IAPR TECHNICAL PAPER SERIES

RANDOM UTILITY SIMULATION OF SPATIAL ECONOMICS

John E. Abraham
and
J.D. Hunt

Department of Civil Engineering and
Institute for Advanced Policy Research
University of Calgary

October 2007

Technical Paper No. TP-07009

Institute for Advanced Policy Research
University of Calgary
Calgary, Alberta
Canada

http://www.iapr.ca
iapr@ucalgary.ca

Correspondence: John Abraham, jabraham@ucalgary.ca and J.D. Hunt, jdhunt@ucalgary.ca

Random Utility Simulation of Spatial Economics

John E. Abraham and J. D. Hunt

University of Calgary, 2500 University Dr. NW
Calgary, AB, T2N 1N4, Canada
{jabraham,jdhunt}@ucalgary.ca

Abstract. Using random utility theory for production and location decisions can lead to a consistent and general modeling approach with a number of distinct advantages. The paper provides a brief overview of Walrasian economic theory, input-output models, and random utility theory. It then shows historically how these have been integrated together into spatial planning models. Finally the use of these spatial planning models as a platform for future research into fully dynamic agent based spatial economic models is described.

Introduction

A general system for spatial economic simulation would be useful for research and application in a number of fields: transportation economics, transportation planning, land use planning, urban economics and regional science, among others.

Such a system would require an explicit representation of

1. the costs of transport,
2. the use of land,
3. interactions that occur between activities in locations, that require transport,
4. the heterogeneity of goods and services, and
5. the heterogeneity of agents.

The requirement for an explicit representation of heterogeneity has been shown in transportation demand modeling: for any pair of locations there seems to be a non-zero probability of a trip occurring between them – even between the most distant pairs of locations. The agents at each location must offer something unique that attracts trips to them; otherwise the more distant agents would choose a closer agent to interact with.

Transportation modelers have, for the most part, settled on Random Utility Theory (Ben-Akiva and Lerman, 1985) as the appropriate technique for representing heterogeneity in travel, and consumer demand modelling has more tentatively adopted random utility theory (Brown and Walker, 1989; Anderson et al, 1992) to address specific issues.

Over the past three decades many spatial planning simulation models have adopted random utility location choice theory, and integrated it into traditional economic models. These integrated simulations have been commercially successful and practical for policy analysis, but there have been a number of theoretical inconsistencies in the integration of the two theories. Recent work seems to overcome most of these inconsistencies, and current work seeks to extend these simulations into the realm of agent based simulations of spatial interactions.

Theories

Walrasian view

The standard microeconomic "grand view" of an economy has the following distinguishing factors (Katzner, 1988)

1. **Categories of goods and services:** All goods in the same category are treated as homogeneous, and are measured in the same units.
2. **The world is a collection of homogenous categories of households and establishments that are optimizers.**
3. **A market and a single price for each category of good and service:** Goods are bought and sold at a price. At any time there is one price (and one market) for each category of good. Within a market there are a large number of small players.
4. **Equilibrium assumption (short run and long run separately):** For most goods there is an equilibrium between supply and demand, the exceptions being "capital" goods such as money capital, machinery, and land improvements. There is the option of using a long run equilibrium, where prices are a function of input costs (profit is zero) and for households, income equals expenditures, or a short run equilibrium, where short run supply curves and expenditure functions are specified as additional inputs. (Rutherfod and Paltsev, 1999)
5. **Usually no need for modelling processes:** The equilibrium price is established through some process that doesn't involve trade, and then all trade occurs at that equilibrium price, so that the process for determining the equilibrium price does not need to be explicitly represented.
6. **There is no cost to entering a market.**
7. **The same information is available to all participants.**

This standard microeconomic view is the basis for computational general equilibrium (CGE) models, which are a system of equations describing the economy based on the flows of goods between firms and between firms and households and the production and expenditure functions of firms and households.

Input Output

An input-output model is consistent with the predominant Walrasian view of economics, but emphasizes the quantity of transactions. It generally adds the following restrictive assumptions:

1. constant (Leontief-style) production functions (often specified as direct relationships between and among households and industries, i.e. a “square” matrix of coefficients, rather than through the intermediate categories of goods and services.)
2. a fixed ratio of production for exports to production for internal consumption, and
3. assumption of a long-run equilibrium.

The fixed production functions remove the assumptions of optimizing behaviour.

