Urban Studies, Vol. 32, No. 8, 1995 1361–1367

Further Evidence on the Accuracy of Residential Vacancy Chain Models

Philip C. Emmi and Lena Magnusson

[Paper first received, October 1994; in final form, May 1995]

Summary. Residential vacancy chain models simulate the transfer of vacant housing opportunities among sectors of an urban housing market. The Markov model simulates forward-reaching chains in response to vacancy initiations. The Leontief model simulates backward-reaching chains in response to vacancy absorptions. Each simulates residential mobility among housing sectors as a by-product. The accuracy of these models was assessed in earlier work by the authors, using 1975–80 data from Gävle, Jönköping and Västerås, Sweden, to project intra-urban residential mobility in each town during the 1980–85 period. Using log-linear analysis to compare projected moves with observed moves, they found projection errors ranging from 3–12 per cent. In this paper, data from the 1985–90 period are used first to repeat these assessments over the subsequent time-period and then to extend the projection period from 5 to 10 years. Projection errors range from 5–11 per cent for the 5-year period and from 8–18 per cent for the 10-year period. Both the Markov and the Leontief models perform equally well. Models with more homogeneous housing sector definitions produce more consistent results.

1. Introduction

This paper presents further evidence on the accuracy of residential vacancy chain models. These are a class of multisectoral models used to simulate the intersectoral transfers of vacant housing opportunities. Each yields, as a serendipitous by-product, a simulation of intra-urban residential mobility. The Markov version of the vacancy chain model is used to assess the impacts of new construction on intra-urban residential mobility. The Leontief version is used for housing needs assessment. Their use and design has been discussed in earlier articles (Emmi and Magnusson, 1994a and 1995).

One of the most important issues surrounding the use of such models is the degree to which the basic parameters of these models remain stable through time. The best way to address this issue is to calibrate a model with data gathered during one time-period, use it to project residential mobility during the next time period and compare projections with second period observations. However, the stability issue is often confused with a related issue concerning the internal homogeneity of system sectors. As explained in Emmi and Magnusson (1993), the question of parametric stability can not be clearly

Philip C. Emmi is in the Department of Geography, Room 270, Orsen Spencer Hall, University of Utah, Salt Lake City, Utah 84112, USA; and Lena Magnusson is in the Institute for Housing Research, Uppsala University, P.O. Box 785, S-801-29 Gävle, Sweden.
The authors would like to acknowledge the support of this research by the Swedish Association of Building Officials (SABO) and the Swedish Association of Counties (Kommunforbundet).

0042-0980/95/081361-07 © 1995 Urban Studies

addressed until one has defined system sectors so as to ensure their internal homogeneity. This having been done, evidence on the question of parametric stability can be interpreted more clearly.

Emmi and Magnusson (1994a) and more particularly Magnusson (1994) present the only instances where this procedure has been followed for vacancy chain models. They consider residential mobility and vacancy transfers in the three Swedish municipalities of Gävle, Jönköping and Västerås. Using models calibrated on each of these cities with data collected during the period 1975–80, they project intersectoral household moves during the 1980–85 period. Comparisons between projected and observed moves are made using log-linear analysis in the manner described by Knoke and Burke (1980). These comparisons indicate a high degree of stability in system parameters over the 5-year period. (See Table 1).

Since then, data have become available for a third time-period—the 1985–90 period. With this new data set, tests for stationarity may be repeated over the next 5-year period and extended over a 10-year period. The purpose of this paper is to report on the results of these extended tests of parametric stability.

2. Basic Concepts

Residential vacancy chain models rest on the assumption that mobility within housing markets is limited to the set of currently available housing opportunities. Thus vacant housing opportunities are used as the models' active entities and serve as their basic units of analysis. The models further presume the delineation of a metropolitan housing market from its rural hinterland and the definition of a set of internally homogeneous housing sectors or sub-markets.

Vacant housing opportunities are accounted for by reference to each of three system states. These include initiation, transiency and absorption. Vacancy initiation is associated with discrete housing inventory and household demographic change events. These include new construction, the sub-division of single units into multiple units, the conversion of non-residential structures into residential units, the out-migration of households and household death or dissolution. These discrete change events occur in specific housing market sectors and are accounted for on a sectoral basis. They create opportunities in those sectors that are available for occupancy by households from both within and without the local housing market system.

Vacancies available for occupancy are said to be within the system's transient state space. There they are exposed to the possibility of being transferred both to other locations (dwelling units) within the system's transient state space and to the system's absorbing states. The transfer of vacancies within the transient state space is characterised by the sectoral origin and destination of the transfer. This establishes a transpose relationship between intersectoral vacancy transfers and intra-urban residential mobility. For each household that moves from housing sector *j* to housing sector *i*, there is a vacancy transfer in the opposite direction from housing sector *i* to housing sector *j*. Thus the easiest way to get a count of intersectoral vacancy transfers is to transpose a count of intersectoral household moves.

