the random variable \( y_t^* \) in period \( t \).
actual value $y_{T+s}$ is better than any other operational predictor $\bar{y}_{T+s}$ if the probabilities satisfy the condition

$$\Pr(y_{T+s} - \lambda_1 < \hat{y}_{T+s} < y_{T+s} + \lambda_2) \geq \Pr(y_{T+s} - \lambda_1 < \bar{y}_{T+s} < y_{T+s} + \lambda_2)$$  \hspace{1cm} (1)$$

for all possible values of $\lambda_1$ and $\lambda_2$ in a chosen interval $(0, \lambda)$ and for all possible realizations $y_{T+s}$.

A necessary condition that (1) be satisfied for all $\lambda$ is

$$E(\hat{y}_{T+s} - y_{T+s})^2 \leq E(\bar{y}_{T+s} - y_{T+s})^2,$$$}$ (2)

that is, the mean square error of $\hat{y}_{T+s}$ about the actual realization $y_{T+s}$ is a minimum (see Rao (1973, p. 315)). This condition indicates, for example, when a predictor does not satisfy the so-called criterion of highest concentration in (1).

Now sufficient conditions will be given for a criterion weaker than (1) to be satisfied. This weaker criterion is based on the concept of Pitman's nearness (PN) which is defined as follows:

**Pitman's Nearness (PN)**

A predictor $\hat{y}_{T+s}$ is nearer to the value $y_{T+s}$ than another predictor $\bar{y}_{T+s}$ if $\Pr[L(\hat{y}_{T+s}, y_{T+s}) < L(\bar{y}_{T+s}, y_{T+s})] > 1/2$, where $L(\hat{y}_{T+s}, y_{T+s})$ represents a loss in predicting $y_{T+s}$ by $\hat{y}_{T+s}$. We may consider two standard loss functions, namely,

$$L_1(\hat{y}_{T+s}, y_{T+s}) = |\hat{y}_{T+s} - y_{T+s}| \text{ and } L_2(\hat{y}_{T+s}, y_{T+s}) = (\hat{y}_{T+s} - y_{T+s})^2.$$
With these assumptions, Peddada (1985) has proved that if

\[(1)\] \(EL_i(\hat{y}_{T+s}, y_{T+s}) < EL_i(\hat{y}_{T+s}, y_{T+s})\) for \(i = 1, 2;\) \(\text{ (3a)}\)

\[(2)\] \(e_i = Ee_i = EL_i(\hat{y}_{T+s}, y_{T+s}) - EL_i(\hat{y}_{T+s}, y_{T+s})\)

\(< -2.67\) for \(i = 1, 2;\) and

\[(3)\] \(E(e_i - e_i)^j < j!\) \(\text{ for } i = 1, 2; j = 1, 2, \ldots,\) \(\text{ (3c)}\)

where

\[e_i = L_i(\hat{y}_{T+s}, y_{T+s}) - L_i(\hat{y}_{T+s}, y_{T+s}),\]

then \(\hat{y}_{T+s}\) is closer to \(y_{T+s}\) than \(\hat{y}_{T+s}\) in the PN sense.

3. **Fixed Coefficients Approaches**

With these desirable properties of predictors in hand, we next discuss conventional predictors and their statistical properties. The first step in any statistical method of generating predictions is to formulate a statistical model about the possible data generating process. It is usually postulated that the observations on \(y^*_t\) are generated by the following reduced form equation with fixed coefficients:

\[y^* = X\pi + \varepsilon^*,\] \(\text{(4)}\)

where

\[y^* = \text{the (Txl) vector of variables } y^*_t, t = 1, 2, \ldots T;\]

\[X = \text{the (TxK) matrix of sample observations on K "fixed" exogenous variables;}\]

\[\pi = \text{the (Kxl) vector of reduced-form coefficients; and}\]

\[\varepsilon^* = \text{the (Txl) vector of reduced-form disturbances.}\]
It is usually assumed that

\[ E(\epsilon^* | X) = E(\epsilon^*) = 0 \text{ and } E(\epsilon^* \epsilon^{*\prime} | X) = \sigma^2 \nu. \] (5)

If model (4) holds for the post-sample periods \( t = T + s, s = 1,2, \ldots \), also, then for period \( T + s \) the variable \( y_{T+s}^* \) given the vector of regressors will be

\[ y_{T+s}^* = x_{T+s}^\prime \pi + \epsilon_{T+s}^*, \] (6)

where

\( y_{T+s}^* \) = the scalar regressand;
\( x_{T+s}^* \) = the (l x K) vector of prediction regressors; and
\( \epsilon_{T+s}^* \) = the scalar prediction disturbance.

