

8

Monthly Housing Starts

8.1 Introduction

In order to use the housing investment equation in the money GNP sector for forecasting purposes, housing starts have to be explained within the model or else forecast exogenously. The theoretical and empirical work explaining the level of housing starts is still in its infancy [35, 40], and only limited success has been achieved in developing reliable housing starts equations. The approach taken in this study is to treat the housing market as a market that is not always in equilibrium and to estimate supply and demand schedules of housing starts under this assumption. It seems to be a widespread view that the housing and mortgage market is not always in equilibrium,¹ and one of the advantages of the technique used in this chapter is that this view can be tested.

The outline of this chapter is as follows. In the next section the basic model of the housing market is presented and discussed. The technique that has been used to estimate the model is then described in Section 8.3. The technique is based on the work in Fair and Jaffee [20] and Fair [16]. The data are discussed in Section 8.4, and the results of estimating the model are presented in Section 8.5. The chapter concludes with a discussion in Section 8.6 of how the housing starts equations can be used for forecasting purposes.

8.2 A Model of the Housing and Mortgage Market

The housing and mortgage market is a difficult market to specify. The interaction between the financial (mortgage) side of the market and the real side of the market is complex, and it does not as yet appear to be well understood. In this section an attempt is made at a reasonable specification of the housing and mortgage market and of the interaction between the two sides of the market, but a number of simplifying assumptions have been made in order to keep the analysis as tractable as possible. To begin with, the present model is concerned only with the market for *new* houses (i.e., housing starts) and for the mortgage funds associated with these houses.

¹ See, for example, de Leeuw and Gramlich [8], pp. 482–483.

Looking first at the demand side of the market, let HS_t^D denote the demand for housing starts (new houses) during period t . Then the demand schedule for housing starts is taken to be

$$HS_t^D = f(X_t^D, \varepsilon_t^D), \quad (8.1)$$

where X_t^D denotes the vector of variables that determine HS_t^D and where ε_t^D is an error term. The variables that have been included in the X_t^D vector in the present model will be discussed below, but in general the X_t^D vector should include such variables as population, income, the number of houses already in existence, the purchase price of new houses, and the cost of obtaining mortgage funds to finance the purchase of a house (i.e., the mortgage rate).

An important simplifying assumption will now be made concerning the relationship between the demand for housing starts and the demand for the mortgage funds associated with these starts. Let $MORT_t^D$ denote the demand for mortgage funds associated with HS_t^D . Then it is assumed that

$$\frac{MORT_t^D}{HS_t^D} = a_0 + a_1 t, \quad (8.2)$$

where t is a time trend. Equation (8.2) states that the ratio of the demand for new mortgage funds to the demand for housing starts is equal to some constant value plus a time trend. The time trend is designed to pick up any trend increase in the average size of mortgages per housing start. The assumption made in (8.2) is admittedly a highly simplifying one, since the mortgage-fund-housing-starts ratio is likely to fluctuate in the short run in response to such things as the mortgage rate, but for purposes of this study, ignoring these fluctuations may not be too serious. Equations (8.1) and (8.2) imply that the demand for new mortgage funds is, aside from a trend term, merely a function of the variables in X_t^D and the error term ε_t^D .

Turning next to the supply side of the market, let HS_t^S denote the supply of housing starts during period t . Then the supply schedule of housing starts is taken to be

$$HS_t^S = g(X_t^S, \varepsilon_t^S), \quad (8.3)$$

where X_t^S denotes the vector of variables that determines HS_t^S and where ε_t^S is an error term. In general, the X_t^S vector should include such variables as the price of houses, the cost of building houses (materials and supplies plus labor costs), and the cost of short-term credit. Home builders, in other words, are likely to decide how many new houses to build on the basis of the price of houses vis-à-vis their building cost and on the basis of the cost of short-term credit. Note that it is the cost of short-term credit that is likely to affect the supply of housing starts and not the cost of long-term credit, as reflected in,

say, the mortgage rate. Home builders generally need a mortgage commitment from one of the financial intermediaries before they can get short-term loans from commercial banks; but, providing that commitments are available, the mortgage rate associated with these commitments should not directly concern them. The mortgage cost is incurred by the person who buys the house and takes out the mortgage, not by the person who builds the house.

