13. Choice models of travel demand in practice: current trends, problems and welcome research directions

Peter Vovsha

1. INTRODUCTION AND SUMMARY

The chapter contains an overview of some new choice model structures applied in practice, primarily in the framework of regional travel models including a growing number of activity-based models (ABMs). The focus of the chapter is on the new choice dimensions, associated techniques and problems, and possible research directions that would be welcome from the practical perspective. The discussion is structured by the following main topics:

- *Meaningful parameterization of GEV models* In the currently adopted theoretical form of the GEV models of nested logit (NL) and cross-nested logit (CNL) type, nesting coefficients and allocation parameters are not parameterized. Possible variation in these parameters (i.e. variation of the nesting structure itself) should be handled by an *a priori* segmentation. There is a growing number of examples in practice where a parameterization of these model components makes a lot of sense and practitioners went ahead in doing this on an empirical basis. The park and ride (PNR) alternative in a mode choice context represents an appealing example where the allocation coefficients in a CNL structure are naturally dependent on the proportions between the auto leg and transit leg.
- *New choice dimensions* that are far beyond trip frequency, destination, mode and departure time. ABMs are characterized by a large number of new dimensions where conventional choice models look sometimes awkward. As a result new ideas come up and new demands for modeling tools emerge.
- *Parallel convoluted choices* A good example is choices made by several household members in parallel with interactions between

them. Choices of this complexity are not considered in a conventional trip-based model where each trip is essentially modeled separately. Examples considered in the chapter include individual daily activity pattern (DAP) type and commuting mode choices for workers and children in the same household. A salient feature of these choice contexts is that they are largely driven by person characteristics (that's why interactions are ignored in simple models) but there are substantial interactions on top of that. Hence we consider these choices as parallel and convoluted.

- *Multi-dimensional ordered choices* Another particular aspect relates to multidimensional choices where alternatives are combined in a combinatorial way of joint ordered frequency models. These models quickly suffer from a combinatorial "explosion". It is essential to better distinguish between qualitative (set of activities) and quantitative (number of episodes of each activity type) aspects of the choice. An analogy between choice models and continuous models like microeconomic models of consumer behavior is appealing but has not resulted in a comprehensive solution yet.

2. THEORY AND PRACTICE

With the advent of ABMs in US practice in the period 2000–10, demands for more flexible and advanced choice models grew significantly. ABMs include new choice dimensions that are very different from trip mode choice, destination choice, or departure time choice that had been most frequently used as prototypes for travel choices. ABMs are associated with a more general view of travel behavior compared to conventional trip-based models with multi-way linkages between different choice dimensions. ABMs in practice have been following different designs that reflect different assumptions on causality and order of different decision-making steps and corresponding choice contexts (Bradley et al., 2009; Eluru et al., 2010; Davidson et al., 2007; Vovsha et al., 2005a, 2011; Arentze and Timmermans, 2004).

A good example of how underlying model system design can affect the choice context is in using such a variable as car availability in mode choice. In conventional trip-based models, household car ownership is frequently used as one of the key variables explaining mode choice; hence a car ownership model is placed before mode choice model in the model system. In many ABMs, this causality was reversed for work commuting. The joint choice of usual work location and commuting mode is modeled prior to car ownership (Bradley et al., 2009; Vovsha et al., 2011). In fact,

car ownership in the US is largely a function of the dependence on auto for commuting. Car ownership is still used as a variable explaining mode choice for other (non-work) trips. However, even if the household car ownership is known it is just a proxy for actual car availability for the given trip. If the household has fewer cars than drivers and they implement travel tours at the same time, car availability becomes a function of intra-household priorities. If the additional assumption is made that all activity schedules are fixed, a joint model of mode choice for several parallel tours can be applied with an explicit consideration of car allocation within the household (Vovsha and Petersen, 2007). This, however, requires a scheduling model to be placed before joint mode choice in the model chain.

In general, the ABM framework allows for the inclusion of a large number of so-called situational explanatory variables, i.e. variables that are outcomes of the prior models in the chain. This naturally requires more flexible choice models that could handle large multidimensional choice sets and an intricate structure of similarities across alternatives, as well as provide efficient ways to parameterize both observed utilities and disturbances.

Practical application of choice models in regional travel model systems that are used for various project and policy analyses adds new facets to model evaluation. In theory, the primary criterion of choice model quality is a statistical measure of fit (most frequently, likelihood) that corresponds to a good replication of observed choices across multiple segments and dimensions. In practical applications, choice models are scrutinized in some different ways including a logical and intuitive response to changes in input variables. It is a very frequent situation in practice that the best model specification in terms of statistical fit exhibits illogical elasticity with respect to some input variables. Thus, one of the demands that come from the practical side is to have a controlled structure of disturbances that ensures the logical elasticity if it can be imposed a priori. A very frequent adjustment in this regard is to the nesting coefficients in a mode choice model. Very frequently in model estimation, these coefficients take values close to 1.0 (i.e. very weak nesting) while in application, strong nesting is needed to resolve situations that are similar to the "red bus – blue bus" example when several transit services compete in the same corridor. Another frequently applied technique to resolve complicated multi-dimensional choices is to apply complicated component-wise utility expressions that are different from pure generic or alternative-specific coefficients.

In this chapter, we discuss several typical examples of choice situations associated with advanced ABMs in practice and corresponding specification issues. They include parameterized nested structures and combi-

natorial multidimensional choices that tend to "explode" if all possible alternatives are listed mechanically. Specifically, for handling multidimensional choices with large combinatorial choice sets, we consider such techniques as convoluted parallel choices (and the analogy with genetic algorithms) and ordered frequency choices (and the analogy with continuous models).

3. PARAMETERIZED NESTED LOGIT STRUCTURES

The recent developments in GEV model structures allow for a substantial flexibility in terms of the nested structure, number of nesting levels, allocation of alternatives to the nests and utility expressions (Daly and Bierlaire, 2006; Fosgerau et al., 2010). However, the parameters that define the nesting structure, i.e. nesting coefficients and allocation coefficients, are estimated per the entire dataset/segment and cannot be made a function of person or trip variables. There is strong evidence that the nested structure itself (i.e. the assumed correlation pattern between the disturbances) depends on certain variables the same way the (observed) utility function does.

One example relates to the park-and-ride (PNR) mode that mixes features of auto and transit in the mode choice context. A cross-nested logit (CNL) model can handle this mode by splitting it between the auto and transit nests. However, the proportion (i.e., allocation coefficients) will be fixed per each segment in the model estimation. In reality, this proportion is largely a function of the relative length of the auto leg compared to the transit leg. For example, in the New York Metropolitan Region where an ABM has been applied for the last 10 years there are indeed two very different spatial segments: Long Island and eastern New Jersey – from which many residents commute to the Metropolitan Core (Manhattan). By virtue of the very different geography, urban structure and travel conditions, the observed PNR types are very different as shown schematically in Figure 13.1.

On the Long Island side, travel conditions most frequently dictate a short auto access leg followed by a long train ride. On the New Jersey side, traffic conditions allow for a relatively long auto access leg followed by a short train ride. The Long Island PNR type can be classified as "almost transit" mode with auto access. The New Jersey PNR type can be better classified as "almost auto" mode with remote parking and auxiliary transit leg. Consequently, in forming the CNL structure it would be reasonable to assign different allocation coefficients reflecting the differential degree of similarity of PNR to auto and transit modes across the observations.
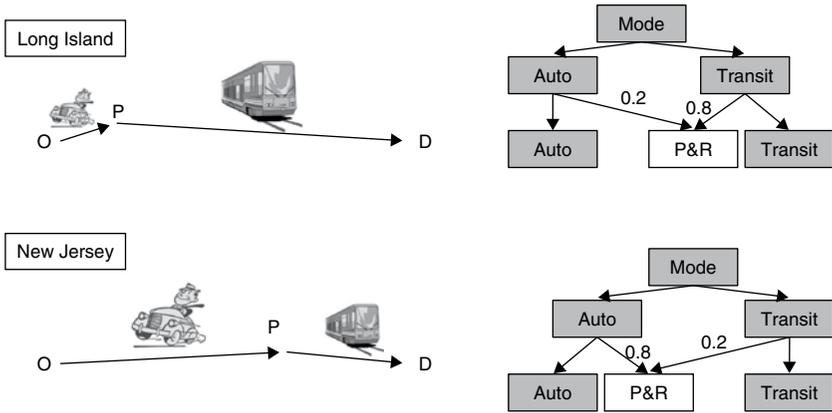
Figure 13.1   Different park-and-ride types

However, it is impossible to estimate a model of this type with the existing software and methods. It is of course possible to segment the travel market (and model parameters) a priori. However, it is an effective solution only if the market consists of a few clear sub-markets. In reality, of course, we observe a large gray area between two extreme cases described above.

