FORECASTING THE HOUSING SECTOR AND MORTGAGE MARKET: A REVIEW OF METHODOLOGY

FINAL REPORT

SOO-BIN PARK AND N. HARVEY LITHWICK

CARLETON UNIVERSITY

MARCH 1991

The work reported herein has been carried out under contract with the Canada Mortgage and Housing Corporation. We thank Mark Burchinshaw and Leonard FitzPatrick for their many valuable comments and suggestions. We also thank Peter Torraville and Geoffrey Oliver for able research assistance. Peter Torraville did the major part of the computer work involved in the project.

EXECUTIVE SUMMARY

FORECASTING THE HOUSING SECTOR AND MORTGAGE MARKET: A REVIEW OF METHODOLOGIES

In econometrics and time series analysis substantial effort has been devoted to developing forecasting methodologies in recent years. As a result a forecaster now has several models to use for forecasting. They include the conventional simultaneous-equation econometric model (SEM), the univariate autoregressive integrated moving average (ARIMA) model, the transfer function model, and the vector autoregressive (VAR) model. The primary objective of this study is to review these models and examine their strengths and drawbacks as a tool for forecasting the Canadian housing sector and mortgage market.

Regardless of which model one uses for forecasting, there are common issues in its construction and maintenance:

(1) Specification of equation(s) in the model—guided by the data and, in the case of an econometric model, also by relevant economic theory;
(2) Availability and collection of the data;
(3) Initial estimation and evaluation of the model; and
(4) Validation of the estimated model—diagnostic checking, *ex post* forecasting, simulation, and tracking tests of the model.

One has to address these issues very carefully if one wants to build and maintain a reliable forecasting model.

In Chapter 2 we have considered the conventional econometric model. It was perhaps the most popular forecasting tool until the 1970s. The major distinguishing feature of the econometric model is that a model builder can appeal to economic theory for model specification. Thus one considers the presumed interdependence and causality relationships between the variables, and imposes structure on data sets by specifying how variables may be related to one another. One may then be able to produce superior forecasts by utilizing *a priori* information on the interaction among the variables in the model.

Forecasting wth an econometric model in practice is not based on the model and data alone. A forecaster using an econometric model makes judgemental "add factor" adjustments to model parameters as well as the model forecasts. It is because of the subjective adjustments (1) that forecasts based on an econometric model cannot be duplicated by others and (2) that their reliability as measured by the standard errors cannot be computed.

There are several reasons why econometric approach to forecasting may not yield reliable forecasts:

(1) One requires forecasts of the exogenous variables to obtain forecasts of the endogenous variables. Inaccuracy in the forecasts of the former cause the conditional forecasts of the latter to be inaccurate;

(2) Some variables could be erroneously regarded as being exogenous;
(3) In dealing with unobservable variables like permanent income and expected inflation, one often utilizes their crude proxies in the model; and
(4) Economic theory seldom provides a model builder with a guide for dynamic specification.

Also reviewed in Chapter 2 is the theoretical framework of the housing sector and mortgage market model. The housing sectors of the major Canadian macroeconometric models are briefly reviewed in the context of this conceptual framework. Availability and quality of the housing and mortgage data are also critically reviewed.

Chapter 3 examines the autoregressive integrated moving- average (ARIMA) model as a forecasting tool. In contrast to the econometric model an ARIMA model requires little economic theory for its specification; it explains a time series in terms of its own past values and current and past error terms. However simple the approach may be, several researchers have found that, even without being judgementally adjusted, ARIMA forecasts perform as well as those judgementally adjusted forecasts based on large elaborate econometric models. Moreover, unlike the forecasts based on an econometric model, one can easily compute standard errors of ARIMA forecasts so that a precision measure can be attached to point forecasts. An ARIMA model of Canadian housing starts that we have built is seen to be remarkably accurate in *ex post* forecasting.

As a forecasting tool, the univariate ARIMA model shares the following strengths with other time series models:

(1) It generates very accurate forecasts at least for the short run;
(2) It is possible to estimate an ARIMA model on data up to any point in the past and then produce forecasts so that one can check easily the forecasting accuracy of the model; and
(3) It imposes none of the controversial theoretical restrictions that the conventional econometric model may contain.

In Chapter 4 we investigated the transfer function model as a forecasting tool. Forecasting with an ARIMA model yields forecasts of a single series without using information contained in other related series. In many forecasting situations in which other variables systematically influence the series to be forecast, one can build a transfer function model that contains more than one time series and introduce explicitly the dynamic characteristics of the series.

We have built a transfer function model of Canadian housing starts with the housing price index as an independent input variable. However, *ex post* forecasts from the transfer function model were not as accurate as those from the ARIMA model when judged in terms of such criteria as the mean absolute percentage errors (MAPE) or the root mean square percentage errors (RMSPE) of forecasts. Two reasons can be put forward for poor performance of the model relative to the ARIMA: (1) although the housing price variable does influence housing starts, it may not contain information about the variation in the starts variable above and beyond that already imbedded in the starts and (2) the forecasts of housing price variable based on its ARIMA model may not have been very accurate. As the forecasts of endogenous variables require forecasts of exogenous variables in the SEM, the forecasts of the output variable are conditional on the forecasts of input variables in the transfer function model.

In Chapter 5 we have considered the technique of vector autoregression (VAR) suggested by Sims for macroeconomic modelling. As the name implies, the VAR model regresses each variable in the system on the past values of its own and as well as other variables.

Like the ARIMA and transfer function models, the VAR model is formed from the regularity in the movements of time series and do away with the theoretical restrictions on the econometric models. Difficulties experienced in the conventional modelling of the housing and mortgage markets indicate that, despite or perhaps because of its *a*theoretical nature of approach, the univariate ARIMA and VAR models are very strong competitors in forecasting to the conventional econometric model.

Most VAR models of the Sims type in macroeconometrics use several variables and include several lagged values for each variable. The number of parameters to be estimated in a VAR model increases very rapidly as one increases the number of variables and/or the order of the model. Hsiao's refinement on the Sims VAR or the Bayesian procedure of the VAR modelling should partially mitigate the number-of-parameters problem in the VAR modelling. We have built different types of VAR models of Canadian housing starts and mortgage approvals. When judged by *ex post* forecasting accuracy, the Bayesian VAR seems to outperform the Sims and Hsiao types of VAR models in forecasting.

As a competitor to the conventional econometric forecasting, the VAR method has several advantages.

(1) It generates very accurate forecasts at least for the short term and maybe for the medium term;
(2) Unlike the SEM and the transfer function model it does not employ a dubious exogeneity definition; and
(3) It provides a conceptually straightforward and remarkably simple method of yielding forecasts that do not assume any particular values of exogenous variables.

It is well-recognized in the literature that VAR models forecast well in the short run but their ability to forecast may quickly deteriorate so that the conventional models offer superior predictions further in the future. Forecasting results reported in this report indicate that time series models forecast the housing starts remarkably accurately up to the twelve-month horizon.

Little is available in the literature on the comparisons of the long-term forecasting performance between the conventional econometric and time series models. Forecasts based on the time series models and the econometric models tend toward the mean of the series to be forecast as the forecast horizon increases. However, the SEM may outperform the VAR in the long-term forecasting, benefiting from the judgemental adjustments of the forecasters.

TABLE OF CONTENTS

1. Introduction 1

2. Forecasting with an Econometric Model 4

2.1 The Simultaneous-Equation Econometric Model 4

2.2 Forecasting with an Econometric Model 4

2.3 The Housing Sector Model 11

2.3.1 Economic Theory of Housing 11

2.3.2 Framework of the Macro Housing Sector Model. 14

2.3.3 Framework of the Micro Market Model 20

2.3.4 The Data 24

3. Forecasting with a Univariate ARIMA Model. 28

3.1 An ARIMA Model of Housing Starts. 28

3.2 Forecasting with an ARIMA Model. 32

4. Forecasting with a Transfer Function Model 48

4.1 A Transfer Function Model of Housing Starts 48

4.2 Forecasting with a Transfer Function Model 55

5. Forecasting with a Vector Autoregression Model 69

5.1 The Sims VAR System 69

5.2 The Hsiao VAR System. 70

5.3 The Bayseian VAR System 71

5.4 Forecasting Performance of VAR Models 74

6. Conclusions 83

References. 88

1. INTRODUCTION

The housing and mortgage markets are characterized by cyclical fluctuations and disequilibrium. Volatility of the housing sector is readily seen from the annual growth rate in the housing starts. It ranged from -26.4 to 26.0 per cent in Canada over the period from 1961 to 1989 while the growth rate in real GDP varied from -3.65 to 7.14 per cent. It is also widely recognized that the national housing market in an economy consists of many segmented local markets which may be in disequilibrium in the short run due to immobile housing stock and other factors. Such characteristics partly explain why activities in the housing and mortgage markets in Canada are difficult to forecast well.

The primary objective of this study is to review major forecasting methods that may be employed in forecasting the housing sector and mortgage markets in Canada at the macroeconomic level. Most forecasting activities in economics in recent years have been at the macroeconomic level using simultaneous-equation econometric models. Scepticism has been expressed about the value of a large-scale macroeconomic model as a forecasting tool in the short run; see, for example, Stekler (1968) and Sims (1980). Outside the economics discipline, on the other hand, Box and Jenkins (1970) have popularized a class of time series models as a tool for forecasting. Time series models have become a strong competitor to the conventional econometric forecasting.

In this study we concern ourselves with the assessment of major model-based forecasting methods. One can classify models into two groups: those based on the econometric models and those on time series models. The models in the first group include:

(1) the multiple linear regression model, which is based on the relationship between one dependent (or effect) variable and a set of independent (or cause) variables as suggested by economic theory; and
(2) the simultaneous-equation econometric model (SEM), which consists of structural equations describing the relationships among the endogenous and predetermined variables as the relevant economic theory would suggest.

On the other hand, the second group includes several different models.

(1) The univariate autoregressive integrated moving average (ARIMA) model, often referred to as the Box-Jenkins model, is used for forecasting a single time series from its own past history.
(2) The transfer function model relates an (output) time series, to be forecast, to a set of (input) variables. This model enables a time series to be forecast not only from its own past history but also from the past history of other related variables. This model resembles the multiple linear regression.

(3) The vector autoregressive integrated moving average (VARIMA) model is an extension of the univariate ARIMA model, representing a multiple time series with mutual interaction or feedback. Its special case is the vector autoregression (VAR).

The econometric approach in its pure form appeals to economic theory in building an econometric model, and may be termed "theory based." Data is used at the stage of model estimation. On the other hand, the time series approach relies heavily on the data in model building, and is thus "data based." Of course, the time series analyst does rely on economic theory for suggestions as to which variables could be jointly modelled in the transfer function or VAR modelling.

Regardless of whether one uses an econometric or time series model for forecasting, there are common issues in the construction and maintenance of a model:

(1) Specification of equations in the model—guided by the purpose of the model to be served and, in the case of an econometric model, also by the relevant subject matter consideration;
(2) Assembly of the necessary data;
(3) Initial estimation of the unknown parameters in the model; and
(4) The validation of the model—diagnostic checking, simulation, and tracking tests of the estimated model.

One has to address these issues very carefully if one wants to obtain reliable forecasts from the estimated model.

Recognizing the regression model as a special case of an econometric model, the SEM together with the three time series models are considered in this report. Of these four models two are of a system type and the remaining two are of a single-equation type. The conventional SEM or VAR model is of a system type in the sense that it contains two or more equations in it. The econometric model contains as many equations as there are endogenous variables, and yields joint forecasts of all endogenous variables given the future values of predetermined variables. On the other hand, the VAR model treats all variables in the system as being endogenous and generates joint forecasts of all variables in the system given their current and past values.

The univariate ARIMA and transfer function models are of a single-equation type. But one can use them to model the housing and mortgage markets by building a system in which each equation deals with only one endogenous variable. A successful example of this approach is provided by Mills and Stephenson (1987) in a time series forecasting system of the UK money supply.

The plan of this report is as follows. Chapter 2 examines the simultaneous-equation econometric model as a forecasting tool, and formulates the conceptual framework for the economic model of the housing sector and mortgage market. It also reviews the housing sectors of selected Canadian macroeconometric models and examines the data availability.

Chapters 3, 4, and 5 are concerned with the time series models. In Chapter 3 we discuss the ARIMA model and present the forecasting results based on a multiplicative seasonal ARIMA model of Canadian housing starts. Chapter 4 deals with the transfer function model; it also presents a transfer function model of housing starts with the housing price indexes as an input variable. Chapter 5 is concerned with three different types of the VAR model: the Sims, Hsiao and Bayesian types. It also compares the housing starts forecasts generated from different types of bivariate VAR models of housing starts and mortgage approvals. Some concluding remarks are contained in Chapter 6.

2. FORECASTING WITH AN ECONOMETRIC MODEL

Since the pioneering work of Tinbergen (1939) and Klein (1950) on macroeconomic modelling, the simultaneous-equation econometric model (SEM) has become a very important tool for forecasting and policy analysis. The 1960s and 1970s saw the rapid development of macroeconomic models, aided by the development of econometric methods, enormous advances in computing power and the availability of good quality economic data. This chapter examines the SEM as a forecasting tool.

In section 2.1 we briefly describe the standard simultaneous-equation econometric model. Section 2.2 discusses major distinguishing features of the forecasting approach with an econometric model. In Section 2.3 we provide the conceptual framework for the housing sector of an economy and briefly review the housing sectors of selected macroeconomic models and the national housing models. Section 2.4 examines the availability and quality of data on the housing and mortgage market variables.

2.1 THE SIMULTANEOUS-EQUATION ECONOMETRIC MODEL

Economic theory often hypothesizes that a set of endogenous variables are jointly interdependent and simultaneously determined by a set of predetermined variables. Endogenous variables can be thought of as the variables of interest, while predetermined variables are lagged values of endogenous variables and exogenous variables which are supposed to be determined outside the model.

Typically, a *simultaneous-equation econometric model* is represented by a system of G structural equations containing G endogenous variables, $Y_1, Y_2, \ldots, Y_G$, and K predetermined variables, $X_1, X_2, \ldots, X_K$. In practice, many structural equations are nonlinear in variables because different functions such as ratios and logarithms of some of the variables appear in the model. A general expression of the i-th structural equation is

$$f_i(\mathbf{y}_t, \mathbf{x}_t; \theta_i) = u_{ti} \qquad i = 1, \ldots, G; \qquad t = 1, \ldots, T, \tag{2.1}$$

where f_i is a known general function, θ_i is a vector of structural coefficients in the i-th equation, $\mathbf{y}_t' = (Y_{t1}, \ldots, Y_{tG})$ and $\mathbf{x}_t' = (X_{t1}, \ldots, X_{tK})$ are G-, and K-vectors of observations on endogenous variables $Y_1, \ldots, Y_G$ and predetermined variables, $X_1, \ldots, X_K$, respectively, and u_{ti} is the random disturbance in the i-th equation, all at time t. The θ's are called structural coefficients. Economic theory provides *a priori* information about the θ coefficients so that they are "identified."

When the SEM is dynamic and its predetermined variables include lagged values of endogenous variables, it specifies how the time paths of the endogenous variables are generated by the time paths of the exogenous variables over time. Suppose that the first G predetermined variables in the model are one-period lagged endogenous variables and the remaining $K - G$ predetermined variables are exogenous so that we may partition the vector of observations on the predetermined variables at time t as $\mathbf{x}'_t = (\mathbf{y}'_{t-1} \vdots \mathbf{z}'_t)$. The SEM in (2.1) can then be written as

$$f_i(\mathbf{y}_t, \mathbf{y}_{t-1}, \mathbf{z}_t; \theta_i) = u_{ti} \qquad i = 1, \ldots, G; \qquad t = 1, \ldots, T, \tag{2.2}$$

An econometric model such as (2.1) or (2.2) involves an explicit representation of the presumed causality or feedback relationships among the variables in the model. It is with this specification that economic theory plays a major role in econometric modelling. In the econometric approach to modelling and forecasting the forecaster's major concern is to formulate a model which is derived from economic theory. However, the "theory-based" approach in its pure form may not be practical for several reasons.

Economic theory may provide competing hypotheses. For "nested hypotheses" we may choose a particular hypothesis on the basis of standard statistical tests on the parameters on which we have hypotheses. A good deal of research effort has been spent recently on the choice between competing hypotheses of the non-nested type in regression models. See, e.g., Pesaran (1974), Fisher and McAleer (1981), and McAleer and Pesaran (1986). In practice, however, the model builder cannot express with confidence a preference for one model over any other.

Theory may suggest as relevant variables for which no data are available on them. For example, variables like permanent income and expected inflation may be considered as determinants of asset demand for houses. In dealing with such unobservable variables one often utilizes a crude, theoretically unjustified procedure to incorporate their proxies into the model.

Theory seldom provides the forecaster with a guide as to the formulation of dynamic structure in the model. Econometricians have recognized the importance of dynamic structure in modelling and spent a great deal of research effort on the introduction of lags into the econometric models. Their effort has produced a wide range of alternative lag structures including the geometric and rational distributed lag models. However, little theoretical justification seems to exist for imposing particular lag structures in practice, and an appeal to the data is necessarily made for help.

Economic theory appears has less to say about the time series structure of error terms. Typically in macroeconometric modelling a white noise error structure is assumed at the stage of model specification although tests of this assumption are subsequently made. In fact, error terms are added on to structural equations for estimation after a model structure is developed in deterministic terms.

The typical treatment of time series error specification in econometric work has been the first-order autoregression

$$u_t = \rho u_{t-1} + \epsilon_t,$$

where ϵ_t is "white noise." Greatly stimulated by recent development in time series analysis, model builders have considered higher order autoregressive and moving-average error structures in the context of the regression model and SEM.

2.2 Forecasting with an Econometric Model

Once a SEM is specified, it can be consistently estimated by a single-equation method such as the two-stage least squares (2SLS) or by a system method such as the three-stage least squares (3SLS). If the model size is large and contains many predetermined variables, these methods may not be applicable because of a limited sample size. We discuss in this section how an estimated model is used for econometric forecasting.

A. Forecasting with a SEM.

Klein (1968) defines forecasting as "the attempt to make scientific statements about nonsample situations on the basis of relationships determined from sample observations." Two types of forecasts can be distinguished: *ex ante* and *ex post*. For *ex ante* forecasts, one assembles as much relevant information as possible and extrapolates into the unknown future before an event unfolds. For *ex post* forecasts, one uses relevant information to forecast data already in existence. *Ex post* forecasting is a useful tool for diagnostic checking of the fitted model.

Two other types of forecasts can be distinguished: *unconditional* and *conditional*. The unconditional forecast is formulated as an unqualified forecast of the future event. On the other hand, the conditional forecast is a qualified forecast based on the assumption that some exogenous events simultaneously occur. Econometric forecasting is concerned with the *ex ante* conditional forecasts.

Suppose that at time T, called the forecast origin, h-step ahead forecasts of the variables of interest are required for some future time $T+h$. If the model is static and all its predetermined variables are exogenous, and if the values of the predetermined variables to be assumed at $T+h$ are specified, the forecasting problem becomes one of estimating the values that the endogenous variables of the model will assume at time $T+h$. Thus one can obtain the forecast of $\mathbf{y}_{T+h}$ given $\mathbf{x}_{T+h}$ by setting the future structural error terms $\mathbf{u}_{T+h}$ to zero and solving numerically the following nonlinear system:

$$f_i(\hat{\mathbf{y}}_{T+h}, \mathbf{x}_{T+h}; \hat{\theta}_i) = 0, \qquad i = 1, \ldots, G \tag{2.3}$$

for the forecasts $\hat{\mathbf{y}}_{T+h}$. The vector $\hat{\theta}_i$ in (2.3) is an estimate of θ_i, the vector of structural coefficients. Most of the statistical properties of the forecasts $\hat{\mathbf{y}}_{T+h}$ thus obtained are not known.

If the system is dynamic, the one-step ahead forecast of $\mathbf{y}_{T+1}$ at the origin T can be obtained by numerically solving

$$f_i(\hat{\mathbf{y}}_{T+1}, \hat{\mathbf{y}}_T, \mathbf{z}_{T+1}; \hat{\theta}_i) = 0, \qquad i = 1, \ldots, G \tag{2.4}$$

for $\hat{\mathbf{y}}_{T+1}$, where $\hat{\theta}_i$ is an estimate of θ_i in (2.2), $\mathbf{z}_{T+1}$ stands for the vector of the future values of the exogenous variables at time $T+1$, and the expected value of the future error $u_{T+1,i}$ is set to zero.

The h-step ahead forecasts can be obtained from the following recursion:

$$f_i(\hat{\mathbf{y}}_{T+h}, \hat{\mathbf{y}}_{T+h-1}, \mathbf{z}_{T+h}; \hat{\theta}_i) = 0, \qquad i = 1, \ldots, G \tag{2.5}$$

Thus h-step ahead forecasts are conditional on the current value $\mathbf{y}_T$ of endogenous variables, and the future values $\mathbf{z}_{T+1}, \ldots, \mathbf{z}_{T+h}$ of the exogenous variables up to time $T + h$.

B. Forecasting with an Econometric Model in Practice.

Forecasting with an econometric model in practice is based not just on the estimated model and the data. Evans, Haitovsky, and Treyz (1972) describe three distinct steps involved in econometric forecasting: forecasting future values of exogenous variables, adjustments made to individual equations, and the model solutions.

Future values of most exogenous variables are unknown at the time of forecasting and have to be forecast as well. The problem of forecasting endogenous variables has been translated into that of forecasting exogenous variables.

Future values of some exogenous variables may be known at the time forecasts are made; policy variables belong to this category. For other exogenous variables one may have to obtain their forecasts *judgementally*. One may also attempt to forecast them on the basis of all information available at the time of forecasting. A set of available information is the past history of the variables. If that history spans over a reasonably long period of time, one could employ a time series model to forecast the future values of the exogenous variables.

Adjustments are made to individual equations of the estimated model by subjectively adjusting the intercept terms and/or other parameter estimates. These "add factors" can be justified on the grounds that they incorporate into the forecasts any information/judgement on the factors that the forecaster has not explicitly brought into the model. Good information may be available about exogenous shifts and structural breaks—changes in legislation and institutional arrangements may be known in advance so that the forecaster modifies the model.

Finally, the (estimated and adjusted) model is solved to obtain forecasts of the endogenous variables. However, if the forecaster has an *a priori* assessment of the likely range of future values, and if the forecasts obtained as the model solution fall outside the range, the forecaster may modify them, either directly or by readjusting the fitted structural equations and resolving the model.

Degree of judgemental adjustments in forecasting varies from one forecaster to another and also depends on the state of the economy at the time forecasts are made. According to McNees (1988), the weight for judgmental adjustment in econometric forecasting in the U.S. varies from 20% for Chase Econometrics to 30% for of Wharton Econometric

Forecasting Associates. Such heavy use of judgmental adjustment in economic forecasting can be justified on the grounds that the forecaster takes into account all the information available at the time of forecasting and improves the accuracy of the forecasts. The quality of the forecast will then reflect the quality of the information used regardless of whether or not such information has been incorporated into the model building or through the add factor.

Until the 1970s the simultaneous-equation econometric model was perhaps the most popular forecasting tool. Major distinguishing feature of the econometric forecasting approach is that economic theory plays a role in model specification. Forecasters tend to build a large econometric model with the belief that small models are unrealistic and incapable of answering many important questions.

An econometric model has several advantages as a forecasting tool:

(1) It is a useful device by which current and past values of endogenous and exogenous variables as well as forecasts of exogenous variables are systematically transformed into forecasts of endogenous variables;
(2) As long as the model parameters are believed to remain unchanged in the future, forecasts can be made as far into the future as required given the forecasts of future exogenous variables;
(3) It also leads to forecasts of related variables that are consistent with one another since they must jointly satisfy the restrictions of the model, particularly its identities; and
(4) It may also be used for purposes other than forecasting. For example, the SEMs cam be used to assess the effects of various policy alternatives by simulating the time-paths of endogenous variables under particular assumed time-paths of the exogenous variables.

On the other hand, drawbacks of an econometric model as a forecasting tool include the following:

(1) It tends to incorporate many "incredible" restrictions on the parameters particularly when the model is large;
(2) Time series problems such as autocorrelated structural disturbances cannot be systematically taken into account and often ignored; and
(3) Standard errors as a measure of forecast accuracy cannot be calculated because forecasts are subjectively adjusted by add- factoring.

There are several sources of forecast errors associated with econometric forecasting.

(1) The most obvious sources of errors are the error terms themselves in the model equations. It is sensible to set the future values of the disturbance terms to zero, their expected value. But their realized values will be different from their expected values used for generating the forecasts of endogenous variables.
(2) Forecasts of the exogenous variables are required to solve the model to obtain the forecasts of the endogenous variables. Errors in forecasting the exogenous variables

outside the SEM will undoubtedly affect the forecast accuracy of the endogenous variables.

(3) Sampling errors in the parameter estimates will surely affect the forecast accuracy.
(4) The model itself is subject to specification error. There is no "true" model.
(5) Judgemental adjustment of forecasts is a source of errors even if it is meant to reduce the size of errors.

Many researchers have tried to assess the performance of econometric forecasting relative to alternative forecasting methods. A common approach to evaluate an econometric model is to regard it and a univariate ARIMA or other naïve model as competitors, and compare their forecast performance directly. The findings in the early 1970s indicate that the econometric models, unaided by judgemental adjustment, have not done well. Cooper (1972) compared one-quarter ahead post-sample forecasts of autoregressive models with those from seven quarterly U. S. macroeconometric models, and found that the former generally had smaller forecast errors than the latter. Nelson (1972) also found that ARIMA models outperformed the FRB-MIT-Penn model of the U.S. economy in *ex post* forecasting outside the sample period.

The forecasting performance of macroeconometric models has improved over the years. A more recent comparison by Longbottom and Holly (1985) between the forecasting abilities of ARIMA models and the London Business School model of the U.K. economy has shown that for many variables the SEM outperforms the ARIMA models in both *ex ante* and *ex post* forecasts though for some variables the time series model is better. With the econometric models of the U.S. economy McNees (1988) has found that for key macroeconomic variables the judgementally adjusted econometric forecasts have outperformed the univariate ARIMA models.

The heavy use of judgemental adjustment in macroeconometric forecasting poses a serious problem when one wants to assess the relative performance of alternative forecasting methods. The ability of a model to forecast well must be distinguished from the ability of an experienced and knowledgeable forecaster, aided by a well-built model, to forecast well. It appears that a forecaster, aided by an econometric model, can do better than the simple ARIMA models for a longer term forecast horizon.

2.3 The Housing Sector Model

2.3.1 Economic Theory of Housing

In this and the following sections we formulate a conceptual framework of the housing market within which the housing sector in macroeconometric models and national housing models can be assessed, in particular for forecasting purposes. The ultimate framework will reflect our understanding of the current conception of this sector, based on our review of recent literature in the field of housing economics.