Finally, let $MORT_t^S$ denote the supply of new mortgage funds during period t . Then the supply schedule of mortgage funds is taken to be

$$MORT_t^S = h(Z_t^S, \eta_t^S), \quad (8.4)$$

where Z_t^S denotes the vector of variables that determines $MORT_t^S$ and where η_t^S is an error term. The variables that have been included in the Z_t^S vector in this study will be discussed below; but in general the vector should include such variables as deposit flows into financial intermediaries, the mortgage rate, and interest rates on competing assets. Since mortgages are supplied primarily by financial intermediaries, deposit flows into these intermediaries should have a positive effect on the supply of mortgages. Also, for a given flow of deposits, financial intermediaries are likely to put more of the flow into the mortgage market the higher is the mortgage rate relative to other interest rates.

The demand and supply sides of the housing and mortgage market differ in that the people who demand new houses are essentially the same people who demand mortgage funds, whereas the people who supply (build) new houses are in general not the same people who supply mortgage funds. There are thus three groups of people or institutions under consideration: the consumers, the home builders, and the financial intermediaries. If the housing and mortgage market were always in equilibrium, then it would be the case that:

$$HS_t = HS_t^D = HS_t^S, \quad (8.5)$$

and

$$MORT_t = MORT_t^D = MORT_t^S, \quad (8.6)$$

where HS_t is the actual number of housing starts during period t and $MORT_t$ is the actual value of new mortgage funds during period t . In equilibrium, the purchase price of houses would clear the housing side of the market, as in (8.5), and the mortgage rate would clear the mortgage side of the market, as in (8.6). Note that the assumption made in (8.2) above implies that in equilibrium,

$$\frac{MORT_t}{HS_t} = a_0 + a_1 t.$$

If the housing and mortgage market is not always in equilibrium, then (8.5) and (8.6) obviously do not always hold, and the question arises as to how the disequilibrium aspects of the market should be specified. In this study the specification is as follows. It is first assumed that the actual ratio of new mortgage funds to housing starts is always equal to $a_0 + a_1t$. It was seen above that, given (8.2), the ratio is equal to $a_0 + a_1t$ in equilibrium; and it is now assumed that the actual ratio is equal to $a_0 + a_1t$ even if the market is not in equilibrium. Because of this assumption, the supply of mortgages from the financial intermediaries in (8.4) can be translated into an equivalent supply of housing starts. The equivalent supply is $MORT_t^S/(a_0 + a_1t)$. There are thus two supply schedules of housing starts under consideration—the supply schedule from the home builders and the supply schedule from the financial intermediaries. It is finally assumed that the observed quantity of housing starts is equal to the minimum of the *ex ante* demand and supply schedules:

$$HS_t = \min \left\{ HS_t^D, HS_t^S, \frac{MORT_t^S}{a_0 + a_1t} \right\}. \quad (8.7)$$

Equation (8.7) implies that there are three possible constraints in the housing market. Either demand is the constraint (HS_t^D is the minimum) so that home builders and financial intermediaries go unsatisfied at prevailing prices, or supply from the home builders is the constraint (HS_t^S is the minimum) so that demanders and financial intermediaries go unsatisfied, or supply from the financial intermediaries is the constraint ($MORT_t^S/(a_0 + a_1t)$ is the minimum) so that demanders and home builders go unsatisfied. It appears to be commonly accepted that most of the “supply” constraint in the housing market comes from the financial sector, and thus as a simplifying approximation in this study, HS_t^S is assumed always to be greater than or equal to the minimum of HS_t^D and $MORT_t^S/(a_0 + a_1t)$. This assumption simplifies matters in that the supply schedule of home builders in (8.3) does not have to be specified. In a more detailed study of the housing market it would, of course be desirable to specify and estimate the home builders’ side of the market as well.

What remains to be done, then, is to specify equations (8.1) and (8.4). With respect to equation (8.1), the demand for housing starts is assumed to be a function of (1) population growth and trend income, both of which are approximated by a time trend; (2) the number of houses in existence or under construction during the previous month; (3) the mortgage rate lagged two months; and (4) seasonal factors.

Let H_t denote the number of houses in existence or under construction during month t and let HS_t continue to denote the number of housing starts during month t . Then H_t is approximated as follows. It is assumed

that the number of houses removed (i.e., destroyed) each month is constant from month to month, which implies that

$$HS_t = H_t - H_{t-1} + b_0, \quad (8.8)$$

where b_0 is the constant number of removals each month. Equation (8.8) then implies that for any base period 0:

$$H_t = H_0 + \sum_{i=1}^t HS_i - b_0 t, \quad (8.9)$$

where H_0 is the number of houses in the base period. In other words, the number of houses at the end of month t is equal to the sum of past housing starts less the sum of past removals, the sum of past removals being approximated by a time trend, as implied by the assumption in equation (8.8).