Random Utility

The traditional Walrasian economic paradigm is aspatial – the only spatial component is the separation of the model region from the external economy. In addition, the traditional paradigm assumes homogeneity in categories. Random utility theory has been used to relax both of these assumptions.

Random Utility Theory assumes heterogeneous actors and/or heterogeneous goods. Each actor is assumed to maximize its utility. Its utility is assumed to consist of two components – a deterministic component that can be calculated based on model equations, and a stochastic component reflecting the uniqueness of individuals and situations that varies according to a distribution.

The dominant application of random utility theory in spatial modelling is through *discrete choice* models, representing the choice of one alternative over others in a set of alternatives. In the *logit model* the stochastic component of the utility is assumed to vary with an independent Weibull distribution, which leads to the following closed-form equation for the probability of choosing one option *i* out of a set of options *J*:

$$P_i = \frac{e^{\lambda U_i}}{\sum_{j \in J} e^{\lambda U_j}} \tag{1}$$

where U_i is the deterministic component of the utility of alternative *i*, and λ is inversely related to the variance of the stochastic component.

A larger stochastic term in the utility function represents more variation in the utility of the individual options. Random utility theory assumes that actors choose the best of the available options, and so a larger variation leads to an expectation of an actor finding a better option. This is reflected in the expected maximum utility (or *composite utility*) of choosing one option from a set of alternatives:

$$U = \frac{1}{\lambda} \ln \left(\sum_{j \in J} e^{\lambda U_j} \right) \tag{2}$$

Note the composite utility is inversely proportional to λ and so is proportional to the variance in the stochastic term in the utility of the options.

In spatial economic representations, random utility is used to represent spatial choice. This can represent, for example, an establishment's choice of where to purchase inputs, a household's choice of home location, or an individual's choice of where to go to school. The model region is divided into polygons called "zones", and each zone is an element of the set of alternatives in the logit model.

Random utility theory and the logit model offer a number of advantages: a continuous response to policy inputs, easily estimated through maximum likelihood techniques, explicit valuation of situational and policy attributes, and closed form equations for calculating consumer surplus and producer surplus.

Random Utility Spatial Models

MEPLAN style spatial IO

For over 25 years, a form of spatial modelling called "spatial input-output" modelling, or the "Martin Centre Model" (Hunt and Simmonds, 1993), has been operationalized in land-use/transport models for urban planners and transportation planners. These models use the input-output framework, but introduce an element of spatial choice using logit models. These have been encapsulated into the commercial software packages TRANUS and MEPLAN.

In these models, the logit model equation 1 and the law of large numbers is used to apportion inputs amongst zones. For industry, the factors and intermediate demand required by the industry located in one zone is apportioned to be produced amongst all zones, assuming that the portions are equal to the calculated probabilities. For households, the final demand for goods and services by households in one zone is allocated to be produced in other zones based, again, on the probabilities in equation 1.

These models use the square input-output model, where the role of categories of goods is not explicitly represented. The ownership of land and capital by households is not generally represented – households are assumed to supply labour only, and land and capital are represented separately, as fixed inputs (vertical supply curves).

The price of each sectors' output is allowed to vary by zone, and that price is a component of the utility function for that zone (the other principle component being transport costs.) Overall production costs are assumed to be the weighted average of the prices plus transport costs for each of the zones that are used for inputs. A "long range" equilibrium is assumed – with prices equal to the sum of these weighted average costs.

The framework is usually made "quasi-dynamic" by computing a point equilibrium for a series of time steps, and

1. assuming floorspace quantities are fixed in each zone at the end of each time step,
2. using fixed transport costs and transport disutilities in computing the equilibrium (then updating them in the next time step using a transport demand model), and
3. adding an "inertia" term to the utility of producing in any one zone

The Martin Centre model was an important development in bringing economic modelling to transportation planning and urban planning, and for extending input-output modelling theory to space, and it continues to be used for practical policy analysis (e.g. Abraham and Hunt, 1999). But it does have some theoretical inconsistencies (Abraham, 1998):

1. It uses the logit model formula for allocating space to zones, but does not follow through with adopting random utility theory. In particular, it uses the law of large numbers (assuming portions are equal to probabilities) even when categories are small, and it uses weighted average costs, instead of the expected maximum utilities derived from integrating over the range of possibilites, as the basis for calculating long-run equilibrium prices.
2. It adopts the equations associated with a long-run equilibrium, yet uses a short-run calculation of transport costs and of floorspace supply.