Goldberger (1962) points out that for model (4), an appealing set of assumptions is one which allows the prediction disturbance to be correlated with the sample disturbances. Therefore, we shall assume

\[
E(\epsilon_{T+s}^* | x_{T+s}^*) = E(\epsilon_{T+s}^*) = 0, \quad (7a)
\]

\[
E(\epsilon_{T+s}^* \epsilon_{T+s}^{*\prime} | x_{T+s}^*) = \sigma^2, \text{ and} \quad (7b)
\]

\[
E(\epsilon_{T+s}^* \epsilon^* | X, x_{T+s}^*) = w, \quad (7c)
\]

where \( w \) is the Tx1 vector of covariances of the prediction disturbance with the vector of sample disturbances.\(^5\) Thus, implicit in the use of model (4)

---

\(^5\) If equation (1) represents an autoregressive model, then \( X \) consists of the lagged values of \( y \) and \( w \) can be equal to 0. Alternatively, the vector \( w \) can be zero if equation (1) represents a regression model with serially uncorrelated error term.
for prediction is the assumption that the actual value of the regressand for period T+s will be a drawing from a distribution with mean \( x_{T+s}^* \) and variance \( \sigma^2 \).

The second step in any prediction method is to estimate \( \pi \). The vector \( \pi \) can be estimated by one or more of the following procedures:

**Estimation Procedures**

1. The least squares procedure;
2. The generalized least squares procedure based on an estimated error covariance matrix;
3. A fully or partially restricted reduced form procedure that fully or partially accounts for the connection between \( \pi \) and the coefficients of a structural model;
4. A Bayes procedure;
5. Bayes-like procedures; and
6. Robust procedures.

Let \( \hat{\pi} \) denote any one of these estimators.

The final step is to use an estimated \( \hat{\pi} \) in predicting the value \( y_{T+s} \).

Even though any or all of the estimation procedures (1)-(6) can be considered for sequential estimation, to save some computations we may, if possible, want to first find the most efficient of all the above six estimators and then use it in a sequential estimation. Furthermore, all the criteria of estimation used by the above estimation procedures may not be compatible with the criterion (1) or Pitman's nearness, as we show in the following section.
A Minimum Variance Predictor

It has been shown by Goldberger (1962) that the minimum variance linear "unbiased" predictor of \( y_{T+S} \) is

\[
\hat{y}_{T+S} = x_{T+S}^\prime \hat{\pi} + w^{-1}y^{-1}(y^* - X\hat{\pi})
\]

(8)

if \( \hat{\pi} = (X'V^{-1}X)^{-1}X'V^{-1}y^* \) is the minimum variance linear unbiased estimator of \( \pi \).

It is also shown by Goldberger (1962) that

\[
E(\hat{y}_{T+S} - y^*)^2 \leq E(x_{T+S}^\prime \hat{\pi} - y^*)^2
\]

(9)

with the equality holding when \( w = 0 \).

A difficulty with the minimum variance predictor (8) is that \( w \) and \( \sigma^2V \) are usually unknown. The inequality (9) may be reversed if \( \hat{y}_{T+S} \) represents an operational predictor of the form

\[
x_{T+S}^\prime \hat{\pi} + w^{-1}V^{-1}(y^* - X\hat{\pi})
\]

(10)

based on some sample estimates of \( w \), \( \sigma^2 \), and \( V \), and \( \hat{\pi} \) represents an operational estimator

\[
(X'V^{-1}X)^{-1}X'V^{-1}y^*
\]

(11)

---

6 A predictor \( \hat{y}_{T+S} \) is said to be "unbiased" if \( E \hat{y}_{T+S} = E y^* \).

7 The minimum variance linear unbiased predictor of an element of a vector variable following a vector autoregressive (VAR) model also has the same form. For example, if equation (4) is an autoregressive model imbedded in a VAR model, then \( x_{T+S}^\prime \) includes the lagged values of \( y_{T+S} \), \( w \) represents the contemporaneous covariances between the dependent variable of equation (4) and the other dependent variables included in the VAR model, \( \sigma^2V \) represents the contemporaneous covariance matrix of these other variables, and \( y^* \) and \( X \) consist of the current and lagged values of the other variables respectively.
or a restricted estimator based on \( \hat{\nu} \), provided the second-order moments of (10) and \( \hat{\nu} \) are finite (see Rao (1967, 1975)). In other words, an operational version of (8) may not have the minimum variance (about \( y^* \)) and may even have infinite variance. The estimator (11) may also have infinite variances. The predictors with infinite variances violate condition (2). They may not even be good in the PN sense.