With respect to seasonal factors, the housing starts series does have a pronounced seasonal pattern in it, due in large part to the weather, and in an attempt to account for this pattern eleven seasonal dummy variables were included in the equation.² An alternative approach would have been to use the seasonally adjusted housing starts series that is published by the Department of Commerce, but the Department of Commerce does not adjust the series for the number of working days in the month. This causes the month-to-month changes in the seasonally adjusted series to be more erratic than is really warranted. In an attempt to account in this study for the influence of the number of working days in the month on the number of housing starts for that month, a working-day variable was included in the equation. The variable was constructed by adding up all of the weekdays in the month less any holidays that fell on these days. The holidays were excluded in the following manner. One day was always excluded for January, September, November, and December, and one day was also excluded for May and July unless May 30 or July 4 respectively fell on a Saturday. The data on this variable, denoted as W_t , are presented in Appendix A.

The demand schedule for housing starts is thus taken to be

$$HS_t^D = \sum_{I=1}^{11} d_I DI_t + d_{12} W_t + b_1 H_{t-1} + b_2 t + b_3 RM_{t-2} + \varepsilon_t^D, \quad (8.10)$$

where DI_t is the seasonal dummy variable for month I , $b_2 t$ is the trend term,

² Dummy variable 1 being equal to one in January, minus one in December, and zero otherwise; dummy variable 2 being equal to one in February, minus one in December, and zero otherwise; and so on. A constant term was included in the equation, which is the reason why only eleven dummy variables were included. The values for December were set equal to minus one instead of zero so that the seasonal factors could be more readily identified from the estimates of the coefficients of the dummy variables.

and RM_{t-2} is the mortgage rate lagged two months. Using the definition of H_t in equation (8.9), equation (8.10) becomes

$$HS_t^D = \sum_{t=1}^{11} d_t DI_t + d_{12} W_t + (b_1 H_0 + b_1 b_0) + b_1 \sum_{i=1}^{t-1} HS_i + (b_2 - b_1 b_0)t + b_3 RM_{t-2} + \varepsilon_t^D, \quad (8.11)$$

which introduces the constant $b_1 H_0 + b_1 b_0$ in the equation and changes the interpretation of the coefficient of the time trend. The data that have been used to estimate equation (8.11) will be discussed below.

It should be noted that the purchase price of houses has not been included as an explanatory variable in the demand equation. Theoretically the price of houses (or, more specifically, the price of houses deflated by some general price index) should be included in the equation, but this was not done for the work here because of the difficulty that would be involved in forecasting the price of houses exogenously. To the extent that the influence of the (relative) price of houses on the demand for housing starts is not picked up by the time trend in equation (8.11), the equation is misspecified, but for short-run forecasting purposes this misspecification is not likely to be too serious. It should also be noted that various lagged values of the mortgage rate were tried in the work below, and the mortgage rate lagged two months gave the best results for the demand equation.

With respect to equation (8.4), the supply of new mortgage funds is assumed to be a function of (1) lagged deposit flows into Savings and Loan Associations (SLAs) and Mutual Savings Banks (MSBs), (2) lagged borrowings by the SLAs from the Federal Home Loan Bank (FHLB), (3) the mortgage rate lagged one month, and (4) seasonal factors. Let DSF_t denote the flow of private deposits into SLAs and MSBs during month t , and let DHF_t denote the flow of borrowings by the SLAs from the FHLB during month t . Various lags and moving averages of DSF and DHF were tried in the work below, and the best results were achieved by using the six-month moving average of DSF lagged one month (denoted as $DSF6_{t-1}$) and the three-month moving average of DHF lagged two months (denoted as $DHF3_{t-2}$). The results were not very sensitive, however, to slightly different specifications. The six-month moving average of DSF has the advantage of eliminating the monthly fluctuations in the series due to the quarterly interest payments by the SLAs and MSBs and the switching of funds at the beginning of each quarter. The current and various lagged values of the mortgage rate were also tried in the supply equation, and the one-month lagged value gave the best results. Seasonal factors were assumed to enter the supply equation in the same way in which they entered the demand equation. The supply of new

mortgage funds is thus taken to be

$$MORT_t^S = \sum_{i=1}^{11} d'_i DI_t + d'_{12} W_t + c_1 DSF6_{t-1} + c_2 DHF3_{t-2} + c_3 RM_{t-1} + \eta_t^S. \quad (8.12)$$