A. The Basic Theory.

All theoretical and empirical research in housing is now based on a common understanding of the basic elements of housing economics. That basic theory is articulated in Smith (1974), Arnott (1987), and Olsen (1987). The following is a brief review of those elements.

i). *Demand*: The demand for "housing" as a commodity is viewed as a stock demand derived from the flow demand for specific housing services from that stock. In this sense, housing is treated within the mainstream of capital theory. Because of the many attributes of housing that contribute to those services (size, quality, etc.), demand for the stock is generally taken to be the demand for a single commodity providing a composite of those attributes. (Some recent work has stressed the demand for those attributes individually via hedonic indexes). Moreover, because housing is both an asset and a consumer good, the need to explain demand fully must reflect these dual motives. For renters, the demand is dealt with as the demand for a consumer good. However, the alternative of ownership is always available.

ii). *Supply*: Historically, the supply of housing was assumed to be new construction because it constitutes a flow. Within the newer stock-flow conception, the actual supply is understood to be the flow of services from the existing stock. New construction plays a role by adding to the stock, albeit with a lag, based on the usual profitability considerations (expected price relative to total construction plus land costs). Hence stock adjustment models are typical. Because of the fixity in the housing stock, the supply of housing services is inelastic in in the short run. In the long run, the elasticity is much greater because new construction increases the stock of housing services.

iii). *Nature of the Market*: To the extent that the housing stock is assumed to be homogeneous, the market is assumed to clear efficiently, and the price in the market is assumed to reflect the equilibrium or market clearing price.

iv). *Finance*: The market for housing must be viewed in close proximity to the market for finance. That is because the size of the outlay necessary to acquire the stock requires debt financing via mortgages in most cases. Most models incorporate interest rates explicitly within the demand function. Further, the costs and availability of mortgage credit would affect the housing market.

B. Unresolved Issues in the Basic Theory.

A number of issues are not fully resolved by the basic theory. They include:

i). *Demand*: How to deal with the whole question of uncertainty, and of expectations particularly with respect to price and interest rates, remains a contentious but important matter.

ii). *Supply*: It is now recognized that there are crucial short run changes on the supply side, involving quantity adjustments (vacancies/crowding), conversions, demolitions, upgrading, etc. that must be accounted for, both in determining price and in understanding tenure choices.

iii). *Price*: The pricing of the stock is problematic because only a small proportion of the stock is transacted in any year.

iv). *The Market*: Because of the differentiated nature of the commodity, in terms of type (single, multiple), location, and so forth, there is some question as to whether it can be treated as a homogeneous commodity in a unitary market. In fact, most theorists treat the situation as a complex of closely related, interdependent markets.

v). *Equilibrium*: The nature of lags, on both the demand and the supply side, suggest that equilibrium prices will rarely be observed. This is likely the case within a locationally unique housing market (an urban area or Census Metropolitan Area), and certainly the case nationally.

vi). *Public Policy*: Housing has come to be viewed as a social good, indeed a right. It therefore commands significant attention by policy-makers at all levels. Moreover, the limited tax base of municipalities encourages them to pursue policies for housing designed to augment their revenues, often in ways that distort the housing market. These policies are rarely introduced into housing models.

vii). *Finance*: The "traditional" view of housing supply as equivalent to new construction had direct consequences for the treatment of financial markets. What was hypothesized was a link between selected financial variables (mortgage rates, approvals, etc.) and the housing starts. The nature of the mortgage markets was never formally modelled. Indeed, the complexity of that market has no doubt deterred such efforts. Moreover, under the newer stock adjustment paradigm, this view of causality is inappropriate. Clearly, there is much room for theoretical advance in this area.

None of these qualifications require a change in the basic theoretical structure of the housing market, but they do impose on analysts the need to elaborate the basic model with much greater care than would be required for most other markets.

C. Macro vs. Market Approaches: Some Contrasts.

For very practical reasons, the elaboration on this basic model by housing economists has in fact proceeded on two quite distinct planes.

The first of these planes is essentially macro-forecasting focused, driven by the need to predict key variables in the housing sector that are of primary interest to economic policy-makers (starts, level of investment, mortgage rates, prices and rentals, vacancies, etc.). As such, this literature is concerned less with how the housing market works, and more with aggregate housing outcomes. In Canada, there have been few if any macro models explicitly dedicated to forecasting housing outcomes at the macro level. What we do have are segments of larger macro models devoted to explaining key housing outcomes. Explaining those outcomes tends to be largely eclectic, however, with only the most simplistic theoretical foundations because the size of the larger models limits the number of variables and equations in the housing sector. The emphasis is necessarily highly aggregative, almost exclusively at the national level. Finally, because housing markets are spatially disjoint, the aggregate data used in these models tend to reflect non-comparable and *disequilibrium* circumstances. Interpreting the results can be problematical in consequence.

Equilibration is achieved via an unspecified black box of stock adjustments and transformations. In addition, complex capital market adjustments as well as public policies (national and local) that impact on the housing market tend to be handled in a somewhat cavalier fashion.

As we shall see in the next section, there have been important theoretical advances over the past five years that permit refinements to be made to this basic model, making it of potentially greater use for housing analysis even at this aggregate level. Of particular interest has been the analysis of what is going on inside the black box. Improved understanding of the role of capital markets and of public policy are also available.

The second plane of housing analysis focuses more on the functioning of ***housing as a market process.*** The work here tends to be less comprehensive, focusing on ***particular*** outcomes, such as type of unit, tenure choice, etc. Moreover, ***location*** of the housing assumes much greater importance, as does the crucial question of ***market interdependence.*** Because of this micro focus, the models are generally of a ***long-run equilibrium*** nature. It should be added that the specificity of this micro work makes it difficult to apply the findings on a national basis. The models are much more data intensive. Problems of quality of data, lack of comparability between markets, timeliness, and analytical complexity, etc. render these models of limited use for macro-forecasting. Systematic application of the models to all housing markets is only rarely pursued.

Nevertheless, such models, rooted in the work of Muth (1988), permit much more detailed analysis of how the housing market works, and in particular have encouraged diverse investigations into what lies within the black boxes, via stock adjustments, including repairs, demolitions, etc.

The problem for forecasters (and ultimately for policy-makers) lies in reconciling these two planes. Ideally, we would like a consistent modelling approach which covers both of them, meeting the practical needs of the macro models, and satisfying the theoretical requirements of the less aggregative models. Some work in this area has commenced, but essentially at a relatively abstract level [Carruthers (1989)]. In practice, reconciliation, if pursued at all, tends to take place on the basis of casual empiricism only.

In the following section we report on recent developments (theoretical and empirical) in each of these areas, and to see where more formal reconciliation might begin to take place. We begin with an elaboration of the basic framework of the macro model. Recent developments in macro-modelling will then be reviewed in the context of this model. In other words, we shall consider how they improve the conceptual rigour of the model. We then turn to an analysis of recent developments in the micro sphere that, in our judgement, hold out promise for improving the macro model.

2.3.2 Framework of the Macro Housing Sector Model

A. The State of the Art in Canada.

Our search of the literature began with a review of existing macroeconometric models in Canada. We were fortunate in being able to draw on very competent assessments [Foot (1985) and Grady (1985)]. What is clear is that because housing is a small component of those models, little attention has been paid to building a theoretically rigorous housing sector. In an important sense, this is a serious failing for the macro models themselves, because housing has a major impact on key macro variables because of the contribution of housing to macro instability, as well as the crucial link between the monetary and real sides of the economy that takes place within the housing sector.

All housing sector models are based on essentially the same basic housing paradigm briefly described in the preceding section, namely stock demand impacting on existing fixed supply to determine price, with supply adjustment (primarily new construction) in response to price.

Using this paradigm, Foot concluded that the housing sector in existing Canadian macro models does not build upon micro foundations in a serious way. Specifically, these models neglect portfolio analysis and are weak on interest rates, market segmentation, the resale market, the index of housing prices selected, and the major lags that are inherent in this sector.

It is useful at this stage to provide a more detailed assessment of the housing sector in Canadian macro models. Grady provides excellent summaries of them [pp. 253-293]. Virtually all are almost exclusively preoccupied with estimating the level of investment in residential construction. He divides the models into two broad categories:

(A). Essentially, housing starts are linked to mortgage approvals. Note, however, that this approach gets the direction of causality in the housing market wrong; normally housing starts would precede mortgage approvals.

(1) Candide 2.0 (Economic Council of Canada): Demand (function of mortgage approvals, demographic variables, and the relative cost of owning as opposed to renting) and supply (vacancy changes, government policy) determine single and multiple starts.
(2) RDX2 (Bank of Canada): Demand is a function of availability of mortgage funds.
(3) FOCUS (University of Toronto): Demand is a function of availability of mortgage funds.

(B). Housing starts are modelled directly;

(1) Based only on demand considerations;
   i) TIM (Informetrica): Demand driven stock adjustment model.
   ii) DRI (Data Resources of Canada): Demand driven stock adjustment model.

(2) Considering both demand and supply;
   i) RDXF (Bank of Canada): Price (MLS) responds to the gap between demand supply (starts).
   ii) QFS (Department of Finance): Similar to RDXF.
   iii) CHASE (Chase Econometrics): Demand (real effective purchasing power and mortgage rate) and supply (sales price relative to construction costs) determine starts per capita.
   iv) MTFM (Conference Board of Canada): Demand, supply (profitability of starts) and government assistance.

Grady's negative conclusion based on his comprehensive evaluation is worth noting:

"Housing is a problem sector. ...It may well be that the only viable option is to start from scratch and build a structural model of the housing market that will reflect the behaviour of all agents and clearly specify demand and supply in all sub-markets." [p. 237]

More recently, the Department of Finance has introduced a new macro model, Canadian Economic and Fiscal Model (CEFM), which is described in some detail by Stokes (1987). Although based on the same basic stock-flow paradigm, the model goes some way to incorporate key improvements indicated in the literature. In particular, the supply side is enhanced to include a specific production function for new housing; the crucial links between the financial and real sides of the markets are enhanced; expectations are incorporated explicitly; and both short run (disequilibrium) and long run (equilibrium) outcomes are considered.

The basic structure of the model is as follows:

(1) Two markets are identified (new and existing housing) and two types (single and multiple);
(2) In the short run, stock is fixed. Production of new housing is explicitly considered as addition to stock;
(3) A traditional explanation of stock demand is provided.

Given fixed stock, the price for existing housing, and expected price appreciation can be derived. (Rent estimation is more difficult because of rent control.)

(1) Starts respond to the supply of new housing, and depend on profitability, itself a function of expected selling price and cost of construction;
(2) Stock includes demolitions and completions (a function of starts plus time); and
(3) Construction expenditures are then a function of new housing (current plus past starts) plus other costs.

Clearly this is a much more satisfactory model of behaviour in the housing market.

Our own analysis indicates that the supply side, especially for new housing, remains particularly weak in most models (except CEFM), neglecting key issues such as production functions, key cost elements, and technological changes. In addition, the process of stock adjustment is simplistic in virtually all models.

In no model did we find a satisfactory submarket for finance. In other words, despite its importance, the *behaviour* of the mortgage market has not been introduced endogenously into housing models. The reasons are clear. It might well be simply impossible to specify in a few key equations what is a very complex market dominated by oligopolistic institutions that engage in both price and quantity management, and model-builders may have been discouraged from devoting significant efforts to this subsector. [Hatch(1975) provides a somewhat dated but valuable description of the Canadian mortgage market.]

The only other attempt to model the housing sector on an aggregative basis has been that of Clayton (1987). In reality, this is essentially a forecasting exercise, and the "model" is rather mechanistic. It consists of the following elements:

(1) Derivation of household projections from population projections, estimating households by age, type, then by tenure and ultimately by dwelling type. Regional projections are also provided. The method of analysis is the cohort method, and the variables are not endogenous to a housing sector model. Rather, trends are identified in an admittedly subjective manner. This holds as well for tenure choice, which is therefore not a function of explicitly modelled market outcomes. The same is true for dwelling type.
(2) Estimate replacement demand by tenure and dwelling type. These numbers are crude because there are no good data.
(3) Estimate level of vacancies. Again crude estimates only.
(4) Projection of requirements for new housing by tenure and dwelling type, based on new demand (1), replacement demand (2) and changes in vacancy (3).

Clearly, this exercise falls at the outer limits of macroeconomic forecasting. It includes virtually no traditional economic variables, such as prices, incomes and mortgage rates. Significantly, there is no reference to any macro models or even theoretical literature in the text. Because it takes a long view, it might be argued that these traditional macro variables are of less use, because they are concerned with shorter run developments. But it is difficult to accept that economic variables play no role in the long run, or more accurately, that they can be incorporated implicitly through the subjective judgements regarding weighting factors and so forth.

However, the methodology might also be taken as a reflection on the utility of current models for practical forecasting purposes. In any event, the study is of no particular use in explaining the workings of the housing sector, either in macro or in market terms.

B. Developments Elsewhere in Macro forecasting Models.

Other than the CEFM there is little evidence of important advances in macro modelling of the housing sector in Canada. Work in other countries has proceeded, however, with explicit forecasting housing models well beyond the level of what appears to have been achieved in Canada.

One innovative study by Williams (1984) for Australia makes use of some of the advances in the literature. An intermediate term model with 13 equations takes into explicit account the links between portfolio choice, housing prices and new construction. A generalized asset adjustment model provides a good explanation of the asset price of the existing stock (and associated land). New construction is a response to asset prices, and the elasticity of total starts with respect to asset price is found to be 1.4, with important policy implications regarding restricting housing prices. A number of serious data gaps preclude completely satisfactory estimation, however. Finally, and significantly, the author is of the opinion that:

"The model developed in this paper could be expected to yield richer results if estimated for regions. The concept of an average value of dwellings in Australia has obvious limitations. [p. 153]."

A number of other essentially market (as opposed to forecasting) models, but with macro modelling implications are reviewed in a later section. However, one important macro-forecasting study by Goodman and Gabriel (1987) is worth particular notice, as it may well represent the state of the art in such models.

Their study is a specific investigation into the reasons for the poor forecasting performance of U.S. residential construction activity in the 1980s. Their emphasis is on the importance of structural changes that have occurred, and the fact that they are omitted from most models. They proceed to evaluate the key influences on residential construction, taking into account these structural changes as well.

In summary, they find that:

(1) New financing instruments, such as ARMs, have had a modest impact;
(2) Affordability and expectations have had a positive impact;
(3) Changes in federal tax treatment had a positive impact;
(4) Tax-exempt bond financing of mortgage credit had a positive impact but it had largely displaced conventional mortgages;
(5) Demographic changes, especially the decline in household formation in the early 1980s had a negative impact, but this variable rebounded when household growth resumed later in the decade; and
(6) Regional economic development changes that favoured the higher cost Northeast led to a reduction in construction activity.

They also find that traditional explanatory variables were important, especially income growth and interest rates. *The reason the forecasts were poor was that these variables were not correctly forecast.*

C. A Modified Paradigm.

In this section, we propose to provide a modification of the traditional model [as summarized by Foot (1985)] to incorporate several of the newer ideas emerging from the literature, some of which have already been introduced into the CEFM model and the foreign models.

The revised framework is shown in Figure 2.1.[1] The following are the basic elements of the modified paradigm:

(1) The demand for housing services is a function of prices, (permanent) income, demographic factors, financial variables (mortgage rates and conditions), policy variables, and key macro forces (inflation, employment).

(2) The demand for housing services is translated into a demand for the stock of housing necessary to provide those services.

(3) The supply of stock in the short run is the existing net stock. This stock consists of the existing gross stock, plus improvements (rehabilitation) less removals (demolitions, depreciation). New investment is not equivalent to the immediate supply of housing services, but a lagged response of stock augmentation to signals provided in the short term housing market. The process by which new investment works its way into actual housing stock proceeds from approvals to starts to completions.

(4) New housing investment also impacts directly onto macroeconomic variables (aggregate demand) and monetary variables because of the scale of this activity.

[1] Figures and tables are at the end of each chapter.

(5) The interaction between demand and supply of housing stock constitutes the process of shorter run stock adjustment. Mismatches in the adjustment process produce vacancies, and households adjust via mobility. The result is a filtering of the stock through demanders, some of whom in turn change their tenure status, from owners to renters and the reverse. Prices and rents are determined in response to these adjustments, although they rarely reflect equilibria.

(6) The mortgage market is treated endogenously so as to (try to) predict somewhat more rigorously the role of key financial variables.

This modified paradigm differs from the traditional one primarily in its emphasis. For example:

(1) It proposes to give much greater emphasis to the stock adjustment process, and to the various transactions that underlie that process. This modification is essential if we seek to link developments in existing housing markets to the traditionally more heavily emphasized new construction activity. Because of their interdependence in the stock-flow model, this linkage is vital. Such a framework points to more emphasis on estimating key variables in the stock adjustment process, such as vacancies, renovation activity, and removals. It also emphasizes the nature of transactions and price developments for existing homes than has occurred to date.

(2) Our framework also underscores the unique position of capital markets. This has the intention of encouraging new models to incorporate much of the new thinking about housing as an asset, and the impact on the demand for housing of influences that work their way through capital markets, such as inflation, tax changes, and so forth. The rapidly changing structure of capital markets has altered significantly the impact of those variables, which requires that models of the housing sector, which are so dependent on financial variables, be capable of reflecting those changes.

(3) Finally, the very substantial literature on tenure choice, as well as its policy importance, points to attempting to incorporate that decision directly within the framework of the housing model.

It is this modified framework that has guided our review of the various literatures that follows. We shall modify it further after our more comprehensive review of the micro literature in the following section.

Conceptually, the major problem with the housing sector in all the available Canadian macroeconometric models is that the link to the micro foundations of housing economics is weak at best [Grady (1985), 280-82]. In addition, there is no analysis below the national level of aggregation. This problem is acute in the housing sector, because the various submarkets will necessarily be in diverse stages of disequilibrium in the short run, because there is no trade or rapid factor flow mechanism for short run adjustment between markets.

Aggregating over these markets in the long run is less problematical, as there is mobility of persons and fungible capital over time. But these models are by nature short run, and the housing sector is necessarily inadequately handled as a result.

2.3.3 Framework of the Micro Market Model

The level of our understanding of housing markets, and the state of the art in theoretical and empirical terms was well summarized by Smith, Rosen and Fallis (1988). Their focus, however, was primarily on market oriented research, and only peripherally on macro models.

Their goal was to see how new contributions have addressed special problems that are intrinsic to the unique characteristics of the housing sector. They stressed that they were not able to subsume all the diverse research within one general model. They did observe increasing overlap between models, suggesting an eventual convergence in our understanding of the nature of housing markets. This fact is important, for it confirms that at this stage it is very unlikely that a general model of the housing sector can be constructed for macroeconomic forecasting purposes that in any fundamental sense represents the working of the housing market. Rather, the macro models will continue to be eclectic and not theoretically satisfying for the foreseeable future.

The key findings of their review may be briefly summarized:

(1) There has been a shift from demand and supply for housing services, to stock demand and supply.
(2) The emphasis on stock permits greater focus on stock adjustments (maintenance, filtering), and permits an explicit consideration of disequilibrium situations (vacancies), the consequences for the prices of the stock, and thereby, the impact on new construction.
(3) There has also been growing emphasis on the heterogeneity of the stock via characteristics and the use of hedonic indexes to measure those characteristics.
(4) The fact that housing choice spans time periods has been explicitly considered.
(5) The demand for housing as an investment as well as a consumption good, requiring portfolio analysis has been addressed.
(6) Finally, work on general equilibrium models has continued, but these remain highly theoretical, abstract and hard to solve. In other words, they are still not too applicable for macro-modelling.

With respect to macro considerations, they pointed to several areas of advance. One involves work on the impact of financial variables (on the instability of construction and hence on costs), with emphasis on institutional factors and information imperfections in addition to interest rates. Another focus has been on the impact of inflation, directly on the demand for housing, and indirectly on the mortgage market (the tilt problem).

Our own extensive review of work since 1985 basically confirms these impressions. However, we have chosen to summarize the literature on the basis of our modified framework, to provide more explicit focus to model-builders and forecasters.

On the *demand* side, debate has continued over the correct elasticity coefficient. Harmon (1988) finds that for alternative measures of permanent income, the results remain quite stable. However, the meaning of those coefficients has been contested, due to problems with grouped data [Rosenthal (1986)], or with the mixing of multiple decisions, such as household formation, tenure choice and quantity demanded, into the demand equation. Denton, Robb and Spencer (1986) have developed a model of expenditure that includes tenure choice. Dynarski and Sheffrin (1985) stress that transitory income must also be considered, because it has been found to play a key role in explaining tenure choice and homeowner mobility. In addition, the debate over stock vs. flow demand continues. A recent empirical study by Glennon (1989) attempts to integrate this debate into the housing as consumption plus investment framework, by relating stock demand to the latter, and flow demand to the former. Muth (1986) presents a strong case for adaptive over rational expectations.

Most of the recent work on housing demand has focused on demographic variables. One important idea that bears consideration is the argument by Mutchler and Krivo (1989) that demographic variables are in fact dependent on housing stock, particularly via household formation, as well as the reverse, and hence the treatment of demography must become endogenous to housing models. Certainly for long run models, this is an important consideration. It would also appear from several studies, that the demands of the elderly and of the young for housing will require much greater analysis than has been accorded thus far, given their growing weight in the housing market.

The consequences of inflation for housing demand, and especially for estimates of income elasticity, remains an active question, especially with the evolution of portfolio choice models. Pelser and Smith (1985) have developed a model considering the impact of inflation on tenure choice as well, based on a user cost model. Surprisingly little work has taken place on the impact on demand of taxes other than national levies. Given the proliferation of local development charges by school boards and municipalities, this area will require careful study, as the amounts are large, as will be the impacts on housing.

The *supply* side of the housing market continues to receive much less attention, perhaps reflecting the underdeveloped supply side in many macro models. The production or cost function for new construction is most neglected. Despite the growing importance of land costs in housing prices, the role of land in macro models remains invisible. Similarly, other inputs, such as labour and materials costs, infrastructure, productivity, etc. remain peripheral topics. Much more analysis has focused on new residential construction and investment outlays. A recent study by Topel and Rosen (1988) begins to address the issue, by offering a model based on dynamic marginal cost pricing. Little new work was discovered on the important macro question of economic instability and its impact on housing investment, despite the volatile swings in the economy since the early 1970s.

Consistent with the growing recognition that the flow of housing services depends on the size and composition of the available net stock of housing, much research has been undertaken to help define and measure the volume of *net stock available out of the existing gross stock*. Hendershott and Smith (1988) have provided a general framework by stressing the role of inventory adjustment as a way to examine previously unrecorded additions to the housing stock. Much detailed analysis has gone into additions (other than construction) to and removals from gross stock. The role of renovations and rehabilitation has been studied, primarily from a theoretical perspective, although some empirical work [Boehm and Ihlanfeldt (1986)] is available. The removals end has focused on improving estimates of depreciation [Malpezzi, Ozanne and Thibodeau (1987)], and on housing mortality.

The growing emphasis on the stock-flow model has also made it imperative that the means by which the stock is allocated be more fully understood. The process of *stock adjustment* has been approached from a variety of perspectives. One looks at the process in its entirety, by exploring models of adjustment in the housing market. A recent long run model of the U.S. housing market [Gahvari (1986)] computes an explicit intertemporal elasticity of substitution between housing and other consumption goods. Turner and Struyk (1985) employ the Urban Institute Market Simulation Model, with its 10 year forecasting framework, to project the long-term effects of U.S. policy. An explicit model of the resale market has been introduced by Rosen and Smith (1986) which features stock demand and stock supply impacting on prices.

Less attention has been focused on the behaviour of players in the housing market as part of the adjustment process. However, the actual behaviour of sellers of stock has been explicitly analyzed in Sweden [Aguilar and Sandelin (1984)]. Work on transactions more generally has introduced search and bidding models. The housing market as a contracts market has been considered, adding new insights to the adjustment process. In addition, the role of intermediaries in these transactions has come under scrutiny [e.g., Jud and Frew, (1986)].

Because of the lags in adjustment in the housing market, many of the models of stock adjustment have had to explicitly consider quantity as well as price adjustments. A key variable in this process is vacancy rates. One creative way to deal with this issue, via vacancy transfer modelling, has been proposed by Emmi (1984). Of course, adjustments are also made by households, through their changing of residence, or mobility. There has been little work explicitly on that subject in North America, although Quigley (1987) has analyzed the impact of financial variables on mobility. Some research in Holland has stressed mobility, which will undoubtedly be more systematically examined here in the future.

A very useful way of linking vacant units and mobile households within the housing market adjustment process is through filtering. Baer and Williamson (1988) have developed a model that offers a unifying framework for this research. They also identify areas for future research on this topic. The importance of filtering for policy is highlighted by Weicher and Thibodeau (1988).

An important mechanism that has profound impact on stock adjustment is the choice of tenure. For policy purposes, distinguishing the situation of owners from renters is of great importance. Both considerations have helped promote a voluminous literature on tenure choice. Krumm (1987) demonstrates rather convincingly that such choices must be seen in an intertemporal context, and that evidence based on one particular point in time are misleading. In addition, the relevant income and wealth variables influencing tenure choice have been debated rather extensively. Finally, several studies have pointed to the condominium phenomenon as a useful intermediate stage between the traditional transition from renting to owning.

The role of *finance* continues to be explored. Most of the empirical research has examined the impact of key financial variables rather than the working of the mortgage market itself. For example, estimates of the elasticity of supply of mortgage lending and the impact on housing have been computed by Muth (1986). Credit rationing is found to be important in the short run, but not in the long run. Little advance on both the descriptive works of the 1970s [Anderson and Ostas (1977)] or the theoretical-empirical works of that period [e.g., Smith and Sparks (1970)] has been noted. The latter is of importance because it articulates a model of the mortgage market, relating mortgage rates directly to NHA and bond rates, and indirectly to portfolio adjustments of financial institutions. A more recent study by Jaffee and Rosen (1979) exploring the efficacy of stabilization policies was also based on a model of particular interest. Rather than a single equation model with a single "price" variable, they introduce both the contract rate and other mortgage terms. Also, they compare partial to instantaneous adjustment processes, and find that the former performs better.

Thus, there are elements upon which a full model of the mortgage market could be constructed and inserted into a larger housing model. To date, we have found no evidence that this has been done. The changing role of financial variables is explored in important papers by Friedman (1989) and Kahn (1989), which emphasize institutional and other changes in monetary variables, that make historical relationships extremely unreliable for predicting the future. The close interdependence between financial and housing markets, both of which tend to be in disequilibrium, has been examined by Goodwin (1986).

Specific housing *market outcomes* are analyzed in several important articles. Hendershott (1988) attributes the increase in house prices in the 1980s primarily to negative productivity growth and increased ownership to rising income. The impact of vacancy rates is found to be less than previously thought by Gilderbloom (1986).