The same difficulty may arise if \( \hat{\nu} \) represents a partially or fully restricted reduced-form estimator. The conditions for the finiteness of the moments of partially restricted reduced-form estimators in the normal case are given in Swamy and Mehta (1981, 1982), and Swamy, Mehta and Iyengar (1983). If the predictors of \( y_{T+S}^* \) based on the estimation procedure (3) need satisfy conditions (3a)-(3c), then Swamy and Mehta's modifications, that guarantee the existence of moments of all orders in small samples, are necessary. Using these modifications, whether the predictor (10) has a smaller variance about \( y^*_{T+S} \) than the predictor \( x_{T+S}^* \hat{\nu} \) depends on how far \( E\hat{\nu} \) is from \( \nu \) and on the precision of the estimators \( \hat{\nu} \) and \( \hat{\nu}^{-1}(1/\delta^2) \).

A result due to Rao (1967) shows that for a matrix \( Z \) of maximum rank such that \( X'Z = 0 \), the estimator (11) will have bigger variances than the least squares estimator \( (X'X)^{-1}X'y^* \) if and only if \( X'VZ \) is sufficiently close to a null matrix. In this case the predictor \( x_{T+S}^*(X'X)^{-1}X'y^* \) may have a smaller variance than the predictor (10) or the predictor \( x_{T+S}^* \hat{\nu} \). Thus, it is not possible in practice to recognize an operational predictor with minimum variance in small samples.

---

8 This comment does not apply if we assume that \( \nu=0 \) because in this case the second term in (10) will be zero.
Asymptotically Efficient Predictors

If the estimation period is sufficiently long, asymptotic theory may apply and predictors (10) and (8) may have the same limiting distributions. This can occur even when \( \hat{\pi} \) represents a partially or fully restricted reduced-form estimator. It is possible that the asymptotic variance of the predictor (10) is smaller if \( \hat{\pi} \) represents a robust estimator than if \( \hat{\pi} \) represents the estimator (11) or a restricted reduced-form estimator. The comparisons of asymptotic variances are relevant for moderate sample sizes if the estimators of \( \pi \) are convergent in the \( r \)th mean. Such estimators are developed in Swamy, Mehta and Iyengar (1983).

On the basis of either the exact finite sample distribution theory or the asymptotic theory, a universally preferred choice among the estimation procedures (1)-(6) is not possible.

4. Does Sequential Estimation Necessarily Improve Forecasting Accuracy?

Suppose that one of the estimation procedures (1)-(6) is used in sequential estimation and the corresponding predictors of \( y_{T+S} \) are computed, and further assume that the corresponding predictors possess finite variances. Suppose also that these variances decrease as the estimation period increases. What is the correct interpretation of the predictors of \( y_{T+S} \) obtained in this sequential estimation procedure?

We can answer this question by using the following standard result in probability theory. Let \( \tilde{y}_{T+S} \) and \( \bar{y}_{T+S} \) be two operational predictors of \( y_{T+S} \) with finite means \( \mu_1 \) and \( \mu_2 \), finite variances \( \sigma_1^2 \) and \( \sigma_2^2 \), and distribution functions \( F_1 \) and \( F_2 \) respectively. Suppose that both \( \tilde{y}_{T+S} \) and \( \bar{y}_{T+S} \) are based on the same formula but \( \tilde{y}_{T+S} \) uses a bigger sample than
so that $\sigma_1^2 \leq \sigma_2^2$. Then according to a standard result in probability theory,

$$F_1(y_{T+s} + \mu_1) - F_1(-y_{T+s} + \mu_1) \geq F_2(y_{T+s} + \mu_2) - F_2(-y_{T+s} + \mu_2)$$

(12)

for each possible value $y_{T+s}$ implies that $\sigma_1^2 \leq \sigma_2^2$. However, the converse proposition is not true. The inequality $\sigma_1^2 \leq \sigma_2^2$ implies that inequality (12) is true for at least one value $y_{T+s}$ but not necessarily for all possible values of $y_{T+s}$ (see Rao (1973, p. 96)). Thus, if we have two operational predictors with finite second-order moments, then it is not necessarily true that the predictor with a smaller variance will take values around the actual value $y_{T+s}$ with a higher probability than the predictor with a larger variance.

These general arguments can now be applied to the specific choice of two sets of predictions, based on the same predictor using different samples. We can therefore conclude the following about "rolling:" if we reestimate a model using all past data prior to each forecast period to obtain one-step ahead forecasts to reduce the variance, then the one-step ahead forecasts will not necessarily be closer to the realized values of the forecasted variable than the multi-step ahead forecasts, even though the former are based on more observations and hence may have smaller variances than the latter.

