Simulating A Local Housing Market: An Empirical Test[#]

Terry Robertson*

I. INTRODUCTION

The housing segment of the national economy is an important sector in most national forecasting models.¹ While there has been little effort made to model the housing portion of a local economy, it seems extremely unlikely that indices applicable to national housing markets would be related to local indices on a consistent basis. Local market analysis has thus far been confined to either cross section studies or analyzing results from models developed to predict the prices of individual units.

Two national housing models and the housing segment of eight models of the national economy were the focus of review by G. Fromm.² He found numerous shortcomings, notably the failure of national models to reflect local market imperfections. Fromm concluded that the national models he studied had proven unreliable for forecasting, since they were generally lacking variables encompassing the effects of:

- (1) mortgage rates reflecting fees and points;
- (2) federal housing subsidies;
- (3) institutional equity participations;
- (4) housing starts expressed both in units and dollars;
- (5) changes in construction costs over time;
- (6) commercial bank mortgages
- (7) vacancy rates;
- (8) dissaggregation of single and multiple housing units;
- (9) public housing; and
- (10) local market fluctuations.

Another model reviewed by Fromm was that of R. C. Fair.³ Fair used constraints to reflect the influence of credit availability and the lag time in production to the response in demand for housing. Although such disequilibrium effects are widely held to be important, the results were no better than if equilibrium had been assumed. Fair also commented that equations for estimating housing starts have had very limited success.

A more recent study by Von Furstenberg and Herr concluded that short

*Associate Professor of Finance, University of Oklahoma.

[#]This study was partially funded by a grant from the metropolitan Dade County Planning Department, Dade County, Florida.

term levels of housing demand were directly influenced by the supply of mortgage credit, while housing starts were influenced by an excess supply of or demand for living units.⁴

The purpose of this research is to develop an econometric model of the Dade County, Florida housing market which will: (1) identify and measure the influence of major determinants of housing supply and demand in Dade County, and (2) provide a predictive tool to help planning agencies, lending institutions, and other participants in the local housing market foresee market performance based on various assumptions. This first section briefly examined the literature on housing models. The model is described and tested in Section II. Section III contains the summary and conclusions.

II. FORMULATION OF THE MODEL

The Dade County model contains several variables suggested by Fromm.⁵ The mortgage rate used is the effective rate reflecting fees and points; housing starts are estimated both in units and dollar amounts; per capita time deposits in institutions are used to reflect area wealth and availability of credit; the housing market is segmented into the single family, condominium, and apartment components; and public housing starts are an exogenously determined policy variable that will affect the private rental market. The model of the Dade County housing market is based on quarterly rather than annual data.⁶ The general form of the model is:

$$(1) \quad BY_t + CZ_t = U_t;$$

where: B = a nonsingular $\underline{G} \times \underline{G}$ matrix of coefficients of the endogenous variables;

Y_t = a vector of $\underline{G}$ endogenous variables in time period t ;

C = a $\underline{G} \times \underline{K}$ matrix of coefficients of the exogenous variables;

Z_t = a vector of $\underline{K}$ exogenous variables in time period t ; and

U_t = a vector of $\underline{G}$ random errors in time t .

Each equation within the model is in the form:

$$(2) \quad Y = a + a_1x_1 + a_2x_2 + \dots + a_nx_n;$$

where: Y = the dependent endogenous variable;

$a_1 \dots a_n$ = the coefficients for the independent endogenous or exogenous variables 1 thru n ; and

$x_1 \dots x_n$ = the independent endogenous or exogenous variables 1 thru n.

The formulation is simultaneous and incorporates feedback loops for co-determining relationships among the exogenous and endogenous variables. Since the model is specified as being linear, the reduced form equation can be mathematically derived for expected outcomes. However, as models become larger this manipulation becomes quite cumbersome. The Gauss-Seidel iterative technique provides an alternative solution to approximate the reduced form estimates.⁷

It is assumed that the market can achieve approximate levels of equilibrium through adjustments in prices, vacancies, and housing starts. From a planning standpoint, it is more important to predict a normal level of housing starts to maintain equilibrium, or a period of time for a market to return to an approximate state of equilibrium. Thus the housing starts portion of this model is formulated with "expected" or "warranted" starts given prices, rents and population changes rather than a prediction of actual starts, which Fair indicated has had limited success. The "starts" segment of the model is specified as:⁸

$$(3) \quad STF = STFA/SFNAVE;$$

$$(4) \quad STAGA = B_0 + B_1 (STA) + B_2 (YA);$$

$$(5) \quad STA = B_0 + B_1 (ST);$$

$$(6) \quad STFA = STA - STAGA;$$

$$(7) \quad ST = STF + STAG;$$

where: STF = single family housing starts;

STFA = value of single family housing starts;

SFNAVE = single family new house, average price;

STAGA = value of multi-family starts;

STA = value of all residential starts;

YA = Dade county total personal income;

ST = all residential starts; and

STAG = multi-family unit starts.