PECAS spatial IO

The PECAS model is a more complete and consistent integration of Walrasian economic principles with random utility location choice. PECAS is described more fully in Hunt and Abraham, 2005, and Abraham and Hunt, 2007.

PECAS uses the rectangular make and use matrices to represent production and expenditure functions more fully. Thus there is an explicit representation of the transaction of goods and services from supplier to demander. PECAS uses a logit model to allocate the interactions in space. The allocations are done to 'exchanges', which are submarkets for a particular good within the regional market. The price at the exchange is a term in the utility function, with a positive coefficient for sellers and a negative coefficient for buyers. The prices of goods at exchanges are adjusted to clear the market at each exchange – leading to a price landscape for each commodity.

PECAS uses equation 2 to measure the attractiveness of selling (or buying) a good to (or from) the set of available exchanges. These attractiveness functions for all relevant goods influence the choice from amongst different “Technology Options” available to each industry or household category, causing establishments or households to consume or produce more of those goods that are easier to come by or sell. The set of attractiveness functions are combined in a way consistent with the theory of the production or expenditure function (Anderson *et al*, 1992), to measure the overall attractiveness of a location as a business or residence location.

PECAS can be considered a type of additive nested logit model with three levels (Abraham and Hunt, 2007):

1. The choice of where to locate,
2. The choice of the quantities of production or consumption, conditional upon the chosen location of the establishment or household, and

3. The choice of where to buy or sell (or otherwise acquire or divest) various inputs and outputs, conditional upon both the location of the establishment or household and on the quantities produced or consumed.

The full use of equations like equation 2 to ensure consistency between these three levels in PECAS means that PECAS adopts random utility theory more completely than the Martin Centre models. PECAS is based on the assumption that aggregate flows and stocks are made up of individual actors and individual transactions, each unique. This is critically important in this type of modelling because the primary advantage of urbanization is the wide range of goods and services available to those living and doing business in the city (Jacobs, 1969). The logit model equation 2 shows that the composite utility increases when there is a wider range of options available.

In PECAS, the amount of each industry or household category is specified, and that quantity of activity is allocated to zones, and production functions, and flows between locations, using the nested logit formulation. The prices in each exchange zone are adjusted to clear the markets for commodities consistently with the nested logit allocation process. Floorspace (and, possibly, other capital items) are given vertical supply curves in each zone. Thus the PECAS equilibrium is a short run equilibrium, since there is no assumption that prices must equal costs, that profit must equal zero, nor that income must equal expenditures.

PECAS is a quasi-dynamic model, and uses four strategies to move through time. The first three strategies are the same ones described above for the Martin Centre models (floorspace development as a separate dynamic model; stepping through time jointly with a transportation route and mode equilibrium model; and using inertia terms.) PECAS also requires a dynamic representation of the change in industry size. The log sum (equation 2) from the top level location choice model is a measure of the overall welfare or attractiveness of the region to an industry or household category. This top level index is a measure of region wide consumer or producer welfare and can be used to guide a long term dynamic model of migration and industry growth and decline.

PECAS is an aggregate model, and so uses the law of large number to allocate quantities using equation 1.

Dynamic Economic Models

Dynamic approach with timeless behaviour

The PECAS model is a consistent integration of random utility theory for transaction/production/location choice with a general equilibrium model of an economy. The system of equations in PECAS is solved by adjusting the prices at exchange locations until each local submarket clears.

A dynamic model similar to PECAS can be implemented by:

1. solving the PECAS allocation equations given a set of initial prices at the exchanges,

2. calculating the *excess demand* (shortage or surplus) at each exchange, and
3. adjusting the price at each exchange over time based on the excess demand at that exchange.

This is consistent with the dynamic approach to Walrasian market theory in which an equilibrium system can be converted to a dynamic system by adding only a formal specification of how prices change (see Katzner, 1988, p 258). Thus, the "complete dynamic system, then, contains two parts – the economic, market excess demand equations; and the dynamic or Walrasian adjustment rule" (p. 264). The Walrasian approach assumes that behaviour is timeless or instantaneous, but reacts to prices that change over time. PECAS could be operationalised as a Walrasian dynamic model by replacing the price search algorithm with a price update algorithm, and appropriately considering the stability of the resulting system.