More formally, comparing the inequality (12) with (1) shows that the criterion of minimum variance ("unbiasedness") prediction only satisfies a necessary condition for maximizing the probabilities of intervals around the mean values of predictors (see (12)) while the criterion of minimum mean square error about actual realizations satisfies a necessary condition.
for maximizing the probabilities of intervals around actual realizations (see (1)). Therefore, predictors should satisfy condition (2). If predictors satisfying the criterion of highest concentration do not generally exist, then we should satisfy condition (2)(minimum mean square error) as closely as possible.

How do we nearly satisfy condition (2)? As explained by Chipman (1976, pp. 612-613), the conditional expectation of the random variable $y^{*}_{T+s}$ given $y$ will have the minimum mean square error about $y^{*}_{T+s}$ within a wide class of functions of $y$. If we restrict this class to linear functions or if we assume that $y^{*}_{T+s}$ and $y^{*}$ are jointly normally distributed random variables satisfying equations (6) and (4) respectively, then the conditional expectation of $y^{*}_{T+s}$ given $y$ is the same as the predictor (8) with $\pi$ replaced by $\pi$ (see Chipman (1976, pp. 603-604)). Thus, the conditional expectation of $y^{*}_{T+s}$ given $y$ nearly satisfies condition (2), and it can be evaluated fairly accurately, provided model (6) and assumptions (5), (7a)-(7c) are true.\footnote{Unfortunately, it is not possible to establish the truth of any logically valid model (see Swamy, Conway and von zur Muehlen (1985)).} Another difficulty is that if we believe a priori that model (6) and assumptions (5), (7a)-(7c) are true, then the conditional mean of $y^{*}_{T+s}$ given $y$ involves $\pi$, $w$ and $\sigma^2V$ which are usually unknown. The result (12) shows that any procedure of estimating these unknown quantities that attempts to reduce the variance of an estimator of $\pi$ does not necessarily lead to better predictors of $y^{*}_{T+s}$.

Minimum Mean Square Error Predictors

An appropriate way to overcome this difficulty is to estimate models using a smaller mean square error criterion. We can consider two cases:
normal and non-normal distributions of the errors. For the case in which a multivariate normal mean under quadratic loss is to be estimated, a paper by Natarajan and Strawderman (1985) establishes the existence of two-stage sequential estimators that are better, both in risk (mean square error) and sample size, than the usual estimator of a given fixed sample size. When $e$ in equation (1) is normal, the problem of estimating $\pi$ under quadratic loss becomes identical to the case discussed by Natarajan and Strawderman (1985). Given any sample size $T$, we can find a two-stage sequential estimator of $\pi$ truncated at $T$, with a positive probability of stopping earlier and mean square error lower than that of an estimator of $\pi$ based on all the $T$ observations. The predictor of $y_{T+s}$ based on a Natarajan-Strawderman-like estimator of $\pi$ can have smaller mean square error than any of the predictors of $y_{T+s}$ based on an estimator of $\pi$ that uses all the observations available up to $T$ or $T+s-1$.

Where normality is not appropriate, the advantages of a median or a trimmed mean or a robust estimator relative to the mean are well-known (see Lehmann (1983, pp. 352-388)). The predictors of $y_{T+s}$ based on these estimators of $\pi$ may have smaller mean square errors than predictors of $y_{T+s}$ based on other estimators of $\pi$ that use all the sample observations. Thus, there is no theoretical justification for reestimating models of the type (4) using all past data prior to each forecast period either in the normal case or in a nonnormal case.\textsuperscript{10}

\textsuperscript{10} If, in fact, the slope coefficients of equation (4) change over time, then it is incorrect to stack the observations as

$$[y', y_{T+1}', ..., y_{T+n}'] = [X', x_{T+1}', ..., x_{T+n}']' \pi + [e', e_{T+1}', ..., e_{T+n}']'.$$

In other words, the estimators of fixed coefficients are inappropriate when the slope coefficients vary over time. Appropriate estimators for a time-varying parameters model are given in Swamy and Tinsley (1980) and two applications in Resler, Barth, Swamy and Davis (1985) and Swamy, Kennickell and von zur Muehren (1986).


