Volume Title: The Detroit Prototype of the NBER Urban Simulation Model

Volume Author/Editor: Gregory K. Ingram, John F. Kain, and J. Royce Ginn

Volume Publisher: NBER

Volume ISBN: 0-870-14258-5

Volume URL: http://www.nber.org/books/ingr72-1

Publication Date: 1972

Chapter Title: Status of the Modeling Effort

Chapter Author: Gregory K. Ingram, John F. Kain, J. Royce Ginn

Chapter URL: http://www.nber.org/chapters/c3506

Chapter pages in book: (p. 163 - 174)

9

STATUS OF THE MODELING EFFORT

IN THE PREVIOUS CHAPTERS we tried to indicate the basis of our optimism about the correctness of the basic design of the NBER model and to convey some appreciation of the many difficulties encountered in attempting to calibrate the model to Detroit. In this chapter we examine further some of the problems encountered in attempting to calibrate the Detroit Prototype and Pittsburgh I, describe the significant modifications in model design incorporated in Pittsburgh I, and outline probable extensions and modifications of the model that will form the basis of Pittsburgh II. Finally, we speculate on longer-run modifications and elaborations of the NBER model and describe some potential applications of the model to public policy.

Problems of Calibrating the Detroit Prototype

It should be apparent that a calibration of sorts was achieved for the Detroit Prototype, even though it fell far short of a satisfactory representation of Detroit's housing market. The principal obstacle preventing a satisfactory calibration was the lack of suitable data on the prices of individual dwelling units. For Detroit, the only available data on housing prices were median values of single-family, detached, owner-occupied homes and median rents of renter-occupied dwellings by Census tract. Even these data were five years older than the housing choices we were attempting to explain.

The resulting estimation problems are enumerated in Chapter 8. Still, if the only problem had been estimation of the gross price coefficients, we probably could have eventually achieved an adequate calibration of the model for Detroit. Unfortunately, the unavailability of prices for individual housing units also meant we could not estimate accurate housing price surfaces by housing submarket. If the submarket demand functions had been "correct," this problem could have been circumvented by running the model over several time periods to produce a consistent set of housing prices. Conversely, if submarket housing prices had been available, we believe we eventually could have calibrated the submarket demand functions by trial and error. But since we had neither a good set of gross price coefficients nor reasonable estimates of the housing price surfaces, the task of achieving satisfactory calibration of the model for Detroit appeared nearly impossible.

In spite of these difficulties, the primitive calibration of the Detroit Prototype permitted us to carry out a number of valuable test runs. These test runs increased our confidence in the basic design concept and provided us with a number of pleasant surprises about model efficiency. Unfortunately, they also increased our pessimism about satisfactorily calibrating the NBER model with the data then available for Detroit. It is worth noting that the TALUS sample also was not adequate enough for an unambiguous definition of housing submarkets. The problem of defining housing submarkets from Census tract statistics, which is discussed in Chapter 8, is minor only in comparison to the problems caused by the unavailability of housing prices for dwelling units.[1]

Because of the serious data problems of the Detroit Prototype, when the more complete Pittsburgh data base became available, we

1. The difficulties encountered were mainly attributable to the aggregate nature of the Census information. The 1960 Census reports only two house prices for each Census tract, the median rent of renter-occupied units and the median value of owner-occupied, single-family, detached units. Yet the Detroit Prototype allows as many as twenty-seven types of unit in each zone. The problem is less severe for the large residence zones used in the NBER model since they typically include many individual Census tracts. Even so, the housing submarket price surfaces obtained by aggregating Census tract medians within the forty-four residence zones used in the Detroit Prototype proved to be unsatisfactory. Substantial modifications of the surfaces estimated from Census data were required to bring these estimates into even rough conformity with more general information on the structure of prices by housing submarket. The resulting estimates of housing prices by submarket lacked authority and were clearly not suitable for calibrating the demand model. These problems were further aggravated because the Census tract median values and rents were for 1960, while the housing choices used to estimate the submarket demand equations were for 1961–65. During this period there were undoubtedly significant changes in relative housing prices in the Detroit area.