A second important segment of the model deals with financing and credit supply. The effective cost of mortgage debt is related to the long-

term bond yields and income while time deposits are related to the implied rent levels for all types of housing. Implied or quasi-rent levels are used as a proxy for wealth and therefore deposits; where quasi-rent is defined as a constant 1/100th of the average price of new units. The individual equations are specified as:

$$(8) \quad \text{MEBR} = B_0 + B_1 (\text{RAA}) + B_2 (\text{ILE}) + B_3 (\text{YA});$$

$$(9) \quad \text{MEFF} = \text{MEBR} (\text{BND});$$

$$(10) \quad \text{DEP} = B_0 + B_1 (\text{RNA}) + B_2 (\text{RAA}) + B_3 (\text{RCNA});$$

where: MEBR = ratio of mortgage rate to bond rate;

RAA = average monthly apartment rent;

ILE = land cost index;

MEFF = effective mortgage rate, new homes;

BND = Moody's AAA bond yield, quarterly average;

DEP = per capita time deposits;

RNA = quasi-rent for all new housing; and

RCNA = quasi-rent for new condominiums.

The last major segment to be discussed is comprised of the price or rent equations for single family units, condominiums, and apartment units. The pricing for each unit type is affected by prices or rents for the other unit types and thus allows for the allocation-purchase decisions on the basis of relative cost. The individual specifications are:

$$(11) \quad \text{SFNAVE} = B_0 + B_1 (\text{PIH}) + B_2 (\text{CCI}) + B_3 (\text{DOH});$$

$$(12) \quad \text{SFUAVE} = B_0 + B_1 (\text{PRI}) + B_2 (\text{DOH}) + B_3 (\text{SFNAVE});$$

$$(13) \quad \text{SCUAVE} = B_0 + B_1 (\text{ILE}) + B_2 (\text{RNA});$$

$$(14) \quad \text{RAA} = B_0 + B_1 (\text{CCI}) + B_2 (\text{SFCUA}) + B_3 (\text{DEP});$$

where: PIH = U.S. consumer price index, all items except shelter;

CCI = construction cost index;

DOH = change in occupied housing units;

SFUAVE = average price, single family used house;

PRI = prime interest rate;

SCUAVE = average price, used condominiums;

RNA = quasi-rent, all new housing; and

SFCUA = dollar sales, used homes and condominiums.

Large-scale models must be tested for reliability. Individual equations may lose the ability to predict when combined into a dynamic framework. Two measures of accuracy based on predicted values compared to actual values are used with the model. The mean-absolute-percentage-error (MAPE) and root-mean-squared-percentage-error (RMSPE) are computed as:

$$(15) \quad \text{MAPE} = \frac{N \sum_{i=1}^N \left| \frac{a_{ij}^p - a_{ij}^a}{a_{ij}^a} \right|}{N} ; \text{ and}$$

$$(16) \quad \text{RMSPE} = \left[\frac{N \sum_{i=1}^N \left| \frac{a_{ij}^p - a_{ij}^a}{a_{ij}^a} \right|^2}{N} \right]^{1/2}$$

where: a_{ij}^p = predicted value of j th variable in time period i ;

a_{ij}^a = actual value of j th variable in time period i ; and

N = number of time periods being simulated.

The acceptable magnitude of error is open to debate. Some authors such as Glickman state that three percent or less is adequate.⁹ That level, however, is for a large scale annual model of an area; the aggregate economy would probably not be as volatile as the housing sector itself. The annualizing of data should also lead to more accurate estimates. This is not necessarily a contradiction of Grebler and Maisel. The fact that annual

data would obscure important shifts does not imply that these shifts would not be more accurately estimated.

Table 1 contains the MAPE and RMSPE statistics for the model over the 1973/1 to 1973/3 time period while Table 2 contains the distribution of the error statistics. The data in Table 2 indicate that approximately 60 percent of the errors for the model are less than 10 percent. The errors associated with the equations for average prices and quasi-rent for single and multi-family units range from .87 to 5.58 percent. The pricing equations seem to be quite accurate, which would be expected from the literature.¹⁰ The

TABLE 1
MAPE and RMSPE Statistics
13 Periods: 1973-01 to 1975-03

	MAPE	RMSPE
YA	.92	1.19
MEBR	3.51	4.14
DEP	2.75	3.74
MOBL	1.73	2.55
RAA	1.72	1.97
SCUAVE	.87	1.12
SCUA	11.12	13.08
SFUAVE	3.56	4.31
SFNAVE	5.58	7.27
SFUA	18.14	21.30
CCI	1.20	1.72
STF	40.13	42.19
ST	11.92	13.23
STAGA	15.10	18.58
STFA	16.81	21.03
STA	12.52	15.11
OH	.63	.73
SNCC	5.99	8.33
SCNA	12.33	22.47
ILE	.17	.21
RNA	18.61	23.86
P2264	.07	.08

TABLE 2
Distribution of MAPE
and RMSPE Errors

Range	MAPE		RMSPE	
	Number	%	Number	%
0 - 5	11	50.0	11	50.0
6 - 10	2	9.1	2	9.1
10 - 15	4	18.2	2	9.1
15 - 20	4	18.2	2	9.1
20 +	1	4.5	5	22.7

Source: Table 1

error statistics for the actual housing starts segment, ranging from 11.92 to 40.13, percent are very large, which would also be expected from the literature.¹¹ Two different definitions of starts are being used; prediction of "warranted" starts versus actual starts.

