NOTES ON THE SPECIFICATION OF AN AGGREGATIVE MODEL
OF METROPOLITAN HOUSING MARKETS

George C. Galster and Jerome Rothenberg

Number 129 May 1974

This paper was written with the collaboration of Richard V. Butler, Rice Institute, and John R. Pitkin, Harvard-M.I.T. Joint Center for Urban Studies. We would also like to acknowledge the contributions made by William J. Stull, Swarthmore College, on a prototype model which forms the foundation of the current study.

This research is supported by National Science Foundation Grant GS-40769. The views expressed herein are the authors' responsibility and do not reflect those of the Department of Economics, the Massachusetts Institute of Technology, or the National Science Foundation.
Addendum

Further Consideration of Long Run Equilibrium Prices

In order to simulate the results of various housing policies with our model and compare their effectiveness it becomes manifest that not only the supplies of units produced but also the change in long run housing prices which accompany such supplies must be considered. In order to apply this latter criterion in our model we must explicitly model the change in long run equilibrium prices in each sub-market, \( \Delta \hat{P}_i \), and thus must consider in more depth the underlying concept of a long run stock supply function, \( L R_{Q_i}^{ST} \) (c.f. page 11). Note that previous discussions dealt in terms of equilibrium market values, \( MV_i = \hat{P}_iHV_i \), since it was assumed for simplicity that the absolute quality or hedonic value of \( L_i \) was fixed. In our model we want to allow for quality improvements, whence changes in equilibrium prices per hedon, \( \Delta \hat{P}_i \), becomes the important variable.

As in the conventional microeconomic formulation of an industry long run supply curve, we view the slope of \( L R_{Q_i}^{ST} \) (i.e. \( \Delta \hat{P}_i \)) as inversely dependent upon the degree to which housing stocks in \( L_i \) respond to a given shift in demand, \( Q^D_i \). But, unlike the conventional model, the degree of response in one \( L_i \) crucially depends upon events in other \( L_{k \neq i} \). These interactions have already been described heuristically, as well as formally in equations (9), (10), (12), (15), and (16). Of course, part of the demand shift in \( L_i \) is, itself, determined by the relative submarket price levels, as modelled in (3), (8), and (11). Thus, it must be emphasized that \( L R_{Q_i}^{ST} \) does not represent an equilibrium price/stock relationship in \( L_i \) which is immutable across SMSA's at a given point in time or within a given SMSA over time; it cannot be viewed in isolation from other \( L R_{Q_k}^{ST} \). Rather, it represents how \( \hat{P}_i \) changes given an initial distribution of stocks and a given pattern of exogenous demand.
disturbances across \([ L_k ]\), which together simultaneously determine how much suppliers' cost will rise, what types of supplies will be forthcoming, how they will be distributed across \([ L_k ]\), to what degree the responses will be carried out in each \( L_k \), and how the demands are altered by the endogenous (price) adjustments.

We propose, then, that decadal changes in long run submarket housing prices be modelled by:

\[
(60-70) \Delta P_i = f \left( \frac{69Q_i}{60-70} \right) \frac{Q_i}{60-70} ST, \Delta Q_k, k = 1, 2, \ldots, t \right)
\]

Where \( 69Q_i \) is the estimated demand in 1969 holding all endogenous \([ P_k ]\) at their 1960 levels (c.f. p. 17), and \( 60-70 Q_i ST \) is the decadal supply response.

Unfortunately, (a) cannot be estimated directly since we have no observations of the left hand variable. Figure 1 makes it clear, however, that at any point in time observed submarket price, \( P_i \), can be defined as the sum of equilibrium price, \( \hat{P}_i \), and transitory or disequilibrium price, \( \tilde{P}_i \) (the vertical distance between \( M^R \) and \( LR^S_i \) in Figure 1), where \( \tilde{P}_i \) is associated with disequilibrium vacancy rates, \( VR_i \), such that:

\[
(60-70) \Delta \tilde{P}_i = f \left( 60-70 VR_i \right)
\]

By taking decadal first differences in observed \( P_i \) it is thus possible to derive an estimable equation:

\[
(60-70) \Delta P_i = f \left( (a), (b) \right)
\]

whence (assuming all functions are linear) the coefficients estimated in (c) implicitly define (a) and (b).