shifted the development of the model to the Pittsburgh metropolitan area. The Pittsburgh home interview survey of 1968 included a far more complete description of sampled dwelling units than the Detroit home survey and also reported the rent or value of sampled dwelling units. Of the housing attributes pertinent to the definition of housing submarkets—structural type, number of rooms, dwelling unit quality, lot size for single-family units, and rent or value—the Detroit survey provided information only on structural type. All other characteristics had to be imputed from Census tract statistics. This procedure was somewhat unsatisfactory for defining housing submarkets, completely unsatisfactory for deriving the housing prices needed to estimate gross prices for use in estimating the submarket demand equations, and hopeless for estimating housing prices by submarket for inclusion in the Detroit Prototype.

Although there were clear gains in shifting the development of the model to Pittsburgh, there were also disadvantages. The most obvious of these was the large amount of data processing required to reproduce the quantities of data needed to estimate the submarket demand equations and to calibrate Pittsburgh I. A less obvious cost, but one which we regard as potentially a more serious disadvantage, stems from the peculiarities of the Pittsburgh area. These peculiarities include the area's extreme topographic irregularities, an unusually dispersed spatial distribution of employment, an industrial employment structure dominated by the steel and metal fabricating industries, and low rates of employment and population growth in recent decades. We are concerned that such factors may have strongly influenced, if not controlled, Pittsburgh's growth and development. The pattern of metropolitan development found in Pittsburgh may, therefore, be unique in more particulars than is desirable for the development of a generic model of urban growth and development. In these terms Detroit was somewhat more suitable.

Characteristics of Pittsburgh I

There are a few important differences between the Detroit Prototype and Pittsburgh I, the operating version of the NBER model now being tested. The most important is the use of neighborhood quality to define housing submarkets in Pittsburgh I. Instead of the 27 housing types used in the Detroit Prototype, Pittsburgh I includes the 40 housing types described in Table 9.1. The model geography of Pittsburgh I includes 20 workplace zones and 50 residence zones, comprising 41 high-quality residence zones and 9 low-quality residence zones. In the Detroit Prototype all 27 housing types could be located in each of the 44 residence zones. In Pittsburgh I each of 20 different housing types defined in terms of structural type, number of rooms, dwelling unit quality, and lot size can be provided in either a high- or low-quality zone.

This change in model design has no significant effect on the estimation of the demand equations. Households are assumed to consider neighborhood quality in a manner analogous to their consideration of the number of rooms or any other housing attribute. In the supply models, however, high-neighborhood-quality units cannot be built in low-quality zones, nor can low-neighborhood-quality units be built in high-quality zones.

The submarket demand equations for Pittsburgh I were estimated by a two-stage procedure similar to that described in Chapter 8.[2] Although it appears that we have made substantial progress in solving the difficult estimation and calibration problems that plagued us in Detroit, we are still far from satisfied with the gross price coefficients used in Pittsburgh I. We are certain, however, that the estimates can be improved, and we are currently engaged in an extensive program of econometric research aimed at improving both the definitions of housing submarkets and the estimates of the submarket demand equations. These revised definitions of housing submarkets and improved estimates of gross price coefficients will form the core of the third version of the NBER model, Pittsburgh II.