To model a continuous range of possible proportions between the auto leg and the transit leg, the allocation parameter should be parameterized and estimated as a function of the corresponding distances (or travel times). A plausible assumption would be that the allocation coefficients be set proportional to distance itself (that is how similarities between network routes are measured in the link-nested logit model (Vovsha and Bekhor, 1998):

$$\alpha_{i=PNR,m=Auto} = \frac{D_{Auto}}{D_{Auto} + D_{Transit}} \qquad (13.1)$$

$$\alpha_{i=PNR,m=Transit} = \frac{D_{Transit}}{D_{Auto} + D_{Transit}} \qquad (13.2)$$

where:

α_{im} = inclusion coefficient for alternative i in nest m,
D_{Auto} = length of the auto leg of PNR,
$D_{Transit}$ = length of the transit leg of PNR.

However, this is an arbitrary setting although a behaviorally appealing one. It would be beneficial to estimate various possible functions for allo-

cation coefficients and explore their statistical validity. For example, auto and transit distance or time can be differentially weighted as is frequently applied in the utility expressions. There is currently no way for a practitioner to estimate such a model. In the research implemented with the New York data, the mode choice model originally estimated as NL was implemented and calibrated as a parameterized CNL. The resulting elasticities in model application were favored by practitioners over the standard NL. However, the real proof of the concept would require a more formal statistical analysis and improvement of the likelihood function over a standard NL and CNL with fixed allocation coefficients.

A different way to tackle this issue is to use a mixed logit formulation with different specifications of parameterized disturbances. However, there are several drawbacks of this approach. First, it is much more complicated than a closed-form CNL with parameterized allocation coefficients. The latter is only slightly more complex in application compared to standard NL or CNL. Secondly, formulation of a mixed logit model that would mimic the parameterized CNL is not straightforward and may result in a complicated structure to estimate (especially if it is combined with a complicated structure of the utility expressions). Thirdly, a mixed logit formulation would somewhat obscure the analysis and interpretation of the results (model elasticities) that otherwise are very transparent with a parameterized CNL. Thus, a closed-form CNL with parameterized allocation coefficients is very welcome in practice.

Another related example reflects the accumulated experience with different nested structures applied for modeling mode choice. There is an ongoing discussion on the grouping of modes into nests including similarities between main (long-haul) modes like auto, bus and rail as well as similarities between transit access sub-modes like walk and auto (assuming PNR). All possible nested structures covering these modes and sub-modes can be found in practice. Consider first possible nesting of the main long-haul modes. Several different structures are applied in practice that can be related to two main dimensions of the population and travel markets – income and distance, as shown in Figure 13 2.

This preliminary meta-analysis can be substantiated by the following behavioral assumptions. Unobserved attributes for bus and rail are shared because they are both public transit modes and differ from the private automobile. However, for higher income travelers who appreciate comfort and convenience (and are willing to pay for it), rail is much more attractive than bus since rail possesses many more comfort and convenience attributes (within the vehicles and at the stations) and, primarily, possibility to spend commuting time productively. This effect logically manifests itself for longer distances more strongly than for shorter distances. As the
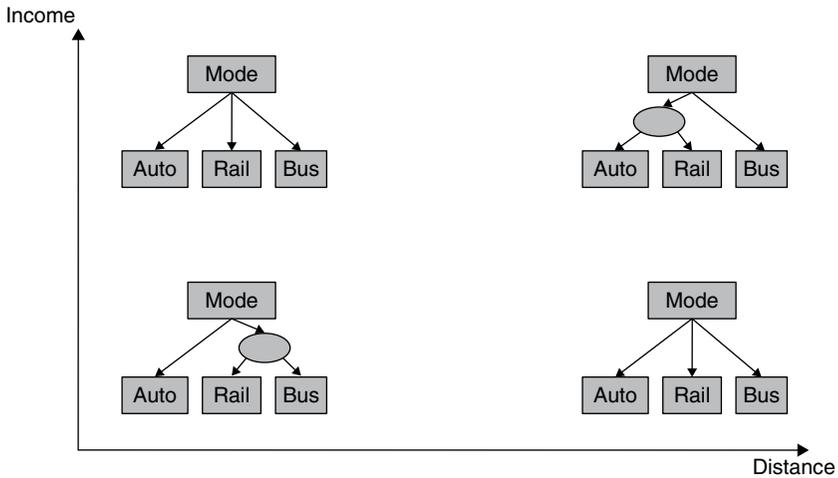
Figure 13.2   Perception of mode similarities – main modes

result of combined income-distance effects the nested structure is transformed along these two dimensions. The lower the average income and the shorter the distance the most frequently the nested structure where transit modes are nested together would be found appropriate. The higher is the income and longer is the distance the more appropriate an alternative structure where auto and rail are nested together would be. It might be close to a simple MNL when only one of the factors is strong (high income combined with short distance or low income combined with long distance).

In practice, the decisions are made specifically for each region and segment. However, it is important to recognize that this "black and white" decision can be better replaced with a parameterized model that would capture different gray areas along the income and distance dimensions. Also, there can be other variables that would affect the nesting that are currently even difficult to recognize since there is no way to estimate a parameterized model; hence multiple trials with model segmentation by different variables are required. This modeling capability of estimating parameterized nesting and allocation coefficients rather than guessing which extreme case would fit better is very welcome in practice.

Another example relates to the way nests are organized in the mode choice model with respect to transit main-haul modes and access sub-modes. There is an ongoing discussion and roughly 50–50 split in modeling practices in the US in this regard. In some models, main-haul modes (bus, rail, etc) are considered as the upper-level nests while access sub-
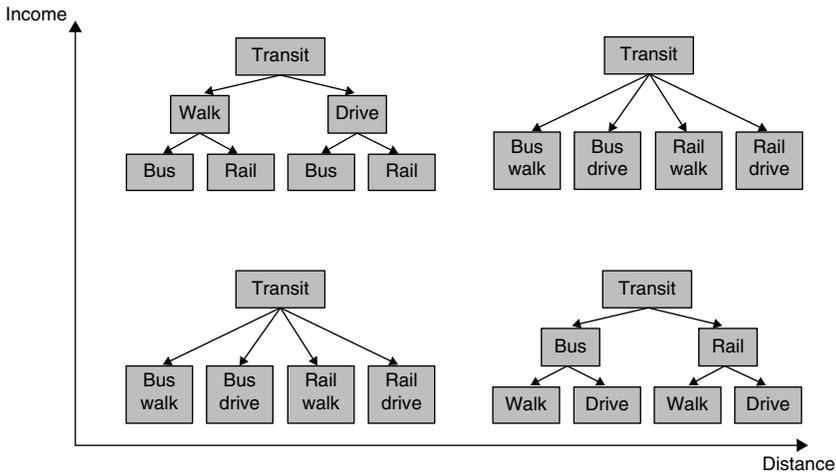
Figure 13.3 Perception of transit mode similarities – main-haul modes and access sub-modes

modes (walk, PNR) are considered as lower-level alternatives. In some other models, the order of nesting levels is reversed as the result of model estimation or calibration. As in the previous example, some meta-analysis of the applied mode choice models along the income and distance dimensions reveals a systematic pattern as shown in Figure 13.3.