The number of vacant units, \( Q_i^{SV} \), and the vacancy rate, \( VR_i \), have now been mentioned several times in the text without explicit consideration of how they are
assumed to function in the market. Our model assumes that an equilibrium vacancy rate, $\hat{VR}_i$, can be defined for each $L_i$ throughout the decade. $\hat{VR}_i$ can vary, however, between submarkets within an SMSA. Given that we posit a specific rate and not a number of vacancies prevails in each $L_i$ it should be clear that $Q_i^{SV}$ must grow as $Q_i^{ST}$ does (c.f. Figure 1), whence any supply response, $\Delta Q_i^{ST}$, may be decomposed into additional supplies which would actually be occupied given $\hat{VR}_i$ (hereafter referred to as "effective" supplies), $\Delta Q_i^{SO}$, and those needed to provide the necessary cushion of extra vacancies, $\Delta Q_i^{SV}$. Since we view the former component as critical for policy evaluation purposes we plan to estimate (12) and (15) with $\Delta Q_i^N$ and $\Delta Q_i^C$ defined in terms of "effective" units, i.e., with total construction and conversion flows being deflated by $(1/\hat{VR}_i)$. Operationally, we assume that $\hat{VR}_i$ can be proxied by the mean of $VR_i$'s across our sample SMSA's in 1960, whence the $VR_i$ variables in (12) and (15) may interpreted as the SMSA-specific differences between actual $VR_i$ and equilibrium $\hat{VR}_i$, which may affect supply responses 1960-61.
The objective of our research is the development of an econometric model of metropolitan area housing markets. Attention will be primarily focused in this paper on the issues involved in specifying our underlying theoretical model in an estimable form. The model is distinguished by its application of the concept that the housing market can be meaningfully partitioned into quality submarkets, by its careful consideration of conversions of existing units between these submarkets, and by its analysis of long-run adjustments. We propose that housing markets form a complex of segmented but interrelated quality submarkets, where there is some scope for independent behavior in each, and also a set of intersectoral repercussions of varying degrees. Quality submarkets are linked in demand by closeness of substitution and in supply by technological transformation possibilities. They can thus be applied to our purposes for analyzing demands for and supplies of both newly constructed and converted units of varying qualities. We concern ourselves here only with an aggregative view of metropolitan housing markets, and do not attempt to explore the intricacies of the spatial distribution of housing within SMSA's. Particular attention has been paid in the model's specification to facilitate simulations of long run housing market impacts of a broad range of public policies, including public construction and demolition, direct demand and supply subsidies, and income redistributions. The purpose of this paper is not to present econometric results but rather to generate discussion and suggestions for modifications of the provisional model outlined below.
I. The Theoretical Model

Although a detailed analysis of our theoretical model of housing markets is not the primary goal of this paper (for such an analysis see [6]) a brief description is mandated before exploring more carefully the problems raised by attempts at its empirical specification. Our modelling of the housing market is founded on the principle that the commodity called 'housing' cannot be treated like most other consumption items, but rather possesses unique characteristics which create fundamental idiosyncracies in its supply response: durability, heterogeneity, and convertibility. Little need be said about the fact that housing can provide valuable services long after its construction, although the critical role of maintenance expenditures in determining this value is often overlooked. Similarly, it is no revelation that housing packages differ markedly in terms of their myriad components: structural and parcel characteristics, neighborhood and public service quality, accessibility to work and amenities, etc. The final feature is simply that a unit originally built to provide a certain type of housing service may be subsequently modified, through renovations, additions, partitions, or changing maintenance activities, to provide substantially different services.

While these unique characteristics of housing are well known, their implications for supply response processes, particularly conversions of existing units, have been frequently overlooked by other housing market researchers. The durability of housing, accounting for the numerical dominance of second-hand over new units, its heterogeneity, which creates a spectrum of housing service types as alternative opportunities for existing units, and the technological possibilities for transforming property toward many of these alternatives, combine to create a form of market supply
distinctive from, but in many respects as important as, new construction. Since the technology, the costs, and the social consequences of this form of supply differ from that of new construction, a model which omits it may seriously misrepresent the working of urban housing markets.

The characteristic of heterogeneity holds another implication central to our present purpose -- different housing packages represent only imperfect substitutes for one another, with degrees of substitutability varying from pair to pair. Thus, the market is likely to contain sets of housing clusters or submarkets, each one containing very close substitutes, and the submarkets themselves differing in the degree of substitutability with one another. The structure of substitutability will influence the pattern of impulse propagation through the entire market -- events spreading rapidly and completely within a submarket, with external repercussions dampening progressively before reaching those of lower substitutability.

Although this concept of segmentation is important since it implies different patterns of competitive interaction than a single, perfectly articulated market, discerning some meaningful pattern in the multidimensional maze of components underlying patterns of substitutability is complicated. The basic approach used here is that the market can be functionally segmented into "quality submarkets". "Quality" reflects the overall market attractiveness of a parcel when all components of the package are evaluated. It is assumed that demanders implicitly evaluate the relative attractiveness of different packages as a function of their components, and that this attractiveness reflects their widespread willingness to trade off more of one dimension for less of another. Analogously on the supply side, suppliers of existing and new units vary components in their packages subject to technologically-
fashion the household chooses $HV_1$ and $Z$ so as to maximize utility subject to its budget constraint:

$$\text{(2)} \quad Y^n = P_1HV_1 + P_2Z = MV_1 + P_2Z$$

where $P_1$ is the price per unit of quality in quality level $L_1$; $P_2$ is the price of $Z$; and $MV_1$ is the annualized market value of, or total annual expenditure on, the bundle of housing services provided in $L_1$. We assume $P$ and $MV$ effectively depend only on $L_1$ and not on idiosyncratic differences between units in $L_1$.

Note that the choice of $HV_1$ in this model differs from the traditional approach insofar as $HV_1$ doesn't represent the optimal quantity of housing quality units chosen by households when faced with their constant price per unit. Rather, households face varying $P_1$ over different, narrowly defined ranges of $HV_1$ (quality submarkets $L_1$). Thus, a household cannot always choose the $L_1$ with lowest $P_1$ since it may be associated with such a large $HV_1$ that the implied budget allocation $MV_1$ would be suboptimal. This choice of an optimal $HV_1$ can be represented by a household demand function:

$$\text{(3)} \quad HV_1^n = f^n(Y^n, [P_1], P_2)$$

where $Y^n$ is the household's permanent income, $[P_1]$ is a vector of per unit housing quality prices at different quality levels, and $HV_1$ is that quality of unit (i.e. quality level $i$) which maximizes $U^n$ given the budget constraint.