Characteristics of Pittsburgh II

A number of working hypotheses embodied in our current studies will undoubtedly determine in large measure the general configuration of Pittsburgh II. First, we are more certain than ever that neighborhood characteristics should be included in the definition of housing

2. Ingram, "Model of a Housing Market," Appendix.

Table 9.1
Summary of Model Dimensions for Pittsburgh I

Households (96 classes)

Characteristics	Class Intervals
Annual income	0–$5,000, $5,001–$7,000, $7,001–$10,000, $10,001 or more
Family size (no. of persons)	1, 2, 3–4, 5 or more
Age of head (years)	0–30, 31–60, 61 or more
Education of head (years)	Less than 12 years, 12 years or more

Industries (11 types)

Type	Description
1	Agriculture, forestry, fishery
2	Mining, petroleum
3	Contract construction
4	Manufacturing
5	Primary metals
6	Transport, communications
7	Wholesale, retail
8	Financial, insurance, real estate
9	Services
10	Government
11	Nonclassifiable, unclassified

Dwelling Units (40 types)

Group	Type[a]	Structural Type	Lot Size (acres)	Number of Bedrooms
1	1,11,21,31	Single family	.25	2 or less
2	2,12,22,32	Single family	.25	3
3	3,13,23,33	Single family	.25	4 or more
4	4,14,24,34	Single family	.50	2 or less
5	5,15,25,35	Single family	.50	3
6	6,16,26,36	Single family	.50	4 or more
7	7,17,27,37	Common wall	.625	2 or less
8	8,18,28,38	Common wall	.625	3 or more
9	9,19,29,39	Small multifamily	.25	-
10	10,20,30,40	Large multifamily	.25	-

a. Types 1–10: sound condition, in good neighborhoods; types 11–20; poor condition, in good neighborhoods; types 21–30: sound condition, in poor neighborhoods; types 31–40: poor condition, in poor neighborhoods.

submarkets. Pittsburgh I contains the modifications in model design and programming needed to accommodate this change, and Pittsburgh II will probably incorporate more neighborhood quality detail than Pittsburgh I. In addition, it seems virtually certain that Pittsburgh II will include both more types of neighborhoods and a larger number of residence zones. For example, we may use two or more neighborhood characteristics to describe neighborhoods.

The variables most likely to be used to describe neighborhoods in Pittsburgh II are: (1) the average quality of dwelling units in the neighborhood, (2) the average socioeconomic status of the residents of the neighborhood, and (3) some measure of neighborhood density, such as net residential density or the percentage of units that are single family. If the average quality of dwelling units in the neighborhood is used to define housing types in Pittsburgh II, it should be possible to make this neighborhood classification scheme endogenous to the model by linking it to the filtering submodel. Similarly, if density is the second neighborhood characteristic used to describe housing types, the classification of neighborhoods by density could be modified over time within the model in response to new construction, transformations, and demolitions in each zone. Such an approach presumably would make the model's bookkeeping far more complex. Therefore, we will have to proceed somewhat further with model development before we will be able to assess the costs of these alternatives in terms of storage and running time. When the costs are known, we must then determine whether such extensions are worth the price.

The use of more types of neighborhood is appealing on several grounds. First, for a given number of residence zones, the use of a larger number of neighborhood types reduces the maximum size of the linear programming problems that must be solved by the market-clearing submodel. Since 90 per cent or more of the running time of the Detroit Prototype and Pittsburgh I are accounted for by the linear programming solutions, and since the largest problems account for nearly all of the program solution running time, the gains from using more neighborhood categories may be quite large. Although Pittsburgh I has more residence zones (50 versus 44), housing types (40 versus 27), and workplaces (20 versus 19), than the Detroit Prototype, the computer running time of its market-clearing submodel is about half that of the Detroit Prototype. This difference in solution time of the two market-clearing submodels results from the stratification of residence zones by neighborhood quality, which reduces the maximum size of the linear programming problems in Pittsburgh I, as well as from other improvements in model design that in many cases make the actual problem size smaller than the maximum problem size.

Perhaps more significant than reduced execution time, however, is the promise that by defining housing types by more detailed neighborhood characteristics we may be able to increase substantially the number of residence zones with little or no increase in model running time. Use of a larger number of residence zones would make the neighborhood definitions more meaningful and allow a much more precise description of the residence price surface. The advantages of greater zonal detail for most planning and policy applications of the model are obvious.

