Network Equilibrium with Activity-Based Microsimulation Models: The New York Experience

Peter Vovsha, Robert Donnelly, Surabhi Gupta (PB)
1 Penn Plaza, 2nd Floor, New York, NY 10001, 212-465-5511, vovsha@pbworld.com

Definition and Theoretical Issues of Equilibrium

Travel demand models and network simulation models are distinct sets of procedures that are combined within in the framework of regional transportation modeling systems. The accuracy of the travel forecasts produced ultimately depends on the quality of both components, as well as the manner in which the travel demand and network simulation are combined in the context of global equilibrium. Conventional 4-step models have numerous limitations compared to more advanced activity-based (AB) micro-simulation (MSM) models, primarily with respect to the internal consistency and detailed behavioral realism. However, one of the remaining advantages of 4-step models is an established theory, as well as an effective set of practical rules for achieving global network equilibrium so that travel time and cost simulated in the networks exactly correspond to the demand (trip tables) generated by the model. The existence and uniqueness of the fixed-point equilibrium solution makes the model system outcome objectively well-defined and independent of the starting conditions.

The reason it is possible to formulate network equilibrium with a 4-step model stems from two fundamental features:

- Simple demand functions incorporating trip distribution, mode choice, and time-of-day choice in a trip matrix format, and in a logit-based form that can be derived from convex entropy-type demand terms.
- Simple deterministic user equilibrium assignment methods that can written as an optimization problem with link-level convex Beckman-type terms.

The entropy-type matrix terms and the Beckman-type link terms can be conveniently combined in one program that serves as the basis for equilibrium formulations. Evans demonstrated this technique first for a combined trip distribution and assignment model [*1*] and it was subsequently extended to other travel dimensions including mode and location choices [*2-4*].

On the other hand, this issue remains less explored and somewhat obscure for AB models. These models have a more complicated analytical structure compared to 4-step models that makes it difficult to derive equilibrium conditions in a rigorous theoretical way. Additionally, implementation of an AB model requires MSM of individual outcomes in a form of "crisp" discrete choices that is very different from the summation of fractional probabilities implemented in conventional models. The application of MSM in AB models involves methods such as Monte-Carlo which creates certain non-continuous variability of outcomes. Algorithms for network equilibrium with aggregate 4-step models are based on averaging trip tables across several iterations; that is a consequence of the continuous nature of the demand model outcome in a fixed structure. The outcomes of AB MSM model application structurally changes from iteration to iteration (same person can implement different sets of tours and trips) and consequently it is difficult to establish any meaningful averaging rule to apply.

This paper documents the results of testing various equilibrium strategies implemented with the New York AB MSM model ("NYBPM"). The purpose of the paper is two-fold:

1. To outline more fundamental research directions and extensions of the network equilibrium theory to cover AB MSM models in a more rigorous way.
2. To present realistic levels of convergence that can be achieved with AB MSM models in practice and with establish practical rules and protocols for using this type of models for different projects and policies.

Practical Ways to Reach Equilibrium with an Activity-Based Model

There are several theoretical issues associated with the achievement of network equilibrium with AB MSM models:

- Running the same model twice with exactly the same set of inputs and the same starting random seed would reproduce the results exactly; the reason for practical non-convergence is that small variations in Level-of-Service (LOS) variables can produce significant local disturbance in the chain of demand-related choices. A naïve way of implementing multiple iterations with feeding back LOS variables does not work with AB models, just as it is generally not effective even with 4-step models.
- There are two types of Monte-Carlo effects which have significant and distinctive impacts on global convergence – one relates to MSM with a fixed structure of modeled agents and choices, and the other relates to MSM where the structure of modeled agents and choices is dependent on the results of MSM of prior choices in the model chain.
- MSM of choices with a fixed structure is the simpler case where convergence can be achieved by simply iterating since averaging the LOS variables would ultimately tend to reproduce not only the aggregate shares but even individual choices.
- MSM of choice model chains with structural impacts can have discontinuity and abrupt responses to even small variations in inputs; in this case it is difficult to predict the level of convergence and to analytically estimate the Monte-Carlo component of variation.

There are two main practical ways to ensure convergence (assuming that a fixed-point equilibrium solution exists):

- **Enforcement**. These methods are specific to MSM and designed to ensure convergence of "crisp" individual choices by suppressing or avoiding Monte-Carlo variability. These methods are currently only at a beginning stage of theoretical foundation, with some empirical strategies being tested.
- **Averaging**. These methods have been borrowed from conventional 4-step modeling techniques, but can be also used with MSM as far as they are applied to continuous outputs/inputs like LOS variables and/or synthetic trip tables generated by the MSM.

Enforcement Methods

Several ways to "enforce" convergence at the individual level have been suggested and tested in practice. From the theoretical perspective they can be broken into three groups:

- Re-using the same random numbers or starting random seeds for certain choices that would ensure that the choice will be replicated if no change occurs to the inputs,
- Gradual freezing of portions of households or travel dimensions from iteration to iteration,
- Analytical discretizing of probability matrices instead of Monte-Carlo simulation.