Upon transfer, each vacancy is also exposed to the possibility of being absorbed. Vacancy absorption is associated with discrete housing inventory and household demographic change events. Housing inventory change events include the destruction or demolition of dwellings, their withdrawal or conversion to non-residential use and the merger of multiple units into fewer units. Housing demographic events include new household formation and household in-migration. Once absorbed, the chain of vacancy transfers is ended and the vacancies disappear from the system.

A double-entry accounting framework is used to keep track of these events. It is designed to ensure the maintenance of a mass balance equality between vacancy inflow and vacancy outflow. That is, it ensures that vacancies initiated within or transferred to a sector (the column sums) equal the vacancies absorbed within or transferred from a sector (the row sums). Violation of the mass-balance equation would imply a long-term increase or decrease in vacancy levels by sector when empirically vacancy levels tend to oscillate around reasonably stable equilibrium values.

Both Markov and Leontief vacancy chain models are based on mass-balance equations governing the accounting of vacant housing opportunities. The Markov model is based on row accounts within the double-entry vacancy accounting framework. Row sums are used to convert counts of vacancy transfers into intersectoral vacancy transition probabilities and vacancy absorption probabilities. These are then used to simulate the forward chain of transfers and the associated pattern of household mobility induced by vacancy initiation events. Thus they may be used to assess the impacts of new construction, outmigration or household death and dissolution on both intra-urban residential mobility and the accommodation of new entrants into the housing market.

The Leontief model is based on column accounts within the double-entry accounting framework. Column sums are used to convert counts of vacancy transfers into intersectoral vacancy input coefficients and vacancy initiation coefficients. These are then used to simulate the backward chains of vacancy transfers and the associated pattern of household mobility induced by vacancy-absorbing events. Thus a Leontief model may be used to assess by sector the implications of housing need associated with in-migration, new household formation and housing demolition for the construction of new dwelling units.

3. Data and Methods

Vacancy transfer counts are based on data from the 1975, 1980, 1985 and 1990 Census of Population made available by Statistics Sweden. Special data files are compiled for the urban areas of Gävle, Jönköping and Västerås. In these, individuals and households are linked through time so that those present in period n can be defined by their housing status in period $n+1$. From these files, vacancy transfer counts are constructed for each urban area covering the three periods, 1975–80, 1980–85 and 1985–90.

In Emmi and Magnusson (1994a) and in Magnusson (1994), Markov and Leontief models are calibrated with 1975–80 data and used to project 1980–85 household moves that are then compared with 1980–85 observed moves. In each case, log-linear analysis is used to measure the degree of fit between observed and projected frequencies. Knoke and Burke (1980, pp. 40–42) describe the techniques used. Observed and projected volumes of intra-urban vacancy transfers are arrayed in a three-dimensional contingency table. Vacancy transfer frequencies in each cell are regarded as dependent variables. The sectors of vacancy *o*rigination (O), the sectors of vacancy *d*estination (D) and the observed versus projected *s*tatus of vacancy transfers (S) are regarded as independent categorical variables. Variation in the model can be attributed to interaction between or among any of these categorical variables.

A model of mutual independence among these variables [O, D, S] identifies the baseline variation in the data. A goodness-of-fit statistic G^2 is computed and noted as a basis of further comparison.

Suppose vacancy transition probabilities and vacancy input coefficients are stationary with respect to time. Then, once having controlled for interperiod differences in the frequencies of transfers by vacancy origins, cell frequencies should be independent of the impacts that vacancy status might have on cell frequencies by vacancy destination. In other words, if parameters are stationary, then vacancy status and vacancy destination will be statistically independent categorical variables. Variation within the data will be completely associated with the relationships that exist between (O and D) and (O and S).

A model of conditional independence is defined to reflect the independence of D and S. This model specifically includes the interaction between vacancy origin and destination. It also includes the interaction between vacancy origin and the observed versus projected status of vacancy transfers, but it explicitly excludes interactions between vacancy destination and status plus the three-way interaction among O, D and S.

A much smaller G^2 statistic is computed for the model of conditional independence. It identifies the variation remaining after the correlation between vacancy origin and destination and vacancy origin and status have been considered. One hundred times the ratio of the smaller G^2 statistic to the larger G^2 statistic is regarded as the percentage variation due to errors in the projection methodology. Errors are largely attributed to non-stationarity in the fundamental matrices of the Markov and Leontief models though they might also be due to non-homogeneous housing sector definitions or to data errors.