<table>
<thead>
<tr>
<th>IFDP NUMBER</th>
<th>TITLES</th>
<th>AUTHOR(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>287</td>
<td>Should Fixed Coefficients be Reestimated Every Period for Extrapolation?</td>
<td>P.A.V.B. Swamy, Garry J. Schinasi</td>
</tr>
<tr>
<td>286</td>
<td>An Empirical Analysis of Policy Coordination in the U.S., Japan and Europe</td>
<td>Hali J. Edison, Ralph Tryor</td>
</tr>
<tr>
<td>285</td>
<td>Comovements in Aggregate and Relative Prices: Some Evidence on Neutrality</td>
<td>B. Dianne Faulls</td>
</tr>
<tr>
<td>284</td>
<td>Labor Market Rigidities and Unemployment: the Case of Severance Costs</td>
<td>Michael K. Gavin</td>
</tr>
<tr>
<td>283</td>
<td>A Framework for Analyzing the Process of Financial Innovation</td>
<td>Allen B. Frankel, Catherine L. Mann</td>
</tr>
<tr>
<td>282</td>
<td>Is the ECU an Optimal Currency Basket?</td>
<td>Hali J. Edison</td>
</tr>
<tr>
<td>281</td>
<td>Are Foreign Exchange Forecasts Rational? New Evidence from Survey Data</td>
<td>Kathryn M. Dominguez</td>
</tr>
<tr>
<td>280</td>
<td>Taxation of Capital Gains on Foreign Exchange Transactions and the Non-neutrality of Changes in Anticipated Inflation</td>
<td>Garry J. Schinasi</td>
</tr>
<tr>
<td>279</td>
<td>The Prospect of a Depreciating Dollar and Possible Tension Inside the EMS</td>
<td>Jacques Melitz</td>
</tr>
<tr>
<td>278</td>
<td>The Stock Market and Exchange Rate Dynamics</td>
<td>Michael K. Gavin</td>
</tr>
<tr>
<td>277</td>
<td>Can Debtor Countries Service Their Debt? Income and Price Elasticities for Exports of Developing Countries</td>
<td>Jaime Marquez, Caryl McNeilly</td>
</tr>
<tr>
<td>275</td>
<td>A Method for Solving Systems of First Order Linear Homogeneous Differential Equations When the Elements of the Forcing Vector are Modelled as Step Functions</td>
<td>Robert A. Johnson</td>
</tr>
</tbody>
</table>

Please address requests for copies to International Finance Discussion Papers, Division of International Finance, Stop 24, Board of Governors of the Federal Reserve System, Washington, D.C. 20551.
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>274</td>
<td>International Comparisons of Fiscal Policy: The OECD and the IMF Measures of Fiscal Impulse</td>
<td>Garry Schinasi</td>
</tr>
<tr>
<td>273</td>
<td>An Analysis of the Welfare Implications of Alternative Exchange Rate Regimes: An Intertemporal Model with an Application</td>
<td>Andrew Feltenstein, David Lebow, Anne Sibert</td>
</tr>
<tr>
<td></td>
<td>1985 (partial listing)</td>
<td></td>
</tr>
<tr>
<td>271</td>
<td>Elections and Macroeconomic Policy Cycles</td>
<td>Kenneth Rogoff, Anne Sibert</td>
</tr>
<tr>
<td>270</td>
<td>Assertion Without Empirical Basis: An Econometric Appraisal of Monetary Trends in ... the United Kingdom by Milton Friedman and Anna J. Schwartz</td>
<td>David F. Hendry, Neil R. Ericsson</td>
</tr>
<tr>
<td>267</td>
<td>The U.K. Sector of the Federal Reserve's Multicountry Model: The Effects of Monetary and Fiscal Policies</td>
<td>Hali J. Edison</td>
</tr>
<tr>
<td>266</td>
<td>Optimal Currency Basket in a World of Generalized Floating: An Application to the Nordic Countries</td>
<td>Hali J. Edison, Erling Vardal</td>
</tr>
<tr>
<td>265</td>
<td>Money Demand in Open Economies: A Currency Substitution Model for Venezuela</td>
<td>Jaime Marquez</td>
</tr>
<tr>
<td>264</td>
<td>Comparing Costs of Note Issuance Facilities and Eurocredits</td>
<td>Rodney H. Mills</td>
</tr>
<tr>
<td>263</td>
<td>Some Implications of the President's Tax Proposals for U.S. Banks with Claims on Developing Countries</td>
<td>Allen B. Frankel</td>
</tr>
<tr>
<td>262</td>
<td>Monetary Stabilization Policy in an Open Economy</td>
<td>Marcus H. Miller</td>
</tr>
<tr>
<td>261</td>
<td>Anticipatory Capital Flows and the Behaviour of the Dollar</td>
<td>Arnold Kling</td>
</tr>
</tbody>
</table>