With growing income, there is a tendency to emphasize the access mode. In the US, PNR constitutes about 50–70 percent of high-income transit commuters. With growing distance, there is an overall tendency to emphasize main-haul mode because it corresponds to the longer leg of the trip compared to the access leg. Superimposition of these two tendencies results in the plausible systematic transformations of the nested structure. Again, with the existing modeling techniques, only certain structures can be explored one at a time. A parameterized nested structure would be very welcome as a tool to explore these effects in more systematic way. It would be highly appreciated in practice since these transformations of the nested structure result in very different model elasticities in forecasting. Practitioners frequently scrutinize model forecasts in a comparative before-and-after fashion and are particularly interested in sources of ridership of a new or improved service. In this regard, a very well calibrated model for the base year may exhibit frustrating mode switching results for a particular geographic market since the nested structure is defined as one-for-all.

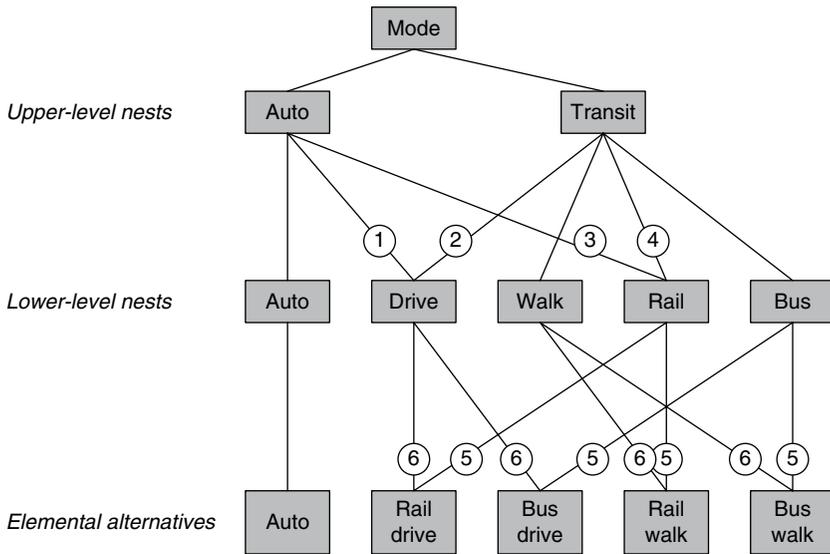The three examples above can be summarized in one hypothetical

Figure 13.4   General CNL mode choice structure combining several effects

parameterized CNL of the structure shown in Figure 13.4. The structure has two nesting levels. Two upper-level nests represent principal differences between auto and transit modes. Five lower-level nests represent differences between mode components or legs including two main-haul transit modes (rail and bus) and two transit access sub-modes (walk and drive, the latter meaning PNR). Five elemental alternatives correspond to possible mode options to implement the entire trip including auto and all possible cross-combinations of two main-haul modes with two access sub-modes.

The CNL structure is formed to address all three effects described above in the following way:

- Allocation coefficients 1 and 2 reflect the PNR split between the auto and transit nest. These coefficients should be parameterized by relative length of auto and transit legs of the PNR trip.
- Allocation coefficients 3 and 4 reflect the rail mode similarity to auto and bus. These coefficients should be parameterized by trip length and income. In general, coefficient 3 should grow with income and distance at the expense of coefficient 4.
- Allocation coefficients 5 and 6 reflect the transit mode similarities. In general coefficient 5 reflects nesting by main-haul mode while coefficient 6 reflects nesting by access sub-modes. It is expected that

with growing distance, coefficient 5 should become greater while with growing income, coefficient 6 should become greater.

In summary, estimation of parameterized nested structures with nesting and allocation parameters being functions of variables (such as income and distance) would be very welcome in practice. Attempts were made to parameterize nesting for particular (Gliebe and Koppelman, 2005) but the corresponding methods and tools have never been generalized and are not available to practitioners. As a surrogate, the author in his practice frequently applied a naïve search strategy based on multiple trials with estimation of utility coefficients with different nesting functions fixed at different levels of variables. It is useful in practice and similar to trials with alternative NL structure that one would have to implement when estimating an ordinary NL. However, this is not a substitution for the true parameterization of the nesting structure and corresponding optimization of the likelihood. Also, the number of trials may quickly become infeasible if the nested structure is parameterized across several dimensions.

4. EXAMPLES OF NEW CHOICE DIMENSIONS

ABMs in practice and individual activity analysis in general are associated with several new dimensions that frequently result in combinatorial explosion of the choice set and an infeasible number of model parameters to estimate if treated in a standard way. The complexity arises for the following typical reasons:

- "Parallel choices" made by several household members with coordination between them that calls for a simultaneous consideration of possible joint alternatives. For example, a parsimonious structure of a choice between major person daily pattern types $d = 1,2,3,4$ when considered for a 6-person household $N = 6$, results in a total number of distinct combinations $\{d_n\}_{n=1,2...N}$ that is equal to $4^6 = 4096$. The current practice to overcome the curse of dimensionality is to restrict the person daily pattern types to 3 aggregate categories and restrict the household size to 5 for simultaneous modeling.
- "Multidimensional ordered choices" of frequencies of activities of different types undertaken by individual during a day or other time horizon. For example, for a set of 10 activities $a = 1,2...10$ with a possible frequency of each $f_a = 0,1,2$ and maximum total number of activities limited to 6 (that is quite a parsimonious and realistic setting) the choice set will include 5193 alternatives even without

distinguishing the chronological order of activities. The current practice to overcome the curse of dimensionality is to aggregate the activity types into 5–6 major categories.

- Simultaneous consideration of tour departure-arrival time combinations or activity start-end time combinations with a fine level of temporal resolution. If the temporal dimension is discretized by, say, 15-min intervals, that results in more than 80 time bins $t = 1, 2 \ldots 80$ in the individual daily schedule. A simultaneous consideration of all feasible combinations for start and end of an activity $\{t_s \leq t_e\}$ would result in $(80 \times 81)/2 = 3240$ alternatives. The current practice to overcome the curse of dimensionality is to limit the temporal resolution to 30 min.

- "Multi-destination choices for trip chains". Destination choices for multi-destination tours in ABMs are normally modeled sequentially. First, the primary destination is modeled and then secondary stops are inserted conditional upon the chosen primary destination. While for work tours this one-way conditionality may be plausible, for the majority of non-work tours, identification of the primary destination is quite arbitrary and a joint choice of multiple destinations is more behaviorally appealing.

All these choice contexts have the following common features:

- An infeasible choice set arises as a result of a combinatorial explosion when a joint multidimensional choice is considered, while the structure might be quite parsimonious within each dimension. In this regard, a choice model that would obviate a direct enumeration of all possible combined alternatives is welcome. The choice situation in many respects is quite similar to network route choice and some techniques of network representation might be borrowed.

- Simultaneous modeling is essential because the choices are interdependent across dimensions and any sequential modeling would impose a problem of a simplified one-way causality. The simplified causality can manifest itself quite strongly in certain policy contexts.

- The utility function can be reasonably broken into components associated with particular dimensions. This opens a way to greatly reduce the number of coefficients to estimate. However, some interaction terms between the dimensions are still needed to address the interdependency. Moreover, for frequency-related dimensions the utility component can be further simplified by considering only the base utility of a single episode while for multiple episodes of

the same type the same utility is assumed (with some satiation effects accounted through a limited number of additional parameters).