From a more macro perspective, where the entire structure of housing prices in the nation is important, it is important to understand how in fact prices in one market impact on others. Rosenthal (1986) examines the spatial interdependence of housing markets, and finds that in the short run, because of the complexity of the linkages, there is little spillover of prices. If this is validated, then the necessity to address markets individually for short term forecasting becomes very strong indeed. A model that explains some 70% of intercity variations in home price appreciation has been introduced by Manning (1986).

2.4 The Data

The feasibility of constructing an econometric model of the housing sector and mortgage market depends not only on the economic theory one may appeal to for its specification but on the availability and quality of the data to be used. Hence our study included an examination of the principal sources of data available for econometric modelling of this sector at the national level.

We grouped the housing sector and mortgage market variables into five categories relating to: demand for stock, supply of new housing, supply of stock, market clearing, and finance. We were particularly concerned with the frequency of availability of the data. We were also concerned with data quality. We ascertained data quality by consulting both with producers of the data, and with users, especially those involved in building and maintaining major Canadian macroeconometric models.

A. Demand for Stock.

The data on the *price of existing dwellings* are available monthly, quarterly, and annually from the Multiple Listing Service (MLS) of the Canadian Real Estate Association. Although very broad in coverage, the proportion of MLS listings varies cyclically. We were unable to ascertain to what extent this variability distorted market price estimates.

For *rents* there are two major sources of data: the Consumer Price Index (monthly and annual) and CMHC survey reports (semi-annual). The former source maintains superior quality control over time, but lacks the large sample size of the latter. In addition, it was suggested that the survey reports provide a better indication of market rents (as opposed to pure price movements). It should also be recognized that there is a general consensus that the rent component of CPI has a downward bias.

Data on *assets* of households are very poor. They are not produced on a regular basis, and coverage is incomplete. The sources are Statistics Canada and FAMEX.

Most surprising is the poor quality of the data on the number of ***households***, which is by definition equal to the stock of occupied dwellings. Annual data is available from HFE/HIFE and the Science, Technology, and Capital Stock Division of Statistics Canada, although they are deemed by users to be highly suspect. The former derives its data from estimates of total population and average household size, while the latter estimates net housing stock, and by making an allowance for vacancies, derives the occupied stock; consequently, they independently provide different estimates of the same variable.

The *property tax* data of households obtained from StatsCan's Provincial Accounts appear to be satisfactory.

B. Supply of New Housing.

The supply of new housing depends very heavily on the profitability of the activity on the part of landlords and builders.

New house price indexes (monthly, quarterly, and annual) are available from StatsCan. CMHC also provides monthly and quarterly data on the average unit selling price of newly completed, single-detached dwellings. Both sources provide a starting point for price data.

With respect to *land, labour, and materials costs* there are annual data provided by the Construction Census of StatsCan although the quality could not be ascertained. In addition, monthly and annual indices of *labour costs* and *materials costs* are also available from StatsCan. The labour cost indices are based on contracted union wage rates which clearly underestimate the costs of labour. The materials cost indices do not reflect the true costs to the builders because different builders receive different rates. The cost indices from StatsCan appear to be of questionable quality.

Annual *tax rate* data for the construction industry appear to be adequate although obtaining data on the residential sector alone is more costly. Annual data on *capital consumption allowances* are limited to the entire construction industry, which is too inclusive to be of use for analyzing the residential subcomponent alone. The quality of annual *profit* estimates is inconclusive, although a great deal of caution was advised when using them. Moreover, only the Construction Census gives a breakdown for the residential subcomponent.

If input costs data are of poor quality, measures of housing activity are of very high quality. Monthly, quarterly and annual data on the *housing starts* and *housing completions* are available from CMHC. It also provides monthly and annual data on *new construction expenditures*. It reports data on the estimated construction costs of new and existing structures but only for those that are NHA financed. Since the value of land is not included, the series underestimate the true construction expenditures. CMHC also publishes annual figures for the housing stock that is supplied through *public programs.*

C. Supply of Stock.

Annual *gross stock* data from StatsCan is viewed by users as extremely problematical, so much so that at least one person argued that as a result of this gap, it is not possible to build an econometric model of the housing sector at this time. It is also to be noted that the construction census data of StatsCan is reported with a 2 year lag.

Annual *removals* data must rely on demolition reports from StatsCan which appear to be extremely suspect due to problems in reporting. Consistency is also a problem due to increasing coverage over time. Most regard the annual survey data on *renovations* as quite poor. Quarterly and annual *depreciation* data at replacement cost are available from StatsCan, and appear to be adequate. Annual *net stock* data suffer from the same shortcomings as the gross stock data.

D. Market Clearing Variables.

CMHC's annual and semi-annual *vacancy* data are biased downwards due to a very tight definition. *Sales* data are available from MLS monthly, quarterly, and annually. They are subject to systematic bias, as was noted under prices of existing dwellings. The alternative source, Teela, is judged to be sloppily compiled and is not well-regarded either by users or producers of data.

Annual data on *tenure* (HFE) and on *type of housing* (CMHC, Construction Census) are generally very acceptable.

E. Finance.

Data on *mortgage rates* (CMHC) are available monthly and quarterly, *mortgage approvals* (CMHC) monthly, quarterly, and annually, *credit availability* (Bank of Canada) monthly and annually, and *carrying charges* (HIFE) annually. They are all considered quite good.

F. Conclusion.

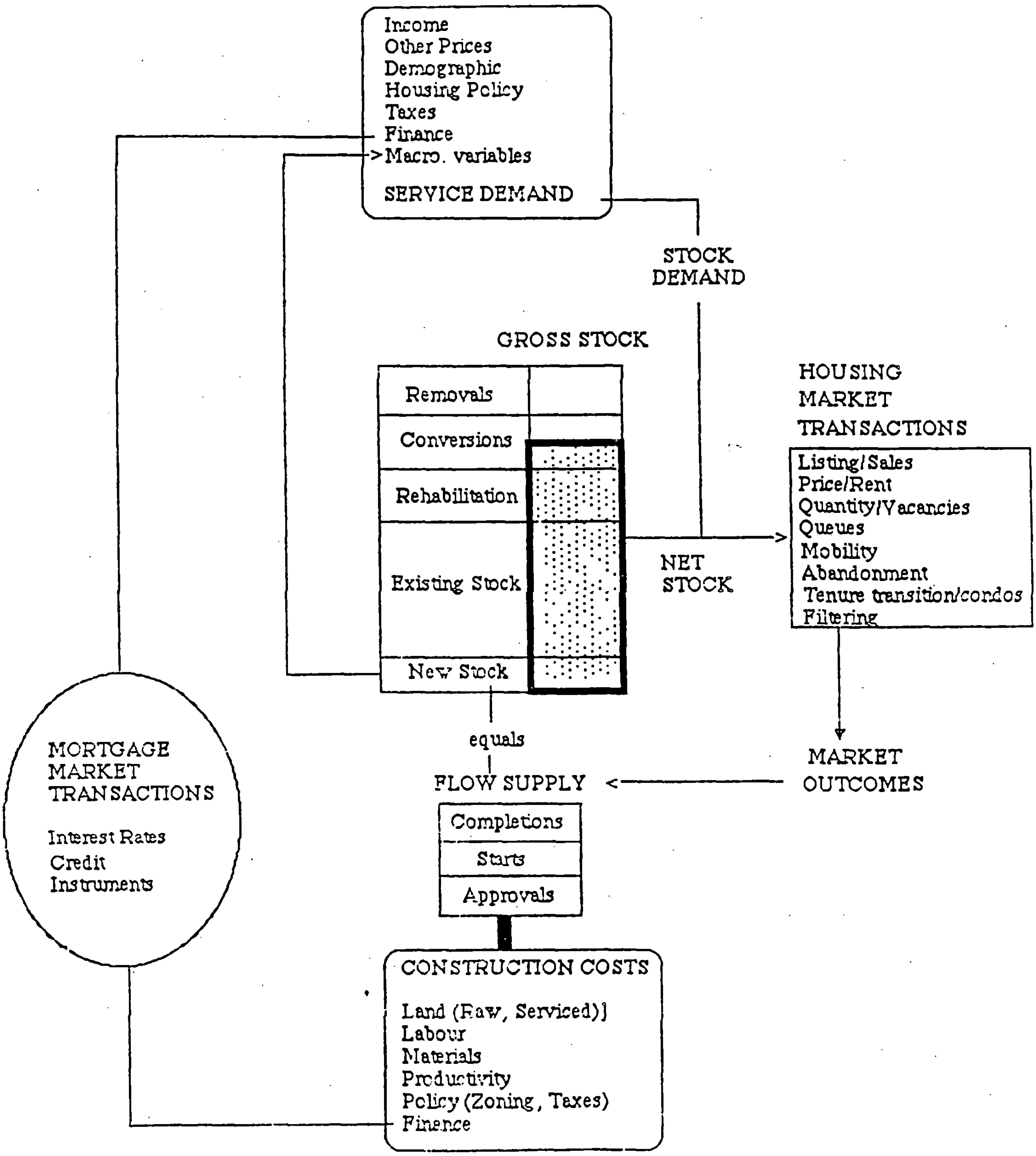
Our examination of data availability reveals that despite some areas of satisfactory data, many variables are seriously deficient in data availability and/or reliability for purposes of econometric modelling. This is especially the case with data that are available annually. The less frequent sources (Census, FAMEX) are more reliable, but can only be used as a check on the other sources.

Data on many key housing and mortgage variables are available on a quarterly or monthly basis, and should provide a basis for constructing a small-scale econometric model of the housing sector and mortgage market. The available data base is clearly inadequate for building a comprehensive quarterly econometric model. Data on many important series including housing stock and several flow variables which affect it–removals, depreciation, additions and alterations–are available only on the annual basis. Even the available annual data are of poor quality. Data availability would certainly constrain the econometric modelling of the housing sector and mortgage market.

The quality of forecasts is affected greatly by the quality and scope of the data used in modelling. It is strongly recommended that CMHC create and maintain a comprehensive data bank of the housing and mortgage market variables. Some series may have to be constructed or estimated from more than one source. For others such as households CMHC may have to start collecting their data.

Figure 2.1

CONCEPTUAL FRAMEWORK OF THE HOUSING MARKET FOR FORECASTING PURPOSES



3. FORECASTING WITH A UNIVARIATE MODEL

A simple but flexible approach to forecasting is to forecast the future values of a variable of interest on the basis of its own past history. One may first fit a time series model to the data and then extrapolate the series into the future using the fitted model. The model is "data based" in that it is specified from its goodness-of-fit to the data series. Unlike the econometric modelling considered in the preceding chapter, economic theory plays little role in model specification.

Of the many classes of time series models which yield extrapolative forecasts, the univariate *autoregressive integrated moving average* (ARIMA) model also referred to as the "Box-Jenkins" model has been found very useful in practice.[2] In this chapter (1) we report on an ARIMA model of Canadian housing starts we have built and (2) discuss the advantages and drawbacks of an ARIMA model as a forecasting tool. Detailed discussions on ARIMA models are available in many books including Box and Jenkins (1970), Vandaele (1983), Granger and Newbold (1986), and Park (1989).

3.1 An ARIMA Model of Housing Starts

A. Multiplicative Seasonal ARIMA Modelling.

The *autoregressive moving average* (ARMA) model of a time series relates the current value of a variable to its own past values and the current and past values of random errors. Thus the ARMA process of order (p,q) is often written as

$$Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + u_t - \theta_1 u_{t-1} - \cdots - \theta_q u_{t-q}, \tag{3.1a}$$

relating Z_t , the value at time t of a time series, to its past values $Z_{t-1}, \ldots, Z_{t-p}$, and the current and past random errors $u_t, u_{t-1}, \ldots, u_{t-q}$. The ϕ_j's and θ_j's are the autoregressive and moving-average parameters, respectively. The u_t's are serially uncorrelated random errors with zero mean and constant variance σ^2, and referred to in time series analysis as "white noise."

For simplicity in notation we introduce a backward shift operator B on the time subscript so that $B\,Z_t = Z_{t-1}$. Thus we may write (3.1a) as

$$Z_t - \phi_1 B Z_t - \cdots - \phi_p B^p Z_t = u_t - \theta_1 B u_t - \cdots - \theta_q B^q u_t \tag{3.1b}$$

or, more compactly, as

$$\phi(B) Z_t = \theta(B) u_t, \tag{3.1c}$$

[2] The literature contains a large number of extrapolative forecasting procedures. Makridakis *et al.* (1982) report on 24 extrapolative time series methods used in a forecasting competition.

where $\phi(B)$ and $\theta(B)$ are polynomials in B of order p and q, respectively:

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

and

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q.$$

The ARMA(p,q) process in (3.1) is assumed to be "(weakly) stationary" in the sense that the mean, variance and autocorrelations do not change over time. The practical consequence of this property is that the mean, variance and autocorrelations of the process can be consistently estimated from a single series of observations.

Common in economic time series like housing starts is a rather smooth trend indicating a slow change in its mean over time. It is often possible to transform such a nonstationary series into a stationary one by differencing it d times:

$$Z_t = (1 - B)^d X_t,$$

where d is a nonnegative integer. If the ARMA(p,q) process Z_t in (3.1) has been obtained by differencing the original X_t series d times, X_t is obtained by summing up the Z_t series d times, and referred to as an *autoregressive integrated moving average* (ARIMA) process of order (p,d,q).

Using the notation for an ARMA model in (3.1), we may define an ARIMA(p,d,q) model by

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d X_t = (1 - \theta_1 B - \cdots - \theta_q B^q) u_t, \tag{3.2}$$

where the u_t's are a white noise process with mean zero and variance σ^2. A slight extension of (3.2) is to allow a constant term δ in it:

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d X_t = \delta + (1 - \theta_1 B - \cdots - \theta_q B^q) u_t, \tag{3.3}$$

Economic time series often displays an exponential trend with gradual change in its spread over time. In such circumstances the logarithmic transformation of the original series may induce stationarity in variance.

Also common in an economic time series is a component called "seasonality." It is any cyclical or periodic variation in a time series that repeats itself with a fixed period. Many monthly or quarterly economic time series have strong seasonal components. The housing starts series has seasonality with the seasonal span of twelve months. Multiplicative ARIMA models are useful in modelling times series in which seasonal variations with a known period of s occur.

If the monthly or quarterly series under consideration trends in annual steps, seasonal differencing may be necessary to induce stationarity in mean. The seasonal differencing

$$X_t - X_{t-s} = (1 - B^s) X_t$$

may have to be applied D times to remove seasonal trend from the series. Regular or nonseasonal differencing of order d, $(1-B)^d$, may be required of the seasonally differenced series to induce stationarity in mean. The ***multiplicative seasonal autoregressive integrated moving average*** model of order p, d, q, P, D, Q, denoted by ARIMA $(p, d, q) \times (P, D, Q)_s$, is written as

$$\phi(B)\Phi(B)(1-B)^d(1-B^s)^D X_t = \theta(B)\Theta(B^s)u_t, \tag{3.4}$$

where $\Phi(B^s)$ and $\Theta(B^s)$ are polynomials in B^s of degree P and Q, respectively, defined by

$$\Phi(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_p B^{Ps}$$

and

$$\Theta(B^s) = 1 - \Theta_1 B^s - \cdots - \Theta_q B^{Qs},$$

and the polynomials $\phi(B)$ and $\theta(B)$ are as defined in (3.1). The ARIMA model in (3.4) is the type that we have built for Canadian housing starts.[3]

The time series plotted in Figure 3.1(a) is the monthly housing starts in Canadian urban centres of 10,000 or more (CANSIM series D849795) from January 1965 to December 1989. The data are in total number of units and unadjusted for seasonal variations. A visual inspection of the time series plot indicates that it has a slowly changing trend in mean and variance as well as a strong seasonal component indicated by peaks and valleys at multiples of 12 months. We have taken the common logarithms of the original series to induce stationarity in variance. The "logged" series are plotted in Figure 3.1(b).

At the stage of identifying an ARIMA model we have used various data analytic tools to arrive at initial guesses of the data transformation, the degrees of differencing to induce stationarity, and the orders of the AR and MA polynomials in the model. Data analytic tools for identification include time series plots, sample autocorrelation function, sample partial autocorrelation function, extended sample autocorrelation function [Tsay and Tiao (1984)], and inverse autocorrelations [Cleveland (1972)].

A model selection procedure based on these tools has been found generally effective particularly when supplemented by the model selection criteria such as Akaike's (1969, 1970) final prediction error (FPE) criterion and Akaike's (1974) information criterion (AIC). Since the ARIMA model is specified on the basis of data alone, however, the model builder often has to resort to his judgment. A brief review of various model selection techniques in time series analysis is given by Shibata (1985).

We have decided to fit an ARIMA model to the 288 observations over the period January 1965 to December 1988, leaving 1989 as a post-sample period to be used for checking the forecasting ability of the fitted model. The sample ACF and PACF of the transformed

[3]Earlier univariate ARIMA modelling applied to the housing market includes the textbook examples in Abraham and Ledolter (1983), Pankratz (1983), and Vandaele (1983). More recently, Sklarz *et al.* (1987) and Puri and Van Lierop (1988) have built ARIMA models to the U.S. monthly housing starts series and found their forecasting performance satisfactory.

series shown in Figures 3.2(a) and (b) indicate nonstationarity in mean. A very large and significant partial autocorrelation at lag 1 and slowly declining autocorrelations at seasonal lags of 12, 24, and 36 suggest differencing of the transformed series both regularly and seasonally once each. This has reduced the total number of observations to $T = 275$.

The sample ACF and PACF of the differenced series in Figure 3.3 suggest a multiplicative ARIMA$(1,1,1) \times (0,1,1)_{12}$ or ARIMA$(1,1,1) \times (3,1,0)_{12}$ model. We write these tentative models as Model 1:

$$(1 - \phi_1 B)(1 - B)(1 - B^{12})LOGHS_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})u_t$$

and Model 2:

$$(1-\phi_1 B)(1-\Phi_1 B^{12}-\Phi_2 B^{24}-\Phi_3 B^{36})(1-B)(1-B^{12})LOGHS_t = (1-\theta_1 B)(1-\Theta_1 B^{12})u_t,$$

where $LOGHS_t$ is the common logarithm of the original housing starts at time t.

Once an ARIMA$(p,d,q) \times (P,D,Q)_s$ model is identified, its parameters can be estimated by the conditional least squares (LS) method. Under the normality assumption of the random errors, one can also obtain *approximate* maximum likelihood (ML) estimates. The approximation arises because of the assumptions to be made about the initial values of the X_t's and u_t's. See Box and Jenkins (1970). *Exact* ML estimates can be computed in a number of ways, suggested by Newbold (1974), Dent (1977), Ansley (1979), and others. Although these estimators are asymptotically equivalent in the sense that they have the same probability distributions in large samples, their estimates can be considerably different from one another.

ML estimates for the tentative Model 1 are summarized in Table 3.1(A).[4] The AR coefficient estimate is significant and satisfies the stationarity. On the other hand, both MA coefficients are significant and satisfy the "invertibility" conditions although $\hat{\theta}_1 = .9192$ is close to the noninvertibility boundary.[5]

No model is "true." Estimated models should be thoroughly checked for adequacy. Tools for diagnostic checking include the residual autocorrelations and partial autocorrelations, the Box-Pierce-Ljung portmanteau statistic for lack of fit and Lagrange multiplier tests against a variety of specific alternatives. A survey of diagnostic tests used in both univariate and multivariate time series modelling is given by Newbold (1983b). Although the Box-Pierce-Ljung statistic was consistent with the hypothesis that the residuals series from the fitted model were like a white noise process, their sample ACF and PACF had a significant coefficient at lag 36 and indicated that Model 2 was perhaps a better specification.

[4] The computing reported in Chapters 3 and 4 are the results from the SCA-UTS program of the Scientific Computing Associates. Their estimation method is based on a conditional maximum likelihood approach. See Liu *et al.* (1986).

[5] An ARMA model is said to be invertible if the series of coefficients in its pure AR representation converges as the time lag goes to infinity. In Model 1 the invertibility conditions are $-1 < \theta < 1$ and $-1 < \Theta < 1$.

Parameter estimates for the tentative Model 2 were

$$\begin{array}{ll} \hat{\phi}_1 = 0.6996 & \text{with } t = 8.60 \\ \hat{\Phi}_1 = -.1621 & \text{with } t = -2.04 \\ \hat{\Phi}_2 = -.0661 & \text{with } t = -0.90 \\ \hat{\Phi}_3 = -.2503 & \text{with } t = -3.85 \\ \hat{\theta}_1 = 0.9003 & \text{with } t = 18.19 \\ \hat{\Theta}_1 = 0.6850 & \text{with } t = 9.82, \end{array}$$

and SEE = .0658. We note that seasonal AR(2) parameter estimate is not significant at $\alpha = .05$.

An additional check for model adequacy was made by imposing constraints. Estimating an ARIMA$(1,1,1) \times (3,1,1)_{12}$ model with no constant and a constraint $\Phi_2 = 0$, we found the SAR(1) parameter estimate to be statistically insignificant. On the principle of parsimony, we chose the ARIMA$(1,1,1) \times (3,1,1)_{12}$ model with constraints that $\Phi_1 = \Phi_2 = 0$:

$$(1 - \phi_1 B)(1 - \Phi_3 B^{36})(1 - B)(1 - B^{12}) LOGHS_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12}) u_t. \quad (3.5)$$

The results of estimating the model in (3.5) are presented in Table 3.1 (B). All coefficient estimates are significant, satisfy the stationarity and invertibility conditions, and the residual standard error is reduced from SEE=.0676 to SEE=.0656. The fitted model is then

$$(1-.7232B)(1+.2119B^{36})(1-B)(1-B^{12}) LOGHS_t = (1-.9174B)(1-.7773B^{12}) u_t. \quad (3.6)$$

The sample ACF and PACF for the residuals of the estimated model (3.6) are shown in Figure 3.4. Autocorrelations and partial autocorrelations remain significant at lag 34. However, the Box-Pierce-Ljung statistics are $Q(12) = 4.9, Q(24) = 12.2$, and $Q(36) = 25.2$, and strongly support the hypothesis that the residuals behave like white noise. The estimated model in (3.6) appears to be adequate.

One may note that the estimate of non-seasonal MA(1) parameter is close to the boundary of the non-invertibility region. In general, many economic time series only tend to become stationary after differencing, and differencing often results in an MA component in which the parameters are very close to the boundary of the non-invertibility region. An ARIMA model in which some of the roots of the MA polynomial are equal to one is perfectly valid and it may not be necessarily sensible to require that the estimates of the MA parameters satisfy the invertibility conditions.

3.2 Forecasting with an ARIMA Model

Once we have built an adequate ARIMA model, we can derive forecasts from it in a mechanical way. Although the mechanics of obtaining forecasts are described below with the basic ARMA model, the same principles apply in dealing with the multiplicative seasonal ARIMA model.

Assume for the time being that the model parameters are known and that we require at time T an h-period ahead forecast of an ARMA process Z_t . If the model holds in the future, then

$$Z_{T+h} = \phi_1 Z_{T+h-1} + \cdots + \phi_p Z_{T+h-p} + u_{T+h} - \theta_1 u_{T+h-1} - \cdots - \theta_q u_{T+h-q} \tag{3.7}$$

At the time T when we make forecasts we may have a set of available information, called the "information set," that can be used for forecasting Z_{T+h}. For example, the information set, denoted $\mathcal{I}_T$, may consist of the full history of the series up to time T: $Z_T, Z_{T-1}, \ldots$.

If the white noise process u_t is normal, the "*optimal*" (in the minimum mean square error sense) forecast at time T of Z_{T+h} conditional on the information set $\mathcal{I}_T$ is simply its conditional expectation

$$f_T(h) = E_T[Z_{T+h}|\mathcal{I}_T] \tag{3.8}$$

where $f_T(h)$ stands for the h-period ahead forecast at time T and $E_T[\cdot]$ denotes the conditional expectation taken at time T. Taking the conditional expectation of (3.7), we write:

$$\begin{aligned} f_T(h) = E_T(Z_{T+h}) &= \phi_1 E_T(Z_{T+h-1}) + \cdots + \phi_p E_T(Z_{T+h-p}) \\ &+ E_T(u_{T+h}) - \theta_1 E_T(u_{T+h-1}) - \cdots - \theta_q E_T(u_{T+h-q}) \end{aligned} \tag{3.9}$$

The conditional expectation of the future Z_t given $\mathcal{I}_T$ is its forecast value while that of the current or past Z_t is equal to its realized value. On the other hand, since the u_t's are serially independent, the conditional expectation of the future u_t is zero while that of the current or past u_t is equal to its realized value.

It is easily seen that the h-step ahead forecast error, $e_T(h)$, is given by

$$\begin{aligned} e_T(h) &= Z_{T+h} - f_T(h) \\ &= (1 + \psi_1 B + \cdots + \psi_{h-1} B^{h-1}) u_{T+h}, \end{aligned} \tag{3.10}$$

where the ψ's are the coefficients obtained when we express the process Z_t in terms of the current and past random errors alone. Thus the h-step ahead forecast error has mean zero and variance

$$Var[e_T(h)] = (1 + \psi_1^2 + \cdots + \psi_{h-1}^2)\sigma^2. \tag{3.11}$$

Unlike the forecasts based on an econometric model, we can easily compute the standard errors of ARIMA forecasts so that a precision measure is attached with point forecasts. If the u_t process is normal, then $e_T(h)$ is normal with mean zero and variance given in

(3.11), and provides a basis for forecast intervals. For example, a 95% forecast interval is given by

$$f_T(h) \pm 1.96[(1 + \psi_1^2 + \cdots + \psi_{h-1}^2)\,\sigma^2]^{1/2}. \tag{3.12}$$

The above discussions on forecasting have been made with an assumption that the full history of the ARIMA process is known. In practice, its full history is not available. That the available data set consists only of T observations, $X_1, \ldots, X_T$ poses two problems. First, forecast errors $u_1, u_2, \ldots, u_T$ are no longer available and have to be estimated. If we build up the sequence of forecasts from the beginning of the data series and if the ARIMA process under consideration is invertible, forecast errors arising from estimated u_t series should be negligibly small. Second, since the full history of the series is not available, the model parameters are unknown and have to be estimated. Thus, the conditional expectations and the forecast variance have to be estimated. Ambiguity in model specification and errors in estimation make forecasts biased and the true value of forecast variance underestimated. In general, however, identification and estimation of an ARIMA model becomes more accurate as the length of the time series to analyze increases.