The equivalent supply of housing starts from the financial sector was defined above to be $MORT_t^S/(a_0 + a_1 t)$. Let HS_t^{FS} denote this equivalent supply. Then $MORT_t^S = (a_0 + a_1 t)HS_t^{FS}$. As a further simplifying assumption, $t \cdot HS_t^{FS}$ will be approximated by $t + HS_t^{FS} + c_0$, where c_0 is a constant. This then implies that

$$HS_t^{FS} = \frac{1}{a_0 + a_1} (-a_1 c_0 - a_1 t + MORT_t^S).$$

Using equation (8.12) and ignoring the $1/(a_0 + a_1)$ multiplier, the equation determining HS_t^{FS} can thus be written

$$HS_t^{FS} = -a_1 c_0 - a_1 t + \sum_{i=1}^{11} d'_i DI_t + d'_{12} W_t + c_1 DSF6_{t-1} + c_2 DHF3_{t-2} + c_3 RM_{t-1} + \eta_t^S. \quad (8.13)$$

In other words, equation (8.12) explaining the supply of mortgage funds can be transformed into an equation explaining the equivalent supply of housing starts from the financial sector. The latter differs from equation (8.12) only in that a constant term and a time trend have been added to the equation. The time trend is designed to pick up any trend in the mortgage-fund-housing-starts ratio.

Equations (8.11) and (8.13) thus determine the demand and supply of housing starts respectively, and the model is closed by equation (8.7), which from the above assumption about home builders can be written

$$HS_t = \min\{HS_t^D, HS_t^{FS}\}. \quad (8.14)$$

The technique that was used to estimate equations (8.11) and (8.13) will now be discussed.

8.3 The Estimation Technique³

In Fair and Jaffee [20] four techniques for estimating disequilibrium markets were developed. Three of the techniques were designed to separate the sample period into demand and supply regimes so that each schedule could be fitted

³ Some of the discussion in this section follows closely the discussion in [20], Section II.

against the observed quantity for the sample points falling within its regime. The fourth technique was designed to adjust the observed quantity for the effects of rationing so that both schedules could be estimated over the entire sample period using the adjusted quantity. The fourth technique has been used in this study, and it will be briefly outlined below. All four of the techniques developed in [20] were used to estimate the present model, and two of the four techniques gave good results. These results are presented and compared in [20], Section III. The fourth technique was chosen to be used in this study because it appeared to be somewhat more suited for forecasting purposes.

The technique used here is based on the following assumption about how prices (or, in this case, interest rates) are determined:

$$\Delta RM_t = q(HS_t^D - HS_t^{FS}), \quad 0 \leq q \leq \infty. \quad (8.15)$$

Equation (8.15) states that the change in the mortgage rate is directly proportional to the amount of excess demand in the market. q equal to zero is the polar case of no adjustment, and q equal to ∞ is the polar case of perfect adjustment. Equation (8.15) is consistent with many theories of dynamic price setting behavior.

Solving equation (8.15) for excess demand yields:

$$HS_t^D - HS_t^{FS} = \frac{1}{q} \Delta RM_t. \quad (8.16)$$

If q can be estimated, then the actual amount of excess demand or supply can be determined directly from the change in the mortgage rate, and thus both the demand and supply schedules can be estimated over the entire sample period. The procedure described below simultaneously estimates q and the parameters of the two schedules.

First consider a period of rising rates. From equation (8.16) it is known that this will be a period of excess demand; and thus, from equation (8.14), the observed quantity will equal the supply. Consequently, the supply function can be directly estimated using the observed quantity as the dependent variable:

$$HS_t = HS_t^{FS}, \quad \Delta RM_t \geq 0, \quad (8.17)$$

where HS_t^{FS} is given in (8.13). Furthermore, because the supply equals the observed quantity, equation (8.16) can be rewritten as

$$HS_t = HS_t^D - \frac{1}{q} \Delta RM_t, \quad \Delta RM_t \geq 0, \quad (8.18)$$

where HS_t^D is given in (8.11). Thus the parameters of the demand function can also be estimated, using the observed quantity as the dependent variable,

as long as the change in the mortgage rate is included in the equation as an implicit adjustment for the amount of rationing.

In periods of falling rates essentially the same principles apply. The supply and demand functions will then be estimated as, respectively:

$$HS_t = HS_t^{FS} - \frac{1}{q} |\Delta RM_t|, \quad \Delta RM_t \leq 0, \quad (8.19)$$

and

$$HS_t = HS_t^D, \quad \Delta RM_t \leq 0. \quad (8.20)$$

Indeed, the system of equations (8.17) to (8.20) can be reduced to a single demand equation and a single supply equation, each to be estimated over the entire sample period, by making the appropriate adjustment for the change in the mortgage rate:

$$HS_t = HS_t^D - \frac{1}{q} |\Delta RM_t|, \quad (8.21)$$

where

$$|\Delta RM_t| = \begin{cases} \Delta RM_t & \text{if } \Delta RM_t \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

and

$$HS_t = HS_t^{FS} - \frac{1}{q} \backslash \Delta RM_t \backslash, \quad (8.22)$$

where

$$\backslash \Delta RM_t \backslash = \begin{cases} -\Delta RM_t & \text{if } \Delta RM_t \leq 0 \\ 0 & \text{otherwise} \end{cases}.$$

It is apparent that equation (8.21) is equivalent to the two demand equations (8.18) and (8.20) and that equation (8.22) is equivalent to the two supply equations (8.17) and (8.19).