- An intricate structure of correlations is associated with different choice dimensions. The structure of correlations normally directly relates to the choice dimensions but is not entirely defined by them. For example, for a 2-dimensional choice of tour start and end times there are two aspects of similarities associated with similar start times and similar end times. However, in addition to that, there are strong similarities along the derived activity duration (calculated as the end time minus start time).

There are appealing analogies between the underlying discrete choice model and some other modeling tools and approaches:

- *Microeconomic models of consumer behavior under constrained time and/or monetary budgets*   In this context, bringing optimization in can overcome the horrendous dimensionality of all possible combinations and create the logical linkage across dimensions (Jara-Diaz, 2007; Konduri et al., 2011). It is specifically appealing for choices that are naturally ordered along particular dimensions like activity frequency choices. Each activity can be viewed as a product in the consumption bucket while its frequency corresponds to the product quantity. Budget constraints would correspond to the observed maximum total number of activity episodes although this analogy is not straightforward. Nor is it straightforward how monetary budget constraints could be brought into this model since some activities (and associated travel cost) can be counted as expenditures while some other ones like work actually generate income. For spatial choices like multi-destination trip chain choice, constraining is required from the time-space perspective.
- *Continuous representation of naturally ordered dimensions*   like time (duration models) or frequency (analogous to amount of consumption of a certain product). The latter can be subject to satiation effects meaning that a marginal utility of each next activity of the same type decreases. Continuous representation is especially appealing if instead of the number of episodes total activity durations are considered, leading to a continuous time-allocation model combined with a discrete choice (to undertake or not at all) for each activity (Bhat, 2008; Pinjari and Bhat, 2010). Naturally, activity duration and travel times for all out-of-home activities during a day total to 24 hours minus in-home activity duration.

- *Structural equation models*   These models can efficiently connect several continuous dependent variables between them and with a large number of continuous and/or discrete endogenous explanatory variables (Golob and Recker, 2003). However, an inherent structural limitation of the structural equation approach is that the dependent variables cannot be discrete.

5.   CONVOLUTED PARALLEL CHOICES

This category of choices involves several household members undertaking activities in parallel with interactions between them; thus modeling these persons and activities independently would not be justified. Examples of parallel convoluted choices include person daily activity pattern (DAP) type, commuting mode choices for workers and children in the household, and many others. The salient feature of these choices is that they are largely driven by person and household individual characteristics (that's why interactions are ignored in simple models). For example, the person DAP is strongly affected by person worker status, student status, age, gender, income, presence of children, etc. (Bowman, 2009). Work and school commuting mode choice are strongly affected by commuting distance, available modes, car ownership, and mode time and cost differences (Vovsha and Petersen, 2007).

However, there are substantial interactions between the persons on top of that, which makes the choices dependent on each other. For example, choice of certain DAP types by some household members affect the choices of DAP of the other household members due to an added activity of joint participation (Bradley and Vovsha, 2005). This specifically relates to non-mandatory activities undertaken at home or out-of-home as shown in Figure 13.5 and analysed in (Vovsha et al., 2005b). Mode choices for workers and school children in the US are closely intertwined since more than 40 percent of school children are escorted by the parents to school (Vovsha and Petersen, 2007).

Some common features of these choices can be mentioned. First, a joint choice formulation across all dimensions "explodes" and results in an infeasible number of alternatives in model estimation and application. Sampling can be applied in model estimation and application but the sampling theory and methods for calculation of correction factors are well established for the multinomial logit (MNL) model only. However, these choices are characterized by a highly differential structure of similarities between alternatives; hence a simple MNL is not always adequate. Another solution that is frequently applied in
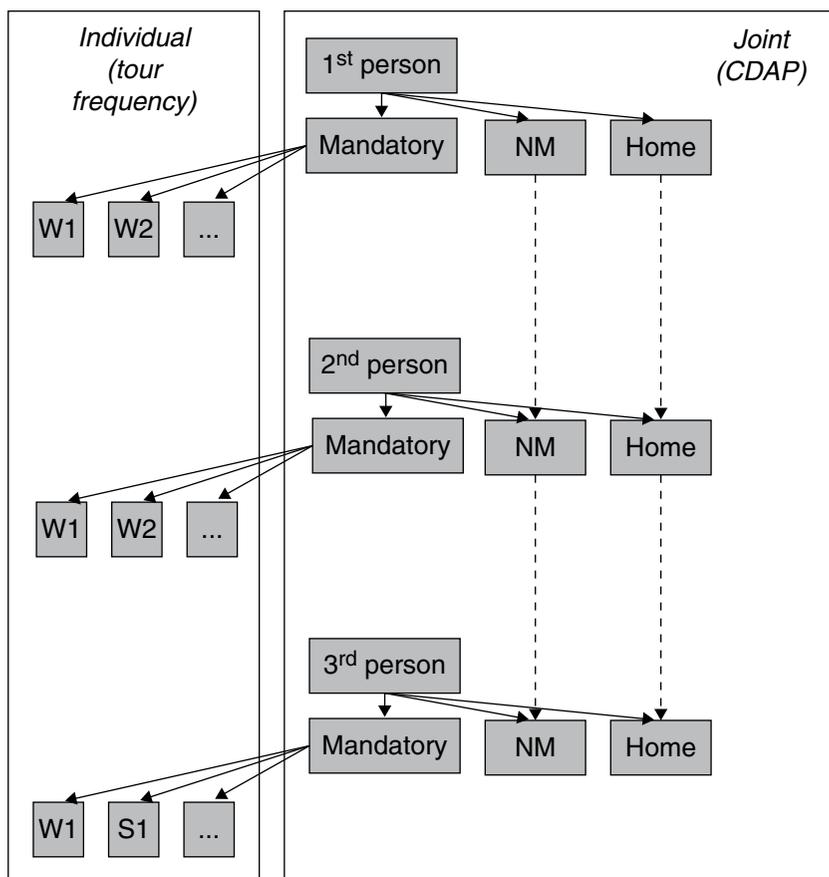
Figure 13.5   Parallel choices of daily activity pattern types by household members

operational ABMs in practice is to process these choices sequentially. However, the order of choices (or dimensions) has to be established a priori and is largely arbitrary. This is a model system design aspect that is especially difficult to resolve for closely intertwined and inter-dependent choices. In addition to that, sequential modeling of choices might result in a model system that is deficient in terms of sensitivity to policies. If there is a policy that specifically targets one of the choice dimensions (for example, improvement of school bus service that affects school mode choice) the other choice dimensions (for example, mode choice for workers) will not be affected if they are modeled prior to the

school mode choice. In reality, there is an obvious sensitivity because the necessity of school escorting affects work mode choice, but this will be lost in a sequential structure.

We will consider a method where a multidimensional choice is processed by dimensions in an iterative fashion. This structure obviates the curse of dimensionality of the joint model but preserves interrelationships between the different choice dimensions that are lost in a pure sequential decomposition.

Without a loss of generality but for the sake of clarity we will consider first an interaction between two persons (for example, household heads). We assume that both are choosing a daily pattern type from the same set $i, j \in C$ where i relates to the first person denoted as m and j relates to the second person denoted as n. An underlying joint choice specification (that might be preferable but is computationally infeasible) would have an (observed) utility function of the following general structure (Gliebe avd Koppelman, 2005; Bradley and Vovsha, 2005):

$$W_{mi,nj} = V_{mi} + V_{nj} + \Delta_{mi,nj} \tag{13.3}$$

The corresponding "true" joint choice model in a MNL form would be written as

$$P_{mn}(ij) = \frac{\exp(W_{mi,nj})}{\sum_{ij\cdots}} = \frac{\exp(V_{mi} + V_{ni} + \Delta_{mi,nj})}{\sum_{ij\cdots}} \tag{13.4}$$

If intra-household interactions are negligible the corresponding term $\Delta_{mi,ni} \approx 0$ and the joint model can be decomposed into two independent MNLs:

$$P_{mn}(ij) = \frac{\exp(V_{mi} + V_{ni})}{\sum_{ij\cdots}} = \frac{\exp(V_{mi}) \times \exp(V_{nj})}{\sum_{ij\cdots}} = P_m(i) \times P_n(j) \tag{13.5}$$

However, in the general case $\Delta_{mi,ni} \neq 0$ and such decomposition is impossible. In behavioral terms, when dealing with daily activity patterns, strong positive correlations are observed for choosing the same pattern, for example, staying at home together for the entire day $\Delta_{mi,ni} > 0$. This effect is associated with an added utility of participation in the same activity. However, in some cases, we have observed compensatory effects with $\Delta_{mi,ni} < 0$, for example for household maintenance activities like shopping, banking, or escorting children.