The fitted model (3.6) was used to forecast the housing starts series in 1989. Multiplying out (3.6) and rearranging terms, we can write the h-step ahead forecast function at origin T as

$$\begin{aligned} E_T(LOGHS_{T+h}) = {} & 1.7232E_T(LOGHS_{T+h-1}) - .7232E_T(LOGHS_{T+h-2}) \\ & + E_T(LOGHS_{T+h-12}) - 1.7232E_T(LOGHS_{T+h-13}) \\ & + .7232E_T(LOGHS_{T+h-14}) - .2119E_T(LOGHS_{T+h-36}) \\ & + .3651E_T(LOGHS_{T+h-37}) - .1532E_T(LOGHS_{T+h-38}) \\ & + .2119E_T(LOGHS_{T+h-48}) - .3651E_T(LOGHS_{T+h-49}) \\ & + .1532E_T(LOGHS_{T+h-50}) + E_T(u_{T+h}) \\ & - .9172E_T(u_{T+h-1}) - .7773E_T(u_{T+h-12}) \\ & + .7129E_T(u_{T+h-13}) \end{aligned} \tag{3.13}$$

in terms of conditional expectations. If we want the forecasts in the original units rather than their logged values, we have only to take the anti-logarithm of the forecasts of the $LOGHS_t$ series.

Table 3.2 presents a 12-month forecast profile together with a set of 95% confidence forecast intervals at the origin December 1988 for the 12 months of 1989. Since they have been generated from the model in (3.6) based on the 1965-1988 data, they are strictly post-sample forecasts. Figure 3.5 contains the plots of these forecasts. It is remarkable that even with all the changes and fluctuations in the housing starts in 1989, the 12-month forecast profile traces the actual values in 1989 remarkably well. Further, the forecasts did not deteriorate at all with the forecast horizon.

Actual forecast errors ranged in absolute value from a mere .081 % for the 10-month-ahead forecast for October 1989 to 1.737% for the 1-month-ahead forecast for January

1989. For the sample period January 1965 to December 1988 the standard deviation of the transformed series was .1584 while that of the residual series was .0656. When the series to forecast contains a large noise component, the forecasts are judged to be remarkably accurate. We conclude that the ARIMA model in (3.6) fits the housing starts data of Canada remarkably well.

Since 1974 Nelson (1984) has been forecasting three major U.S.macroeconomic variables including GNP using ARIMA models and has been providing benchmark forecasts for comparison with forecasts from large-scale econometric models. McNees (1988) finds that although Nelson's forecasts were not judgementally adjusted at all, their performance was comparable to those judgementally adjusted forecasts of the prominent commercial forecasters.

It is widely recognized that simple ARIMA models perform as well in short term forecasting as those judgementally adjusted forecasts based on large elaborate econometric models. We have also found that the ARIMA model which we have built in the present study forecasts Canadian housing starts remarkably well up to the twelve-month horizon.

As one can readily see from the preceding discussion of an ARIMA modelling and forecasting, its strengths as a forecasting tool include:

(1) It generates very accurate forecasts at least for the short run;
(2) It provides a remarkably simple method of obtaining forecasts;
(3) It is possible to build an ARIMA model using data up to any point in the past and then produce forecasts so that the forecasting accuracy of the model can be easily checked;
(4) Dynamics of the variables are identified more precisely;
(5) Error structures are more readily identified and incorporated into modelling;
(6) Seasonality in time series is brought into modelling in a systematic manner; and
(7) The standard errors as a measure of forecast accuracy can be calculated at each forecast horizon.

It is not, therefore, surprising that an ARIMA model does perform remarkably well for forecasting, at least in the short run or when economic activity is relatively stable.

On the other hand, the univariate ARIMA forecasting method has the following weaknesses:

(1) Forecast accuracy may fall off quickly as the forecast horizon increases;
(2) Nonlinearities of the economy may not be easily picked up by the ARIMA model; and
(3) It may not do well in times of structural change or when economic activity is unstable due to exogenous shocks such as wars or unexpected policy changes.

Forecasting ability of an ARIMA or other time series models comes from the dependence of successive observations in a time series. As the forecast horizon increases, the degree of correlation between the variable to be forecast and the observed data series tends to decrease, and the forecasting accuracy of an ARIMA model deteriorates over longer forecasting horizons.

A sudden change in structure in times of exogenous shocks such as a labour dispute or oil embargo is the most difficult type of nonstationarity to handle in time series modelling. The change may produce a short term transient effect or a long term change in the model structure. Hillmer (1984) shows how one outlier can affect several consecutive forecasts unless adjusted. Change, Tiao and Chen (1988) discuss ARIMA modelling in the presence of outliers. A variation of the ARIMA model, put forward by Box and Tiao (1975) and called "intervention analysis," can be used when certain known events have affected the time series being forecast.

It is clear that ARIMA models provide a very simple framework for forecasting time series in which the specification of a model is based on the data alone. This may be a useful methodology particularly when it is difficult to identify the main factors determining the variable to forecast and model their relationship.

Table 3.1

(A) SUMMARY OF A UNIVARIATE MODEL OF HOUSING STARTS
(In Common Logarithm)

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
LOGHS	RANDOM	ORIGINAL	$(1-B^{1})(1-B^{12})$

PARAMETER LABEL		VARIABLE NAME	VALUE	STD ERROR	T VALUE
1	THETA1	LOGHS	.9192	.0370	24.83
2	THETA12	LOGHS	.8462	.0340	24.91
3	PHI1	LOGHS	.7118	.0670	10.63

TOTAL SUM OF SQUARES	.731151E+01
TOTAL NUMBER OF OBSERVATIONS	288
RESIDUAL SUM OF SQUARES.	.124591E+01
R-SQUARE	.821
EFFECTIVE NUMBER OF OBSERVATIONS . .	274
RESIDUAL VARIANCE ESTIMATE	.454712E-02
RESIDUAL STANDARD ERROR.	.674323E-01

(b) SUMMARY OF A UNIVARIATE MODEL OF HOUSING STARTS
(In Common Logarithm)

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
LOGHS	RANDOM	ORIGINAL	$(1-B^{1})(1-B^{12})$

PARAMETER LABEL		VARIABLE NAME	VALUE	STD ERROR	T VALUE
1	THETA1	LOGHS	.9174	.0421	21.78
2	THETA12	LOGHS	.7773	.0435	17.86
3	PHI1	LOGHS	.7232	.0734	9.86
4	PHI36	LOGHS	-.2119	.0621	-3.42

TOTAL SUM OF SQUARES	.731151E+01
TOTAL NUMBER OF OBSERVATIONS	288
RESIDUAL SUM OF SQUARES.	.102034E+01
R-SQUARE	.831
EFFECTIVE NUMBER OF OBSERVATIONS . .	238
RESIDUAL VARIANCE ESTIMATE	.428716E-02
RESIDUAL STANDARD ERROR.	.654764E-01

Table 3.2

FORECAST PROFILE WITH 12-MONTH HORIZON
(In Common Logarithm)

MONTH 1989	LOWER LIMIT	FORECAST VALUE	UPPER LIMIT	ACTUAL VALUE	FORECAST ERROR	PERCENTAGE ERROR
Jan	3.936	4.064	4.192	4.136	.072	1.737
Feb	3.786	3.951	4.116	3.990	.039	.982
Mar	3.846	4.032	4.218	4.085	.053	1.290
Apr	3.997	4.196	4.395	4.240	.044	1.046
May	4.094	4.303	4.512	4.282	-.021	-.489
Jun	4.057	4.273	4.489	4.240	-.033	-.779
Jul	4.000	4.222	4.444	4.215	-.007	-.156
Aug	3.998	4.225	4.453	4.190	-.036	-.850
Sep	3.959	4.191	4.423	4.213	.022	.531
Oct	3.959	4.194	4.430	4.198	.003	.081
Nov	3.982	4.221	4.461	4.193	-.028	-.662
Dec	3.918	4.161	4.404	4.140	-.021	-.515

Note: Forecasts are for the 12 months in 1989 based on the data upto and including December 1988.

Figure 3.1 (A) Housing Starts in Level

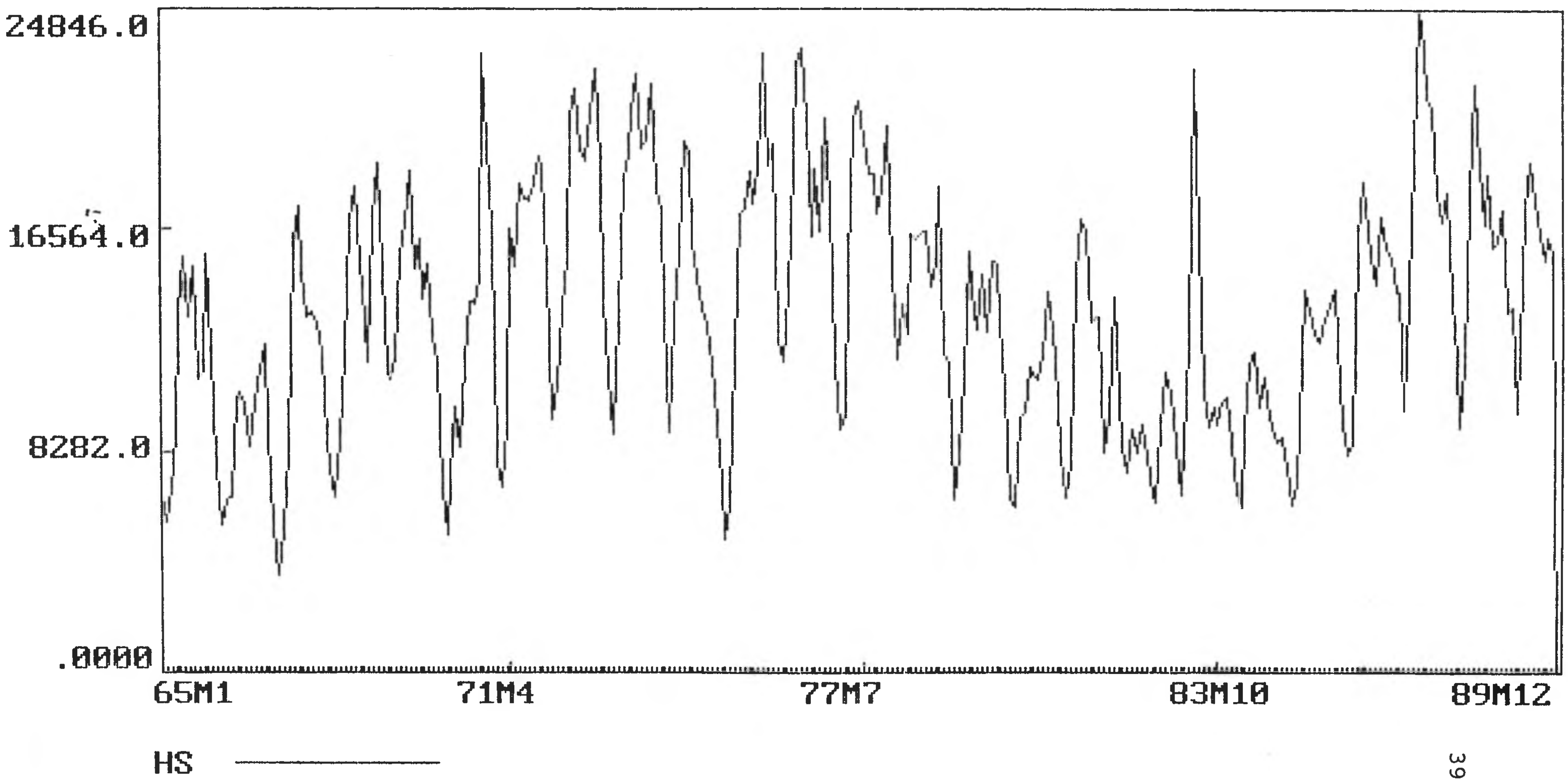


Figure 3.1 (B) Housing Starts in Logarithm

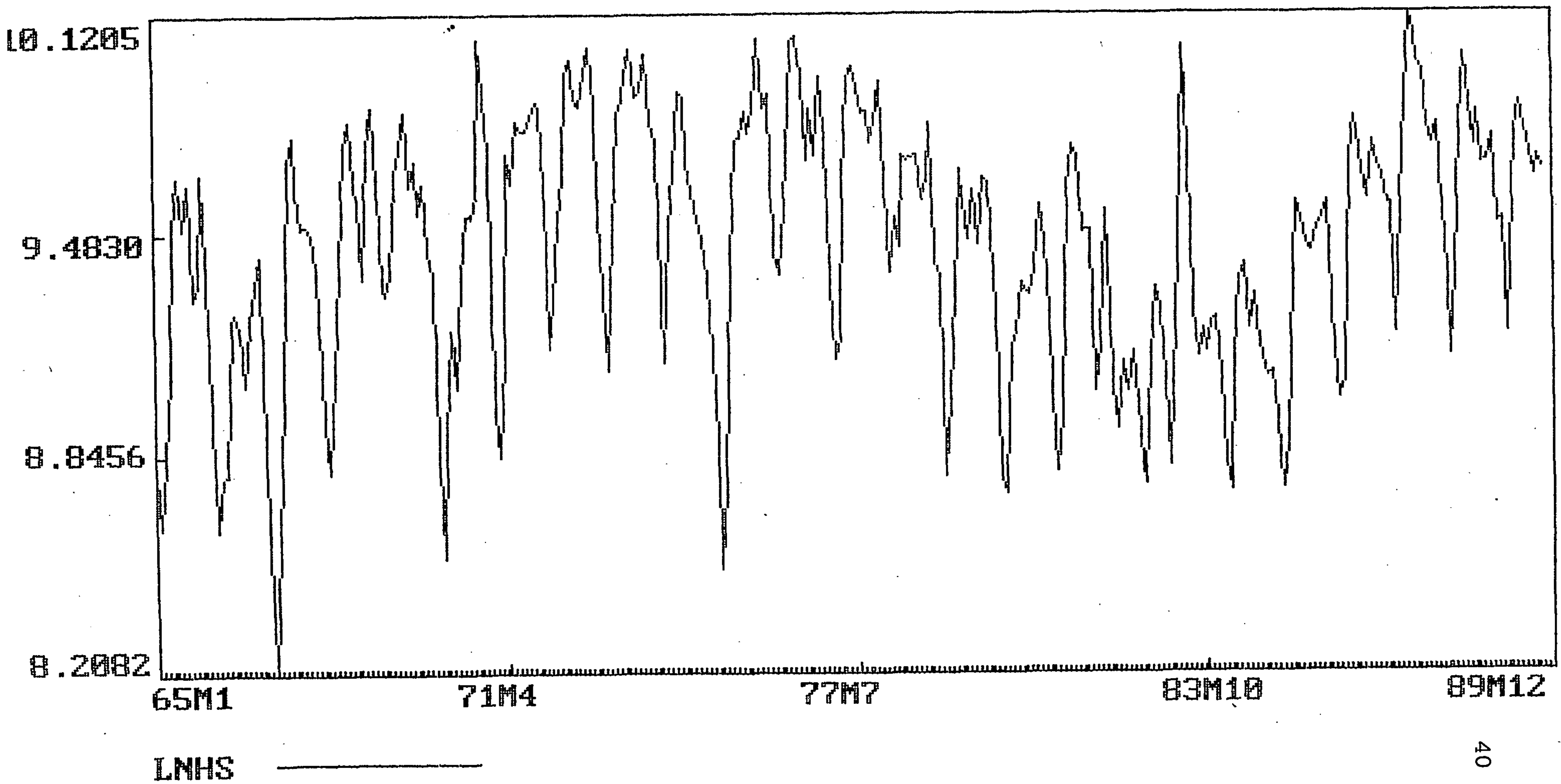


Figure 3.2 (A)

SAMPLE AUTOCORRELATIONS OF HOUSING STARTS
(In Common Logarithm)

TIME PERIOD ANALYZED	1 TO 288
NAME OF THE SERIES	LOGHS
EFFECTIVE NUMBER OF OBSERVATIONS . . .	288
STANDARD DEVIATION OF THE SERIES . . .	.1593
MEAN OF THE (DIFFERENCED) SERIES . . .	4.1014
STANDARD DEVIATION OF THE MEAN	.0094
T-VALUE OF MEAN (AGAINST ZERO)	436.8426

1-12	.79	.47	.22	.13	.14	.15	.11	.06	.09	.28	.52	.62
ST.E.	.06	.09	.10	.10	.10	.10	.10	.10	.10	.10	.10	.11
Q	181	245	259	264	270	276	280	281	283	307	389	504
13-24	.47	.22	.02	-.03	.01	.03	-.00	-.03	.01	.20	.44	.54
ST.E.	.12	.13	.13	.13	.13	.13	.13	.13	.13	.13	.13	.14
Q	572	587	587	587	587	587	587	588	588	601	662	753
25-36	.40	.13	-.07	-.12	-.08	-.07	-.11	-.16	-.12	.05	.27	.36
ST.E.	.14	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15	.15
Q	803	808	810	815	817	819	823	831	835	836	861	904

```
          -1.0  -.8  -.6  -.4  -.2  .0   .2   .4   .6   .8  1.0
            +----+----+----+----+----+----+----+----+----+----+
  1     .79                       +  IXX+XXXXXXXXXXXXXXXXX
  2     .47                      +   IXXX+XXXXXXXX
  3     .22                     +    IXXXXX
  4     .13                     +    IXXX +
  5     .14                     +    IXXX +
  6     .15                     +    IXXXX+
  7     .11                     +    IXXX +
  8     .06                     +    IX   +
  9     .09                     +    IXX  +
 10     .28                     +    IXXXX+XX
 11     .52                     +    IXXXX+XXXXXXXX
 12     .62                    +     IXXXXX+XXXXXXXXX
 13     .47                    +     IXXXXX+XXXXXX
 14     .22                    +     IXXXXX+
 15     .02                    +     IX    +
 16    -.03                    +    XI     +
 17     .01                    +     I     +
 18     .03                    +     IX    +
 19     .00                    +     I     +
 20    -.03                    +    XI     +
 21     .01                    +     I     +
 22     .20                    +     IXXXXX+
 23     .44                    +     IXXXXX+XXXXX
 24     .54                   +      IXXXXXX+XXXXXX
```

Figure 3.2 (B)

SAMPLE PARTIAL AUTOCORRELATIONS OF HOUSING STARTS
(In Common Logarithm)

1-12	.79	-.41	.05	.18	.01	-.03	-.09	.10	.21	.38	.21	-.05
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.28	-.10	.01	.04	-.01	-.07	.06	.05	.01	.13	.17	-.02
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	-.17	-.16	.00	.01	-.02	-.08	-.03	.01	-.04	-.04	.12	-.02
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

```
            -1.0 -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1    .79                         +  IXX+XXXXXXXXXXXXXXXXX
   2   -.41                  XXXXXXX+XXI  +
   3    .05                         +  IX +
   4    .18                         +  IXX+XX
   5    .01                         +  I  +
   6   -.03                         + XI  +
   7   -.09                         +XXI  +
   8    .10                         +  IXX+
   9    .21                         +  IXX+XX
  10    .38                         +  IXX+XXXXXX
  11    .21                         +  IXX+XX
  12   -.05                         + XI  +
  13   -.28                     XXXX+XXI  +
  14   -.10                         +XXI  +
  15    .01                         +  I  +
  16    .04                         +  IX +
  17   -.01                         +  I  +
  18   -.07                         +XXI  +
  19    .06                         +  IXX+
  20    .05                         +  IX +
  21    .01                         +  I  +
  22    .13                         +  IXXX
  23    .17                         +  IXX+X
  24   -.02                         +  I  +
  25   -.17                        X+XXI  +
  26   -.16                        X+XXI  +
  27    .00                         +  I  +
  28    .01                         +  I  +
  29   -.02                         + XI  +
  30   -.08                         +XXI  +
  31   -.03                         + XI  +
  32    .01                         +  I  +
  33   -.04                         + XI  +
  34   -.04                         + XI  +
  35    .12                         +  IXXX
  36   -.02                         + XI  +
```

Figure 3.3 (A)

SAMPLE AUTOCORRELATIONS OF TRANSFORMED HOUSING STARTS
(In Common Logarithm, d = D = 1)

```
                                              1       12
DIFFERENCE ORDERS. . . . . . . . . . . . . (1-B  ) (1-B   )
TIME PERIOD ANALYZED . . . . . . . . . . .  1  TO    288
NAME OF THE SERIES . . . . . . . . . . . .          LOGHS
EFFECTIVE NUMBER OF OBSERVATIONS . . .               275
STANDARD DEVIATION OF THE SERIES . . .             .0916
MEAN OF THE (DIFFERENCED) SERIES . . .            -.0003
STANDARD DEVIATION OF THE MEAN . . . .             .0055
T-VALUE OF MEAN (AGAINST ZERO) . . . .            -.0589
```

1-12	-.11	-.07	-.04	-.04	-.12	.03	.05	-.00	.00	.10	.05	-.50
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
Q	3.6	5.1	5.5	6.0	10.2	10.4	11.1	11.1	11.1	14.1	14.8	88.0
13-24	-.05	.04	.03	.01	.08	-.05	-.03	.07	.01	-.00	.00	.13
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
Q	88.6	89.1	89.3	89.4	91.2	92.1	92.3	94.0	94.0	94.0	94.0	99.1
25-36	.04	-.02	-.07	-.01	-.00	.06	-.01	.01	-.02	-.12	.06	-.24
ST.E.	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08	.08
Q	99.5	99.6	101	101	101	102	102	102	102	107	108	126

```
        -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
           +----+----+----+----+----+----+----+----+----+----+
 1   -.11                        XXXI  +
 2   -.07                        +XXI  +
 3   -.04                        + XI  +
 4   -.04                        + XI  +
 5   -.12                        XXXI  +
 6    .03                        +  IX +
 7    .05                        +  IX +
 8    .00                        +  I  +
 9    .00                        +  I  +
10    .10                        +  IXXX
11    .05                        +  IX +
12   -.50              XXXXXXXXXX+XXI  +
13   -.05                       +  XI   +
14    .04                       +   IX  +
15    .03                       +   IX  +
16    .01                       +   I   +
17    .08                       +   IXX +
18   -.05                       +  XI   +
19   -.03                       +  XI   +
20    .07                       +   IXX +
21    .01                       +   I   +
22    .00                       +   I   +
23    .00                       +   I   +
24    .13                       +   IXXX+
```

Figure 3.3 (B)

SAMPLE PARTIAL AUTOCORRELATIONS OF TRANSFORMED HOUSING STARTS
(In Common Logarithm, d = D = 1)

1-12	-.11	-.09	-.06	-.06	-.15	-.02	.02	-.01	-.00	.09	.09	-.49
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.21	-.07	-.03	-.07	-.10	-.09	-.05	.05	-.01	.10	.10	-.12
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	-.09	-.04	-.08	-.07	-.02	-.03	-.08	.06	.03	-.09	.12	-.33
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

```
             -1.0  -.8  -.6  -.4  -.2  .0   .2   .4   .6   .8 1.0
              +----+----+----+----+----+----+----+----+----+----+
   1   -.11                         XXXI  +
   2   -.09                         +XXI  +
   3   -.06                         + XI  +
   4   -.06                         +XXI  +
   5   -.15                        X+XXI  +
   6   -.02                         +  I  +
   7    .02                         +  IX +
   8   -.01                         +  I  +
   9    .00                         +  I  +
  10    .09                         +  IXX+
  11    .09                         +  IXX+
  12   -.49                XXXXXXXXX+XXI  +
  13   -.21                       XX+XXI  +
  14   -.07                         +XXI  +
  15   -.03                         + XI  +
  16   -.07                         +XXI  +
  17   -.10                         +XXI  +
  18   -.09                         +XXI  +
  19   -.05                         + XI  +
  20    .05                         +  IX +
  21   -.01                         +  I  +
  22    .10                         +  IXXX
  23    .10                         +  IXX+
  24   -.12                         XXXI  +
  25   -.09                         +XXI  +
  26   -.04                         + XI  +
  27   -.08                         +XXI  +
  28   -.07                         +XXI  +
  29   -.02                         + XI  +
  30   -.03                         + XI  +
  31   -.08                         +XXI  +
  32    .06                         +  IX +
  33    .03                         +  IX +
  34   -.09                         +XXI  +
  35    .12                         +  IXXX
  36   -.33                    XXXXX+XXI  +
```

Figure 3.4 (A)

SAMPLE AUTOCORRELATIONS OF RESIDUALS

```
TIME PERIOD ANALYZED . . . . . . . . . . 51  TO    288
NAME OF THE SERIES . . . . . . . . . . .           RY
EFFECTIVE NUMBER OF OBSERVATIONS . . .            238
STANDARD DEVIATION OF THE SERIES . . .          .0642
MEAN OF THE (DIFFERENCED) SERIES . . .          .0000
STANDARD DEVIATION OF THE MEAN . . . .          .0042
T-VALUE OF MEAN (AGAINST ZERO) . . . .         -.0094
```

1-12	-.01	.00	.04	.03	-.04	.04	.03	-.01	-.07	.04	.04	-.07
ST.E.	06	.06	.06	.06	.06	.07	.07	.07	.07	.07	.07	.07
Q	.0	.0	.4	.6	.9	1.3	1.5	1.5	2.8	3.2	3.6	4.9
13-24	-.06	.04	.01	.01	-.01	-.09	-.08	.07	-.03	.01	.04	.03
ST.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
Q	5.9	6.2	6.3	6.3	6.3	8.5	10.2	11.3	11.6	11.6	12.0	12.2
25-36	.07	-.02	-.05	.02	-.02	.03	-.05	.01	-.06	-.16	.05	-.03
ST.E.	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07	.07
Q	13.5	13.6	14.3	14.4	14.4	14.7	15.5	15.5	16.6	24.2	24.9	25.2

```
            -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1   -.01                           +  I  +
   2    .00                           +  I  +
   3    .04                           +  IX +
   4    .03                           +  IX +
   5   -.04                           + XI  +
   6    .04                           +  IX +
   7    .03                           +  IX +
   8   -.01                           +  I  +
   9   -.07                           +XXI  +
  10    .04                           +  IX +
  11    .04                           +  IX +
  12   -.07                           +XXI  +
  13   -.06                           +XXI  +
  14    .04                           +  IX +
  15    .01                           +  I  +
  16    .01                           +  I  +
  17   -.01                           +  I  +
  18   -.09                           +XXI  +
  19   -.08                           +XXI  +
  20    .07                           +  IXX+
  21   -.03                           + XI  +
  22    .01                           +  I  +
  23    .04                           +  IX +
  24    .03                           +  IX +
```

Figure 3.4 (B)

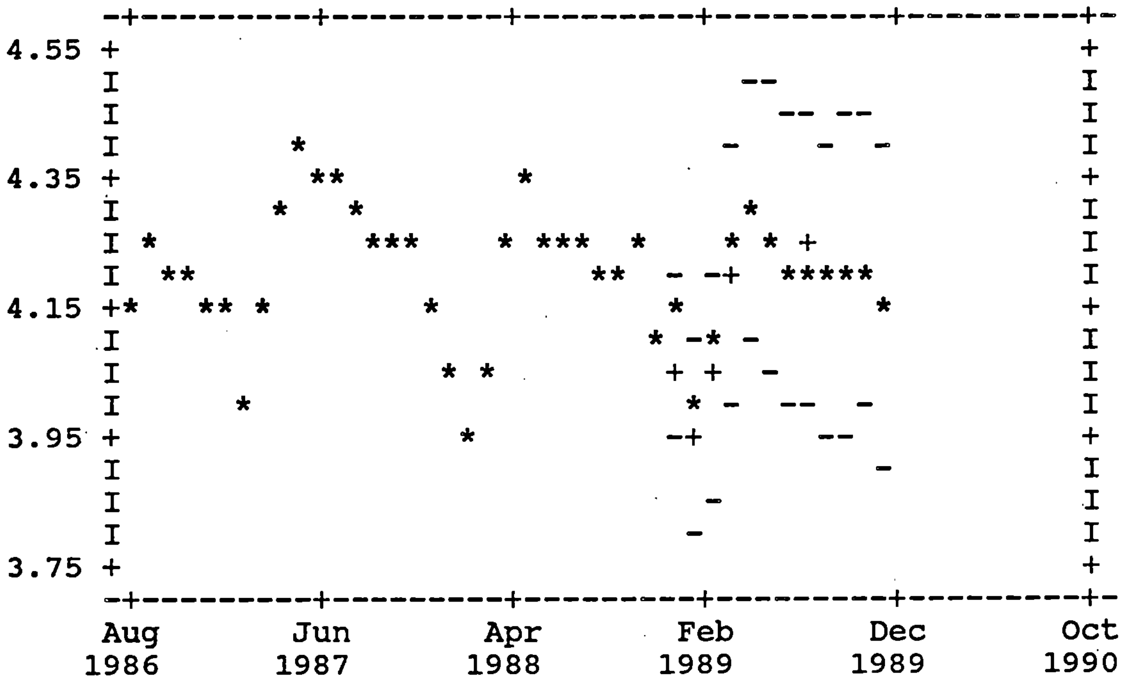
SAMPLE PARTIAL AUTOCORRELATIONS OF RESIDUALS

1-12	-.01	.00	.04	.03	-.04	.04	.02	-.01	-.07	.03	.04	-.07
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
13-24	-.07	.03	.03	.02	-.02	-.10	-.07	.08	-.04	.01	.05	.04
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06
25-36	.08	-.03	-.07	.02	.01	.02	-.09	.02	-.04	-.17	.05	-.05
ST.E.	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06	.06

```
            -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
              +----+----+----+----+----+----+----+----+----+----+
  1  -.01                           +  I  +
  2   .00                           +  I  +
  3   .04                           +  IX +
  4   .03                           +  IX +
  5  -.04                           + XI  +
  6   .04                           +  IX +
  7   .02                           +  IX +
  8  -.01                           +  I  +
  9  -.07                           +XXI  +
 10   .03                           +  IX +
 11   .04                           +  IX +
 12  -.07                           +XXI  +
 13  -.07                           +XXI  +
 14   .03                           +  IX +
 15   .03                           +  IX +
 16   .02                           +  IX +
 17  -.02                           +  I  +
 18  -.10                           XXXI  +
 19  -.07                           +XXI  +
 20   .08                           +  IXX+
 21  -.04                           + XI  +
 22   .01                           +  I  +
 23   .05                           +  IX +
 24   .04                           +  IX +
 25   .08                           +  IXX+
 26  -.03                           + XI  +
 27  -.07                           +XXI  +
 28   .02                           +  IX +
 29   .01                           +  I  +
 30   .02                           +  I  +
 31  -.09                           +XXI  +
 32   .02                           +  I  +
 33  -.04                           + XI  +
 34  -.17                          X+XXI  +
 35   .05                           +  IX +
 36  -.05                           + XI  +
```

Figure 3.5

FORECAST PROFILE WITH 12-MONTH HORIZON



Legend: -: upper limit for a 95% interval forecast
-: lower limit for a 95% interval forecast
+: point forecast
*: actual value

4. FORECASTING WITH A TRANSFER FUNCTION MODEL

We have seen in the preceding chapter that forecasts with a univariate ARIMA model of a time series are based on its own history alone. In many forecasting situations the series to be forecast may be systematically influenced by other series. In such circumstances one may be able to make more accurate forecasts by taking into account the past history of other related series as well.