III. SUMMARY AND CONCLUSIONS

The model confirms the findings of Fair regarding the inability to predict housing starts; however the prediction of prices/rents was found to be relatively accurate. This would be expected for several reasons. Prices tend to react fairly rapidly to both local and national economic conditions. Starts, on the other hand, are a function of individual developer expectation about market conditions in the future, that is, when units are completed. Further, individual developers may not consider current vacancies to be important for their start decision since they may believe their unit unique and able to capture a sufficient portion of the market to "sell out" the project. Lastly, it would be more important from a planning standpoint to predict a normal or expected level of starts each period rather than actual starts.

While the error statistics in the model exceed those in other modeling efforts, the model has demonstrated its ability to predict directional changes in the endogenous variables. The model was able to distinguish changes in market direction within one quarter of when the change actually occurred.

Both lenders and planners can utilize simulation models of this type. Planners are normally faced with providing information to local government zoning bodies when requests for rezoning are received. The necessity for additional land in various residential zoning classifications can be substantiated by the model. Lending institutions may use a model of this type in order to test macromarket conditions for housing development.

While projects accepted would still require submarket studies, as currently performed, general analysis at an aggregate level would possibly eliminate overbuilding such as that which occurred in South Florida.

FOOTNOTES

¹See F. Strobel and W. Toal, "Econometric Models: What They Are And What They Say For 1971," *Monthly Review*, Federal Reserve Bank of Atlanta, March, 1971, pp. 42-51.

²G. Fromm, "Econometric Models of the Residential Construction Sector," in *National Housing Models*, ed. R. Bruce Ricks, Lexington, Mass., D. C. Heath, 1973, pp. 125-153.

³R. C. Fair, "Methods of Estimation for Markets in Disequilibrium," *Econometrica*, 40: 497-514 (May 1972); "Disequilibrium in Housing Models," *Journal of Finance*, 27: 207-21 (May 1972); and "Monthly Housing Starts," in *National Housing Models*, ed. R. Bruce Ricks, Lexington, Mass., D. C. Heath, 1973, pp. 69-92.

⁴G. VonFurstenberg and B. Herr, "The Normal Demand for Housing Completions and the Stock Adjustment Process in Housing Starts," *Journal of Economics and Business*, 27: 131-40 (Winter 1975).

⁵Fromm, *op. cit.*

⁶L. Grebler and S. J. Maisel, "Determinants of Residential Construction: A Review of Present Knowledge," Impacts of Monetary Policy, studies prepared for the Commission on Money and Credit, Englewood Cliffs, New Jersey, Prentice-Hall, 1973.

⁷See N. Glickman, "Regional Forecasting and Policy Analysis: A Macroeconometric Approach." Paper presented at 1973 meeting of Econometric Society, New York; and R. Preston, "The Wharton Annual and Industry Forecasting Model," *Studies in Qualitative Economics*, no. 7, 1972.

⁸The coefficients for the structural equations were estimated using OLS regression. The r^2 standard errors, Durbin-Watson statistic, f-value, and t-values for each equation are available upon request.

⁹Glickman, *op. cit.*

¹⁰Fair, *op. cit.*

¹¹Fair, *op. cit.*

REFERENCES

- Frank de Leeuw, "The Demand for Housing: A Review of Cross Section Evidence," *Review of Economics and Statistics*, 53: 1-10 (February 1971).
- J. E. Gelfand, "Mortgage Credit and Lower-Middle Income Housing Demand," *Land Economics*, 46: 163-70 (May 1970).
- David S. Huang, "A Study of the Market for New Housing Units," *1969 Proceedings of the Business and Economic Statistics Section*, Washington, D.C., American Statistical Association, pp. 702-12.
- Jobs, People, Land*, Bay Area Simulation Study (BASS), Special Report No. 6, By the Staff, Center for Real Estate and Urban Economics, Institute of Urban and Regional Development, Berkeley, University of California, 1968.
- J. H. Kalchbrenner, "A summary of the Current Financial Intermediary, Mortgage and Housing Sectors of the FRB-MIT-Penn Econometric Models," in *National Housing Models*, ed. R. Bruce Ricks, Lexington, Mass., D. C. Heath, 1973, pp. 93-124.