Parallel choices normally do not have satiation effects since the individual alternatives $i \in C$ are associated with unique daily pattern types that are not naturally ordered compared to multidimensional frequency choices. Thus, a continuous representation or any analogy with a microeconomic model is not appealing in this case. It should also be mentioned, that the matrix of interaction terms $\Delta_{mi,nj}$ normally has a diagonal form with non-zero terms corresponding to the same patterns for both persons while most off-diagonal interactions are less significant.

Since the joint choice is infeasible, some ABMs in practice relied on a sequential modeling process where some household members are modeled first and some others are made conditional upon the choices made by the previously modeled household members. We assume again an MNL structure and take into account that a full logsum calculation cannot be implemented (that would be computationally equivalent to a joint model specification). This can be formalized as the following breakdown of the joint probability into a marginal (for the first person) and conditional (for the second person) part:

$$P_{mn}(ij) = p_m(i) \times P_n(j|i) \frac{\exp(V_{mi})}{\sum_{i\cdots}} \times \frac{\exp(V_{ni} + \Delta_{mi,nj})}{\sum_{j\cdots}} \qquad (13.6)$$

It should be noted that the interaction term can only be incorporated in the conditional probability formula for the second person while the marginal probability for the first person is calculated independently of the second person choices. It also should be mentioned that despite some similarities, (Equation 13.6) is not equivalent to (Equation 13.4). In the general case, (Equation 13.4) cannot be broken into a product of marginal and conditional probabilities without using logsums in the marginal probability calculation. (Equation 13.6) is a different model that assumes that the first person makes decisions independently of the second person while the second person adjusts his/her behavior to the choices made by the first person.

An important difference between (Equation 13.4) and (Equation 13.6) is that in (Equation 13.4), a change in the utility of the second person V_{nj} or interaction terms $\Delta_{mi,nj}$ will affect the marginal choice probabilities of the first person $P_m(i)$. In (Equation 13.6), this effect is lost.

This one-way conditionality is a very strong limitation that manifests itself in certain policy contexts. A good practical example that is of primary importance in many metropolitan areas in the US is a two-way linkage between commuting mode choices made by the household heads and mode choices for school children – see Figure 13 6. Approximately 30–40 percent of school children are driven by one of the parents to school (Vovsha and Petersen, 2007). When the commuting choice is considered,
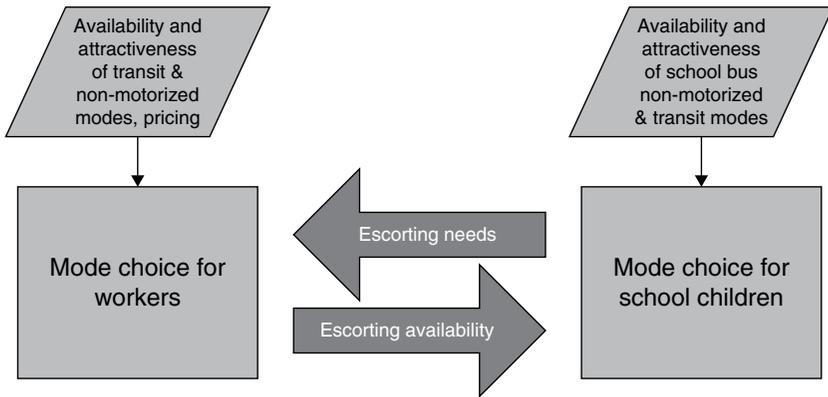
Figure 13.6   Two-way causal linkages between commuting mode choices of workers and school children

this frequently results in a strong transit-averse behavior. In modeling terms, this means that when mode choice for workers is modeled the school mode choice should be defined first as an important constraint. On the other hand, if a very convenient transit option is offered to one of the workers it might result in rearranging the escorting responsibilities within the household (the other worker could take this task). This means that the workers' mode choice should come first. In reality, these choices are closely intertwined and should have been modeled simultaneously but this would result in an infeasible choice set size. Putting these choices into any sequence would always lead to a restricted sensitivity to some policies. If the school mode choice is modeled first the model will probably be too conservative in shifting workers to transit when a good transit option is offered. If the worker mode choice is modeled first the model will not be able to recognize that such a policy as improving the school bus might make many workers more receptive to commuting by transit.

In the sequential framework (Equation 13.6), to make both choices sensitive to the same policy one has to parameterize both utility components V_{mi} and V_{nj} with largely duplicative variables (like using both work and school level-of-service variables by different modes in both work and school choices) since the logsum linkage cannot be established. This is also impractical.

The solution that we would like to discuss has been found empirically and applied in several ABMs in practice on an empirical basis (Bradley and Bowman, 2011). It can be formalized as the following iterative procedure shown in Figure 13.7.
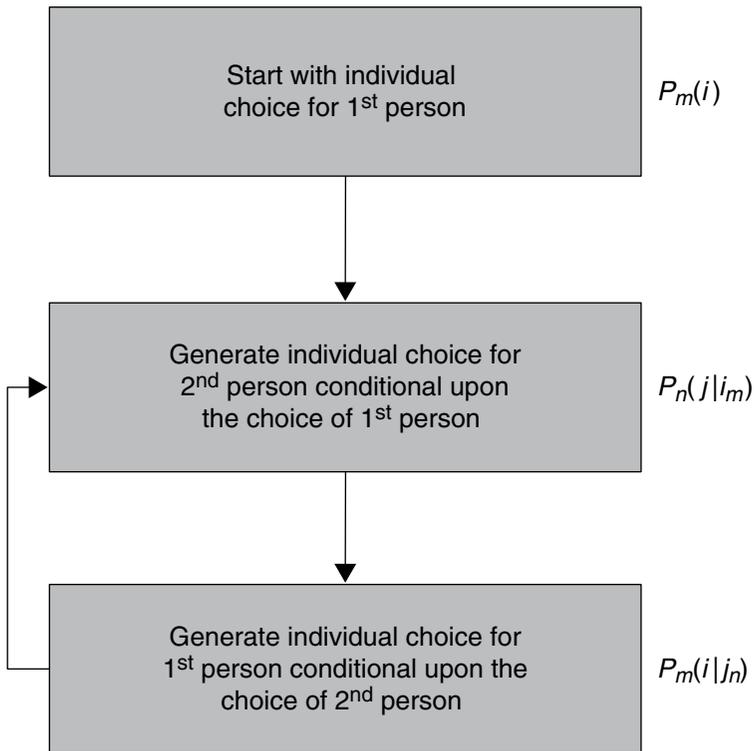
Figure 13.7   Parallel convoluted choices: naïve iterative microsimulation

Step 1: Arbitrarily start with the first person and apply a simplified marginal choice model without interactions:

$$P_m(i) = \frac{\exp(V_{mi})}{\sum_{i \cdots}} \qquad (13.7)$$

Step 2: Apply a conditional model for the second person with intra-household interactions:

$$P_n(j|i_m) = \frac{\exp(V_{nj} + \Delta_{mi,nj})}{\sum_{j \cdots}} \qquad (13.8)$$