**Individual Builder and Converter Supply Behavior**

Both new construction and conversion supply responses of individual suppliers are controlled by changes in the underlying cost and revenue functions which are defined for the supplier over the various quality submarkets. These functions are discussed below.
Annual costs of a new unit which the mth builder may construct in quality level \( i \) may be expressed:

\[
C^N_m(L_i) = CK^m(L_i) + CR^m(L_i)
\]

where \( CK \) and \( CR \) are the annual flows of capital (mortgage amortization of construction costs, interest carrying costs, equity capital cost), and recurrent (property tax, maintenance, repair and operations) expenditures, respectively. \( C^N(L) \) increases monotonically with quality. Neglecting capital gains or losses, the annual cost which the \( r \)th owner of an existing unit would incur by converting it from its current level \( L_1 \) to \( L_k \) is:

\[
C^C_{1r}(L_k) = CK^r(L_0) + CK^r(L_k - L_1) + CR^r(L_k)
\]

where \( CK(L_k - L_1) \) is the annualized flow of capital costs expended by the conversion process and \( CK(L_0) \) is the "sunk" amortized capital costs incurred when the unit was built in its original \( L_0 \). Conversion capital costs vary depending on the initial quality of the unit and on the direction of conversion. Upward conversion to higher quality is analogous to new construction investment. Downward conversion with investment (partitioning, e.g.) is more costly than downward conversion without investment ("running down" via deferred maintenance, e.g.) but is less costly than upward conversion for most quality destinations envisioned.

The vector of \( MV_1 \)'s defines a revenue function for suppliers, \( R(L_i) \), which monotonically increases with quality level and is independent of how a unit arrives at \( L_i \). Recall that \( R(L_i) \) is not necessarily a measure of quality -- the imperfect substitutability of units across quality levels implies \( P \) and \( L \) are capable of independent changes. Here, as before, we assume that a single "aggregate" \( MV_1 \) is meaningfully defined for each \( L_i \).

Given this specification of cost and revenue functions we suppose that each of \( m \) existing and potential builders of new units are confronted by two
simultaneous supply decisions at a given moment in time: at which \( L_1 \) to build units and how many to build there. The answer to the former question is determined by scanning all \( L_1 \), assessing the rate of return in each, \( I^m(L_1)_{[Q^ST]} = R(L_1)/C^m(L_1) \) (where the \( R \) and \( C \) functions are defined for \( [Q^ST] \), a vector of the current housing stocks in all \( L_1 \)): and selecting the \( L_1 \) offering the highest \( I \). The answer to the latter is determined by the availability of capital to the construction industry, \( IN \) (i.e., the rate of return in the target \( L_1 \) relative to alternative, nonhousing investment opportunities). Individual builders are assumed to be price takers and unable to affect factor prices by their purchases. Thus, the probability of the \( m \)th builder constructing new units in \( L_1 \) at some point in time may be expressed as:

\[
(6) \quad PR^m_m = f_1([I^m_k \ k = 1, 2, \ldots ], IN)
\]

Behavior of individual converters of existing units is also characterized by the attempt to procure maximum rates of return. Assuming the \( r \)th existing unit (which is currently in \( L_1 \)) is earning \( I^F(L_1) > R(L_k)/C^FR(L_k) > IN \), there will be no conversion incentive. But if rates of return are larger for at least some \( L_k \) there will exist pressures on the \( r \)th owner to convert to \( L_k \), the strength of the pressure being positively related to the relative \( I^F(L_k) \) vs. \( I^F(L_1) \).

These pressures will be dampened by higher \( IN \) in the following manner. As \( IN \) rises some owners who previously would have found it profitable to convert will now find that the additional out-of-pocket expenses which would be entailed in the conversion, \( C_F(L_k - L_1) + CR(L_k - L_1) \), could bring a higher rate of return elsewhere in the economy than the marginal gain of \( I(L_k) - I(L_1) \). An additional effect is possible here. \( IN \) may prove so attractive compared to all rates of return in housing that owners may find it more profitable to divert their current \( CR(L_1) \) expenditures into other
Investments if I(L₁) produced by such expenditures, even when combined with sunk costs C(KL₀), is inferior. This would mean, of course, the abandonment of the given unit and its effective removal from Q₁. More formally, this process can be viewed as a conversion from L₁ to the "abandonment" submarket, Lₐ.

In light of the foregoing analysis, the probability of the nth owner converting his unit from L₁ to L_k (including Lₐ) may be expressed as:

(7) \( PR_{1k} = f_{1k}(I_{1k}^f, k = 1, 2, \ldots, \text{IN}) \)

where \( I_{1k}^f \) is the rate of return available to r by converting from L₁ to L_k.

Aggregate Demander and Supplier Behavior

The individual behavioral patterns just presented will now be combined to provide models for aggregate demand and supply behavior.

Individual household demands such as (3) can be aggregated into a stock demand in L₁ for an arbitrary population. Given the population's income distribution, its utility functions, and relative housing submarket and non-housing prices, each household will choose a unit at a particular quality level (i.e. will consume a particular \( H^V \)), thereby determining the total number of units demanded in each of the different L₁:

(8) \( Q^D_{1i} = \sum_{n} H^V_k \text{ if } k = 1 \)

Changes in the stock of housing due to new construction at some point in time can be viewed as the aggregate result of the myriad of individual builder L₁-choice/numbers-of-units-to-build decision as summarized in (6):

(9) \( \Delta Q^{SN}_{1i} = f_{1i}(I_{1k}^u, k = 1, 2, \ldots, \text{IN}) \)
where \( \Delta_0^{SN} \) is the total number of new units constructed in \( L_1 \) in a given period.