Equations (8.21) and (8.22) can thus be estimated directly, given the specifications of HS_t^D and HS_t^{FS} in (8.11) and (8.13) respectively, but two problems occur in the estimation. One problem is that the same coefficient $1/q$ appears in both equations. The second problem is the likelihood of simultaneous equation bias due to the endogeneity of $|\Delta RM_t|$ and $\backslash \Delta RM_t \backslash$. The introduction of equation (8.15) above makes RM_t an endogenous variable, and even though RM enters with a lag in (8.11) and (8.13), RM_t still enters in equations (8.21) and (8.22) through the $|\Delta RM_t|$ and $\backslash \Delta RM_t \backslash$ variables.

These two problems are heightened in the present case by the fact that the error terms e_t^D and η_t^S , which enter equations (8.21) and (8.22) respectively, are assumed to be serially correlated.

Ignoring the fact that $1/q$ appears in both equations, the problem of simultaneous equation bias can be handled in the manner described by Fair and Jaffee [20]. Essentially the two-stage least squares technique can be used, but the step function characteristic of ΔRM_t and ΔRM_t makes the application of the technique somewhat more complicated than usual. In addition, if the error terms are serially correlated, the technique described in Chapter 2 (and in more detail in Fair [17]) must be used in place of the standard two-stage least squares technique. Ignoring the problem of simultaneous equation bias, the constraint across equations can be taken into account by using the technique developed in Fair [16]. This technique is designed for the estimation of models with restrictions across equations and serially correlated errors. In Fair and Jaffee [20], both of these techniques were used to estimate the present model, and both yielded reasonable results. Since techniques are not yet available for dealing with simultaneous equation bias and restrictions across equations at the same time, it is not clear theoretically which technique should be used. One sacrifices efficiency to gain consistency, while the other gains efficiency at a cost of consistency. The decision was made in this study to *ignore possible simultaneous equation bias and use the second technique to account for the restriction across the two equations*. This technique is somewhat easier to use than the other one, and this is the main reason for its use here.

It should be pointed out that the technique used here is based on the assumption that the error terms in the two equations (i.e., e_t^D and η_t^S in (8.11) and (8.13) above) are each first order serially correlated, but are uncorrelated with one another. While it may not be too unrealistic to assume that the demand and supply error terms are uncorrelated, it may be unrealistic to assume that the error terms in equations (8.21) and (8.22) are uncorrelated. This is because HS_t may be measured with error. If HS_t is measured with error, this same error will be included in both (8.21) and (8.22), and thus the error terms in the two equations will be correlated. To the extent that this is true, the technique used here loses efficiency by not taking the correlation into account.

8.4 The Data

The data that have been used to estimate the demand and supply equations are presented in Table 8-1. All of the variables listed in the table are seasonally

Table 8-1. List and Description of the Variables Used in the Monthly Housing Starts Sector.

HS_t	= Private Nonfarm Housing Starts in thousands of units.
RM_t	= FHA Mortgage Rate series on new homes in units of 100 (beginning-of-month data).
$DSLA_t$	= Savings Capital (Deposits) of Savings and Loan Associations in millions of dollars.
$DMSB_t$	= Deposits of Mutual Savings Banks in millions of dollars.
DSF_t	= $(DSLA_t + DMSB_t) - (DSLA_{t-1} + DMSB_{t-1})$.
DSF_6_t	= Six-month moving average of DSF.
$DHLB_t$	= Advances of the Federal Home Loan Bank to Savings and Loan Associations in million of dollars.
DHF_t	= $DHLB_t - DHLB_{t-1}$.
DHF_3_t	= Three-month moving average of DHF.
W_t	= Number of working days in month t .
DI_t	= Dummy variable I for month t , $I = 1, \dots, 11$.

unadjusted. Data on HS_t are currently published in *Economic Indicators*,⁴ and data on the three deposit variables and the mortgage rate are currently published in the *Federal Reserve Bulletin*. Data on the RM_t series were not directly available for January 1959 through April 1960, and the figures used here were constructed from an FHA series on the average of new and existing conventional mortgage rates. The data on RM_t and W_t are presented in Appendix A for the January 1959 to December 1969 period. The other data used in this chapter are easily obtainable from *Economic Indicators* or the *Federal Reserve Bulletin*.