Transaction simulation

Any theory that relies on a "price update" function still requires an average price for a good, and relatively smooth responses for the excess demand function. As models move to more spatial accuracy (more zones), and more temporal accuracy (more time steps) inevitably smaller and smaller submarkets will be represented. The "market" will no longer be representable by a landscape of average prices, because any such averages will be over such small areas of space and time that most will contain only 1 or 0 transactions. The market needs to be explicitly represented as a large number of individual transactions, each unique in time and space as well as in any stochastic error term associated with unobserved heterogeneity in random utility theory.

A method of representing markets as individual transactions has been developed. It was first described in Abraham and Hunt, 2001. The technology coefficients associated with industrial categories and household categories are divided into lists of desired transactions. Each of these desired transactions may involve a transaction between two parties. The transaction process is divided into three parts:

1. An offer to purchase or sell a quantity of goods or services in a particular place at a particular price,
2. An acceptance of an offer, and
3. An update of expectations.

Instead of using equation 1 to calculate the probability of accepting or making an offer, randomly generated attributes are used to describe the offer, and specific functional forms are specified for how the attractiveness of an offer to another party changes with the random attributes. Thus the heterogeneity of goods and services is represented directly, and the random component of the utility function is no longer required. Equation 2 is also not necessary – the maximum utility is directly chosen; the expected maximum is not required.

Global search process In the PECAS model described previously the nested and additive nature of the logit model means industries and household categories make consistent choices between the three levels of PECAS: the exchange location choice (travel to interact with other agents), technology (quantity of various interactions), and home location. For example, home location choice takes into account the full range of interactions with other agents that are possible from each home. The PECAS model takes into account the full range of options by calculating various forms of equation 2 and adding them together in specific and careful ways (Abraham and Hunt, 2007). In a dynamic simulation model with an explicit representation of time not all of the options can be explicitly considered – it would require too much computational power for models of real cities or regions. Further, a full consideration of all options is not realistic. No real establishments or households can explicitly evaluate all of the combinations of options regarding home location, technology option or lifestyle, and exchange location and partner for every interaction. A more realistic search process must be developed.

There is concern that the search process does not get stuck on local maxima. For instance, the current home location might always appear optimal unless the agent also considers large changes to the pattern of destinations and interactions as well.

Current research involves using modified genetic algorithms to properly consider the full interactions of the various choices in the PECAS model, but in an agent based transaction simulation framework. Individuals use genetic algorithms to try to find the best option for them, but to manage the size of the computational problem the cross-over process uses patterns of spatial interaction that have been successful for other similar agents (peers). This algorithm is behavioural, in that many people do imitate their peers and learn from what has been successful for others.

Assumptions The assumptions in the transaction simulation framework are much less restrictive than the assumptions in the earlier models. The framework is much more realistic. Yet that realism does not come without a cost: the conditions under which such a system will possess an equilibrium state have not been shown, nor whether the equilibrium state is stable locally or globally. The reaction responses of the various actors in the system could lead to a chaotic or divergent system. (Although real systems could be divergent or chaotic, it is an immense simplification of modelling to work with a mathematical representation that is convergent.)

The assumptions are still different than reality, however. For instance, for many goods, real offers are vague, and a period of bi-lateral negotiation occurs between a potential buyer and seller. In oligopolistic markets players may be involved in some game-theory approach to interacting and setting offer prices. The framework implies that participants' only knowledge of the market is the list of current and previous offers and transactions, and that an offer is a public offer that can be accepted by anyone. Any proof of stability is likely to assume that there are a large number of small players.

It is important to note that there is no need for any averaging in the simulation. Individual actors could use some averaging in their decision making, if there is some behavioural reason to believe that individuals consult or calculate average conditions before making decisions – but such averaging is a behavioural enhancement, not an essential part of the framework. Thus, model inputs can be provided using any geography and time scale, and model outputs can be summarized using any geography and time scale. Zones and time steps become conveniences for arranging inputs and outputs – and are no longer integral to the modelling itself.