Here \( I(L_1) \) must be interpreted as the rate of return available to builders in \( L_1 \) on the average. As such we implicitly recognize that individual construction firms may have somewhat different technologies and cost functions. These differences, coupled with builders' varied perceptions and degrees of information about market alternatives mean that all builders will probably not chose the same target \( L_1 \) to supply. Rather, in the aggregate, the distribution of new construction responses across \( L_1 \) will be determined by relative \( I \)'s and the total construction level by those \( I \)'s relative to \( I_N \).

The aggregate supply formulation for converters differs from that for builders because the costs of converting to other submarkets in response to potential profit improvements diverge dramatically depending on the current \( L_1 \) of the units in question. Thus, even if within each \( L_1 \) all potential converters had identical perceptions and cost functions we would expect a divergence in target \( L \)'s between converters in different initial \( L \)'s.

Differences in intra-\( L_1 \) perceptions and costs only add, of course, to the probabilistic nature of the aggregate specification. The pattern of aggregate conversion responses, then, is not only a function of relative submarket (and nonhousing) rates of return, but also of the distribution of potentially convertible units across \( L_1 \), \( [O^{ST}_{1}] \), i.e. the relative numbers of units facing different elements in the vector of conversion costs:

\[
(10) \quad \Delta_0^{SC} = f_{1}(O^{ST}_{1}, [I_{1} \quad k = 1, 2, \ldots], I_N)
\]

where \( \Delta_0^{SC} \) is the net number of conversions into \( L_1 \) from all sources in a given period, and \( I_{1} \) is the average rate of return obtained by converting from \( L_k \) to \( L_1 \).
Aggregate Demand-Supply Interaction

Given the foregoing specifications of demander and supplier behavior the interactions between them on an aggregated basis as contemplated in our model can now be characterized.

Consider the situation at some point in time described by a series of vectors showing for each \( L \) the aggregate demand function \( [\alpha_1^D] \), occupied stock \( [\alpha_1^{SO}] \), vacant stock \( [\alpha_1^{SV}] \), and total stock \( [\alpha_1^{ST}] = [\alpha_1^{SO}] + [\alpha_1^{SV}] \).

Current technology and construction input and capital prices define the underlying new construction cost function, \( C^N \), over all submarkets, as well as the family of conversion cost functions specifying conversion costs to all quality levels from each initial level \( L_k \), \( [C^C_k] \). Now if associated with those vectors there is a set of submarket market values such that the rates of return in all submarkets, \( [I_k] \), are comparable to \( TN \), and \( [\alpha_1^{SO}] \) and \( [\alpha_1^{SV}] \) are such that \( [\alpha_1^{SV}] \) reaches some normal, "frictional" level of vacancies (vacancy rate \( [\check{V}_1^N] \)), equilibrium is attained in the market with this equilibrium market price vector, \( [MV_1] \). As long as the cost and revenue functions don't change, suppliers (builders or converters) in any \( L_i \) will only replace retirements from the stock or those units which, through deterioration, have changed quality unintentionally.

Of course, recent history has not been characterized by stagnant revenue (i.e., demand) functions but rather by ones which have generally grown (i.e., shifted right graphically) with population and incomes. The pattern of aggregate supply response to such a demand shift can be best introduced with the help of Figure 1, showing hypothetical demand-supply-price interactions in an arbitrary \( L_1 \) where all units have the same \( MV_1 \).

Suppose we're at equilibrium when there occurs a once-and-for-all shift of demand from \( Q_1^D \) to \( Q_1^{D*} \). This shift creates an "excess demand", \( X_1 = Q_1^{D*} - \alpha_1^{SO} \).
wherein at the original $\hat{N}_1$ more households are willing to consume $MV_1$, i.e. occupy more units in $L_1$, than the existing stock and normal vacancy rates in $L_1$ permit. The initial response to this shift will be a combination of higher prices, $MV_1 > \hat{MV}_1$, and lowered vacancy rates, $VR_1 < \hat{VR}_1$, let's say as shown by path $MVR$, since suppliers don't respond instantaneously with more units.

Now this rising $MV_1$ and falling $VR_1$ creates for owners of existing units in $L_1$ more revenues and (since their $CK$ is "sunk" cost and $CR$ has not changed) higher rates of return relative to the initial equilibrium levels. As $I(L_1)$ rises it will elicit both new construction and conversion supply responses, whose proportions are a function of the relationship between $C^N(L_1)$ and $C^C_k(L_1)$, all $k$. Ignoring, for a moment, the interactions between submarkets, these responses will simultaneously increase $O_1^{ST}$ and lower $I(L_1)$ by both reducing $R(L_1)$ and (as suppliers' purchases bid up capital and, to a smaller extent, maintenance factor input prices) increasing $C^N(L_1)$ and $C^C_k(L_1)$ until a new equilibrium is reached at $O_1^{ST}$, $O_1^*$, and $MV_1^*$, where $X_1 = 0$.

The locus of $MV$ generated by such demand shifts represents the long run stock supply function in $L_1$, $LRST_1$.