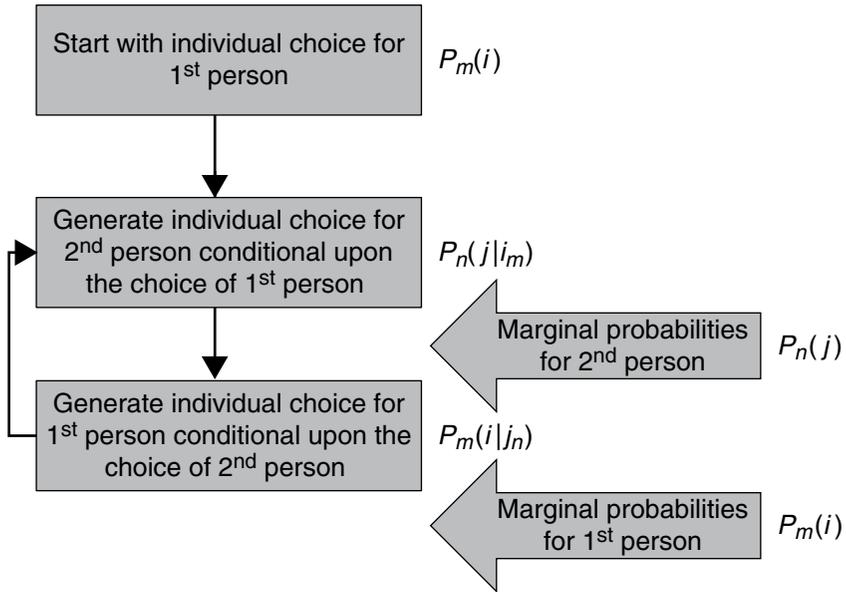Step 3: Apply a conditional model for the first person with intra-household interactions:

Figure 13.8   Parallel convoluted choices: fractional probability version

$$P_m(i|j_n) = \frac{\exp(V_{mi} + \Delta_{mi,nj})}{\sum_{i \cdots}} \qquad (13.9)$$

Step 4: Go to step 2 until convergence in terms of joint probabilities is observed.

This approach is convenient to implement in an individual micro-simulation setting when the fractional choice probabilities after steps 1–3 are converted into a "crisp" choice for each individual. The next choice probabilities are conditional upon this single alternative chosen for the previously modeled individual. Otherwise, too many fractions have to be stored after each step for each individual. In ABMs in practice, a full synthetic population (millions) is modeled explicitly without weighting, thus the random Monte Carlo variability effects are insignificant.

However, if computational efficiency is not an issue and the theoretical properties of the original fractional-probability model are to be explored there is a fractional-probability analogue that does not involve the Monte Carlo step as shown in Figure 13.8.

Step 1: Arbitrarily start with the first person and apply a simplified marginal choice model without interactions:

$$P_m(i) = \frac{\exp(V_{mi})}{\sum_{i \cdots}} \qquad (13.10)$$

Step 2.1: Apply a conditional model for the second person with intra-household interactions for each share with respect to the choice of the first person:

$$P_n(j|i_m) = \frac{\exp(V_{nj} + \Delta_{mi,nj})}{\sum_{j \cdots}} \qquad (13.11)$$

Step 2.2: Calculate marginal shares with respect to the choice of the second person:

$$P_n(j) = \sum_i P_n(j|i_m) \times P_m(i) \qquad (13.12)$$

Step 3.1: Apply a conditional model for the first person with intra-household interactions:

$$P_m(i|j_n) = \frac{\exp(V_{mi} + \Delta_{mi,nj})}{\sum_{i \cdots}} \qquad (13.13)$$

Step 3.2: Calculate marginal shares with respect to the choice of the first person:

$$P_m(i) = \sum_j P_m(i|j_n) \times P_n(j) \qquad (13.14)$$

Step 4: Go to step 2.1 until convergence in terms of marginal and conditional probabilities is achieved.

The fractional-probability methods can be applied for each individual or segment of identical individuals. The microsimulation approach is always applied for each individual and it is essential to obtain variation in choices across identical individuals despite the fact that they share exactly the same probabilities.

The approach was originally applied in some choice contexts without iterations between the choice dimensions and without analysis of the convergence (Bradley and Bowman, 2011). It should be noted that an ABM is applied iteratively with the network simulation model to achieve global

demand-network equilibrium. Thus, it is deemed that achieving convergence between the dimensions of the joint choice within each global iteration is not essential. However, it is an important research question whether this type of convoluted choice model would produce a convergent solution equal to the underlying joint model (Equation 13.4).

Multiple simulations with this process implemented by the author have shown that it converges to the joint model solution in quite an efficient way although the number of iterations required is a strong function of the number of alternatives, distribution of probabilities and the number of interaction terms. Since the entire process is implemented in a microsimulation fashion, convergence is observed in terms of aggregate shares of joint alternatives over a large synthetic population. However, this is an indication that a theoretical convergence for a fractional-probability model could be proven. To the best of our knowledge this approach has never been explored before and the proof of convergence for convoluted choice models has never been produced. It is interesting if the convergence can be proven at least for simple recursive formulations of this type when conditional (we can call them transition probabilities) are derived from the joint model:

$$P_n(j|i_m) = \left[\frac{P_{mn}(ij)}{\sum_j P_{mn}(ij)} \right] \qquad (13.15)$$

$$P_m(i|j_n) = \left[\frac{P_{mn}(ij)}{\sum_i P_{mn}(ij)} \right] \qquad (13.16)$$

The procedure described above relates to the following convex mathematical program with linear constraints formulated with respect to joint probability $x_{mn}(ij)$:

$$\min_{\{x_{mn}(ij)\}} \sum_{ij} x_{mn}(ij) \times \ln\left[\frac{x_{mn}(ij)}{P_m(i)} \right] \qquad (13.17)$$

Subject to:

$$x_{mn}(ij) = P_n(j|i_m) \times \left[\sum_j x_{mn}(ij) \right] = \left[\frac{P_{mn}(ij)}{\sum_j P_{mn}(ij)} \right] \times \left[\sum_j x_{mn}(ij) \right], \alpha_{ij} \qquad (13.18)$$

$$x_{mn}(ij) = P_m(i|j_n) \times \left[\sum_i x_{mn}(ij) \right] = \left[\frac{P_{mn}(ij)}{\sum_i P_{mn}(ij)} \right] \times \left[\sum_i x_{mn}(ij) \right], \beta_{ij} \qquad (13.19)$$

$$\sum_{ij} x_{mn}(ij) = 1 \qquad (13.20)$$

It can be shown that there is a unique solution to this problem that can be written by equating the Lagrangian derivatives to zero. However this solution can be found immediately from (Equation 13.18–Equation 13.19). These equations can be rewritten in the following general way:

$$x_{mn}(ij) = P_{mn}(ij) \times \gamma_i \qquad (13.21)$$

$$x_{mn}(ij) = P_{mn}(ij) \times \lambda_j \qquad (13.22)$$

From which it can be seen that all balancing factors (γ_i, λ_j) have to be equal. Combining this condition with (Equation 13.20) one can obtain that all balancing factors (γ_i, λ_j) have to be equal to one which means that the result of this balancing will be equal to the underlying joint probabilities.

The fact that an iterative application of convoluted choice models (Equation 13.8–Equation 13.9) yields the same results as the joint model (Equation 13.14) opens a way to resolve many issues associated with ABMs. First, the partial models for each dimension are much simpler than the joint model while the latter in many cases is infeasible. The price that is paid for the model decomposition is that multiple iterations are needed but an ABM has many other reasons for iterative application and some of them can be effectively combined. This process in many instances resolves the numerous chicken-and-egg problems associated with putting sub-models in a predetermined sequence. In behavioral terms, it ensures a realistic two-way causality in model responses to different policies without overloading each particular utility with multiple intricate variables that relate to the other choice dimensions.