Re-Using the Same Random Numbers / Seeds

One of the reasons for instability of the demand MSM that manifests itself even if the LOS variables converge, is that the MSM generates “crisp” choices from and on the top of the probabilities generated by choice models. This is usually done by generating a random number for each choice and relating this number to the choice probabilities (analogous to Monte-Carlo roulette). Even if the probabilities become constant, the Monte-Carlo variability alone causes a random fluctuation of the individual MSM results, although the aggregate results are quite stable at a certain reasonable level of aggregation.

To avoid the Monte-Carlo variability, random numbers associated with each choice can be generated in advance, stored and re-used when the model is applied for different scenarios and across different iterations. It is enough to store a seed that would automatically generate the same sequence of random numbers. Application of this strategy is, however requires a structural stability of the agents and their decision chains embedded in the model structure. In comparatively simple model structures, the list of simulated agents (households, persons, tours, and trips) and their choice alternatives are fixed from iteration to iteration and only associated choices probabilities fluctuate with changes in network times and costs. In this structure, it can be shown that convergent probabilities (as function of convergent LOS variables) would ensure convergence of the individual choices. The same random number will be always applied to the same agent and choice dimension. Examples of structurally stable MSM models include tour mode & destination choice with a fixed set of generated tours.

However, this strategy becomes problematic for more complex decision chains where structural impacts of prior choices are on the subsequent choices in the model chain. For example, the daily activity pattern model that generate tours (i.e. create a list of tours by type for every person) is in itself a simulation model. At each iteration, it may generate in a different set of tours for the same person. The subsequent mode & destination model would then be applied for a different set of agents; thus, freezing a seed for each person will not help in mode and destination choice. It would generate the same sequence of random numbers for each person but these random numbers would be applied to different agents and associated choices.

Theoretically, structurally stable decision chains can be ensured by considering the maximum possible number of agents created at each stage of MSM and reserving a random seed for each of them. For example, if the maximum modeled number of work tours per day generated by a worker is two, we could create a placeholder for the 1st and 2nd tour random seed, and just not use it if the tour was not actually created. This, however, might create quite a huge system of placeholders in advanced AB models since they include numerous structural components, for example, such as multiple stops on each half-tour.

Re-using random seeds has been applied for the SFCTA model as well as for the special restricted version of the MORPC model developed for the FTA New Starts projects.

Gradual Freezing of Portions of Households or Travel Dimensions

This is a set of empirical procedures that is based on a predetermined strategy of progressively freezing certain portions of the simulated agents over the course of global iterations (i.e. fixing the corresponding choice outcomes until the end of model application). In general, decisions are required for the following components of this strategy:

- Principal sequence of agents to freeze that can be, for example:
  - Subsets of households with all related choices.
  - Certain travel dimensions with all households considered (like tour generation, destination choice, time-of-day choice, mode choice, etc).
- Steps in progressing through global iterations; for example, one can envision 6 global iterations with freezing additional 20% of households at each iteration, starting from the third iteration. More, specifically, this would result in the following strategy:
  - 100% of households simulated in the 1st iteration
  - 100% of households in the 2nd iteration (i.e. all households are re-simulated)
  - 80% of households in the 3rd iteration (20% of households are frozen)
  - 60% of households in the 4th iteration (another 20% of households are frozen)
  - 40% of households in the 5th iteration (another 20% of households are frozen)
  - 20% of households in the 6th iteration (another 20% of households are frozen)
- Principles for choice of the frozen and re-simulated households:
  - Purely random (with some possible geographic stratification).
  - Based on some criterion that reflects on "unstable travel conditions"; for examples households/persons/tours/trips that are characterized by a high level of congestion would be better to re-simulate multiple times.

Gradual freezing is always effective. It does not mean, however, that a true fixed-point solution is achieved. It relies on the reasonability of the technical strategy that can be established only after multiple trials. In some cases, such as model application for FTA New Starts projects, certain travel dimensions can be fixed across all compared scenarios that simplify the choice of strategy [*5*].

Analytical Discretizing of Probability Matrices

Analytical discretizing represents a method for converting fractional-probability outputs of choice models into "crisp" choices as an alternative to the Monte-Carlo technique.

Analytical discretizing has two major advantages over the Monte-Carlo technique:

- **Full replication** of the model outcome with fixed inputs, i.e. if we run the discretizing procedure several times with over the same choice model with fixed input variables, the results will be identical while the Monte-Carlo technique is characterized by inherent variability (so-called "Monte-Carlo error") of the results.
- **Logical elasticity** of the aggregate model outcome with respect to the input variables that is identical to the elasticity of the parent choice model. The Monte-Carlo technique does not guarantee logical elasticity and fixing seeds for random number does not help in this respect. The expected responses of the parent choice model can be "eaten" by the Monte-Carlo error that will make the model response illogical.