One should realize that our focusing on one particular $L_1$ as if it were independent of others was done only for expository purposes. The presumed underlying responses are considerably more complex. For instance, as demand and prices rise in $L_1$ one would expect some households to begin choosing other $L_{k\neq 1}$ as their optimal quality levels, thus bidding up prices there, too. Conversions into $L_1$ will alter the stock and vacancy characteristics of submarkets out of which the units were converted, thereby upsetting equilibrium there and changing the future ability of these levels to supply converted units. Similarly, new construction in $L_1$ will, by affecting prices of inputs used in
housing construction and maintenance, alter the cost functions facing builders and converters in all other submarkets as well. These interactions between submarkets' demands, supplies, prices, and costs form the heart of our model.

This model of demand-shift/supply response is obviously simplified by the omission of various dynamic adjustments, expectations, information costs, search behavior, lag patterns, etc. For example, suppliers' responses are undoubtedly based not only on current I's but also on expected future levels. These perceptions, in turn, are shaped by anticipated demand pressures, cost changes, and responses of competing suppliers. Furthermore, suppliers may be unable to react without considerable lags and possibilities for overshooting, thereby increasing the interactive complexities. While our simplifications are unfortunate, Census data upon which our empirical specification is based does not permit the formulation of a full dynamic model within which these processes have importance. Nevertheless, we feel that our long run approach can be used effectively to evaluate policy alternatives even while suppressing explicit consideration of shorter run dynamic interactions.

II. The Empirical Specification

This section will suggest an empirical specification of the above theoretical model which can be econometrically estimated using primarily Census data from 1960 and 1970 for a cross sectional sample of 35 SMSA's differing widely in size, geographic location, and composition of housing stock. The reliance upon Census data presents certain problems, of course, the foremost of which involve the aggregate nature of the observations and the lack of intra-decade observations. On the other hand, such a specification should provide an answer to the question of whether meaningful long run analyses of metropolitan housing markets can be made on such an aggregate level using commonly available data. The answer to this question has obvious significance for national housing policy-makers. A different specification of our model using individual unit data gathered in Des Moines, Iowa, is currently being developed in order to explore those processes not visible from the aggregated viewpoint, especially some of the dynamic complexities alluded to earlier.
The empirical specification to be discussed takes as given that the housing market in both 1960 and 1970 in each of our SMSA's has been meaningfully partitioned into 10 quality/tenure submarkets, \( q_{ij} \), \( i = 1, \ldots, 5, j = 1 \) if owner, 2 if renter occupied) via hedonic indexes we have developed especially for this purpose based on 1960 Census data [1], [2]. These indexes give us our invariant yardstick of housing quality over the 1960-70 decade called "hedonic value", \( MV_{ij} \), which is functionally independent of actual market values, \( MV_{ij} \).

**Aggregate Demand Specifications**

The intention here is to predict the total number of units demanded (i.e. the occupied stock) in each quality level of both tenures in 1970 for each SMSA, \( Q_{ij}^{70} \), using a demand function in accordance with the theory summarized in equation (8). Our specification is somewhat more complicated than that analysis implied, however, in that the tenure as well as the quality level choice of households is considered. We propose a single demand equation for each submarket which represents a reduced-form solution to a recursive bloc within the given submarket equation system which explains net household formation, tenure choice, and quality level choice:

\[
Q_{ij}^{70} = f_{ij}\left(\frac{P_{kj}}{P_{z}}, k = 1 - 5, \frac{Y_{ij}}{P_{z}}, \frac{MV_{ij}}{MV_{ij}}, MR\right)
\]

where \( P_{kj} = MV_{kj}/HV_{kj} \) the average price per "hedon" or "unit quality" in submarket \( kj \); \( P_{z} \) is an index of relative prices of all non-housing goods; \( Y_{ij}/P_{z} \) is the number of families in a given SMSA having a real income "appropriate" to the \( ij \) in question (e.g. for the highest quality owner submarket probably only the number in the upper third of the real income distribution would be important); \( [DM] \) is a series of Census demographic characteristics expressed in terms of numbers of families within particular
categories like education, age, family size, race, etc.; and $M_R$ is an index of relative level of mortgage interest rates for the region in which the given SMSA is located.  

It should be obvious that in this specification most variables perform double-duty. For instance, the $[P/P_z]$ vector helps determine household formation (by proxying costs of getting into the housing market relative to those of non-housing consumption) as well as quality level choice (relative submarket prices within a tenure category as in equation (3)). $Y_{4j}/P_z$ also influences household formation (by proxying economic conditions), tenure choice (as a factor shaping preferences for the privacy and status facets of owner occupancy) and quality level choice (via our demand theory). The $[DM]$ vector focuses on "life-cycle" factors influencing both household formation and tenure choice. The significance of these factors is well established, most recently by Kain and Ouiglev [5]. $M_R$ serves as a time-independent proxy for inter-SMSA differences in mortgage costs, which affect tenure choice.

Aggregate Supply Specifications

Unlike the previous equation which predicts the demand for units at some given point in time, the supply equations estimate long run supply responses, i.e. decadal changes in the stock from both construction and conversion sources. The basic element driving both construction and conversion responses in our model is demand shifts. We posit that if an arbitrary housing market is in equilibrium in 1960 the cumulative excess demand pressures ($X_{4j}$) eliciting supply responses in different $L_{4j}$ (by raising I's) can be approximated by the horizontal shifts in the $n_{4j}^P$ functions during the decade. In other words, while we have no information about the intra-decade dynamic adjustments and interactions their net result can be modelled analogously to Figure 1 as a long run supply response to a
"once-and-for-all" demand shift equal in magnitude to that observed over the decade.