In real applications, of course, the "true" joint model is unknown since it is infeasible to estimate across all dimensions simultaneously. Thus, only partial models (Equation 13.8–Equation 13.9) can be estimated and in practice it does not guarantee that the interaction terms $\Delta_{mi,nj}$ will come out identical since the estimation for partial models is normally implemented independently. Although it is possible to estimate both partial models simultaneously (not a joint model) and enforcing the interaction term to be the same, it might actually be behaviorally more appealing to allow for different interaction to happen depending on the order of choices and underlying causality.

It is still unexplored if this process would converge with any interaction terms and to what solution.

The technique of convoluted choices has been recently extended beyond

the parallel choice framework. One of the promising directions is to use this technique to resolve the long-term debate of ordering mode and destination choices, which has been exacerbated by considering time-of-day choice as an additional important dimension that led to six possible nesting (ordering) schemes (Bradley and Bowman, 2011). In a typical applied ABM at the tour-level we have about 20 modes and mode combinations ($m = 1,2 \ldots 20$), 840 departure-arrival time alternatives ($t = 1,2 \ldots 840$) and 4000 primary destination locations ($d = 1,2, \ldots 4000$). Treatment of all joint alternatives in a simultaneous framework in infeasible and putting these choices in any arbitrary order a priori is questionable. In practice, there can always be realistic situations for each of the six possible causality schemes – see Table 13.1.

Table 13.1   Possible order of mode, destination and time-of-day choice models

Order of choices	ABM where applied	Example of underlying behavioral mechanism
Dest→TOD→Mode	Columbus, Atlanta, Bay Area, San Diego, Chicago	Work with fixed schedule, business meeting, doctor appointment, theater, escorting children
Dest→Mode→TOD	New York Sacramento, Denver, 4-step models in US	Major shopping in a predetermined (usual) store for a non-working person
TOD→Dest→Mode	San-Francisco, Portland, 4-step models in Europe, activity scheduling models in academia	Discretionary activity after work by worker with fixed schedule
TOD→Mode→Dest	4-step models in Europe	Grocery shopping by a worker with a fixed schedule in a household with number of adults greater than number of cars
Mode→Dest→TOD	4-step models in Europe	Grocery shopping by a non-worker in a household with number of adults greater than number of cars
Mode→TOD→Dest		Discretionary trip by a non-working person with flexible schedule in a household where number of workers is greater than number of cars and the choice of a car passenger in the worker carpool is considered

5. MULTI-DIMENSIONAL FREQUENCY CHOICE AND ANALOGY WITH CONTINUOUS MODELS

This is a typical choice context in ABMs where simultaneous choice of frequencies $\{f_a = 0,1,2,3, \ldots, F_a\}$ of different activities by type $a = 1,2,3, \ldots, A$ is considered. These types of choices are currently handled by a chain of choice models and/or by aggregation of activities where the integrity of the entire choice context is largely lost. Simultaneous modeling is essential because of time-space and budget constraints that tie the choices made for different activities together. However, it may result in an infeasible dimensionality. For example, if we constrain $F_a = 3$ and consider $A = 10$ (which are quite realistic dimensions) we obtain $4^{10} = 1,048,576$ alternatives.

It should be mentioned that frequency in each dimension represents an ordered choice with satiation effects. It is currently handled by a component-wise utility form where each activity has a base utility component that is repeated f_a times plus an additional satiation penalty for each additional activity of the same type. Taking into account the satiation effects and also general time-space and budget constraints it is intuitive to seek for a different approach to handle such a choice situation rather than enumerating each and every combination explicitly.

There is a great deal of analogy of the model for ordered activity frequency with a model where the corresponding amount of activity would be measured as a continuous variable (it is reasonable to assume that the number of activity episodes is correlated with a continuous variable like total activity duration although the correlation is not perfect). The corresponding models that consider a continuous amount of activity include microeconomic models (Jara-Diaz, 2007), structural equations models (Golob and Recker, 2003) and multiple discrete-continuous choice models (Bhat, 2008). These models are based on different techniques but share some common features that are relevant to the activity frequency choice. First, these models can address interdependency of frequencies across dimensions and include a rich set of explanatory variables, hence ensuring a reasonable policy response. Secondly, these models do not "explode" the same way as a conventional choice model since the number of alternatives is not multiplied by dimensions. Rather, they are linearly sensitive to the number of activities and a number of activities like $A = 10$ is not prohibitive at all. Thirdly, however, instead of the discrete number of activity episodes (like 0, 1, 2 or 3 shopping trips with corresponding probabilities) these models return (average) continuous values for frequencies (like 1.6 shopping trips) or durations (like 2.3 hours for shopping).

A microsimulation ABM is based on discrete choices, not fractional probabilities or average values. Thus, within this framework it is essential

to generate a probability distribution of frequencies $P(0), P(1), P(2)\ldots$ rather than an average number like 1.6. Avoidance of aggregation biases and realistic representation of variation across situational variables (i.e. other choice outcomes) are fundamental advantages of an ABM that are based on the realistic representation of the variation of behavior across microsimulated individuals. This cannot be achieved if probabilistic choice models are replaced with deterministic continuous models.

This brings us to a discussion on possible approaches to probabilistic discretizing of continuous outcomes of these models. Two methods can be considered. First, an independent choice model for each dimension conditional upon the (average) continuous variable can be applied. Since this model is applied separately to each dimension it results in some loss of integrity at the final discretizing stage. Secondly, an entropy-maximizing randomization of continuous values can be applied. We first formulate this model for one-dimensional and two-dimensional choices that can be further generalized to handle a multidimensional case.

The probabilistic discretizing method has to address two aspects that are very different from any form of bucket rounding or deterministic discretizing:

- Generate feasible discrete solutions in an efficient way avoiding complete enumeration and evaluation of all possible combinatorial alternatives and their probabilities.
- Assign probability to each solution (and estimate it in a consistent way).

Consider first a simple case of discretizing a single continuous variable for which an average value $\widetilde{f}$ and some estimate for variance σ^2 are known. The average value can be generated by a continuous model (microeconomic, structural equations, or a simple regression). The variance is supposed to be known a priori from the data used to estimate the model. There can be an additional auxiliary model estimated that parameterizes variance as a function of mean and some other variables. Then the probabilistic discretizing problem can be formulated as the following entropy-maximizing program with respect to probabilities associated with discrete values:

$$\max_{\{p(f)\}} \sum_f P(f) \times \ln[P(f)] \qquad (13.23)$$

Subject to:

$$\sum_{f=1}^{F} f \times P(f) = \widetilde{f}, (\lambda) \qquad (13.24)$$

$$\sum_{f=0}^{F} P(f - \tilde{f})^2 \times p(f) = \sigma^2, (\theta) \qquad (13.25)$$

$$\sum_{f=0}^{F} P(f) = 1, (k) \qquad (13.26)$$

$$P(f) \geq 0 \qquad (13.27)$$

The (unique) solution for this convex problem can be easily found by forming a Lagrangian and equating partial derivatives to zero:

$$P(f) = \exp(k) \times \exp(\lambda f) \times \exp[\theta(f - \tilde{f})^2] \qquad (13.28)$$

The dual variables λ, θ and k can be calculated by iterative fitting or by a more efficient optimization algorithm.