An extension of the univariate ARIMA model to consider two or more series jointly is the "transfer function model." In this chapter we (1) briefly describe a transfer function model of Canadian housing starts that we have built, and (2) discuss the advantages and drawbacks of the transfer function model as a forecasting tool. Discussions on transfer function models are found in several books including Box and Jenkins (1970), Granger and Newbold (1977), and Makridakis *et al.*.

4.1 A Transfer Function Model of Housing Starts

A. The Transfer Function Noise Model.

Let Y_t be housing starts and X_t housing price, both at time t. A realistic formulation of the relationship between between housing starts and the housing price may require that many lagged values of the housing price be included in the model. A change in the housing price may have an effect on the starts which is distributed over a number of periods, and a realistic formulation of the relationship may require that many lagged values of housing price be included in the model. If the lag pattern persists through time, the current value of housing starts is seen as the sum of effects from current and past housing prices. A dynamic relationship between the two variables may be represented by

$$\begin{aligned} Y_t &= v_0 X_t + v_1 X_{t-1} + \cdots + v_m X_{t-m} + e_t \\ &= v(B) X_t + e_t, \end{aligned} \tag{4.1}$$

where $v(B) = v_0 + v_1 B + \cdots + v_m B^m$, the v_j's are fixed parameters, and the e_t's are serially uncorrelated random errors with mean zero and variance σ^2. In the times series literature X_t and Y_t are often referred to as the "input" and "output" series, respectively, while the polynomial $v(B)$ as the *transfer function*. The v_j's in $v(B)$ are called the *impulse response weights*.

In practice, the maximum lag m may have to be fairly large to provide an adequate representation of the relationship between the two variables X and Y. Several suggestions have been made on possible constraints to impose on the lag structure, thereby achieving a more parsimonious representation. One suggestion is that we approximate the polynomial $v(B)$ by a ratio of two polynomials in B, and write (4.1) as

$$Y_t = \frac{\omega(B)}{\delta(B)} X_t + e_t, \tag{4.2}$$

where

$$\omega(B) = \omega_0 + \omega_1 B + \cdots + \omega_s B^s$$

and

$$\delta(B) = 1 - \delta_1 B - \cdots - \delta_r B^r.$$

This is referred to as the *rational distributed lag model.* In parallel to the stationarity concept of an ARMA model, the model in (4.2) is said to be "stable" if all roots of the polynomial $\delta(B)$ are greater than one in absolute value. This condition ensures that the values of X_t in the distant past give negligible influence on Y_t.

Suppose that the error term e_t in (4.2) is an ARMA(p,q) process of the form

$$\phi(B)e_t = \theta(B)u_t, \tag{4.3}$$

where u_t is a white noise process and $\phi(B)$ and $\theta(B)$ are as defined in ARIMA models. Substituting (4.3) into (4.2), we write the model as

$$Y_t = \frac{\omega(B)}{\delta(B)}X_t + \frac{\theta(B)}{\phi(B)}u_t. \tag{4.4}$$

The model is called the *transfer function noise model*, or the *transfer function model* for short. In many applications the Y_t, X_t, and e_t series may be integrated processes.

Since housing starts and housing price are seasonal with a period s of 12 months, a seasonal transfer function model is to be developed by including seasonal operators in the noise. As for univariate models it is found necessary to difference the series seasonally as well as nonseasonally to induce stationarity in e_t. Further, the transfer function $v(B)$ may include seasonal components. Thus (4.4) is extended to the form of

$$Y_t = \frac{\omega(B)\Omega(B^s)}{\delta(B)\Delta(B^s)}X_t + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)}u_t. \tag{4.5}$$

We now report on the results of building a transfer function model of Canadian housing starts with the housing price series as an input variable.[6] This formulation of the transfer function model is consistent with the housing paradigm discussed in Chapter 2; in the

[6] We have also attempted to build a transfer function model of Canadian housing starts with the conventional 5-year mortgage interest rates as the input variable using the data from 1981 to 1989. The estimation results indicate that mortgage interest rates do *not* affect housing starts. This finding is consistent with the Stansell-Mitchell (1985) study of the U.S. housing markets. They have examined the causal relationship between six different variables which describe mortgage rates and terms and single family housing starts in the U.S. Their results indicate that neither credit rationing nor mortgage rates affect housing starts. They have found some evidence that housing prices have a causal relationship to housing starts, possibly reflecting speculative motives. Other work on transfer function modelling of the housing market includes Hillmer and Tiao (1979), Wang and Ma (1981), and Puri and Van Lierop (1988).

short run, stock demand impacts on the existing fixed supply of stock to determine price and the housing industry responds to price through new construction.

We have used monthly single-family housing starts in Canadian urban centres of 10,000 or more (CANSIM index D849796) and the new housing price indexes (CANSIM index D636200) as the output and input series, respectively. As the price index series is available since 1981, the sample data consists only of 108 monthly observations from January 1981 to December 1989. As in Chapter 3 on ARIMA modelling, 12 observations of 1989 were set aside for post-sample forecasting, and the remaining 96 observations were used for transfer function modelling. The length of the series may not be long enough for efficient modelling since both series contain seasonal variations.

As with univariate ARIMA modelling an iterative procedure of identification, estimation, and diagnostic checking is used in the transfer function modelling. Model identification is more difficult in the transfer function modelling, and different tools are required for identification than those used in ARIMA modelling.

Identification is concerned with (1) the estimation of the transfer function $v(B)$, (2) the identification of the ARIMA model for the noise term e_t, and (3) the determination of the polynomials $\delta(B)$ and $\omega(B)$.

The impulse response weights v_j's in the transfer function $v(B)$ can be estimated in two different ways. Box and Jenkins (1970) suggest that a univariate ARIMA model of the form

$$X_t = \frac{\theta_x(B)}{\phi_x(B)}\alpha_t = T(B)\alpha_t \tag{4.6}$$

is built for the input variable X_t. The model is then inverted, i.e., $T^{-1}(B)X_t = \alpha_t$, and the "prewhitened" residual series α_t is obtained. One then applies the prewhitening filter $T^{-1}(B)$ to the output series Y_t to obtain the filtered series $\beta_t = T^{-1}(B)Y_t$. The filtered β_t series is not necessarily white noise.

The *cross correlation function* (CCF) between α_t and β_t is proportional to $v(B)$. Thus the sample CCF between $\hat{\alpha}_t$, the residuals from the fitted ARMA model of X_t, and $\hat{\beta}_t$, the Y_t series filtered with the estimated $T(B)$, should provide an estimate of $v(B)$. Although $\hat{v}(B)$ thus obtained is not an efficient estimate, it provides the basis for identifying the transfer function $v(B)$. So now we turn to the ARIMA modelling of housing starts.

B. Univariate ARIMA Modelling of Housing Starts.

Visual inspection of the time series plot in Figure 4.1 indicates that housing starts series is nonstationary. We have taken the common logarithm of the series to induce stationarity in variance. A plot of the logged series and its sample ACF suggest that regular and seasonal differencing of order 1 each is required to induce stationarity in mean.

The sample ACF and PACF of the differenced series suggest the a seasonal MA(1) as a tentative model:

$$(1-B)(1-B^{12})LOGSHS_t = (1-\Theta_{12}B^{12})u_t, \tag{4.7a}$$

where $LOGSHS_t$ stands for the common logarithm of single-family housing starts at time t. The model was estimated both with and without an intercept term, and the one without was chosen as the intercept term was insignificant.

Estimation results are shown in Table 4.1(A). The fitted model is

$$(1-B)(1-B^{12})LOGSHS_t = (1-.9548_{12}B^{12})u_t, \tag{4.7b}$$

The moving-average parameter is estimated to be 0.9548, which is very close to the non-invertibility boundary.[7] Diagnostic checking based on the residuals was satisfactory except for a significant autocorrelation and partial autocorrelation at lag 7.

The seasonal MA(1) model was further checked by overfitting and obtaining the following seasonal ARMA(1,1) model:

$$(1+.4134B^{12})(1-B)(1-B^{12})LOGSHS_t = (1-.3765B^{12})u_t, \tag{4.8}$$

with $R^2 = .936$ and SEE = .05484. Both coefficient estimates seemed reasonable and were statistically significant. No autocorrelations and partial autocorrelations were significant although they were large at lags 3 and 7. Although the model appeared to be adequate, it did not forecast as well as the seasonal MA(1) model for the 12 months of 1989. The forecasts from the latter model are given in Table 4.4(B) and are used as the basis for checking the forecasting performance of the transfer function model of housing starts.

C. Univariate ARIMA Modelling of Housing Price Indexes.

Visual inspection of the time series plot in Figure 4.2 indicates that the housing price index series is nonstationary. To induce stationarity in the series we took its common logarithm and applied regular differencing twice to the logged series.

Examination of the sample ACF and PACF from the differenced series we have identified the input series to be a regular ARIMA (1,2,1) model:

$$(1-\phi_1 B)(1-B)^2 PRICE_t = (1-\theta_1 B)\alpha_t, \tag{4.9a}$$

where $PRICE_t$ is the common logarithm of the original housing price index.

[7]For estimation of an ARIMA model the TSP and SHAZAM programs use a full ML method with the so-called backcasting while the SCA-UTS uses an approximate ML method even if its "exact" method is specified as an option for estimation. Differences in parameter estimates from different computer programs for the same model (4.7) and the same data are remarkable: $\hat{\Theta}_{12}$ and its standard error are 0.9548 and 0.0739 for SCA-UTS, 0.8595 and 0.0390 for TSP, and 0.8598 and 0.0376 for SHAZAM.

ML estimates of the model in (4.9a) are summarized in Table 4.2. The fitted model is

$$(1-.5899B)(1-B)^2 PRICE_t = (1-.9010B)\alpha_t. \tag{4.9b}$$

Both parameter estimates are significant and fit is very good as indicated by $R^2 = .998$. Autocorrelations and partial autocorrelations were significant at lags 6 and 16.

The following multiplicative ARIMA model is one of several additional models estimated:

$$(1-.6768B)(1+.6038B^{12})(1-B)(1-B^{12})PRICE_t = \alpha_t. \tag{4.10}$$

Both parameter estimates were significant and satisfied the stationarity conditions. The fit was very good when judged by $R^2 = .998$. Autocorrelations were large, if not significant, at several lags and partial autocorrelation was insignificant at lag 12. However, its post-sample forecasting for the 12 months of 1989 was inferior to that of the model in (4.9). Therefore, the ARIMA(1,2,1) model is used to filter the input and output series.

D. Transfer Function Modelling of Housing Starts.

The sample CCF of the prewhitened input series and the filtered output series lacked the clarity we would need for identifying the transfer function. We have, therefore, resorted to an alternative identification method.

We may estimate the linear model (4.1) with a relatively large value of m by least squares. Since e_t is not necessarily a white noise process, Liu and Hanssens (1982) suggest to include a model of autoregressive disturbance terms of the form $(1-\phi_1 B)e_t = u_t$ or $(1-\phi_1 B)(1-\Phi_1 B^s)e_t = u_t$ to improve the efficiency in the least squares estimation of the impulse response weights. They call this approach the *linear transfer function method*. Using the Liu and Hanssens method, we tentatively identified a transfer function model as

$$(1-B)(1-B^{12})LOGSHS_t = (v_0 + v_1 B + \cdots + v_4 B^4)(1-B)^2 PRICE_t,$$

where the preliminary estimates of the v's were $\hat{v}_0 = 2.38$, $\hat{v}_1 = .49$, $\hat{v}_2 = .69$, $\hat{v}_3 = -.63$, and $\hat{v}_4 = 2.83$. Only two estimates $\hat{v}_0$ and $\hat{v}_4$ were significant.

Once an estimate of $v(B)$ is available, we can identify an ARMA model of the noise component by examining the sample ACF and PACF of the estimated noise series:

$$\hat{e}_t = Y_t - \hat{v}(B)X_t,$$

where $Y_t = (1-B)(1-B^{12})LOGSHS_t$ and $X_t = (1-B)^2 PRICE_t$. We have tentatively identified the multiplicative ARIMA model:

$$e_t = (1-\theta_1 B)(1-\Theta_1 B^{12})u_t$$

as the noise process.

Once the orders of the transfer function and the noise component are specified, the fully identified model can be estimated efficiently by maximizing the corresponding likelihood function. When the identified model was estimated by the ML method, the t-statistics for the three numerator parameters $\hat{\omega}_1$, $\hat{\omega}_2$, and $\hat{\omega}_3$ were not significant. These parameters could be deleted from the model with little loss in explanatory power.

The transfer function model to estimate is then

$$(1-B)(1-B^{12})LOGSHS_t = (\omega_0 + \omega_4 B^4)(1-B)^2 PRICE_t + (1-\theta_1 B)(1-\Theta_1 B^{12})u_t. \tag{4.11a}$$

Estimation results of this model are summarized in Table 4.2. The fitted model is

$$\begin{aligned}(1-B)(1-B^{12})LOGSHS_t &= (2.3845 + 2.8852B^4)(1-B)^2 PRICE_t \\ &\quad + (1+.2138B)(1-.94431B^{12})u_t. \end{aligned} \tag{4.11b}$$

Note that all parameter estimates are significant and the model fit is good when judged by $R^2 = .938$.

The estimated model was checked for adequacy. Residuals from an adequate model should behave like a white noise series. The sample ACF and PACF of the residuals and the Ljung-Box Q statistics are useful tools for diagnostic checking. The sample CCF between the residuals and prewhitened X_t series ($\hat{\alpha}_t$) is also a useful diagnostic tool. If the model is adequate, sample cross correlations should be insignificant at all lags.

The sample CCF between the pre-whitened input and the residuals is shown in Table 4.3 and plotted in Figure 4.3. No cross correlations are significant at the 0.05 level of significance, and indicate that the housing price index is an exogenous input as required. Sample ACF and PACF of the residuals are shown in Figure 4.4(A) and (B), respectively. No coefficients are significant, and the Ljung-Box Q statistics are consistent with the hypothesis that the residuals are white noise. Study of the sample ACF, PACF, and CCF does not reveal any strong remaining structure.

Some further insight into the model may be gained by examining the correlation matrix of the parameter estimates. Although high correlations also arise as a result of the particular data series being modelled, they are also consistent with model misspecification or parameter redundancy. We find that all correlation coefficients are small, the largest being 0.25 between $\hat{\omega}_0$ and $\hat{\omega}_4$. We conclude that the fitted transfer function model is adequate.

As a final point we examine what has been gained by use of a transfer function model over a univariate model for housing starts. The univariate model is

$$(1-B)(1-B^{12})LOGSHS_t = (1-.9548B^{12})u_t$$

from Table 4.1(A). This compares in structure to the noise component of the transfer function model

$$e_t = (1 + .2138B)(1 - .9443B^{12})u_t.$$

The residual standard error for the ARIMA model is .05678 while that for the transfer function model is .05371. Although the reduction in the residual standard error is not large (about 5 per cent), the transfer function model explains a higher proportion of the variability in the housing starts data.

4.2 Forecasting with a Transfer Function Model

Once we have obtained an adequate transfer function model, we can use it for forecasting the future values of the output variable. Although the mechanics of generating forecasts are described briefly below using the basic transfer function model with stationary input and output variables, the same principle should apply to the multiplicative seasonal transfer function model with integrated input and output variables.

Assume for the time being that the model is fully known and that at time T we require an h-period ahead forecast of Y_t. The "known" transfer function model is of the form

$$Y_t = \frac{\omega(B)}{\delta(B)} X_t + \frac{\theta(B)}{\phi(B)} u_t \tag{4.12a}$$

and can be written as

$$\delta^*(B) Y_t = \omega^*(B) X_t + \theta^*(B) u_t, \tag{4.12b}$$

where $\delta^*(B)$, $\omega^*(B)$, and $\theta^*(B)$ are polynomials in B. If the model holds in the future, we can write Y_{T+h} from (4.12b) as

$$\begin{aligned} Y_{T+h} = {} & \delta_1^* Y_{T+h-1} + \cdots + \delta_{r*}^* Y_{T+h-r^*} + \omega_0^* X_{T+h} - \cdots - \omega_{s*}^* X_{T+h-s^*} \\ & + u_{T+h} - \theta_1^* u_{T+h-1} - \cdots - \theta_{q*}^* u_{T+h-q^*}. \end{aligned} \tag{4.13}$$

As with the ARIMA model, the conditional expectation of Y_{T+h} given the information set $\mathcal{I}_T = \{Y_T, Y_{T-1}, \ldots; X_T, X_{T-1}, \ldots\}$ is its minimum mean square error forecast at time T if the u_t series are normal. Thus we can compute the h-step ahead optimal forecast by taking the conditional expectation of (4.13):

$$\begin{aligned} f_T(h) = E_T(Y_T + h) = {} & \delta_1^* E_T(Y_{T+h-1}) + \cdots + \delta_{r*}^* E_T(Y_{T+h-r^*}) \\ & + \omega_0^* E_T(X_{T+h}) - \cdots - \omega_{s*}^* E_T(X_{T+h-s^*}) \\ & + E_T(u_{T+h}) - \theta_1^* E_T(u_{T+h-1}) - \cdots - \theta_{q*}^* E_T(u_{T+h-q^*}), \end{aligned} \tag{4.14}$$

where $f_T(h)$ stands for the h-period ahead forecast and $E_T[\cdot]$ the conditional expectation, both at time T.

Note that (4.14) includes the conditional expectation of the X_t series. Thus the forecast of Y_t requires that of X_t.

Given that the X_t series is an ARIMA process, its conditional expectations or forecasts, $E_T(X_{T+h}) = \hat{X}_{T+h}$, can be computed as described by (3.9) in Sections 3.2.

If the X_t and u_t processes are normal, then the h-step ahead forecast error $e_T(h)$ is normal, and an interval forecast can be obtained. For example, a 95% interval forecast is given by

$$f_T(h) \pm 1.96 SE[e_T(h)], \tag{4.15}$$

where $SE[e_T(h)]$ is the standard error of forecast and depends on the model parameters as well as the variances of X_t and u_t. As the forecast horizon h increases, forecast variance in (4.19) increases and so does the width of an interval forecast.

The fitted transfer function model is

$$(1-B)(1-B^{12})LOGSHS_t = (2.3845 + 2.8852B^4)(1-B)^2 PRICE_t + (1+.2138B)(1-.9443B^{12})u_t$$

and that the input model is

$$(1-.5899B)(1-B)^2 PRICE_t = (1-.9010B)\alpha_t.$$

Multiplying out the transfer function model, we write the h-step ahead forecast function at origin T as

$$\begin{aligned} E_T(LOGSHS_{T+h}) = {} & E_T(LOGSHS_{T+h-1}) + E_T(LOGSHS_{T+h-12}) \\ & - E_T(LOGSHS_{T+h-13}) + 2.3845E_T(PRICE_{T+h}) \\ & - 4.7690E_T(PRICE_{T+h-1}) + 2.3845E_T(PRICE_{T+h-2}) \\ & + 2.8852E_T(PRICE_{T+h-6}) + E(u_{T+h}) \\ & + .2138E(u_{T+h-1}) - .9443E(u_{T+h-12}) \\ & + .2019E(u_{T+h-13}), \end{aligned} \tag{4.16}$$

where $E_T[\cdot]$ stands for conditional expectation. Conditional expectations of the price series are obtained from its univariate ARIMA model as

$$\begin{aligned} E_T(PRICE_{T+h}) = {} & 2.5899E_T(PRICE_{T+h-1}) + 2.1798E_T(PRICE_{T+h-2}) \\ & - .5899E_T(PRICE_{T+h-3}) - E_T(\alpha_{T+h}) \\ & - .9018E_T(\alpha_{T+h-1}). \end{aligned} \tag{4.17}$$

Forecasts and their related statistics obtained from (4.16) are shown in Table 4.4(A) for lead time $h = 1, 2, \ldots, 12$. They correspond to forecasts made at December 1988 for the twelve months of 1989. Standard errors of forecasts, actual values, and forecast errors are in the next three columns. Forecast errors as percentage of actual values range in absolute value from .446 for October 1989 to 4.503 for June 1989. Although the turning points were reasonably well forecast, forecasts were consistently above the actual values. We should note that in computing these forecasts of housing starts, forecasts of the housing price index obtained from the ARIMA(1,2,1) model in (4.17) were used as the input.

For comparison, we have generated twelve-month ahead forecasts of housing starts using a univariate ARIMA model in (4.7):

$$(1-B)(1-B^{12})LOGSHS_t = (1-.9548B)u_t.$$

Forecasts and their related statistics from this model are presented in Table 4.4(B).

Define the mean absolute percentage error of forecasts and the root mean square percentage error of forecasts, respectively, as

$$MAPE = \frac{\sum_{i=1}^{n} 100|(A_t - F_t)/A_t|}{n}$$

and

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{n}[100(A_t - F_t)/A_t]^2}{n}},$$

where F_t and A_t are the forecast and actual values and n is the number of forecasts. MAPE's are 19.22 and 12.50 for the transfer function and ARIMA forecasts, respectively, while RMSPE's are 56.89 and 42.95, respectively. Clearly, the transfer function model did not forecast as well as the ARIMA model.

We note in Table 4.4(B) that the standard errors of forecasts are larger for the transfer function model than for the univariate model for all lead times except for $h = 1$. We would expect that the transfer function model should improve short-term forecasting performance by taking account of the additional information contained in the input series. However, the forecasts of the input must be used rather than the actual observations at longer leads.

Forecasts based on a transfer function model should be more accurate than those based on a univariate model (1) if the input variables explain a significant proportion of variation in the output variable and (2) if the input variables can be forecast accurately. Although the fitted transfer function (4.16) explains a higher proportion of housing starts than the univariate ARIMA model, the influence of the housing price on the starts does not seem to be strong enough to yield forecast improvement through the transfer function modelling.

As a forecasting tool the transfer function model has the following advantages:

(1) It generates accurate forecasts at least for the short run;
(2) Using information contained in other related series, the transfer function model can generate more accurate forecasts than the univariate ARIMA models; and
(3) It is possible to estimate a transfer function model using data up to any point in the past and produce forecasts at any horizon so that forecasting accuracy of the model can be checked.

On the other hand, a few drawbacks have to be considered.

(1) Forecasting with a transfer function requires forecasts of the input variables. Forecast errors in the input variables may make the forecasts from the transfer function model very inaccurate.
(2) Like the conventional econometric model the transfer function model requires that the input variables be exogenous. If the input variables are not exogenous, transfer function modelling and forecasting are invalid.