Log-linear analysis used by Emmi and Magnusson (1994a) on models with six identical but not necessarily homogeneous housing sectoral definitions shows projection errors to range from 3–12 per cent. Log-linear analysis used by Magnusson (1994) on models with 8–10 more nearly homogeneous housing sectoral definitions shows projection errors to range from 2–5 per cent. (Non-homogeneous housing sectors can be more accurately defined as two or more smaller sectors each with its own unique vector of vacancy transition probabilities. All reasonable sub-divisions of homogeneous sectors have transition probabilities that are as alike as possible given data density limits. Emmi and Magnusson (1993) explain why lower degrees of homogeneity increase projection error.) Since the 1975, 1980 and 1985 data from Statistics Sweden contain a very low proposition of data fields with missing data errors, these results give preliminary impressions of the 5-year predicative accuracies available with vacancy chain models in several different field settings where data errors are minimal.

4. Current Findings

With the addition of data on vacancy transfers across a third time-period (1985–90), Markov and Leontief model-based projections can be made over the 1985–90 period using models calibrated with data from both the 1980–85 and the 1975–80 periods. The accuracy of these projections can then be compared to accuracy of 1980–85 projections made with models calibrated on 1975–80 data. This can be done using both the housing sector definitions reported in Emmi and Magnusson (1994a) and the definitions reported in Magnusson (1994).

Results based on the less internally homogeneous six-sector definitions reported in Emmi and Magnusson (1994a) are presented in Table 1. Table 1, Part A, shows the percentage errors originally reported in that earlier article for projected moves during 1980–85, based on models calibrated with 1975–80 data. Table 1, Part B, shows the percentage errors for projected moves during 1985–90, based on models calibrated with 1980–85 data. Table 1, Part C, shows the percentage errors for projected moves during the 1985–90 period, based on models calibrated with data from a period 10 years earlier.

If data errors are approximately the same across the three time-periods and if the internal homogeneity of housing sectoral definitions is adequate, then a comparison of Part A and Part B results should indicate whether parametric stability has increased or decreased through time. A comparison of Part C results with those of Parts A and B indicates how projective accuracy deteriorates as one projects further into the future.

The average projection error among the six figures shown in Part A is 6.2 per cent. For Part B, this figure is 7.4 per cent. The new projections are about as accurate as the prior results. The overall differences seem to be a marginally higher projection error on average and a marginally higher variation among error terms. These differences are thought to be due both to an increase in data error and to an increase in parametric instability during what is generally regarded as the most turbulent of the three observation periods.

Table 1. Percentage variation between observed and predicted moves attributed to projection error using models with six housing sectors

Part	Projection period	Calibration period	Model	Gävle	Jönköping	Västerås
A	1980–85	1975–80	Markov	9.3	4.6	3.5
			Leontief	11.8	4.8	3.0
B	1985–90	1980–85	Markov	6.4	9.2	7.3
			Leontief	6.4	10.6	4.7
C	1985–90	1975–80	Markov	13.0	13.9	11.3
			Leontief	18.5	13.5	7.7

Table 2. Percentage variation between observed and predicted moves attributed to projection error using models with 8–10 relatively more homogeneous housing sectors

Part	Projection period	Calibration period	Model	Gävle	Jönköping	Västerås
A	1980–85	1975–80	Markov	8.1	5.3	4.9
			Leontief	7.2	6.1	5.0
B	1985–90	1980–85	Markov	7.3	7.6	7.8
			Leontief	7.1	9.4	5.4
C	1985–90	1975–80	Markov	15.3	14.0	11.1
			Leontief	14.7	14.1	9.0

If the errors can be attributed exclusively to parametric instability (a questionable proposition), then the rate of parametric change appears to have slowed in Gävle and increased in Jönköping and Västerås. (However, the findings presented in Table 2 constitute better evidence with which to consider this conjecture.)

The average projection error among Part C results is 13.0 per cent. This figure is as would be expected given the doubled projection period of Part C results.

Part B results justify continued confidence in the accuracy of vacancy chain models over projection periods up to 5 years long. The magnitudes of the errors noted in Part C indicate mixed results. The range of results varies from quite good to marginally bad. Caution must be urged when using a 10-year projection horizon with sectors that are not fully homogeneous.

Results based on the more homogeneous housing sector definitions reported in Magnusson (1994) are presented in Table 2. The internal organisation of that table is the same as Table 1. However, the use of 8–10 internally homogeneous housing sector definitions permits an interpretation of projection errors as being more exclusively due to parametric instability or data errors and not to errors induced by non-homogeneous housing sector definitions.