8.5 The Results

Equations (8.21) and (8.22) were estimated by the above technique for the June 1959 to December 1969 period, with the following results:⁵

⁴ Actually, the HS_t series was discontinued in December 1969. Beginning in 1970 the breakdown of private housing starts into farm and nonfarm was no longer made. The number of nonfarm housing starts was always a small fraction of the total number of housing starts, and the decision was made by the Department of Commerce to discontinue the breakdown into farm and nonfarm. This change does not affect the work in this study, but for future purposes the published figures on total housing starts will have to be adjusted downward slightly.

⁵ The steel and automobile strikes appeared to have little effect on the level of housing starts, and so no observations were omitted from the period of estimation for the housing starts equations because of the strikes.

$$\begin{aligned}
 HS_t = & \sum_{i=1}^{11} \hat{d}_i DI_t + 2.70 W_t + 112.95 - .0709 \sum_{i=1}^{t-1} HS_i + 8.48 t \\
 & \quad (4.63) \quad (2.46) \quad (2.27) \quad (2.31) \\
 & - .127 RM_{t-2} - .412 \Delta RM_t / \\
 & \quad (1.54) \quad (2.81)
 \end{aligned} \tag{8.23}$$

$$\hat{\rho} = .841$$

$$(17.54)$$

$$SE = 8.98$$

$$R\Delta^2 = .790$$

127 observ.

$$\begin{aligned}
 HS_t = & \sum_{i=1}^{11} \hat{d}_i DI_t + 2.84 W_t - 49.22 - .164 t + .0541 DSF6_{t-1} \\
 & \quad (4.42) \quad (1.75) \quad (2.63) \quad (8.07) \\
 & + .0497 DHF3_{t-2} + .100 RM_{t-1} - .412 \Delta RM_t \\
 & \quad (5.27) \quad (2.67) \quad (2.81)
 \end{aligned} \tag{8.24}$$

$$\hat{\rho} = .507$$

$$(6.64)$$

$$SE = 8.30$$

$$R\Delta^2 = .822$$

127 observ.

$\hat{d}_1 = -34.44$ (12.52)	$\hat{d}_6 = 19.84$ (7.22)	$\hat{d}'_1 = -34.38$ (14.21)	$\hat{d}'_6 = 20.69$ (8.54)
$\hat{d}_2 = -33.72$ (11.46)	$\hat{d}_7 = 15.16$ (5.56)	$\hat{d}'_2 = -38.85$ (14.36)	$\hat{d}'_7 = 12.03$ (5.14)
$\hat{d}_3 = -9.67$ (2.87)	$\hat{d}_8 = 11.97$ (4.27)	$\hat{d}'_3 = -7.33$ (2.83)	$\hat{d}'_8 = 8.46$ (3.24)
$\hat{d}_4 = 18.62$ (5.47)	$\hat{d}_9 = 8.55$ (2.91)	$\hat{d}'_4 = 20.97$ (7.88)	$\hat{d}'_9 = 6.57$ (2.57)
$\hat{d}_5 = 23.72$ (7.76)	$\hat{d}_{10} = 11.61$ (3.85)	$\hat{d}'_5 = 36.68$ (11.20)	$\hat{d}'_{10} = 10.01$ (3.83)
	$\hat{d}_{11} = -4.88$ (1.53)		$\hat{d}'_{11} = -7.74$ (3.16)

$\hat{\rho}$ in equations (8.23) and (8.24) denotes the estimate of the first order serial correlation coefficient. The R -squared is again the R -squared taking the dependent variable in first differenced form and is a measure of the percent of the variance of the change in HS_t explained by the equation. Note that

because of the constraint that has been imposed on the model, the estimate of the coefficient of ΔRM_t in (8.23) is the same as the estimate of the coefficient of ΔRM_t in (8.24).

The dummy variables are in general highly significant in equations (8.23) and (8.24), which indicates the pronounced seasonality in the series. The working-day variable, W_t , is also significant in the equations, and thus the number of working days in a month does appear to influence the number of housing starts for that month. All of the other coefficient estimates in the two equations are of the expected sign, and all but the estimate of the coefficient of RM_{t-2} in (8.23) and the estimate of the constant term in (8.24) are significant. The time trend has a positive effect in the demand equation (8.23) and a negative effect in the supply equation (8.24), and the mortgage rate (RM_{t-2} or RM_{t-1}) has a negative effect in the demand equation and a positive effect in the supply equation. The time trend is expected to have a positive effect in the demand equation, since it is mainly proxying for population growth and trend income. The deposit flow variables are highly significant in the supply equation, and the housing stock variable is moderately significant in the demand equation. The fact that the time trend and the mortgage rate have opposite effects in the two equations (using the same dependent variable) certainly supports the hypothesis that (8.23) represents a demand equation and (8.24) a supply equation.