Conclusions

This paper reviewed a number of spatial market representations. The spatial market representations varied in their complexity and completeness:

- The MEPLAN style "Martin Centre" models combine input-output theory with a spatial choice model, giving a comparatively simple computational equilibrium model with a spatial component. The spatial choice model isn't completely consistent with random utility theory, and the use of the long-run equilibrium concept is inconsistent with the use of time steps, fixed floorspace supply, and iteration with a transport model. Nonetheless, such models have achieved a level of orthodoxy in practical planning, with stable software implementations and a number of individuals and firms experienced in building and using such models, and a track record of large influences on policy decisions.
- The PECAS model, being similar to the Martin Centre models, but adopting Random Utility Theory more completely, removing the inappropriate long-run equilibrium assumption, and specifically representing the role of goods and services in the economy.
- An extension of PECAS into a dynamic model, by adopting a price-update paradigm.
- A transaction based simulation, broadly consistent with the above approaches but supporting infinite spatial and temporal detail.

As one moves down the above list, the models become:

- more theoretically consistent,
- more realistic behaviourally,
- less proven in practice and in research, and
- supporting (but not necessarily requiring) more detailed models.

Each model seems appropriate for some uses, given the current state of research.

It is hoped that, in time, it will be proven that the transaction simulation paradigm can have stable behaviour consistent with the integration of a computable general equilibrium model and random utility theory, but can also provide a richer and more realistic representation of spatial economics. This will require some further research (Miller *et al*, 2002), and should benefit from research currently underway in biological systems.

The current status of this research stream already contains substantial advances in **spatial economic models**, and models of interactions between individuals and firms. A direct **model of markets** has been added to the system through the offer-accept discrete-event simulation. This stable and practical platform has provided a basis for further work in spatial economic modeling, by the adoption of **evolutionary algorithms** that embody learning and adaptation. When those evolutionary algorithms are successfully integrated into the spatial models, the result will be a simulation that represents economic network formation.

Finally, it is important to note that the predecessor models have been used for substantial policy analysis for decades. Extending these models using agent based approaches will bring those approaches into an established realm of practical policy analysis.

References

1 Abraham, J.E., 1998, "A review of the MEPLAN modelling framework from a perspective of urban economics", University of Calgary, Department of Civil Engineering Research Report No. CE98-2, available at http://www.ucalgary.ca/~jabraham/MEPLAN_and_Urban_Economics.PDF

2 Abraham, J.E. and J.D. Hunt, 1999, "Policy Analysis using the Sacramento MEPLAN Land Use Transportation Interaction Model", Transportation Research Record 1685:199-208 and presented at the 1999 annual meeting of the Transportation Research Board

3 Abraham, J.E. and J.D. Hunt, 2001, "Dynamic Microsimulation of Heterogeneous Spatial Markets", presented at the Workshop on Economics with Heterogeneous Interacting Agents, University of Maastricht, The Netherlands, June 7-9 2001 (abstract refereed)

4 Abraham, J.E. and J.D. Hunt, 2007, “Random utility location/production/exchange choice, the additive logit model, and spatial choice microsimulations”, accepted for publication in *Transportation Research Record* and presented at the 2007 Annual Meeting of the Transporation Research Board.

5 Anderson, SP, de Palma, A., and Thisse, J, 1992, *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge, MA.

6 Ben-Akiva, M., and Lerman, S.R., 1985, *Discrete Choice Analysis*, MIT Press, Cambridge, MA

7 Brown, B.W., and Walker, M.B., 1989, "The Random Utility Hypothesis and Inference in Demand Systems", Econometrica 57:815-829

8 Hunt, J.D. and J.E. Abraham, 2005, Design and implementation of PECAS: A generalized system for the allocation of economic production, exchange and consumption quantities. In: *Foundations of Integrated Land-Use And Transportation Models: Assumptions and New Conceptual Frameworks*, Lee Gosselin and Doherty editors

9 Hunt, JD and Simmonds, DC 1993, "Theory and application of an integrated land-use and transport modelling framework" Environment and Planning B 20:221-244.

10 Jacobs, J, 1969, "The Economy of Cities", Random House (Vintage Books), NY.

11 Katzner, D.W., 1988, *Walrasian Microeconomics*, Addison-Wesley, Reading, MA

12 Miller, E.J., Hunt, J.D., Abraham, J.E., Salvini, P.A., 2004, "Microsimulating Urban Systems", *Computers, Environment and Urban Systems* **28**:9-44.

13 Rutherford, T. and Paltsev, S, "From an Input-Output Table to a General Equilibrium Model: Assessing the Excess Burden of Indirect Taxes in Russia", unpublished draft, Department of Economics, University of Colorado.