Average projection error for Part A is 6.1 per cent. For Part B, it is 7.4 per cent. For Part C, it is 13.0 per cent. These are very nearly the same as the average errors in Table 1. Yet the results using models with more internally homogeneous sector definitions vary much less than those of Table 1. Apparently, sector definitions that are more internally homogeneous yield projection results that are equally accurate but considerably more consistent.

Within Table 2, the average projection

error for Part B (7.4 per cent) is moderately larger than for Part A (6.1 per cent). Part B results also vary more than do Part A results. Again, these differences are thought to be due both to an increase in data error and to an increase in parametric instability. Part C results are again an expected function of Part A and Part B results.

Evidence of important interperiod differences in parametric stability is weaker in Table 2 than in Table 1. Earlier assertions that the rates of parametric change have slowed in Gävle but increased in Jönköping and Västerås are still supported by the data in Table 2, but less strongly than before.

Average projection error using the Markov model differs negligibly from average projection error using the Leontief model. The Markov model gives marginally more consistent results that vary over a narrower range (see Emmi and Magnusson, 1993). This could be due to the fact that housing sector definitions were developed for the Markov model, not the Leontief model. That is, sector definitions were designed to minimise variation among transition probabilities within each sector and not to minimise variation among vacancy input coefficients. This could mean that sectors are marginally less homogeneous for the Leontief model and thus marginally less reliable. Otherwise, both models perform equally well.

5. Conclusions

Part A and Part B results from both Tables I and 2 confirm our continued confidence in the quality of 5-year projections. They also confirm our confidence in both the Markov model and the Leontief model. However, the magnitudes of Part C errors require us to repeat a cautionary note concerning the use of 10-year projections. This is particularly so when one contrasts these tests results with typical field applications of vacancy chain models. There the vectors of vacancy-initiating events (for the Markov model) or vacancy absorbing events (for the Leontief model) must be projected, too. The projection errors in these vectors will interact with the errors noted here to yield results that will be even less accurate. The combination of these two errors might easily double the projection errors noted above. Results would still be very acceptable over a shorter 5-year projection period. But over a 10-year projection period, combined projection errors of 25 per cent or more would be common. These would have to be used with considerable caution and could serve only to indicate the most general of trends.

The contrast between Part A and Part B results leads one to speculate on the differences between periods in the rates with which social system change occurs. Assume that 80 per cent of the error in Part A is due to parametric instability (20 per cent due to data error), this implies a 5.0 per cent rate of change between 1975–80 and 1980–85. Similarly, to assume that two-thirds of the error in Part B is due to parametric instability, implies a 6.0 per cent change between 1980–85 and 1985–90. A common assumption is that change during the latter period was much greater than change during the former period. However, the magnitudes of change and the difference in the rates of change (5.0 versus 6.0 per cent) are both very small. These small changes indicate considerable stability in the fundamental matrices of the Markov and Leontief models. These, in turn, imply great stability in the underlying patterns of inter-group relations in society particularly as they are expressed in urban housing market operations. The common perception of social change must relate to something quite different from what we are measuring here for these results indicate a degree of social system stability that few would have expected.

The possible relationship between system transformation and model performance is an interesting one. The possible relationship between the rate of system growth and the relative performance of the Markov model versus the Leontief models is also of interest. Vacancy chain models should be tested on cities that are known to have experienced considerable social structural realignments, on cities that have grown very rapidly and on cities that are in a state of severe decline. Vacancy chain models applied to such situations would help us to understand better the impacts of these kind of changes on the ways in which housing markets work.

References

EMMI, P.C. and MAGNUSSON, L. (1993) Intrasectoral homogeneity and the accuracy of multisectoral models, *Annals of Regional Science*, 27, 343–362.

EMMI, P.C. and MAGNUSSON, L. (1994a) The predictive accuracy of residential vacancy chains, *Urban Studies*, 31, pp. 1117–1132.

EMMI, P.C. and MAGNUSSON, L. (1994b) *Residential opportunity as a basic unit of analysis in housing studies*. Paper presented at the *European Network for Housing Researchers 1994 Conference*, 29 August–2 September, Glasgow.

EMMI, P.C. and MAGNUSSON, L. (1995) Opportunity and mobility in urban housing markets: concepts, models calibrations and tests, *Progress in Planning*, 43(1), pp. 1–88.

KNOKE, D. and BURKE, P. (1980) *Log-linear models*, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-020. Sage Publications: Beverly Hills.

MAGNUSSON, L. (1994) *Omflyttning på den Svenska Bostadsmarknaden: En Studie av vakanskedjemodeller* [Residential Mobility in the Swedish Housing Market: A Study using Vacancy Chain Models]. Geografiska Regionstudier 27, Kulturgeografiska Institutionen vid Uppsala Universitet, Uppsala.