The estimate of the coefficients of ΔRM_t and ΔRM_t in (8.23) and (8.24) is of the expected negative sign and is significant. The significance of the estimate indicates that the housing market is not always in equilibrium and that rationing does occur. When equations (8.23) and (8.24) were estimated separately without imposing the constraint (by the standard Cochrane-Orcutt technique), the estimate of the coefficient of ΔRM_t in (8.23) was $-.408$ and the estimate of the coefficient of ΔRM_t in (8.24) was $-.438$. These compare with the restricted estimate of $-.412$. It is remarkable that the unconstrained estimates are so similar, which perhaps provides further support to the view that rationing does occur in the housing market.

The estimate of the serial correlation coefficient is larger in the demand equation (.841) than it is in the supply equation (.507), and the fit of the demand equation is somewhat worse than that of the supply equation (SE = 8.98 vs. 8.30).

A number of other variables were tried in the two equations, especially in equation (8.24); and some of these results should be mentioned. First, different lags of the mortgage rate were tried in the two equations, and while RM_{t-2} and RM_{t-1} gave the best results in (8.23) and (8.24) respectively, the results were not substantially changed when slightly different lags were used. Theoretically, of course, it is not the absolute size of the mortgage rate that should matter, but the size of the mortgage rate relative to rates on

alternative assets. A number of yield differential variables were tried in the equations, but with no success. While theoretically not very satisfying, it definitely appeared to be the absolute level of rates that mattered and not rate differences.

As mentioned above, different lags of the deposit flow variables in (8.24) were tried, and the ones presented in (8.24) gave the best results. Deposit flows into Life Insurance Companies and Commercial Banks were also tried in (8.24), but these flows added almost no explanatory power to the equation. Deposit and mortgage *stock* variables of the SLAs and MSBs were also tried in (8.24), and again with no real success. The flow variables always dominated the stock variables, which probably indicates that the adjustment of SLAs and MSBs to changing deposit conditions is fairly rapid. The flow variables of the SLAs and MSBs were also tried separately in (8.24), and the coefficient estimates were close enough so that it was decided to consider only the sum of two flow variables. Notice also that in (8.24) the coefficient estimate of DHF_{t-2} is nearly the same as the coefficient estimate of $DSF6_{t-1}$. The lag seemed to be slightly different for the DHF3 variable than for the DSF6 variable, however, and it was decided to treat these two variables separately.

Finally, the mortgage holdings of the Federal National Mortgage Association (FNMA) was tried as an explanatory variable in equation (8.24), but with no success.⁶ Both stock and flow variables were tried and various moving averages and lags were tried, and none of these variables were significant. Most of the time the estimates were even of the wrong sign. The results in this study thus indicate that for policy purposes, the Federal Home Loan Bank lending activity (as reflected through $DHF3_{t-2}$ in (8.24)) has much more of an effect on the level of housing starts than does the activity of FNMA.⁷ These results are, of course, not conclusive, since the level of aggregation is so high, but they do seem to indicate the importance of the FHLB relative to FNMA. It should be noted, however, that not even the FHLB will have an effect on housing starts if demand and not supply is the constraint.

8.6 The Use of the Housing Starts Equation for Forecasting Purposes

There are two basic ways in which equations (8.23) and (8.24) can be used for forecasting purposes. One way is to treat ΔRM_t as exogenous. Assuming

⁶ In 1968 FNMA was split into two groups (the new FNMA and the Government National Mortgage Association), but in this study the two groups were treated as one.

⁷ Jaffee [28] in a detailed study of the mortgage market has found that the activity of FNMA has little effect on the total stock of mortgages, which is consistent with the conclusions reached in this study.