While, as noted before, we have generally abstracted from most dynamic considerations our empirical specification does implicitly make an assumption about the time path of supply response. We assume that the typical response to an $X_{ij} > 0$ (of normal magnitude) in year $t$ consists of a relatively small marginal gain in $Q_{i}^{'ST}$ in year $t + 1$ as suppliers hesitate to gather information and assess the significance of the $X_{ij}$, draw up plans for the new buildings or conversions, and begin work on the projects. These projects are completed in year $t + 2$ and any additional new supplies as well as the $X_{ij}$ return to zero by the end of year $t + 2$. Under this supposition, then, only the demand shifts occurring from 1960 - (early) 1969 will be relevant for the observed 1960-70 elicited supply response, since supplies forthcoming from any continued demand pressures 1969-70 would, presumably, not appear in the decadal totals.

Two additional introductory remarks may be made before proceeding to our provisional specifications of aggregate decadal construction and conversion supplies. These remarks follow easily from our foregoing theoretical discussions. The initial point is that our concept of distinctive and competing forms of construction and conversion supplies implies that over a span as long as a decade the responses must be modelled interactively. Specifically, a larger degree of one type response should provide a disincentive for the other by reducing the $I_{ij}$ levels that would be available to the other, ceteris paribus. Analogously, construction of units by the public sector during the decade will dampen the demand pressures facing both private builders and converters. Secondly, recall that the previous discussion of supply-demand interactions assumed long run equilibrium in 1960 and thus was able to ignore the possibility of excess demand pressures in 1950-60 inducing supply
responses in 1960-61 -- responses that would not be proxied by demand pressures 1960-68 yet still would contribute to the decadal total. We believe altered vacancy rates ($VR_{1j}$) are the initial response to demand changes, whence any degree of (disequilibrium) market tightness in 1960 which would signal supply changes in 1960-61 could be proxied by them in our specification.

With these considerations and the theory of aggregate builder behavior (equation (9)) in mind our provisional specification of new construction supplies over the decade is straightforward:

$$
(12) \quad \Delta_{60-70} N_{1j}^{\text{P}} = f_{ij}(\Delta_{i} C_{1j}, \Delta_{i} G_{1j}, 60 VR_{1j}, [X_{kL}^j, k = 1 - 5, L = 1, 2], 60-70 \Delta CC)
$$

where $\Delta_{60-70} N_{1j}^{\text{P}}$ is the number of new units privately built in $L_{1j}$ 1960-70; $\Delta_{60-70} C_{1j}$ is the net number of units converted into $L_{1j}$ 1960-70; $\Delta_{60-70} G_{1j}$ is the number of units built (directly or indirectly) by the government in $L_{1j}$ 1960-70 (assumed exogenous); $VR_{1j}$ is the vacancy rate in $L_{1j}$ in 1960; $X_{kL}^j$ is a vector showing the cumulative excess demand pressures in all $L_{1j}$ from 1960-69 (to be discussed in more detail below); and $60-70 \Delta CC$ is the change in an index of construction factor input prices 1960-70, an endogenous variable which is specified in (16). 9

The foregoing analyses should make the meaning of these variables clear. New private construction in $L_{1j}$ will be inversely related to the amount of net conversions and/or government construction into $L_{1j}$ since they represent alternative forms of supply which compete away excess $I_{1j}$. Low $60 VR_{1j}$ will capture any residual demand pressures from the previous decade eliciting building responses in $L_{1j}$ 1960-61. $X_{kL}^j$ and $60 \Delta CC$ jointly proxy rates of
return in various submarkets over the decade by denoting shifts in revenue and cost functions, respectively, generated by demand shifts and resultant supply responses. An \( X_{ij} \) term approximates the horizontal shift of the \( Q^D_{ij} \) function, and is formally defined as:

\[
(13) \quad X_{ij} = 69^D_{ij} - 60^D_{ij}
\]

where \( 69^D_{ij} \) is the estimated number of units in \( L_{ij} \) which would be demanded in 1969 if \( P_{kl} \) had remained at 1960 levels. This latter variable is generated by taking the coefficients estimated for (11) and inserting 1968 values for \( Y_{ij} \) (gleaned from non-Census sources), 1969 values for \( P_{kl} \), and 1969 values for \( V_R, P_z, \) and \( [DM] \) (interpolated between Census observations).