The computational efficiency of this approach would be even greater if the use of a larger number of neighborhood types permits us to use a smaller number of housing types within each neighborhood type. This trade-off is evident already from a comparison of the Detroit Prototype and Pittsburgh I. In the former each zone potentially contains 27 structural types, while in the latter no zone may contain more than 20 housing types. Although these lines of development must still be regarded as somewhat tentative, they appear to have considerable promise for improving the realism and operating characteristics of the NBER model.

Definition and use of additional neighborhood types is also a promising method of incorporating housing market discrimination and variations in the level and quality of local public services in the model, since both may be thought of as characterizing neighborhoods in particular ways. That the racial composition of particular neighborhoods strongly influences the locational choices of both black and white households can hardly be doubted. Incorporation of housing market discrimination into the model has a high priority, and racial discrimination could be represented in several ways. One of the approaches we are considering would include the racial composition of the neighborhood as part of the definition of housing submarkets. We would then estimate housing submarket equations for white and nonwhite households similar to the equations included in the current version of the model.[3]

A second technique would limit nonwhite occupancy to certain exogenously specified residence areas. Nonwhite households would be permitted to compete for some residence areas but not for others. This would modify the linear programming solution, affect housing prices by type and residence area, and thereby influence the locational choices of both white and nonwhite households. Nonwhite households would bid up the prices in those residence areas sanctioned for their occupancy and thereby discourage white occupancy in these zones. Because segregation is currently so intense, we anticipate that the use of a simple racial constraint would produce results that are virtually identical to those obtained using separate econometrically estimated demand functions for black and white households.

Yet another approach would represent housing market discrimination by price markups and discounts to white and nonwhite households for housing in particular residence areas. This technique would add premiums and discounts to the gross prices considered by white and nonwhite households in choosing among housing types. In the assignment model certain zones would be more costly to reach for white and for black households. This approach would more closely approximate the formulation of housing market segregation offered by Gary Becker, Richard Muth, and Martin Bailey.[4]

The preceding discussion illustrates some of the techniques that might be used to represent housing market discrimination in the NBER model. We do not expect that Pittsburgh II will include an adequate representation of this phenomenon. But we hope to use Pittsburgh II as a tool for testing several formulations of discrimination to determine which best replicates the behavior observed in U.S. cities. These experiments should greatly increase our understanding of both the causes and consequences of housing market discrimination.

A more elaborate neighborhood classification scheme may also be

3. Kain and Quigley, "Housing Segregation" and unpublished findings; Straszheim, "Demand for Housing Services" and idem, "An Econometric Analysis."

4. Becker, *Economics of Discrimination;* Muth, *Cities and Housing;* and Bailey, "Economics of Zoning and Urban Renewal."

the most promising method of representing variations in local public services, such as schools, in the NBER model. The major obstacles to the introduction of local public services into the model are not computer or programming limitations. Instead it is the lack of any good empirical evidence on the independent effect of public schools or other public services on the housing choices of urban households. This question has been a major issue in NBER econometric studies of the housing market. The findings, however, are quite inconclusive, and more research is required to isolate the determinants of neighborhood quality and their effects on the housing choices of various classes of households. Preliminary investigations in this area suggest that neighborhood attributes are highly correlated. Residence zones with high income levels typically also have good schools, few dilapidated structures, and a low proportion of rental units. Existing data have not permitted us to isolate the effects of public services from other aspects of neighborhood quality.

In addition to probable changes in the definition of housing types and the use of a larger number of residence zones, Pittsburgh II will also incorporate a number of less substantial modifications. First, we anticipate that Pittsburgh II will treat population-serving employment endogenously. The most likely formulation would make the level of population-serving employment depend on the levels of population and basic employment in each zone. At the same time, workers employed in the outlying, predominantly residential zones will be included in the model. Such workers are excluded from both the Detroit Prototype and Pittsburgh I. The two extensions are closely related, since many workers employed in outlying zones are employed in population-serving activities. These workers probably will be allocated to available units in nearby zones before the linear programming model is used to allocate workers employed in major workplaces to the remaining supply of available units.