If more than the first and second moments of the distribution of f are known and need to be controlled they can be introduced in the same way without significant complication of the program. This program can be easily generalized for a multi-dimensional case, for example a two-dimensional case presented below:

$$\max_{\{p(fg)\}} \sum_{fg} P(fg) \times \ln[P(fg)] \qquad (13.29)$$

Subject to:

$$\sum_{fg} f \times P(fg) = \tilde{f}; \sum_{fg} g \times P(fg) = \tilde{g} \qquad (13.30)$$

$$\sum_{fg} (f - \tilde{f})^2 \times P(fg) = \sigma_f^2; \sum_{fg} (g - \tilde{g})^2 \times P(fg) = \sigma_g^2 \qquad (13.31)$$

$$\sum_{fg} (g - \tilde{g})^2 \times (f - \tilde{f}) \times P(fg) = \text{cov}_{fg} \qquad (13.32)$$

$$\sum_{fg} P(fg) = 1 \qquad (13.33)$$

$$P(fg) \geq 0 \qquad (13.34)$$

It is important to recognize the advantages and disadvantages of a discrete choice model compared to a continuous model in the context of ordered frequency choice. The difference is also largely a function of the discrete resolution that is especially appealing in the context of temporal choices. Trip departure time choice with a one-minute temporal resolution becomes essentially continuous while for a trip frequency choice that

operates with large discrete units 0,1,2,3, . . ., 10 continuous representation is too crude. However, the major aspect is not the level of resolution that brings the discrete and continuous space closer to (or father away from) each other. The biggest difference is in assigning probabilities to each frequency, i.e. generating a frequency distribution by a choice model. Continuous models in general cannot do that.

Choice models have the advantage of forming a specific utility expression for each alternative while a continuous model operates with a single expression. A good example of these specifics that might be overlooked by a simple discretizing model based on the first two moments is frequency of escorting trips characterized by a greater probability of even frequencies because of the natural pairing of the trips in most cases (outbound and inbound). Specifics like this can be taken into account by considering higher moments or introducing a priori weights $\widetilde{P}(f)$ in the entropy-maximizing objective function.

This way any complicated discrete choice model with a large number of multi-dimensional frequency alternatives can be replaced by two relatively simple continuous models. The first one would predict average frequency for each dimension in an integrated way. The second one would probabilistically discretize the outcome of the first model. In first experiments of the author, the results look promising, especially in terms of computational efficiency; however, more research is needed in this direction.

7.   CONCLUSIONS

The four-step era created a huge gap between the theory and practice of choice modeling. While tremendous progress has been made in the theory of choice modeling including closed-form logit models (Daly and Bierlaire, 2006; Fosgerau et al., 2010) and mixed-logit models (Hess et al., 2009) the absolute majority of choice models applied in practice still include MNL and NL only. In essence, the four-step framework included only one example of a well-explored choice dimension – mode choice – while all other dimensions such as trip/activity frequency choice, destination choice and time-of-day choice were primarily treated using simplified aggregate techniques.

The ABM framework proved to be a tremendous source and very good platform for experimenting with new choice structures. However, these choices are frequently characterized by an infeasible number of alternatives and a complicated structure of similarities between them. Treatment of these choices represents a challenge where standard methods of decomposition of choices into a sequence result in a disintegrated and behavio-

rally non-appealing structure. The chapter describes examples of parallel convoluted choices and activity frequency choices with some suggestions on new methods of modeling these choices and overcoming the curse of dimensionality. In this regard the analogy with microeconomic models of consumer behavior and multiple discrete-continuous models looks as a promising direction.

Another related aspect is a need for more flexible nested structures where not only the utilities but the nesting and inclusion parameters could be parameterized. The preliminary meta-analysis of different mode choice situations indicated several important ways and dimensions for such a parameterization like trip lengths or personal income. The corresponding models were favorably accepted by practitioners but the parameters have been set by naïve trial-and-error methods since a theoretically consistent estimation of these models cannot yet be implemented in practice.

Progress in any of these directions would be very welcome in practice.

REFERENCES

Arentze, T.A. and H.J.P. Timmermans (2004), 'Albatross – A Learning-Based Transportation Oriented Simulation System', Urban Planning Group/EIRASS, Eindhoven University of Technology.

Bhat, C.R. (2008), 'The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions', Transportation Research B, 42B(3), 274–303.

Bowman, J.L. (2009), 'Historical Development of Activity Based Model Theory and Practice', Traffic Engineering & Control, 50(2), 59–62.

Bradley, M.A and J.L. Bowman (2011), 'Improving the Modeling of Time-of-Day Effects in the PSRC Activity-Based Model: Joint Mode & Time-of-Day Models and Time-Sensitive Logsums', presented at the 13th Transportation Planning Application Conference, TRB, Reno, NV.

Bradley, M. and Vovsha P. (2005), 'A Model for Joint Choice of Daily Activity Pattern Types of Household Members', Transportation, 32(5), 545–571.

Bradley, M., J.L. Bowman and B. Griesenbeck (2009), 'Activity-Based Model for a Medium-Sized City: Sacramento', Traffic Engineering & Control, 50(2), 73–79.

Daly, A. and M. Bierlaire (2006), 'A General and Operational Representation of Generalized Extreme Value Models', Transportation Research Part B, 40(4), 285–305.

Davidson, W., R. Donnelly, P. Vovsha, J. Freedman, S. Ruegg, J. Hicks, J. Castiglione and R. Picado, (2007), 'Synthesis of First Practices and Operational Research Approaches in Activity-Based Travel Demand Modeling', Transportation Research A, 41A (5), 464–488.

Eluru, N., C.R. Bhat, A.R. Pinjari and R.M. Pendyala (2010), 'A Unified Model System of Activity Type Choice, Activity Duration, Activity Timing, Mode Choice, and Location Choice', presented at the 3rd Conference on Innovations in Travel Modeling, TRB, Tempe, AZ.

Fosgerau, M., D. McFadden and M. Bierlaire (2010), 'Choice Probability Generating Functions', Report TRANSP-OR 100605, EPFL.

Gliebe, J. and F. Koppelman (2005), 'Modeling Household Activity-Travel Interactions as Parallel Constrained Choices', Transportation, 32 (5), 449–471.

Golob, T.F. and W.W. Recker (2003), 'Structural equation modeling for travel behavior research', Transportation Research, B - Methodological, 37, 1–25.

Hess, S., M. Ben-Akiva, D. Gopinath and J. Walker (2009), 'Advantages of latent class choice models over continuous mixed logit models', presented at the 12th International Conference on travel Behavior Research, Jaipur.

Jara-Diaz, S.R. (2007), Transport Economic Theory, Elsevier, Holland.

Konduri, K., S. Astroza, B. Sana, R. Pendyala and S.R. Jara-Diaz (2011), 'Joint analysis of time use and consumer expenditure data: examination of two alternative approaches to deriving values of time', presented at the 90th Annual Conference of TRB.

Pinjari, A.R. and C.R. Bhat (2010), 'An Innovative Forecasting Procedure for the MDCEV Model', presented at the 3rd Innovations in Travel Modeling (ITM) Conference, Tempe, AZ.

Vovsha P. and S. Bekhor (1998), 'The Link-Nested Logit Model of Route Choice: Overcoming the Route Overlapping Problem', Transportation Research Record 1645, Washington, D.C., 133–142.

Vovsha, P. and E. Petersen (2005), 'Escorting Children to School: Statistical Analysis and Applied Modeling Approach', Transportation Research Record 1921, Washington, D.C., 131–140.

Vovsha, P. and E. Petersen (2007), 'Incorporating Car-Type Preferences and Intra-Household Car Allocation into Travel Demand Models', presented at the 86th Annual TRB Meeting, Washington, D.C.

Vovsha, P., M. Bradley and J. Bowman (2005a), 'Activity-Based Travel Forecasting Models in the United States: Progress since 1995 and Prospects for the Future'. In Progress in Activity-Based Analysis, H. Timmermans (ed.), Elsevier Science Ltd, Oxford, UK, 389–414.

Vovsha, P., J. Gliebe, E. Petersen and F. Koppelman (2006), 'Comparative Analysis of Sequential and Simultaneous Choice Structures for Modeling Intra-Household Interactions'. In Progress in Activity-Based Analysis, H. Timmermans (ed.), Elsevier Science Ltd, Oxford, UK, 223–258.

Vovsha, P., J. Freedman, V. Livshits and W. Sun (2011), 'Design Features of Activity-Based Models in Practice: Coordinated Travel-Regional Activity Modeling Platform', Transportation Research Record 2254, Washington, D.C. 19–27.