The formulation of aggregate converter behavior as postulated in (10) into a total decadal conversion supply equation proves somewhat more complex than the new construction specification. While the variables \( 60-70^N, 60-70^N, 60^V_{ij}, \) and \( 60-70^C \) play analogous roles, the excess demand terms must be formulated differently. Recall that we posit that the aggregate probability of net conversion inflows into a given \( L_{ij} \) is not only related to the comparative demand pressures (rates of return) in other \( L_{kl} \) but also to the number of units in other \( L_{kl} \) which might serve as potential sources of conversion supply for \( L_{ij} \) and which would be affected by the relative I's. In our empirical specification we therefore see the probability of an individual owner converting from \( L_{kl} \) to \( L_{ij} \) as directly related to

\[
(69^D - 60^D_{ij})/60^D_{ij} - (69^D_{kl} - 60^D_{kl})/60^D_{kl}
\]

with the aggregate probability of such conversions being found by weighting this term by the geometric average of the number of units present in \( ij \) and \( kl \),

\[
(60^ST_{ij} 60^ST_{kl})^{1/2}
\]

(This is, of course, analogous to a "gravity" type model of transportation demands.) Thus, the greater the difference in I's between \( L_{ij} \) and \( L_{kl} \) (i.e. more potential revenues in \( L_{ij} \)), the greater the initial
stock in $L_{ij}$ (i.e. more potential suppliers), and the smaller the initial stock in $L_{ij}$ (i.e. fewer potential suppliers in $L_{ij}$ to other submarkets), the larger the net inflow of conversions into $L_{ij}$ over the decade. This analysis can be formalized by the $X_{ijkl}$ variable:

$$X_{ijkl} = \left( \frac{60_{ij}^{ST} 60_{ij}^{ST}}{1/2 \left( \frac{60_{ij}^{DP} 60_{ij}^{SO}}{60_{ij}^{ST}} - \frac{60_{ij}^{SO} - 60_{ij}^{DP}}{60_{ij}^{ST}} \right)} \right)$$

One additional modification to the conversion supply equations should be made to account for the asymmetry in conversion costs, and therefore responses, between upward and downward conversions. There is no reason to expect the same absolute $\Delta C_{ij}$ due to flows between $L_{ij}$ and $L_{kl}$ regardless of whether $X_{ijkl}$ is some arbitrary positive or equal (absolute) negative value. If, for example, $L_{ij}$ was of higher quality than $L_{kl}$ and $X_{ijkl}$ was positive the conversion response up from $L_{kl}$ (and the resultant $\Delta C_{ij}$) would probably be smaller absolutely than the response down from $L_{ij}$ if the same magnitude of $X_{ijkl}$ was negative, since the upward conversions would likely be more costly and thus would create a smaller increase in rate of return. We therefore add a vector of "asymmetry adjustment factors", $[A_{ijkl}]$, to the model. $A_{ijkl}$ may be specified in two possible ways. It may assume the absolute value of $X_{ijkl}$ when $X_{ijkl} < 0$, and zero otherwise, whence $X_{ijkl}$ assumes only its positive values and is zero otherwise. Alternatively, $A_{ijkl}$ may be viewed as a dummy variable assuming the value one when $X_{ijkl} < 0$ and zero otherwise. We propose to test both possibilities.

Given the foregoing discussion of the independent variables being utilized, our provisional aggregate specification for the net number of units converted into $L_{ij}$ during the decade is:
\[ \Delta 60-70 Q^C_{ij} = f_{ij} (\Delta 60-70 Q^N_{ij}, \Delta 60-70 Q^G_{ij}, 60_{VR_{ij}}, \Delta 60-70 CC_{ij}) \]

where all variables are defined as previously.

**Construction Cost Specification**

Recall from our scenarios described earlier that supply responses are seen to bid up the costs of inputs used by both builders and converters (construction equipment, materials, labor, etc.) and this, in turn will dampen the total increase in stocks. Thus, it is obvious that we must consider \( \Delta 60-70 CC \) in (13) and (15) as an endogenous variable in our system. We posit that the prime influence on CC increases during the decade is the level of construction (private and public) and, to a lesser degree, conversion activity in all \( L_{ij} \) along lines similar to classic "demand-pull" inflation theory. Two factor supply influences may also prove significant. The first is that the same absolute level of housing supply activity may elicit different \( \Delta CC \) responses across SMSA's depending on the capacity of the existing housing factor "industry" at the beginning of the decade, e.g., the scale of local cement firms or the number of laborers and craftsmen attached to the construction industry. This aspect may be captured by the relative housing stock increases during the current decade compared with those during the last, \[ \left[ \frac{\sum_{ij}^L 60_{i,j}^{ST} - \sum_{ij}^L 60_{i,j}^{ST}}{\sum_{ij}^L 60_{i,j}^{ST}} \right] / \left[ \frac{\sum_{ij}^L 50_{i,j}^{ST} - \sum_{ij}^L 50_{i,j}^{ST}}{\sum_{ij}^L 50_{i,j}^{ST}} \right] \].

The other element involves the costs of construction labor. Wages may be pushed up by construction unions even if there is not a surging demand for their services if laborers demand "catch-up" pay increases to meet rising
costs of living, proxied by $\Delta P_z / P_z^0$. Thus, our provisional construction cost change specification is:

$$\Delta 60-70 P_z / P_z^0$$

$$C = \sum_{k=1}^{5} \sum_{L=1,2} \frac{\left(\Delta Q_k L^N + \Delta Q_k L^C \right)}{\Delta Q_k L^ST} \quad k = 1-5 \quad L = 1,2$$

$$\frac{\left(\sum_{L=1}^{L=50} Q_j L_i^ST - \sum_{L=1}^{L=50} Q_j L_i^ST\right)}{\left(\sum_{L=1}^{L=50} Q_j L_i^ST - \sum_{L=1}^{L=50} Q_j L_i^ST\right)}$$

where all variables are defined as above.