Table 4.1

(A) SUMMARY OF AN ARIMA MODEL OF HOUSING STARTS
(Single-Family Units in Common Log, d = D = 1)

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
LNSHS	RANDOM	ORIGINAL	$(1-B^{1})\ (1-B^{12})$

PARAMETER LABEL	VARIABLE NAME	VALUE	STD ERROR	T VALUE
1 THETA12	LOGSHS	.9548	.0739	12.92

TOTAL SUM OF SQUARES	.449533E+01
TOTAL NUMBER OF OBSERVATIONS	96
RESIDUAL SUM OF SQUARES.	.257596E+00
R-SQUARE	.931
EFFECTIVE NUMBER OF OBSERVATIONS . .	83
RESIDUAL VARIANCE ESTIMATE	.322404E-02
RESIDUAL STANDARD ERROR.	.567807E-01

(B) SUMMARY OF AN ARIMA MODEL OF HOUSE PRICE INDEXES
(Base = 1981 in Common Logarithm, d = 2)

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
PRICE	RANDOM	ORIGINAL	$(1-B^{1})\ (1-B^{1})$

PARAMETER LABEL	VARIABLE NAME	VALUE	STD ERROR	T VALUE
1 THETA	PRICE	.9010	.0563	16.01
2 PHI	PRICE	.5899	.1092	5.40

TOTAL SUM OF SQUARES	.126748E+01
TOTAL NUMBER OF OBSERVATIONS	96
RESIDUAL SUM OF SQUARES.	.188733E-02
R-SQUARE	.998
EFFECTIVE NUMBER OF OBSERVATIONS . .	93
RESIDUAL VARIANCE ESTIMATE	.202938E-04
RESIDUAL STANDARD ERROR.	.450487E-02

Table 4.2

SUMMARY OF A TRANSFER FUNCTION OF HOUSING STARTS
(Single-Family Units in Common Log, d = D = 1)

VARIABLE	TYPE OF VARIABLE	ORIGINAL OR CENTERED	DIFFERENCING
LNSHS	RANDOM	ORIGINAL	$(1-B^{1})(1-B^{12})$
PRICE	RANDOM	ORIGINAL	$(1-B^{1})(1-B^{1})$

PARAMETER LABEL		VARIABLE NAME	VALUE	STD ERROR	T VALUE
1	W0	PRICE	2.3845	.9157	2.60
2	W4	PRICE	2.8852	1.0610	2.72
3	THETA1	LOGSHS	-.2138	.1079	-1.98
4	THETA12	LOGSHS	.9443	.0732	12.90

TOTAL SUM OF SQUARES	.449533E+01
TOTAL NUMBER OF OBSERVATIONS	96
RESIDUAL SUM OF SQUARES.	.239441E+00
R-SQUARE	.938
EFFECTIVE NUMBER OF OBSERVATIONS . .	83
RESIDUAL VARIANCE ESTIMATE	.288484E-02
RESIDUAL STANDARD ERROR.	.537107E-01

Table 4.3

CROSS-CORRELATIONS BETWEEN PREWHITENED PRICE INDEXES AND TRANSFER FUNCTION RESIDUALS

TIME PERIOD ANALYZED	14 TO 96	
NAMES OF THE SERIES	RX	RY
EFFECTIVE NUMBER OF OBSERVATIONS . .	83	83
STANDARD DEVIATION OF THE SERIES . .	.0044	.0486
MEAN OF THE (DIFFERENCED) SERIES . .	.0004	.0040
STANDARD DEVIATION OF THE MEAN . . .	.0005	.0053
T-VALUE OF MEAN (AGAINST ZERO) . . .	.9134	.7536

CORRELATION BETWEEN RY AND RX IS .02

CROSS CORRELATION BETWEEN RX(T) AND RY(T-L)

1-12	-.07	.15	-.07	-.16	.09	.22	-.08	.07	-.08	-.07	.03	-.00
ST.E.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
13-24	.17	-.00	-.13	.05	.02	-.07	.09	.05	-.08	-.03	-.15	.02
ST.E.	.12	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13
25-36	-.14	.18	-.00	-.08	-.07	-.10	.05	.10	.09	.05	-.14	.01
ST.E.	.13	.13	.13	.13	.14	.14	.14	.14	.14	.14	.14	.15

CROSS CORRELATION BETWEEN RY(T) AND RX(T-L)

1-12	-.06	.06	-.13	-.03	-.09	.01	-.13	.10	-.18	-.04	-.10	.04
ST.E.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
13-24	.11	.00	-.09	.11	-.11	.12	.03	-.06	.01	.01	-.04	-.02
ST.E.	.12	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13
25-36	.02	-.01	.02	-.09	.06	-.02	.02	.03	.01	-.03	-.00	-.04
ST.E.	.13	.13	.13	.13	.14	.14	.14	.14	.14	.14	.14	.15

Table 4.4

TWELVE-MONTH AHEAD FORECASTS OF HOUSING STARTS
(In Common Logarithm)

(A) Based on a Transfer Function Model

TIME 1989	FORECAST	STD ERROR	ACTUAL VALUE	ERROR	PERCENT ERROR	SQUARED % ERROR
Jan	3.824	.0548	3.770	-.054	-1.444	2.086
Feb	3.737	.0855	3.699	-.038	-1.038	1.078
Mar	3.842	.1076	3.803	-.040	-1.045	1.093
Apr	4.038	.1259	4.285	-.064	-1.612	2.600
May	4.149	.1428	4.067	-.082	-2.022	4.087
Jun	4.197	.1574	4.016	-.181	-4.503	20.275
Jul	4.081	.1706	3.985	-.096	-2.406	5.788
Aug	4.032	.1828	3.998	-.033	-.836	.699
Sep	4.011	.1942	3.943	-.068	-1.721	2.961
Oct	3.952	.2050	3.934	-.018	-.446	.199
Nov	3.942	.2152	3.902	-.040	-1.030	1.061
Dec	3.869	.2249	3.827	-.043	-1.118	1.251

(B) Based on an ARIMA Model

TIME 1989	FORECAST	STD ERROR	ACTUAL VALUE	ERROR	PERCENT ERROR	SQUARED % ERROR
Jan	3.766	.0568	3.770	.004	.104	.011
Feb	3.683	.0803	3.699	.015	.412	.169
Mar	3.768	.0983	3.803	.035	.922	.849
Apr	4.013	.1136	3.974	-.039	-.971	.943
May	4.141	.1270	4.067	-.073	-1.804	3.254
Jun	4.119	.1391	4.016	-.103	-2.558	6.542
Jul	4.026	.1502	3.985	-.041	-1.020	1.041
Aug	4.002	.1606	3.998	-.003	-.081	.007
Sep	3.974	.1703	3.943	-.031	-.795	.632
Oct	3.973	.1796	3.934	-.038	-.973	.946
Nov	3.955	.1883	3.902	-.053	-1.358	1.844
Dec	3.884	.1967	3.827	-.057	-1.501	2.252

Figure 4.1 Single-Family Housing Starts in Common Logarithm

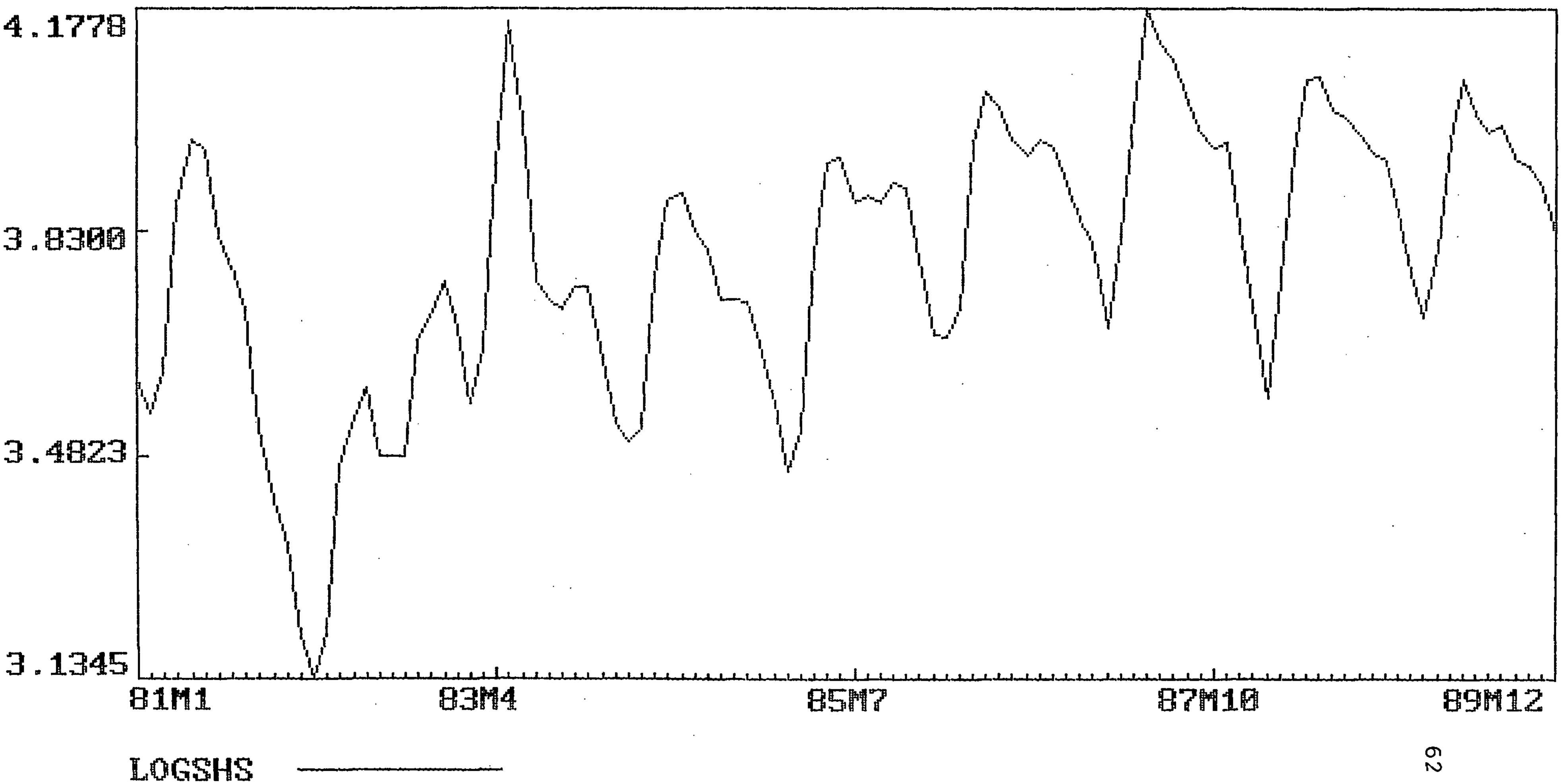


Figure 4.2 Housing Price Index (Base=1981) in Common Logarithm

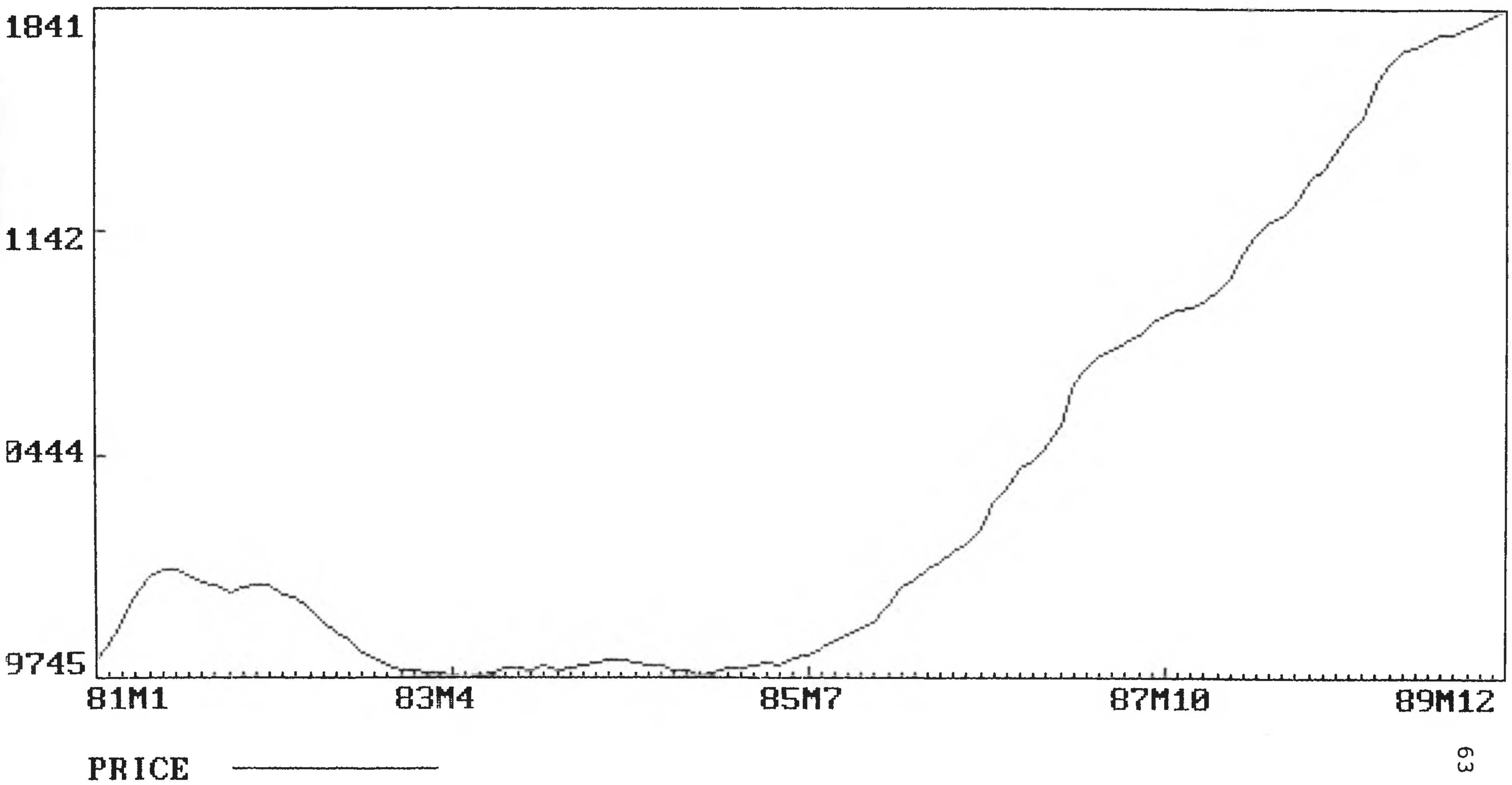


Figure 4.3
CROSS-CORRELATIONS BETWEEN PREWHITENED PRICE INDEXES
AND TRANSFER FUNCTION RESIDUALS

```
         -1.0 -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
          +----+----+----+----+----+----+----+----+----+----+
-24   -.02                   +     I     +
-23   -.04                   +    XI     +
-22    .01                   +     I     +
-21    .01                   +     I     +
-20   -.06                   +    XI     +
-19    .03                   +     IX    +
-18    .12                   +     IXXX  +
-17   -.11                   +  XXXI     +
-16    .11                   +     IXXX  +
-15   -.09                   +   XXI     +
-14    .00                   +     I     +
-13    .11                   +     IXXX  +
-12    .04                   +     IX    +
-11   -.10                   +  XXXI     +
-10   -.04                   +    XI     +
 -9   -.18                   +XXXXXI     +
 -8    .10                   +     IXXX  +
 -7   -.13                   +  XXXI     +
 -6    .01                   +     I     +
 -5   -.09                   +   XXI     +
 -4   -.03                   +    XI     +
 -3   -.13                    + XXXI    +
 -2    .06                    +    IX   +
 -1   -.06                    +   XI    +
  0    .02                    +    I    +
  1   -.07                    +  XXI    +
  2    .15                    +    IXXXX+
  3   -.07                    +  XXI    +
  4   -.16                   +XXXXXI     +
  5    .09                   +     IXX   +
  6    .22                   +     IXXXXX+
  7   -.08                   +   XXI     +
  8    .07                   +     IXX   +
  9   -.08                   +   XXI     +
 10   -.07                   +   XXI     +
 11    .03                   +     IX    +
 12    .00                   +     I     +
 13    .17                   +     IXXXX +
 14    .00                   +     I     +
 15   -.13                   +  XXXI     +
 16    .05                   +     IX    +
 17    .02                   +     IX    +
 18   -.07                   +   XXI     +
 19    .09                   +     IXX   +
 20    .05                   +     IX    +
 21   -.08                   +   XXI     +
 22   -.03                   +    XI     +
 23   -.15                   +XXXXXI     +
 24    .02                   +     I     +
```

Figure 4.4 (A)

SAMPLE AUTOCORRELATIONS OF TRANSFER FUNCTION RESIDUALS

```
TIME PERIOD ANALYZED . . . . . . . . . . 14  TO    96
NAME OF THE SERIES . . . . . . . . . . .           RY
EFFECTIVE NUMBER OF OBSERVATIONS . . .             83
STANDARD DEVIATION OF THE SERIES . . .          .0486
MEAN OF THE (DIFFERENCED) SERIES . . .          .0040
STANDARD DEVIATION OF THE MEAN . . . .          .0053
T-VALUE OF MEAN (AGAINST ZERO) . . . .         .7536
```

1-12	-.06	-.03	-.08	.12	.11	-.04	-.10	-.10	-.11	.04	.06	-.17
ST.E.	.11	.11	.11	.11	.11	.11	.11	.12	.12	.12	.12	.12
Q	.4	.4	1.0	2.3	3.4	3.5	4.4	5.3	6.5	6.7	7.1	10.0

13-24	.02	.08	.10	-.07	-.15	-.01	.03	.18	-.00	-.09	-.12	.04
ST.E.	.12	.12	.12	.12	.12	.13	.13	.13	.13	.13	.13	.13
Q	10.1	10.8	11.9	12.5	15.0	15.0	15.1	18.6	18.6	19.6	21.3	21.5

```
          -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
            +----+----+----+----+----+----+----+----+----+----+
   1   -.06                    +   XXI    +
   2   -.03                    +    XI    +
   3   -.08                    +   XXI    +
   4    .12                    +     IXXX +
   5    .11                   +      IXXX  +
   6   -.04                   +     XI     +
   7   -.10                   +    XXI     +
   8   -.10                   +    XXI     +
   9   -.11                   +   XXXI     +
  10    .04                   +      IX    +
  11    .06                   +      IXX   +
  12   -.17                   +  XXXXI     +
  13    .02                   +      IX    +
  14    .08                   +      IXX   +
  15    .10                   +      IXXX  +
  16   -.07                   +    XXI     +
  17   -.15                   +  XXXXI     +
  18   -.01                   +      I     +
  19    .03                   +      IX    +
  20    .18                   +      IXXXX +
  21    .00                   +      I     +
  22   -.09                   +    XXI     +
  23   -.12                   +   XXXI     +
  24    .04                   +      IX    +
```

Figure 4.4 (B)

SAMPLE PARTIAL AUTOCORRELATIONS OF TRANSFER FUNCTION RESIDUALS

1-12	-0.6	-.03	-.09	.11	.12	-.03	-.08	-.11	-.17	-.00	.09	-.14
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
13-24	.06	.12	.05	-.07	-.17	-.09	-.04	.20	.09	.01	-.08	-.11
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
25-36	-.06	-.12	.00	.04	.06	.13	-.07	.00	.02	-.05	.10	-.10
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

```
             -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1   -.06                       +  XXI    +
   2   -.03                       +   XI    +
   3   -.09                       +  XXI    +
   4    .11                       +    IXXX +
   5    .12                       +    IXXX +
   6   -.03                       +   XI    +
   7   -.08                       +  XXI    +
   8   -.11                       + XXXI    +
   9   -.17                       +XXXXI    +
  10    .00                       +    I    +
  11    .09                       +    IXX  +
  12   -.14                       +XXXXI    +
  13    .06                       +    IXX  +
  14    .12                       +    IXXX +
  15    .05                       +    IX   +
  16   -.07                       +  XXI    +
  17   -.17                       +XXXXI    +
  18   -.09                       +  XXI    +
  19   -.04                       +   XI    +
  20    .20                       +    IXXXXX
  21    .09                       +    IXX  +
  22    .01                       +    I    +
  23   -.08                       +  XXI    +
  24   -.11                       + XXXI    +
```

Figure 4.4 (C)

SAMPLE AUTOCORRELATIONS OF TRANSFER FUCNTION NOISES

```
TIME PERIOD ANALYZED . . . . . . . . . . 14  TO    96
NAME OF THE SERIES . . . . . . . . . .             NT
EFFECTIVE NUMBER OF OBSERVATIONS . . .             83
STANDARD DEVIATION OF THE SERIES . . .          .0805
MEAN OF THE (DIFFERENCED) SERIES . . .          .0043
STANDARD DEVIATION OF THE MEAN . . . .          .0088
T-VALUE OF MEAN (AGAINST ZERO) . . . .         .4857

 1-12   .24   .02   .04   .18   .22   .02 -.15 -.30 -.13   .05 -.05 -.51
ST.E.   .11   .12   .12   .12   .12   .12   .12   .13   .13   .14   .14   .14
Q       5.1   5.1   5.3   8.3  12.5  12.6  14.6  22.9  24.5  24.8  25.1  50.6

13-24 -.12   .04   .07 -.05 -.19 -.08   .10   .25   .07 -.13 -.13   .13
ST.E.   .16   .16   .16   .16   .16   .16   .16   .16   .17   .17   .17   .17
Q      52.1  52.3  52.7  53.0  56.9  57.5  58.6  65.7  66.2  68.1  70.1  72.1

25-36   .04 -.07 -.08 -.07   .10   .11 -.03 -.06   .05   .15   .16 -.10
ST.E.   .17   .17   .17   .17   .17   .17   .17   .17   .17   .18   .18   .18
Q      72.2  72.8  73.6  74.2  75.7  77.3  77.4  78.0  78.3  81.4  85.1  86.5

           -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
             +----+----+----+----+----+----+----+----+----+----+
   1     .24                     +    IXXXX+X
   2     .02                    +     I     +
   3     .04                    +     IX    +
   4     .18                    +     IXXXXX+
   5     .22                    +     IXXXXX+
   6     .02                    +     IX    +
   7    -.15                    + XXXXI     +
   8    -.30                   X+XXXXXI     +
   9    -.13                   +   XXXI      +
  10     .05                   +      IX     +
  11    -.05                   +     XI      +
  12    -.51             XXXXXX+XXXXXXI      +
  13    -.12                  +    XXXI       +
  14     .04                  +       IX      +
  15     .07                  +       IXX     +
  16    -.05                  +      XI       +
  17    -.19                  +  XXXXXI       +
  18    -.08                  +     XXI       +
  19     .10                  +       IXX     +
  20     .25                  +       IXXXXXX +
  21     .07                  +       IXX     +
  22    -.13                  +    XXXI       +
  23    -.13                  +    XXXI       +
  24     .13                  +       IXXX    +
```

Figure 4.4 (D)

SAMPLE PARTIAL AUTOCORRELATIONS OF TRANSFER FUCNTION NOISES

1-12	.24	-.05	.05	.17	.14	-.06	-.16	-.31	-.10	.09	.01	-.43
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
13-24	.24	.07	-.00	-.06	-.12	-.05	.09	.01	-.03	.02	-.13	-.18
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11
25-36	-.04	-.06	.13	.04	.05	.05	-.05	.03	.02	-.07	.01	-.13
ST.E.	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11	.11

```
               -1.0  -.8  -.6  -.4  -.2   .0   .2   .4   .6   .8  1.0
                 +----+----+----+----+----+----+----+----+----+----+
  1    .24                           +    IXXXX+X
  2   -.05                           +   XI    +
  3    .05                           +    IX   +
  4    .17                           +    IXXXX+
  5    .14                           +    IXXXX+
  6   -.06                           +   XI    +
  7   -.16                           +XXXXI    +
  8   -.31                        XXX+XXXXI    +
  9   -.10                           +  XXI    +
 10    .09                           +    IXX  +
 11    .01                           +    I    +
 12   -.43                     XXXXXX+XXXXI    +
 13    .24                           +    IXXXX+X
 14    .07                           +    IXX  +
 15    .00                           +    I    +
 16   -.06                           +   XI    +
 17   -.12                           + XXXI    +
 18   -.05                           +   XI    +
 19    .09                           +    IXX  +
 20    .01                           +    I    +
 21   -.03                           +   XI    +
 22    .02                           +    I    +
 23   -.13                           + XXXI    +
 24   -.18                           XXXXXI    +
```

5. FORECASTING WITH A VECTOR AUTOREGRESSIVE MODEL

In macroeconomic modelling the vector autoregressive (VAR) model, proposed by Sims (1980), has been considered as a viable alternative to the conventional simultaneous-equation econometric model (SEM). One may regard the VAR model as a reduced form in which each endogenous variable is regressed on its own past values as well as the past values of other variables in the system. Further, a VAR model may serve as a practical method of forecasting particularly in the absence of firm *a priori* theory concerning model specification.

Two extensions of the Sims VAR system have been proposed in the literature. One refinement of it, proposed by Hsiao (1981) reduces the number of parameters to be estimated in a VAR model. Unlike the Sims VAR in which all the variables in the system share a common lag, the Hsiao system allows the variables in the system to have different lag lengths in each equation. A second extension to Sims's work is found in the Bayesian approach of Litterman (1979) in which *a priori* information about the model parameters is incorporated into model estimation.

In this chapter we investigate the VAR models. We first illustrate the three types of VAR models by building bivariate VAR systems of Canadian housing starts and mortgage approvals. We then compare their forecasting abilities by examining their *ex post* forecasts.

5.1 The Sims VAR System

Suppose that there are M related time series of interest. A set of joint observations on them at time t may be denoted by an $M \times 1$ vector $\mathbf{x}_t = (X_{t1}, X_{t2}, \ldots, X_{tM})'$. The data at hand consists of T joint observations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T$. A *vector autoregression* of order p is defined by

$$\mathbf{x}_t = \Phi_1 \mathbf{x}_{t-1} + \cdots + \Phi_p \mathbf{x}_{t-p} + \mathbf{u}_t, \tag{5.1}$$

where

$$\Phi_i = \begin{pmatrix} \phi_{11,i} & \cdots & \phi_{1M,i} \\ \vdots & \ddots & \vdots \\ \phi_{M1,i} & \cdots & \phi_{MM,i} \end{pmatrix}; \qquad i = 1, \ldots, p$$

is an $M \times M$ coefficient matrix. The random disturbance term $\mathbf{u}_t$ is a vector white noise with mean zero and covariance matrix Σ.

A typical equation in a VAR(p) system may be expressed algebraically as

$$\begin{aligned} X_{tj} = {} & \phi_{j1,1} X_{t-1,1} + \cdots + \phi_{j1,p} X_{t-p,1} \\ & + \cdots + \phi_{jM,1} X_{t-1,M} + \cdots + \phi_{jM,p} X_{t-p,M} \\ & + u_{tj}, \end{aligned} \tag{5.2}$$

where the number of ϕ coefficients is Mp. Note that all M variables are treated symmetrically in that each variable is determined by its own lagged values and the lagged values of other variables (i.e., cross lags) in the system. Thus the system looks like a set of multiple linear regressions containing the same set of regressors, and each equation may be estimated separately by conditional least squares.