ΔRM_t , to be exogenous, let \overline{HS}_t denote the predicted value of HS_t from equation (8.23), let \widehat{HS}_t denote the predicted value of HS_t from equation (8.24), and let \widehat{HS}_t equal a weighted average of the two predicted values: $\widehat{HS}_t = \lambda \overline{HS}_t + (1 - \lambda) \widehat{HS}_t$. It is easy to show that if the error terms in equations (8.23) and (8.24) are independent and if the desire is to choose λ so as to minimize

$$\sum_{t=1}^T (HS_t - \widehat{HS}_t)^2,$$

then the optimum value of λ is $\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$, where σ_1^2 is the variance of the error term in equation (8.24), σ_2^2 is the variance of the error term in equation (8.23), and T is the number of observations. From estimates of σ_1^2 and σ_2^2 , therefore, an estimate of λ can be used for forecasting purposes. In the present case the estimate of λ is $(8.30)^2/[(8.30)^2 + (8.98)^2] = .46$.⁸ In other words, the predictions from equation (8.23) are weighted slightly less than those from (8.24), since the estimate of the variance of the error term is slightly larger in (8.23).

The other way in which (8.23) and (8.24) can be used for forecasting purposes is to treat ΔRM_t as endogenous. Let \widehat{HS}_t^D denote the predicted value of demand, and let \widehat{HS}_t^{FS} denote the predicted value of supply. \widehat{HS}_t^D is obtained from (8.23) by ignoring the $|\Delta RM_t|$ term, and \widehat{HS}_t^{FS} is obtained from (8.24) by ignoring the $|\Delta RM_t|$ term. [See (8.21) and (8.22).] Then given \widehat{HS}_t^D and \widehat{HS}_t^{FS} , the predicted value of ΔRM_t (denoted as $\widehat{\Delta RM}_t$) can be obtained from equation (8.15), using as the estimate of q the reciprocal of the estimate of the coefficient of $|\Delta RM_t|$ and $|\Delta RM_t|$ in (8.23) and (8.24). $\widehat{\Delta RM}_t$ can be used to compute $|\widehat{\Delta RM}_t|$ and $\widehat{\Delta RM}_t$, and the predicted value of the actual number of housing starts, \widehat{HS}_t , can then be computed as

$$\widehat{HS}_t = \widehat{HS}_t^D - .412|\widehat{\Delta RM}_t|.$$

From equations (8.15), (8.21), and (8.22), it can be seen that the latter expression is the same as $\widehat{HS}_t^{FS} - .412|\widehat{\Delta RM}_t|$. Since RM also enters equations (8.23) and (8.24) with a lag (as RM_{t-1} and RM_{t-2}), in a dynamic simulation or forecast the values for \widehat{RM}_{t-1} and \widehat{RM}_{t-2} can be taken from the $\widehat{\Delta RM}_t$ series.

Treating ΔRM_t as endogenous thus yields one predicted value of HS_t , whereas treating ΔRM_t as exogenous yields two. There are, in other words,

⁸ The question of degrees of freedom has been ignored in this discussion. The estimates of the standard errors in (8.23) and (8.24) have been adjusted for degrees of freedom, whereas the variances that result from the above minimization are not so adjusted. Since the number of variables in equation (8.23) is only one less than the number in (8.24), however, the difference between adjusting or not adjusting for degrees of freedom is trivial.

two independent pieces of information in the system of equations (8.15), (8.23), and (8.24). The decision was made in this study to treat ΔRM_t as exogenous and generate the two predictions of HS_t . Some initial experimentation was done treating ΔRM_t as endogenous, and while the static simulation predictions of ΔRM_t from equation (8.15) were fairly good, the equation was sensitive to dynamic error accumulation and to errors made in forecasting the exogeneous variables. The results seemed to indicate that ΔRM_t could be more accurately forecast exogenously than by the use of equation (8.15).

Given that ΔRM_t is to be taken as exogenous, the question arises as to how the two predictions from (8.23) and (8.24) are to be weighted. The derivation at the beginning of the section suggested that the predictions should be weighted by the estimates of the variances of the error terms in the two equations. The derivation was based on the assumption that the errors in the two equations are uncorrelated. To the extent that the errors are positively correlated, it can be seen that the above minimization approach implies that even more weight should be attached to the equation with the smaller variance. In the limit, if the errors were perfectly correlated, it can be shown that all of the weight should be given to the equation with the smaller variance. The error terms in the two equations are in fact positively correlated (as a regression of one set of error terms on the other revealed), which is probably due in part to errors of measurement in the HS_t series. In spite of this, in the work below the predictions from the two equations have been weighted equally: equation (8.24) with the smaller variance has not been weighted more. For actual forecasting purposes, the better fit of equation (8.24) is somewhat illusory, since the equation includes the two important variables, $DSF6_{t-1}$ and $DHF3_{t-2}$, which must be forecast exogenously. In equation (8.23) the only exogenous variable that is not trivial to forecast is the mortgage rate. On these grounds, then, equation (8.23) should be given more weight, and in the final analysis the simple compromise of treating both equations equally was made.