A more detailed representation of the land market is still another high-priority extension of the model. Monocentric models of urban spatial structure have focused on the land market and ignored the existence of durable stocks of residential capital. The current NBER model focuses mainly on capital stocks and their change. An important element in decisions to alter the residential stock is the cost and availability of land at alternative sites. For example, demolition decisions will be influenced by the price for open land. Local political jurisdictions and their zoning decisions are also relevant to the land market, as are public uses of land such as roadways, parks, and public services. Land prices are determined by competing bids of suppliers of both residential and nonresidential stocks. The representation of the land market in the model will tie together several of these elements. It should be noted, also, that the land market is the traditional focal point of many policies aimed at influencing urban development.

If Pittsburgh II does incorporate a nonresidential land market, it will be quite primitive. In both the Detroit Prototype and Pittsburgh I, all land used for nonresidential purposes must be specified exogenously in each period. For example, the demolition of residential structures required by nonresidential expansion must be exogenously specified. Pittsburgh II will probably include a simple nonresidential expansion function that will remove residential structures in zones where there is substantial expansion of nonresidential land uses and the amount of vacant land available is limited.

Potentially desirable improvements and elaborations of the NBER Urban Simulation Model are quite numerous. The proposed extensions of the model described here are some of the major revisions that we now plan to incorporate into Pittsburgh II, but many other extensions of the model would be desirable as well. In some instances we are fairly certain how to make these additions; others involve major conceptual and estimation problems. The priorities of model improvement and development will depend on the uses to which the model is to be put, the findings of our supporting econometrics studies, and lessons learned in the process of developing and testing Pittsburgh I and II.

Potential Policy Applications

One of the main considerations motivating our efforts to develop an urban simulation model has been our conviction that present models are unsuitable for considering many important issues of urban policy. Therefore, while formulating the underlying design of the

NBER Urban Simulation Model and enriching the detail of its existing versions, we have attempted to provide the model with the capability of dealing with a wide range of policy matters.

For example, our model will be well suited to addressing policy questions that deal with housing market intervention and market imperfections. Since the NBER model simulates markets and includes decision parameters, such questions can be investigated directly. For example, it will be possible to trace the effects of income maintenance or housing allowance programs on specific types of housing by location. Our model should provide some evidence about whether such programs lead primarily to increases in the quality of units consumed by participating families or if much of the benefit of the programs is dissipated in higher prices. Similarly, it should be possible to examine the effects of housing allowance and income maintenance programs on the quality of housing in various parts of the region.

Subsidy programs which operate on the supply side of the market could similarly be evaluated. Programs of this kind would include subsidies for the rehabilitation of slum properties or improvements in construction techniques. Both general subsidy programs and programs aimed at changing the relative costs of different housing types could be considered and an evaluation made of both the direct effects and the effects on competing submarkets. Finally, the model could be used to measure the consequences of market imperfections, such as housing discrimination. If one group of the population is restricted to a few residence zones, the impact of this restriction on housing prices as well as on the residence choices of all population groups could be measured.

The model could similarly be developed to address a wider range of basic issues affecting metropolitan development. These include the measurement of the impact of changes in transportation systems and in patterns of industry location.[5] The explicit linkage between workplace location and residence location in the model implies that different industry location projections will have direct and measurable effects on the price of housing and eventually on the location of new construction. Such projections could include the assessment of

5. Some primarily illustrative attempts to measure the impact of transportation system changes have been made with the present model. See Ginn, "NBER Prototype."

alternatives such as the development of suburban industrial parks, the retention of industry in the central city, and the development of industry in the ghetto.

As suggested elsewhere in this chapter, neither the Detroit Prototype nor Pittsburgh I are calibrated to a standard that is sufficient to allow their use for investigating policy questions in a serious or credible manner. At present we are devoting our resources to model development, but the next version of the model, Pittsburgh II, should be capable of use for some preliminary policy evaluation studies.