**System Considerations**

Some final remarks should be made on the issue of system simultaneity and identification. As should be obvious from the preceding discussion the demand, supply, and cost equations form a large simultaneous system with such endogenous variables as $[P_i j], [\Delta Q_i j^S], [\Delta Q_i j^C], [\Delta Q_i j^N]$ (where $i = 1, \ldots, 5$ and $j = 1, 2$) and $\Delta CC$ appearing as explanatory variables. A total of 31 equations are involved -- one for $\Delta CC$, and one in each of ten quality/tenure submarkets for $Q^D, \Delta Q^C, \Delta Q^N$. Each equation in this system is over-identified, and thus two-stage least squares techniques will be used in estimating the system.
NOTES

1. The concept of "housing submarket" has been employed by other authors but in somewhat different forms than the one used here. The seminal work of Grigsby [4] describes supply changes as a matrix of transitions in single housing characteristics. Sweeney [7] has employed a commodity hierarchy in place of a quality metric to disaggregate the market.

2. For expository purposes we assume in this discussion that idiosyncratic differences in costs facing builders or converters can be suppressed so that we can refer to the aggregate construction and conversion cost functions.

3. Several similarities exist between our model and others currently under development, particularly that of Frank de Leeuw of the Urban Institute [3]. While both models embody similar concepts of supplier/demander motivations and attempt to describe decadal changes in the stocks and prices of units of different qualities in terms relevant to policymakers, there exist significant differences. While de Leeuw assumes the existence of some metric capable of partitioning the market into quality levels our model empirically estimates such a metric (hedonic index) and applies it to this task. This explicit structuring of the market enables us to organize the decisions of suppliers and demanders so that all relevant housing alternatives are weighed by them, whereas the de Leeuw formulation involves myopic market decisions and iterative search processes. Finally, our version is econometrically estimable throughout and does not require the arbitrary assignment of critical parameter values as in the de Leeuw specification.

4. The arbitrary choice of 5 quality levels within each tenure category represents our resolution of the conflict between maximizing submarkets so as to narrow HV ranges vs. minimizing submarkets so as to simplify empirical specifications.

5. Except, of course, 1960. Since the hedonic index was estimated across the same 35 SMSA's on the basis of observed 1960 MV's,

\[
60_{HV_k} = 60_{MV_k} + u_k
\]

where \(u_k\) is the residual of the regression applicable to SMSA \(k\).

It should be mentioned at the outset that the Census does not provide enough cross-tabulated information to directly allow the partitioning of an SMSA's stock into quality levels, then determining such data for these levels as numbers of vacancies, single family structures, owner-occupants in multiple family structures, etc. A special technique has been developed for our project by John Pitkin, Doron Holzer and Kate Gasser, however, which can infer such information from the Census marginal tabulations. Briefly their technique is as follows. Once having chosen a known "target" two-way matrix which was assumed to approximate the desired (unknown) matrix, a linear-programming algorithm minimizes the sum of the proportional differences between the values in the target and the matrix being estimated, subject to the constraint that the estimated column and row values sum to the appropriate (known) value given by the Census marginal tabulation.
6. This variable is generated for our sample SMSA's by a regression equation we have developed which predicts the non-housing component of a medium-income City Worker's Family Budget as defined by the Bureau of Labor Statistics.

7. The 1969 appendage to $Y_{ij}$ is not meant to be confusing but only recognizes the fact 1970 Census data tabulates 1970 housing and demographic characteristics but 1969 incomes. 1969 reported $Y$ is not, of course, equivalent to permanent $Y$ as we would desire but, again, we must resign ourselves to this constraint of our data base.

8. This index was computed from Federal Home Loan Bank Board annual data for specific SMSA's 1965-72. These data showed remarkable inter-SMSA stability in relative mortgage interest rate levels, even though absolute levels varied dramatically over the period. These relative rates, $MR$, were computed as regional averages (Northeast, North-Central, South, West) and our sample SMSA's assigned $MR$'s according to their regional location. Due to their aforementioned stability 1965-72 these $MR$'s were also assumed constant for 1960-64.

9. This index was derived from Dodge construction cost data and calculated so that all SMSA's indexes were comparable, i.e. all were related to a common base period and SMSA. The ideal variable here would be $ACC_{ij}$ where the index is weighted by the particular input ratios typical in the $ij$ level, and not overall $ACC$. Such data were not available, unfortunately.

10. It might be possible to add some dimension of dynamic complexity to the building responses by including an additional variable, 

$$
(69^{D}_{ij} - 60^{SO}_{ij})/60^{ST}_{ijj}
$$

in (12). Extremely large proportional shifts in demand may alter builders' expectations such that an additional amount of speculative construction would be forthcoming during the decade.

11. This approximation of $69^{D}_{ij}$ is not exact, of course, since the position of the demand curves is not only parametrized by $P_{z}$, $Y_{ij}$, [DM], and $MR$, but also by $[P_{kj}$ $k \neq i]$, and we have no information about this last variable in 1968.

12. Certainly the gross number of conversions which require investment is the variable called for here. Unfortunately, our Census data base precludes our observation of anything but net conversions both with and without capital investments, and this may be a poor proxy for the desired variable. For this reason, and the fact that conversions represent a relatively small segment of aggregate construction factor demand, we do not feel the inclusion of a "conversion activity" variable is warranted. Another factor affecting ACC but excluded from our specification is the change in nonresidential building activity. This is omitted since we did not feel the marginal gain from its inclusion justified the added complexity of adding another equation to our model explaining nonresidential building activity, i.e. it certainly could not be considered exogenous, and thus would need to be endogenized.
References