We have built a bivariate VAR model of Canadian housing starts and mortgage approvals. The data used are 252 monthly observations from January 1968 to December 1988 of Canadian housing starts (CANSIM index D84975) and the conventional mortgage loan approvals (CANSIM index D2653). The starts series is the same one as that used in Chapter 3 for ARIMA modelling. Natural logarithms of both series were differenced consecutively and seasonally once each to induce their stationarity.

Although the Sims VAR system is easily estimated by least squares, no specific guide is available for selecting the number of lags to include. Since the variables exhibit strong seasonality, a large value of p is needed for the estimated model to pick up the seasonal effect. However, the degrees of freedom available for estimation deplete rapidly as the order p increases. Using Akaike's (1974) AIC criterion, we have set the order of VAR to be $p = 36$. The number of parameters to be estimated in an equation is then a total of 73 including the constant term.

Estimation results are presented in Table 5.1.[8] Own lag coefficient estimates are found to be statistically significant at lower and seasonal lags while most cross lag coefficient estimates are insignificant even at seasonal lags. Except for those from the recent past the lagged values of mortgage approvals do not appear to help explain housing starts. Forecasts from the fitted VAR model will be discussed in Section 5.4.

5.2 The Hsiao VAR System

Most VAR models used in macroeconometrics include several endogenous variables and several lagged values for each variable. The main disadvantage of VAR models is that the number of parameters to be estimated increases very fast as one increases the number of variables and/or lagged values. Hsiao's (1981) refinement on the Sims VAR mitigates the degrees-of-freedom problem by allowing the lag lengths of different variables in each equation to be different. In fact, there is no *a priori* reason why all the variables have a common lag length in all equations in the model.

Let X_1 be the first endogenous variable in the system, for example, the housing starts series of above. Following Hsiao's modelling strategy, we initially specified the housing starts equation as the autoregression

$$X_{t1} = \phi_{11}(B)X_{t1} + u_{t1}, \tag{5.3}$$

[8] The RATS program, Version 3.1, of VAR Econometrics was used for vector autoregression estimation reported in this chapter.

where $\phi_{11}(B)$ is a polynomial in the operator B. The order p of $\phi_{11}(B)$ was determined by minimizing Akaike's (1969, 1970) FPE over a range of possible orders from 1 to 48:

$$FPE(p) = \frac{T+p}{T-p}\frac{SSR(p)}{T},$$

where $T = 227$ is the number of effective observations after differencing and SSR(p) is the sum of squared residuals from the fitted AR(p) model. The minimum FPE = 0.02935 was obtained when the lag was 36. The intercept term was omitted in the autoregression since it was not significant in the Sims VAR model.

Once the order of autoregression was set, the second variable of the system X_2, the mortgage approvals, was added to the autoregressive equation (5.3) to obtain the bivariate relationship of the form:

$$X_{t1} = \phi_{11}(B)X_{t1} + \phi_{12}(B)X_{t2} + u_{t1}. \tag{5.4}$$

The order of $\phi_{12}(B)$ of X_2 was then determined by minimizing FPE over a range of possible orders from 1 to 48 while the order of $\phi_{11}(B)$ of X_1 was held fixed at the previously determined 36. If FPE in the bivariate regression were not less than that in the autoregression of X_1, then the mortgage approvals variable X_2 would not have helped forecast the housing starts variable X_1 (i.e., X_2 does not "Granger-cause" X_1), and X_2 would have been discarded from the X_1 equation.[9] The minimum FPE = 0.02647 was achieved with the cross lag length of 22. The optimal length of own lags was then re-evaluated by re-estimating (5.4) for a range of possible orders from 1 to 48 while keeping the length of cross-lags fixed at 22. The minimum FPE was found to remain at the same order of 36 for own lags as before.

The regression equation for mortgage approvals was determined similarly. The minimum FPE = 0.05364 was found when the order of own lags was 36 and that for cross lags 8.

The Hsiao system identified above has turned out to have the same length of own lags as the Sims VAR, but, as expected, the cross lags are much shorter in the former than in the latter. The estimation results of the two-equation system by Zellner's SUR method is presented in Table 5.2. As in the case of the Sims VAR, coefficient estimates of own lags are significant at lower and seasonal lags. Many cross lags have been omitted at the model specification stage by the criterion of minimizing FPE. Standard errors of coefficient estimates in the Hsiao VAR are found much smaller than those in the Sims VAR. Overfitting and underfitting the lags in the system indicated that the fitted model was adequate.

5.3 THE BAYESIAN VAR SYSTEM

In the Bayesian approach to VAR modelling forecasters specify a model like (5.1) with a computationally feasible long lag and impose some form of Bayesian priors on all coefficients

[9]Building a VAR model of housing starts and mortgage rates was abandoned because the mortgage rates were found to not "Granger-cause" housing starts.

in the system. We have built two Bayesian VAR models of housing starts and mortgage approvals, and we present them in this section. For surveys of recent developments in the Bayesian VAR modelling see Todd (1984), Doan, Litterman, and Sims (1984), and Litterman (1986b).

In the first Bayesian VAR model of housing starts and mortgage approvals we employ the same dependent variables as in the Sims and Hsiao system, i.e., the stationary differenced series. Since the variables are in differences, we have introduced the priors around the specification of a white noise process with a drift:

$$X_{ti} = \delta + u_{ti}, i = 1, 2. \tag{5.5}$$

Very little is known *a priori* about the drift parameters, and we use noninformative priors for them. When compared to (5.1), the white noise specification in (5.5) implies that all lag coefficients are equal to zero. For the lag coefficients in each equation, therefore, we employ independent normal priors with zero means. Their prior standard deviations are not unique in specification.

The basic type of prior standard deviations would be based on the notion that the more recent values of a variable are more likely to contain useful information for forecasting than the less recent values. Thus the prior standard deviations are specified with the following characteristics:

(1) They decrease as the lags increase; and
(2) They are larger for the own lag coefficients than the cross lag coefficients of other variables in the system.

One can specify prior standard deviations with such characteristics in terms of a few "hyper-parameters":

$$s(i,j,h) = \gamma g(h) f(i,j)(s_i/s_j), \tag{5.6}$$

where $s(i,j,h)$ is the prior standard deviation for the h-th lag coefficient of the j-th variable in the i-th equation; s_i and s_j are the standard errors of autoregressions of variables i and j, respectively; $g(h)$ is the decay parameter for the tightness on lag h relative to the first lag $[g(1) = 1]$; and γ is the overall tightness, which is equal to the standard deviation of the first own lag.

The above basic type of the prior standard deviations applies to variables which are non-seasonal or which have been seasonally adjusted. Since the decay in standard deviation with an increasing lag may dampen the parameter estimates at the important seasonal lags, we have decided to set g(h)=1 for all lags; i.e., the prior standard deviations do not decay with lags. The standard deviations are then

$$s(i,j,h) = \gamma f(i,j)(s_i/s_j) \tag{5.7}$$

and

$$f(i,j) = \begin{cases} 1 & \text{if } i = j \\ \omega & \text{otherwise.} \end{cases}$$

Given the values of the two hyper-parameters γ and ω, the priors and the data are combined to estimate a VAR model of (5.1) using Theil's mixed estimation method or Zellner's seemingly unrelated regression (SUR) procedure. Although the SUR approach yields a gain in efficiency, we have used Theil's method for computational ease.

The hyper-parameters are not specified ***a priori*** but chosen empirically by searching over a range of their possible values and using a forecast criterion. Given a pair of values for γ and ω, the model parameters are estimated, and the fitted VAR is used to make post-sample forecasts. As a selection criterion we have used the mean absolute percentage error (MAPE) or the root mean square percentage error (RMSPE) of one to ten period ahead forecasts.[10]

With a limited search over possible values of γ and ω, the minimum values of MAPE = 7.22 and RMSPE = 8.61 were obtained at $\gamma = 0.5$ and $\omega = .2$. A finer grid search over the γ and ω parameters would have provided a smaller MAPE or RMSPE, but was not attempted due to the time constraints. It was noted, however, that forecasting performance was not overly sensitive to changes in the ω parameter.

Table 5.3 presents the Bayesian VAR model fitted with the values of hyper-parameters as given above. Note that most coefficient estimates are not statistically significant from the prior means of zero except for a few at the first own lags or own seasonal lags. Standard errors of coefficient estimates are found to be smaller than those in the fitted Hsiao VAR.

We have also attempted an alternative Bayesian VAR formulation. Recall that we have transformed the data to induce stationarity in the preceding three VAR models. Such transformation may not be needed in a VAR. Fuller (1976, p.374) shows that no gain in asymptotic efficiency is achieved in autoregression by differencing even when the series are integrated. Furthermore, when different series have different orders of integration, differencing does not provide a satisfactory way of building VARs of nonstationary time series [Harvey (1989), p.469]. Thus we have also built a Bayesian VAR of the series in levels and used the priors on the model parameters around the random walk with drift model:

$$X_{ti} = \delta + X_{t-1,i} + u_{ti}. \tag{5.8}$$

The prior mean of the coefficient of the first own lag is set to be one. The prior means for all other parameters are set to zero.

The seasonality in the series was treated as being deterministic and by including 12 monthly dummy variables in the system. This is an approach advocated by Doan (1990). Noninformative priors are employed for the coefficients of the seasonal dummies. The prior

---

[10] We have used the chosen criterion due to time and computational constraints. One could compute MAPE or RMSPE for each variable using many forecasts at each forecast horizon, and choose the values of the hyper-parameters on the basis of the best overall results. Doan, Litterman, and Sims (1984) suggest that a search be made for the values of hyper-parameters that would minimize the sum of one-step-ahead RMSPE.

standard deviations of other parameters were set to follow a normal harmonic decay so that $g(h) = 1/h$ in (5.5).[11]

The final fitted model chosen after a limited search over the possible values of the hyperparameters γ and ω is presented in Table 5.4. The values of $\gamma = 2.5$ and $\omega = .3$ gave the minimum values of MAPE = 7.06 and RMSPE = 8.83. It is interesting to note that the coefficients of cross lags as well as own lags are significant at the first and/or second lags in both equations. Mortgage approvals do help explain the variations in housing starts. Further, none of the coefficient estimates at seasonal lags are significant. Seasonal variations in the series appear to have been captured adequately by the seasonal dummies.

5.4 Forecasting Performances of VAR Models

Given a VAR model, one can use it to obtain the joint forecasts of all variables in the system. As in the case of univariate ARIMA and transfer function models, the expectations of $\mathbf{x}_{T+h}$ conditional upon the history $\mathcal{I}_T$ of the vector process up to time T is the optimal h-period ahead forecast in the sense of the minimum mean square error:

$$\mathbf{f}_T(h) = E_T[\mathbf{x}_{T+h}|\mathcal{I}_T],$$

where $\mathbf{f}_T(h)$ is the optimal h-step ahead forecast and E_T is the conditional expectation operator, both at time T. We can recursively calculate the h-step ahead forecasts from

$$\begin{aligned} \mathbf{f}_T(h) &= E_T(\mathbf{x}_{T+h}) \\ &= \Phi_1 \mathbf{f}_T(h-1) + \cdots + \Phi_p \mathbf{f}_T(h-p), \quad h = 1, 2, \ldots \end{aligned} \tag{5.9}$$

where $\mathbf{f}_T(h-i) = \mathbf{x}_{T+h-i}$ for $i \geq h$.

If the random errors are a normal white noise process, forecast errors are also normal:

$$\mathbf{x}_{T+h} - \mathbf{f}_T(h) \sim N[0, \Sigma(h)], \tag{5.10}$$

where $\Sigma(h)$ is the forecast error covariance matrix which depends on the Φ_i's among others. Hence one can construct a confidence region for a set of forecasts or for a single point forecast. For example, a 95% confidence interval forecast for a single component of $\mathbf{x}_{T+h}$ can be obtained from (5.10) as

$$f_T^m(h) \pm 1.96\sigma_m(h),$$

where $f_T^m(h)$ is the m-th component of $\mathbf{f}_T(h)$ and $\sigma_m(h)$ is the standard deviation of the point forecast and equal to the square root of the m-th diagonal element of $\Sigma(h)$.

[11] Raynauld and Simonato (1988, 1990) tried to modify the features of the basic Bayesian VAR methodology to account for the seasonal components in the times series. Their so-called "toothsaw" pattern of decay was also tried, but the fitted VAR did not perform as well in forecasting as the one fitted with a simple decay. No decay in standard deviations was also tried. Unlike in the case of a VAR in differences, forecasting ability was not as good as in the case of the reported the Bayesian VAR with decay.

In practice, the parameters of the VAR model are unknown and estimated. We may replace the unknown parameters in the forecast expressions, and the forecast intervals thus obtained are only approximate 95% forecast intervals.

In order to assess the forecasting performance of the VAR models, ARIMA forecasts of housing starts in 1989 were obtained. The ARIMA model fitted with the data from January 1968 to December 1988 is similar to "Model 2" in Chapter 3 except that the AR(1) term is now omitted:

$$(1 + .2383B^{12} + .1227B^{24} + .2557B^{36})(1 - B)(1 - B^{12})LOGHS_t$$
$$= (1 - .2043B)(1 - .5461B^{12})u_t.$$

All coefficient estimates are significant except the seasonal AR(2) parameter estimate. Various diagnostic checks indicated that the fitted model was adequate.

Table 5.5 presents the one- to ten-month ahead forecasts of housing starts (in levels) based on the four VAR models as well as those based on the above ARIMA model. Actual ARIMA forecast errors ranged in absolute value from 0.68% for the ten-month ahead forecast for October 1989 to 11.35% for the eight-month ahead forecast for August 1989. Although the one-month ahead forecast is poor and off by 11%, ARIMA forecasts do not appear to deteriorate as the forecast horizon increases.

The Sims VAR has generated very accurate forecasts for a short-term horizon. After the three-month ahead forecast, its forecasting ability has deteriorated rapidly and yielded forecast errors ranging in absolute value from 9.74 % to 32.47% relative to the actual housing starts. In terms of individual forecast errors, MAPE, or RMSPE, it did not outperform the simple univariate ARIMA model.

The Sims VAR is easy to specify and estimate. Its main disadvantage as a forecasting tool arises from the degrees of freedom problem. When the model includes many parameters to estimate, the resulting loss in efficiency of the parameter estimates gives rise to large standard errors in forecasts. This loss in efficiency explains the poor performance of the Sims VAR in forecasting housing starts of 1989.

The fitted Hsiao VAR model has performed much better than the Sims VAR in forecasting. It has generated very accurate forecasts for the short run horizon up to three months. Although its forecasts deteriorate as the horizon increases, its forecast errors are consistently much smaller than the unrestricted Sims VAR for all forecast horizons from four to ten months.

As we have seen in the preceding section constructing a Hsiao VAR model is more costly in computing than a Sims VAR; the former gives explicit attention to the lag length of each variable in each equation. But Hsiao's approach reduces the degrees of freedom problem in VAR modelling and yields more efficient parameter estimates and more accurate forecasts than the Sims VAR. In terms of MAPE and RMSPE, however, it is not found to have had an advantage over the univariate ARIMA model in forecasting housing starts.

As discussed in Chapter 2 substantial work has been done in comparing *ex ante* macroeconomic forecasts of large-scale SEMs with ARIMA forecasts. Recently, Litterman (1986) and McNees (1986, 1988) have used Bayesian VAR models for forecast comparisons. Litterman finds that forecasts from his small Bayesian VAR model of the U.S. economy are as accurate, on the average, as those from the prominent commercial forecasting services using elaborate large-scale SEMs. McNees has also similar findings that a small Bayesian VAR model of the U.S. economy forecasts very well for some variables, even over long horizons.

The Bayesian VAR models of housing starts that we built performed favourably in comparison with the other two VAR models. In the difference form it has generated forecast errors in absolute value ranging from 0.95% for a 4-month ahead forecast (April 1989) to 12.17% for a one-month ahead forecast (January 1989). Except for the one- and two-month ahead forecasts, the Bayesian VAR has produced forecast errors which are less in absolute value than those from the Hsiao VAR. Further, MAPE and RMSPE were much smaller for the Bayesian VAR than for the Hsiao VAR.

The Bayesian VAR model in the level form performed as well as that in the difference form. Except for the one-month ahead forecast for January 1989, the Bayesian VAR in levels seems to have performed better than that in differences in terms of the size of individual forecast errors as well as MAPE and RMSPE.

One important advantage of VAR models over the conventional SEM is that one can estimate a VAR model using data through any point in the past, produce forecasts as far ahead as desired, and check for forecasting accuracy of the model. This is exactly what we have done in this chapter.

In contrast to the conventional econometric model all types of VAR models considered above share the following strengths as a forecasting tool:

(1) A VAR model does not require judgemental adjustment of model parameters or forecasts;
(2) It does not employ a dubious exogeneity distinction among the variables to forecast;
(3) It provides a conceptually straightforward and simple method of generating unconditional forecasts that do not assume future values of exogenous variables; and
(4) It does not impose controversial theoretical restrictions that the conventional SEM may contain.

The Bayesian VAR model performed favourably in comparison with the ARIMA forecasts. The main disadvantage of the Bayesian VAR over the ARIMA or the Hsiao VAR is that it is by far the most expensive model to construct. Further, the effective treatment of seasonality in time series in the Bayesian VAR modelling is still in the exploratory stage within the current time series literature. When the methodology of treating seasonally develops fully in VAR modelling, the Bayesian VAR forecasts should become a strong competitor to the ARIMA forecasts.

Table 5.1

SUMMARY OF A SIMS VAR MODEL OF HOUSING STARTS AND MORTGAGE APPROVALS

LAG	(1) STARTS EQUATION VARIABLES STARTS	(1) STARTS EQUATION VARIABLES APPROVALS	(2) APPROVALS EQUATION VARIABLES STARTS	(2) APPROVALS EQUATION VARIABLES APPROVALS
1	-.343*	.113*	.232*	-.101
2	-.318*	.173*	-.026	-.154
3	-.280*	.176*	-.050	-.091
4	-.200*	.121*	-.192	-.019
5	-.146	.109	-.115	-.032
6	-.090	.051	-.257	-.034
7	-.066	.040	-.145	.058
8	.007	.011	-.076	-.021
9	-.137	.054	-.005	.137
10	.038	-.115	-.111	.017
11	-.029	.083	-.069	.043
12	-.689*	-.034	-.058	-.581*
13	-.317*	.028	-.043	-.030
14	-.145	.089	.049	-.245*
15	-.174	.104	.030	-.014
16	-.185	.036	-.121	-.093
17	-.036	.066	-.018	-.133
18	-.107	.077	-.043	-.078
19	-.040	-.080	-.016	-.127
20	.141	-.110	-.067	-.214*
21	-.046	-.054	-.153	.021
22	.054	-.110	-.153	-.023
23	-.005	.028	-.149	.181
24	-.289*	-.032	-.130	-.286*
25	-.159	.039	-.051	.005
26	-.031	.023	.020	-.085
27	-.153	.011	.015	.061
28	-.114	-.017	-.024	.034
29	-.045	.024	-.059	.007
30	-.042	.062	.102	-.065
31	-.104	-.027	.077	-.146
32	.036	.019	.112	-.181
33	-.051	.003	.002	-.105
34	-.129	.017	.012	-.097
35	.007	.045	.046	.159
36	-.321*	.011	-.150	-.258*
CONSTANT	.002		-.001	
	R**2=.634		R**2=.534	
	SEE =.1539		SEE =.2256	

Note: The asterisk * indicates that the coefficient estimate is significant at the 0.05 level of significance.

Table 5.2

SUMMARY OF A HSIAO VAR MODEL OF
HOUSING STARTS AND MORTGAGE APPROVALS

LAG	(1) STARTS EQUATION VARIABLES STARTS	(1) STARTS EQUATION VARIABLES APPROVALS	(2) APPROVALS EQUATION VARIABLES STARTS	(2) APPROVALS EQUATION VARIABLES APPROVALS
1	-.388*	.102*	.172*	-.147*
2	-.329*	.181*	-.089	-.150*
3	-.246*	.148*	-.116	-.128
4	-.127	.088	-.169	-.042
5	-.096	.051	-.194*	-.002
6	-.090	-.011	-.323*	-.066
7	-.089	.041	-.214*	.056
8	-.064	.042	-.187*	.032
9	-.178*	.084		.115
10	-.015	-.084		.023
11	-.004	.064		.052
12	-.683*	-.006		-.598*
13	-.332*	-.013		-.088
14	-.125	.085		-.209*
15	-.102	.075		-.027
16	-.085	.013		-.112
17	.004	.037		-.118
18	-.083	.012		-.155*
19	-.060	-.085		-.128
20	.119	-.129*		-.198*
21	-.011	-.074		-.052
22	.049	-.129*		-.060
23	.055			.144*
24	-.254*			-.346*
25	-.133*			-.057
26	-.030			-.073
27	-.119*			.027
28	-.096			.048
29	-.022			-.044
30	-.015			-.036
31	-.085			-.064
32	-.073			-.099
33	-.018			-.066
34	-.129*			-.046
35	-.003			.185*
36	-.294*			-.257*
	R**2=.652		R**2=.518	
	SEE =.1457		SEE =.2122	

Note: The asterisk * indicates that the coefficient estimate is significant at the 0.05 level of significance.

Table 5.3

SUMMARY OF A BAYESIAN VAR MODEL OF
HOUSING STARTS AND MORTGAGE APPROVALS
(IN DIFFERENCES)

LAG	(1) STARTS EQUATION VARIABLES STARTS	APPROVALS	(2) APPROVALS EQUATION VARIABLES STARTS	APPROVALS
1	-.081*	.001	.001	-.029
2	-.041	.005	-.003	-.031
3	-.018	.001	-.002	-.042
4	-.021	-.000	-.002	-.009
5	-.043	.001	.002	-.024
6	.007	-.001	-.006	-.029
7	.009	.002	-.003	.024
8	-.002	.002	-.001	-.000
9	-.027	.002	.005	.048
10	.037	-.003	-.002	.012
11	-.030	.002	.001	-.015
12	-.275*	-.004	-.002	-.238*
13	-.047	-.002	-.010	-.001
14	.022	-.001	.003	-.056
15	.003	.001	.000	.024
16	-.009	-.001	-.000	.005
17	.018	.001	-.000	-.034
18	-.029	.001	.004	.019
19	-.010	-.003	.001	-.004
20	.056	-.004	.006	-.048
21	-.002	.000	-.005	-.020
22	.016	.001	-.003	.024
23	.029	.001	-.002	.075
24	-.023	.002	.002	-.057
25	-.007	.003	.003	.013
26	.011	-.001	-.001	.032
27	-.036	-.003	.001	.008
28	-.002	-.001	-.000	.003
29	.012	-.000	-.005	.009
30	.022	.002	.002	.003
31	-.028	.001	.000	-.029
32	.026	.004	-.002	-.007
33	-.008	-.001	.002	.006
34	-.073	.001	.000	-.019
35	.033	-.002	.005	.063
36	-.134*	-.002	-.007	-.097*
CONSTANT	-.002		-.002	
	R**2=.355		R**2=.296	
	SEE =.1705		SEE =.2316	

Note: The asterisk * indicates that the coefficient estimate is significant at the 0.05 level of significance.

Table 5.4

SUMMARY OF A BAYESIAN VAR MODEL OF
HOUSING STARTS AND MORTGAGE APPROVALS
WITH SEASONAL DUMMY VARIABLES
(IN LEVELS)

	(1) STARTS EQUATION VARIABLES		(2) APPROVALS EQUATION VARIABLES	
LAG	STARTS	APPROVALS	STARTS	APPROVALS
1	.619*	.086	.213*	.825*
2	.011	.144*	-.225*	-.011
3	-.025	-.064	.047	.010
4	.088	-.020	-.088	.038
5	-.041	.071	.039	-.023
6	.030	-.064	-.063	-.009
7	.043	-.004	.081	.082
8	.075	-.012	.023	-.050
9	-.065	-.006	.021	.093
10	.070	-.056	-.000	.079
11	.031	.027	-.004	.059
12	-.001	-.044	-.013	-.015
13	-.032	-.008	-.017	-.012
14	.044	.003	.062	-.125
15	-.017	-.000	.015	.151*
16	-.000	.001	.013	-.059
17	-.053	.011	.026	-.057
18	-.085	-.009	.040	.013
19	-.011	-.025	.002	-.036
20	.090	-.003	-.018	-.059
21	-.075	.023	-.055	.012
22	-.023	.015	-.004	.072
23	.092	.009	.001	.114
24	.035	-.002	-.003	-.082
25	-.012	.001	.006	.020
26	-.055	-.008	-.005	-.004
27	-.032	-.002	-.011	.027
28	.021	.009	-.006	.021
29	.014	.010	-.002	-.011
30	-.015	.010	.020	-.009
31	-.048	.009	.006	-.042
32	-.018	.012	-.005	-.026
33	-.022	.004	.003	-.031
34	-.055	.000	.008	.002
35	.082	-.010	.007	.046
36	-.003	-.006	-.009	-.031
	R**2=.880		R**2=.848	
	SEE =.1246		SEE =.1787	

Note: The asterisk * indicates that the coefficient estimate is significant at the 0.05 level of significance. All 12 seasonal dummies were significant in the starts equation while none of them were significant in the approvals equation.

Table 5.5

ONE- TO TEN-PERIOD AHEAD FORECAST ERRORS OF HOUSING STARTS BASED ON FIVE MODELS

TIME 1989	ACTUAL VALUE	FORECAST ERRORS BASED ON SVAR	HVAR	BVAR (DIFFERENCE)	BVAR (LEVEL)	ARIMA
Jan	13,678	794	1,056	1,664	2,506	1,504
		(5.81)	(7.72)	(12.17)	(18.32)	(11.00)
Feb	9,774	103	-90	845	1,087	459
		(1.05)	(.92)	(8.64)	(11.12)	(4.69)
Mar	12,150	-673	-847	324	1,227	1,066
		(-5.55)	(-6.97)	(2.67)	(10.09)	(8.77)
Apr	17,392	-1,693	-1,179	164	1,643	816
		(-9.74)	(-6.78)	(.95)	(9.44)	(4.70)
May	19,127	-6,210	-3,888	-2,197	814	-1,869
		(-32.47)	(-20.33)	(-11.49)	(4.25)	(-9.77)
Jun	17,697	-3,926	-2,894	-1,106	-477	-1,863
		(-22.19)	(-16.36)	(-6.25)	(-2.69)	(-10.53)
Jul	16,421	-3,773	-3,330	-821	-204	-661
		(-22.98)	(-20.64)	(-5.00)	(-1.24)	(-4.03)
Aug	1,543	-4,635	-3,350	-2,587	-22	-1,757
		(-29.94)	(-21.63)	(-16.71)	(-.14)	(-11.35)
Sep	16,338	-2,879	-963	810	1,526	301
		(-17.62)	(-5.90)	(4.96)	(9.34)	(1.84)
Oct	15,764	-3,244	-1,268	532	622	-108
		(-20.58)	(-8.04)	(3.38)	(3.95)	(-.68)
MAPE		16.79	11.53	7.22	7.06	6.73
RMSPE		19.66	13.53	8.61	8.83	7.72

Note: Figures in parentheses are forecast errors in percentage relative to the actual.

6. CONCLUSIONS

The primary purpose of this study is to review several model-based forecasting methods that may be used for forecasting the housing sector and mortgage markets. The models considered are (1) the simultaneous-equation econometric model (SEM), (2) the univariate autoregressive integrated moving average (ARIMA) model, (3) the transfer function (TF) model, and (4) the vector autoregression (VAR) model.

Regardless of which model one uses for forecasting, there are common issues in its construction and maintenance:

(1) specification of equation(s) in the model guided by the data and, in the case of an econometric model, also by relevant economic theory;
(2) availability and collection of the data;
(3) initial estimation and evaluation of the model; and
(4) validation of the estimated model—diagnostic checking, *ex post* forecasting, simulation, and tracking tests of the model.

One has to address these issues very carefully if one wants to build and maintain a reliable forecasting model.

Until the 1970s the simultaneous-equation econometric model had been used very widely as a macroeconomic forecasting tool. Specification of an econometric model involves explicit representation of the presumed interdependence and causality relationships among the variables in the model. It is with model specification that economic theory plays a role in econometric modelling.

Forecasting with an econometric model in practice is not based on the model and data alone. A forecaster using an econometric model makes judgemental adjustments to model parameters as well as to the model forecasts. It is because of the subjective adjustments that (1) forecasts based on an econometric model cannot be duplicated by any other except the forecaster and (2) the standard errors of forecasts as a measure of forecast reliability cannot be computed for the point forecasts.

There are several reasons why the econometric approach to forecasting may not yield reliable forecasts:

(1) One requires forecasts of the exogenous variables to forecast the endogenous variables. Inaccuracy in the forecasts of exogenous variables causes the conditional forecasts of endogenous variables to be inaccurate;
(2) Some variables could be erroneously regarded as being exogenous;
(3) In dealing with unobservable variables like permanent income and expected inflation, one often utilizes their crude proxies in the model; and
(4) Economic theory seldom provides a model builder with a guide as to dynamic specification. One often includes the lagged endogenous variables in the structural equations on an *ad hoc* basis.

In contrast to the structural econometric approach, ARIMA and other time series models use little economic theory in model specification. Even without being judgementally adjusted, however, ARIMA forecasts have been found to perform as well as those judgementally adjusted forecasts based on large elaborate econometric models. Moreover, unlike the forecasts based on an econometric model, one can easily compute standard errors of ARIMA forecasts so that a precision measure is attached to point forecasts.

As a forecasting tool, the univariate ARIMA model shares the following strengths with other times series models:

(1) It generates very accurate forecasts at least for the short run;
(2) It is possible to estimate an ARIMA model using data up to any point in the past and then produce forecasts so that the forecasting accuracy of the model can be checked; and
(3) It imposes no controversial theoretical restrictions that the conventional econometric model may contain.

Forecasting with an ARIMA model yields forecasts of a single series without using information contained in other related series. In many forecasting situations in which other variables influence the series to be forecast, one can build a transfer function model that incorporates more than one time series and explicitly introduces the dynamic characteristics of the series.

Forecasts based on a transfer function model may be more accurate than those based on a univariate ARIMA model if the input variable explains a significant proportion of the variation in the output variable. Note that, as the forecasts of endogenous variables require forecasts of exogenous variables in the SEM, the forecasts of the output variable in the transfer function model are conditional on the forecasts of input variables. Accuracy of the forecasts of the output variable is, therefore, partly affected by that of the forecasts of the input variables.

As with the ARIMA and transfer function models, the VAR model is formed from the regularity in the movements of time series without appealing to economic theory. Difficulties experienced in the conventional econometric modelling of the housing and mortgage markets indicate that, despite or perhaps because of its *a*theoretical approach, the univariate ARIMA and VAR methods of forecasting are promising competitors to econometric forecasting.

Most VAR models of the Sims type in macroeconometrics use several variables and include several lagged values for each variable. The number of parameters to be estimated in a VAR model increases very quickly as one increases the number of variables and/or the order of the model. Hsiao's refinement on the Sims VAR or the Bayesian procedure of VAR modelling should partially mitigate the number of parameters problem in VAR modelling. The Bayesian VAR seems to outperform the Sims and Hsiao types of VAR models in forecasting.

As a competitor to the conventional econometric forecasting, the VAR method has the following strengths:

(1) It generates very accurate forecasts at least for the short term and maybe for the medium term;
(2) Unlike the SEM and the transfer function model, it does not employs a dubious exogeneity definition; and
(3) It provides a conceptually straightforward and remarkably simple method of yielding forecasts that do not assume any particular values of exogenous variables.

Lawrence Klein (1982) has the view that "VAR models are alright for predictions one quarter ahead, but VAR predictions quickly deteriorate so that the conventional models offer superior predictions further in the future." His view does not necessarily apply to all forecasting situations. Clearly, forecasting results presented in this report indicate that time series models forecast housing starts remarkably well not just up to the three-month horizon but at least up to the twelve-month horizon. Housing starts many months ahead may not be as much dependent on the observed series as those not so many periods ahead. As the forecast horizon increases, therefore, the forecasting ability of an ARIMA model may deteriorate rather quickly.

Little is known about the relative performance of alternative methods in long-tem forecasting. As the forecast horizon increases, forecasts based on ARIMA, VAR or econometric models tend toward the mean of the time series being forecast. It is only because the forecaster makes subjective adjustments of the expected future values of exogenous variables and the mechanical forecasts of the model that econometric models appear to outperform the time series models in long-term forecasting.

When alternatives exist, choices have to me made so that an appropriate forecasting method is used for the specific forecasting situation being considered. In Table 6.1 we present a ranking of the four forecasting methods in terms of a number of major criteria for good forecasts. Some criteria are statistical while others are practical in nature: data requirements, ease of model specification, ease of model estimation, time/cost of modelling, ease of computing forecasts, computation of standard errors of forecasts, incorporation of judgemental adjustments of forecasts, accuracy in short-term forecasting, accuracy in long-term forecasting, and ease of model updating.

This set of criteria is not exhaustive, and one may add other possible criteria. Nor are the criteria necessarily of equal importance in assessing goodness of forecasting methodologies. With the shortcomings of a simple average or sum, however, a ranking of the methods emerges in the order of ARIMA, TF, SEM, and VAR when the ranks are summed over the listed criteria. There is no such thing as the best approach or method. What is important is to understand how various forecasting methods differ from each other so that forecasting users can make a rational choice for their needs.

The Bayesian VAR method is not yet fully developed for dealing with seasonality in modelling and forecasting. On the other hand, the transfer function modelling approach is not very much different qualitatively from the ARIMA modelling. CMHC desires to provide short term as well as longer term forecasts of housing and mortgage market variables at a fairly detailed level. Building a modest- scale econometric model is the first step toward forecasting several key variables at a longer term horizon. ARIMA models can be used for short-term forecasting.

Table 6.1 Ranking of Four Forecasting Methods*

	Models			
Criteria	SEM	ARIMA	TF	VAR
Data requirements	4	1	2	3
Ease of specification	3	1	2	4
Ease of estimation	3	1	2	4
Time/cost of modelling	3	1	2	4
Ease of computing forecasts	4	1	2	4
Computation of standard error of forecasts@	4	2	2	2
Judgemental adjustments of forecasts#	1	2	3	4
Accuracy in short-term forecasting	4	2	2	2
Accuracy in long-term forecasting	1	3	3	3
Ease of updating	3	1	2	4
Sum of ranks	30	15	22	33

Notes: * Ranking of four models is given by an ordinal scale of 1 to 4. The most preferred model is given the highest rank 1 while the least preferred the rank 4.

@ One can but seldom does introduce judgemental adjustments of forecasts in time series models.

Standard errors cannot be computed for SEM forecasts due to their judgemental adjustments.

REFERENCES

Abraham, B. (1985), "Seasonal Time Series and Transfer Function Modelling," *Journal of Business and Economic Statistics*, 3, 356–361.

Aguilar, R., and B. Sandelin (1984), "Realized Supply of Owner-occupied Houses in Sweden: an Exploratory Study," *Scandinavian Housing and Planning Research*, 1, 197–213.

Akaike, H. (1969), "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics*, 21, 243–247.

——— (1970), "Statistical Predictor Identification," *Annals of the Institute of Mathematical Statistics*, 22, 203–217.

——— (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, AC-19, 716–723.

Anderson, P.A., and J.R. Ostas (1977), "Private Credit Rationing," *New England Economic Review*, 24–37.

Arnott, R. (1987), "Economic Theory and Housing," in E.S. Mills (ed.), *Handbook of Regional and Urban Economics, Vol. II: Urban Economics*, Amsterdam: North-Holland.

Baer, W.C., and C.B. Williamson (1988), "The Filtering of Households and Housing Units," *Journal of Planning Literature,* 3, 127–152.

Bagshaw, M. (1987), "Univariate and Multivariate ARIMA versus Vector Autoregression Forecasting," *Proceedings of American Statistical Association Business and Economics Section*, 783–786.

Boehm, T.P., and K.R. Ihlanfeldt (1986), "The Improvement Expenditures of Urban Homeowners: An Empirical Analysis," *American Real Estate and Urban Economics Association Journal*, 14, 48–60.

Boero, G. (1990), "Comparing Ex-ante Forecasts from a SEM and VAR Model: An Application to the Italian Economy," *Journal of Forecasting*, 9, 13–24.

Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-day.

———, and G.C. Tiao (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70–79.

Brady, E.A. (1973), "An Econometric Analysis of the U.S. Residential Housing Market," in R.B. Ricks (ed.), *National Housing Models*, Lexington, MA: Lexington Books.

Carruthers, D. T. (1989), "Housing Market Models and the Regional System," *Urban Studies*, 26, 214–299.

Chang, I., G.C. Tiao, and C. Chen (1988), "Estimation of Time-Series Parameters in the Presence of Outliers," *Technometrics*, 30, 193-204.

Clayton Research Associates (1987), *Medium and Long Term Projections of Housing Requirements in Canada*, Ottawa: CMHC.

Cleveland, W.S. (1972), "The Inverse Autocorrelation of a Time Series and Their Applications," *Technometrics*, 14, 277– 298.

Cooper, R.L. (1972), "The Predictive Performance of Quarterly Econometric Models of the United States," in B.G. Hickman (ed.), *Econometric Models of Cyclical Behaviour*, New York: Columbia University.

De Gooijer, J.G., B. Abrahams, A. Gould, and L. Robinson, "Methods for Determining the Order of Autoregressive-Moving Average Process: A Survey," *International Statistical Review*, 53, 301–329.

Dent, W. (1977), "Computation of the Exact Likelihood Function of an ARIMA Process," *Journal of Statistical Computing and Simulation*, 5, 193–206.

Denton, F., A. Robb, and B.G. Spencer (1986), *Housing and Consumption*, Ottawa: CMHC.

Dicks, M.J. (1990), "A Simple Model of Housing Starts," Bank of England Discussion Paper No. 49.

Doan, T.A. (1990), *User's Manual RATS*, Evanston, IL: VAR Econometrics.

———, R.B. Litterman, and C.A. Sims (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

Dynarski, M., and S.M. Sheffrin (1985), "Housing Purchases and Transitory Income: A Study with Panel Data", *Review of Economics and Statistics*, 67, 195–204.

Eckstein, O. (1981), "Econometric Models for Forecasting and Policy Analysis: the Present State of the Art," in A.R. Sanderson (ed.), *DRI Readings in Econometrics*, New York: McGraw-Hill.

Emmi, P.C. (1984), "On the Duality of Opportunity and Price in a Multi-sectoral Model of an Urban Housing Market," *Environment and Planning*, A, 16, 615–628.

Engle, R.F., C.W.J. Granger, and D. Kraft, "Combining Competing Forecasts of Inflation Using a Bivariate ARCH Model," *Journal of Economic Dynamics and Control*, 8, 151–165.

Engle, R.F., D.M. Lilien, and M. Watson (1985), "A Dynamic Model of Housing Price Determination," *Journal of Econometrics*, 28, 307–326.

Evans, M.K., Y. Haitovsky, and G. I. Treyz (1972), "An Analysis of Forecasting Properties of U.S. Econometric Models," in B.G. Hickman, Ed. (1972), *Econometric Models of Cyclical Behavior*, National Bureau of Economic Research, New York: Columbia University Press.

Fackler, J.S., and S.C. Krieger (1986), "An Application of Vector Time Series Techniques to Macroeconomic Forecasting," *Journal of Business and Economic Statistics*, 4, 71–80.

Fair, R.C. (1979), "An Analysis of the Accuracy of Four Macroeconomic Models, *Journal of Political Economy*, 87, 701–718.

——, and R.J Shiller (1990), "Comparing Information in Forecasts from Econometric Models," *American Economic Review*, 80, 375–389.

Fisher, G. R., and M. McAleer (1981), "Alternative Procedures and Associated Tests for Significance for Non-nested Hypotheses," *Journal of Econometrics*, 16, 103–119.

Foot, D.K. (1985), *Housing Investment in Canadian Macroeconometric Models: An Evaluative Review*, A report prepared for the Research Division, CMHC, Ottawa: Canada Mortgage and Housing Corporation.

Fredland, J.E., and C.D. MacRae (1978), *Econometric Models of the Housing Sector: A Policy-Oriented Survey*, Washington, DC: The Urban Institute.

Friedman, B.M. (1989), "Effects of Monetary Policy on Real Economic Markets Activity" draft, Harvard University.

Fuller, W.A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.

Gahvari, F. (1986), "Demand and Supply of Housing in the U.S.: 1929-1978," *Economic Inquiry*, 24, 333–347.

Glennon, D. (1989), "Estimating the Income, Price and Interest Elasticities of Housing Demand," *Journal of Urban Economics*, 25, 219–229.

Gilderbloom, J.I. (1986), "An Analysis of Intercity Rents," *in P.L. Niebanck (ed.), The Rent Control Debate*, Chapel Hill: North Carolina University Press.

Goodman, J.L., Jr., and S. Gabriel (1987), *Forecasting Housing Construction: Lessons and Puzzles from Recent Years*, Working Paper No. 69, for Economic Activity Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System.

Goodwin, T.H. (1986), "The Impact of Credit Rationing on Housing Investment: A Multi-market Disequilibrium Approach," *International Economic Review*, 27, 445–464.

Grady Economics and Associates Ltd. (1985), *The State of the Art in Canadian Macroeconomic Modelling*, for the Economic Forecasting Division, Department of Finance.

Granger, G.W.J., and P. Newbold (1975), "Economic Forecasting: the Atheist's Viewpoint," in G.A. Renton (ed.), *Modelling the Economy*, London: Heinemann Educational Books.

——— and P. Newbold (1984), *Forecasting Economic Time Series*, New York: Academic Press.

———, P. Newbold, and R. Ramanathan (1984), "Improved Methods of Combining Forecasts," *Journal of Forecasting*, 3, 197–204.

Harmon, O.R. (1988), "The Income Elasticity of Demand for Single-Family Owner-Occupier Housing: An Empirical Reconciliation," *Journal of Urban Economics*, 24, 173–185.

Harvey, A.C. (1981), *Time Series Models*, Oxford: Philip Allan Publishers.

——— (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

———, and P.H.J. Todd (1983), "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study with Comments," *Journal of Business and Economic Statistics*, 1, 299–315.

Hatch, J.E. (1975), *The Canadian Mortgage Market*, Toronto: Ministry of Treasury, Economics and Intergovernmental Affairs, Queen's Printer for Ontario.

Hendershott, P.H. (1988), "Home Ownership and Real House Prices: Sources of Change, 1965–1985", *Housing Finance Review*, 7, 1–18.

——— and M.T. Smith (1988), "Housing Inventory Change and the Adjustment Role of Existing Structures, 1961–1985," *AREUEA Journal*, 16, 364–378.

Hillmer, S. (1984), "Monitoring and Adjusting Forecasts in the Presence of Additive Outliers," *Journal of Forecasting*, 3, 205–215.

Hillmer, S.C., and G.C. Tiao (1979), "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of American Statistical Association*, 74, 652–660.

Howrey, E.P., L.R. Klein, and M.D. McCarthy (1974), "Notes on Testing the Predictive Performance of Econometric Models," *International Economic Review*, 15, 366–383.

Hsiao, C. (1979), "Autoregressive Modelling of Canadian Money and Income," *Journal of American Statistical Association*, 74, 553–560.

——— (1981), "Autoregressive Modelling and Money-Income Causality Detection," *Journal of Monetary Economics*, 7, 85–106.

Jaffee, D.M., and K.T. Rosen (1978), "Estimates of Effectiveness of Stabilization Policies for the Mortgage and Housing Markets," *Journal of Finance*, 33, 933–946.

Jaffee, D., and K. Rosen (1979), "Mortgage Credit Availability and Residential Construction," *Brookings Papers on Economic Activity*, No.2, 333–366.

Jud, D.J., and J. Frew (1986), "Real Estate Brokers, Housing Prices, and the Demand for Housing," *Urban Studies*, 23, 21–31.

Kahn, G.A. (1989), "The Changing Interest Sensitivity of the U.S. Economy," *Economic Review*, 13–34.

Klein, L. R. (1950), *Economic Fluctuations in the United States, 1921-1941*, New York: John Wiley.

——— (1968), *An Essay on the Theory of Economic Prediction*, Helsinki: Yrjö Jahnssonin Säätiö.

——— (1971), "Forecasting and Policy Evaluation Using Large Scale Econometric Models: the State of the Art," in M.D.

Intriligator (ed.), *Frontiers of Quantitative Economics*, Amsterdam: North Holland.

——— (1982), "Comments and Discussion," following Christopher A. Sims, "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, No.1, 164.

Krumm, R (1987), "Intertemporal Tenure Choices," *Journal of Urban Economics*, 22, 263–275.

Litterman, R.B. (1979), "Techniques of Forecasting Using Vector Autoregressions," Working Paper 115. Federal Reserve Bank of Minneapolis.

——— (1980), "A Bayesian Procedure for Forecasting with Vector Autoregressions," Working Paper, MIT.

——— (1984), "Specifying Vector Autoregressions for Macroeconomic Forecasting," Staff Report 92. Federal Reserve Bank of Minneapolis.

——— (1986), "Forecasting with Bayesian Vector Autoregression – Five Years of Experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Liu, L.M., and D.M. Hannssens (1982), "Identification of Multiple Input Transfer Function Models," *Communications in Statistics*, 11, 297–314.

——— *et al.* (1986). *The SCA Statistical System: Reference Manual for Forecasting and Time Series Analysis*, DeKalb, IL: Scientific Computing Associates.

Lütkepohl, H. (1985), "Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process," *Journal of Time Series Analysis*, 6, 35–52.

Makridakis, S. et al. (1982), "The Accuracy of Extrapolation (Time Series Methods: Results of a Forecasting Competition," *Journal of Forecasting*, 1, 111–153.

Malpezzi, S., L. Ozanne, and T.G. Thibodeau (1987), "Microeconomic Estimates of Housing Depreciation," *Land Economics*, 63, 372–385.

Manning, C.A. (1986), "Intercity Differences in Home Price Appreciation," *Journal of Real Estate Research*, 1, 45–66.

McAleer, M., and M. H. Pesaran (1986), "Statistical Inference in Non-nested Econometric Models," *Applied Mathematics and Computation*, 271–311.

McNees, S.K. (1986), "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts," *Journal of Business and Economic Statistics*, 4, 5–23 (with comments).

——— (1988), "How Accurate Are Macroeconomic Forecasts?," *New England Economic Review*, July/August, 15–36.

Mills, T.C., and M.J. Stephenson (1987), "A Time Series Forecasting System for the UK Money Supply," *Economic Modelling*, 4, 355–369.

Mutchler, J.E., and L.J. Krivo (1989), "Availability and Affordability: Household Adaptation to a Housing Squeeze," *Social Forces*, 68, 241–261.

Muth, Richard (1986), "The Supply of Mortgage Lending," *Journal of Urban Economics*, 19, 88–106.

Muth, R.F. (1988), "Housing Market Dynamics," *Regional Science and Urban Dynamics Economics*, 18, 345–356.

Nelson, C.R. (1972), "The Prediction Performance of the FRB-MIT- Penn Model of the U.S. Economy," *American Economic Review*, 92, 902–917.

——— (1984), "A Benchmark for the Accuracy of Econometric Forecasts of GNP," *Business Economics*, 19, 52–58.

Newbold, P (1974), " The Exact Likelihood Function for a Mixed Autoregressive Moving Average Process," *Biometrika*, 61, 423–426.

——— (1983a), "ARIMA Model Building and the Time Series Analysis Approach to Forecasting," *Journal of Forecasting*, 2, 23–35.

——— (1983b), "Model Checking in Time Series Analysis," in A. Zellner (ed.), *Applied Time Series Analysis*, Washington, D.C.: U.S. Bureau of the Census (with comments by Ledolter and Pierce).

Olsen, E. (1987), "Demand and Supply of Housing Service: A Critical Survey of the Empirical Literature," in E.S. Mills (ed.), *Handbook of Regional and Urban Economics, Vol. II: Urban Economics*, Amsterdam: North-Holland.

Park, S.-B. (1989) *Applied Time Series Analysis for Business and Economics*, Mimeo., Carleton University.

Pankratz, A., (1983), *Forecasting with Univariate Box-Jenkins Models*, New York: John Wiley.

——— (1989), "Time Series Forecasts and Extra-Model Information," *Journal of Forecasting*, 8, 75–83.

Pelser, R.B., and L.B. Smith (1985), "Homeownership Returns, Tenure Choice and Inflation," *AREUEA Journal*, 13, 343–360.

Pesaran, M. H. (1974), "On the general problem of model selection," *Review of Economic Studies*, 43, 153–171.

Poskitt, D.S. (1989), "A Method for the Estimation and Identification of Transfer Function Models," *Journal of the Royal Statistical Society*, B, 51, 29–46.

———, and A.R. Tremayne (1981), "An Approach to Testing Linear Time Series Models," *Annals of Statistics*, 5, 974–986.

Puri, A., and J. Van Lierop (1988), "Forecasting Housing Starts," *International Journal of Forecasting*, 4, 125–134.

Quigley, J.M. (1987), "Interest Rate Variations, Mortgage Prepayments and Household Mobility," *Review of Economics and Statistics*, 69, 636–643.

Quinn, B.G. (1988), "A Note on AIC Order Determination for Multivariate Autoregressions," *Journal of Time Series Analysis*, 9, 241–245.

Raynauld, J., and J. Simonato (1988), "Seasonal BVAR Models," *Proceedings of American Statistical Association Business and Economic Statistics Section*, 319–324.

——— (1990), "Seasonal BVAR Models: A Search Along Some Time Domain Priors," Paper presented at the CRDE/*Journal of Econometrics* Conference on Seasonality and Econometric Models.

Ricks, R. B. (1973), *National Housing Models: Application of Econometric Techniques to Problems of Housing Research*, Lexington, MA: Lexington Books.

Rosen, K.T, and L.B. Smith (1986), "The Resale Housing Market", *AREUEA Journal*, 14, 510–524.

Rosen, S., and R. Topel (1986), *A Time-Series Model of Housing Investment in the U.S.*, Working Paper No. 1818, New York: National Bureau of Economic Research.

Rosenthal, L. (1986), "Regional House Price Interactions in the U.K., 1971–1981: A Cross-Spectral Analysis," *Applied Economics*, 18, 1011–1023.

Runkle, D.E. (1987), "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics*, 5, 437–454 (with comments).

Shibata, R. (1985), "Various Model Selection Techniques in Time Series Analysis," in E.J. Hannan, P.R. Krishnaiah, and M.M. Rao (eds.), *Handbook of Statistics*, Vol. 5, New York: Elsevier Science Publishers, 179–187.

Sims, C.A. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

——— (1982), "Policy Analysis with Economic Models," *Brookings Papers on Economic Activity*, No. 1.

Sklarz, M.A., N.G. Miller, and W. Gersch (1987), "Forecasting Using Long-Order Autoregressive Processes: An Example Using Housing Starts," *AREUEA Journal*, 15, 374–388.

Smith, L.B. (1974), *The Postwar Canadian Housing and Residential Mortgage Markets and the Role of Government*, Toronto, University of Toronto Press.

———, L.B., K.T. Rosen, and G. Fallis (1988), "Recent Developments in Economic Models of Housing Markets," *Journal of Economic Literature*, 26, 29–64.

———, and G.R. Sparks (1970), "The Interest Sensitivity of Canadian Mortgage Flows," *Canadian Journal of Economics*, 3, 407–421.

Stansell, S.R., and A.C. Mitchell (1985), "The Impact of Credit Rationing on the Real Sector: A Study of the Effect of Mortgage Rates and Terms on Housing Starts," *Applied Economics*, 17, 781–800.

Stekler, H.O. (1968), "Forecasting with Econometric Models: An Evaluation," *Econometrica*, 36, 437–463.

Stiltner, K.R., and D.R. Barton (1990), "Econometric Models and Construction Forecasting," U.S. Department of Commerce *Construction Review*, 10–20.

Stokes, E. (1987), *Canadian Economic and Fiscal Model (CEFM): Basic Structure and Specification*, Economic Analysis and Forecasting Division, Department of Finance.

Thom, R. (1985), "The Relationship between Housing Starts and Mortgage Availability," *Review of Economics and Statistics*, 67, 693–696.

Todd, R.M. (1984), "Improving Economic Forecasting with Bayesian Vector Autoregression," Federal Reserve Bank of Minneapolis *Quarterly Review*, 18–29.

Topel, R., and S. Rosen (1988), "Housing Investment in the United States," *Journal of Political Economy*, 96, 718–740.

Tsay, R.S., and G.C. Tiao (1984), "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models," *Journal of American Statistical Association*, 79, 84–96.

Turner, M.A., and R.J. Struyk (1985), *Urban Housing in the 1980s: Markets and Policies*, Washington, DC: Urban Institute Press.

Vandaele, W. (1983), *Applied Time Series and Box-Jenkins Models*, New York: Academic Press.

Wang, M., and H. Ma (1981), "Multiple Time Series Models for Housing and the Mortgage Rate," *Proceedings of the American Statistical Association Business and Economic Statistics Section*, 418–423.

Weicher, J.C., and T.G. Thibodeau, T.G. (1988), "Filtering and Housing Markets: An Empirical Analysis," *Journal of Urban Economics*, 23, 21–40.

Williams, R.A. (1984), "An Australian Housing Model," *Economic Record*, 60, 143–155.

Zellner, A., and F. Palm (1979), "Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, 2, 17–54.

Zumpano, L.V., P. Rudolph, and D.C. Cheng (1982), "The Demand and Supply of Mortgage Funds and Mortgage Loan Terms," *AREUEA Journal*, 14.