It is interesting to note, that in real terms, discretizing in application is the opposite of choice model estimation.. In the estimation procedure, discrete outcomes are given in the form of observed choices and the fractional probabilities are generated in order to replicate the observed choices as closely as possible. In the applied discretizing procedure, the fractional probabilities are given by the core choice model and "crisp" choices are generated in order to replicate the modeled probabilities as closely as possible. Thus, discretizing can be thought of as restoring the observations that would be most plausible for the given choice probabilities.

We introduce the following notation:

$n \in N$	=	observations in the model estimation, realizations in the model application,
$i \in C_n \in C$	=	choice alternatives available for each observation / realization taken from universal set of alternatives,
$\delta_{in} = (0,1)$	=	Boolean indicator on the observed / modeled choice of alternative for each observation / realization.
$P_n(i)$	=	modeled choice probability for each alternative and observation / realization.

It is assumed that in both estimation and application of the model the "crisp" choices and fractional probabilities are subject to the logical constraints:

$$\sum_{i \in C_n} \delta_{in} = 1, \quad \sum_{i \in C_n} P_n(i) = 1 \text{ for all observations / realizations.} \tag{1}$$

The choice model **estimation** is done by maximizing the (log) likelihood function over choice probabilities (parameters of the choice model) while the observed choices are given:

$$\max_{\{P_n(i)\}} L = \sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln P_i(i). \tag{2}$$

The discretizing procedure in the model **application** is done by maximizing the entropy function over the "crisp" choices while the fractional probabilities are given by the core choice model:

$$\max_{\{\delta_{in}\}} E = -\sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln \frac{\delta_{in}}{P_n(i)} = \sum_{n \in N} \sum_{i \in C_n} \delta_{in} \ln P_n(i). \tag{3}$$

The discretizing approach considers the whole matrix of probabilistic outcomes of the core choice model and tries to find the structurally closest matrix of discrete numbers that also fits to the marginal totals of original matrix. Rows of the matrix correspond to observations / realizations. Columns of the matrix correspond to the choice alternatives. The marginal totals are readily interpreted. Row totals are all equal to 1 by the condition (1). The column totals correspond to the aggregate shares of alternatives. For a logit choice model with a full set of alternative-specific constants, the aggregate shares predicted by the model are equal to the observed shares for all alternatives:

$$\sum_{n \in N} P_n(i) = \sum_{n \in N} \delta_{in} = A_i . \tag{4}$$

It can be seen that likelihood optimization (2) and discretizing (3) refer to the same objective function under the same set of constraints (1, 4) but the maximum is achieved with respect to the different subsets of variables. The optimization problem associated with the model estimation (maximize (2) given (1)) in the MNL case results in the convex problem that can be solved by the steepest descending method. The constraint (1) is guaranteed by the form of the choice model and constrained (4) is guaranteed by the alternative-specific constants. The optimization problem associated with the discretizing (maximize (3) given (1) and (4)) results in a linear programming (LP) problem of the so-called transportation type. An important property of the transportation problem is that discrete marginal totals guarantee a discrete solution.

Several properties of the discretizing procedure should be mentioned:

- Without constraints (4), maximization of (3) would result in the trivial choice of the alternative with maximum probability for each observation to be converted to the "crisp" choice.
- When the discretizing procedure is applied for the choice model outcomes for the same set of observations that was used in the model estimation, the value of the objective function can only be improved versus the likelihood achieved in the model estimation. The observed "crisp" choices form one of the possible solutions in the feasible region of the LP problem associated with discretizing.
- The better the core model is in terms of likelihood function (i.e. the closer is the modeled probabilities to the observed choices) then the closer the discretizing outcome will be to the observed choices. Indeed, the (log) likelihood function has a theoretical maximum value of zero that corresponds to an ideal model with probabilities equal to the observed choice indicators. The closer the estimated model is to this ideal, the less room left for the further improvement of the likelihood in the discretizing procedure.
- The discretizing procedure guarantees unbiasedness of the solution in aggregate sense by virtue of the constraint (4).

Averaging Methods

Averaging is a universal tool that can be applied for continuous outcomes of any iterative process. If a fixed-point solution exists, an averaging strategy like Moving Successive Averages (MSA) will always find it, although it might require multiple iterations [*3*]. If the equilibrium can be formulated as an optimization problem in view of the assumed simplicity of the demand functions, much more effective analytical procedures than MSA can be applied [*4*]. However, if the demand model is too complicated to be written as an explicit optimization problem, MSA represents the only viable option to ensure a fixed-point solution. MSA has many possible technical variations. Any sequence of numbers S_k would suffice if it satisfies two basic conditions:

$$\lim_{k\to\infty} S_k = 0 \text{ and } \lim_{k\to\infty}\left(\sum_{m=1}^{k} S_m\right) = \infty \,. \qquad \textbf{(5)}$$

In particular, the following MSA modifications are frequently used:

- $S_k = 1/k$ (literally corresponds to the term MSA)
- $S_k = 1/\sqrt{k}$ (may exhibit a faster convergence if the starting point is far off)

The essence of MSA is to smooth up the outcome of an iterative procedure where n denotes iteration, in the following way:

$$D_k = (1 - S_k)D_{k-1} + S_k \widetilde{D}_k , \quad \textbf{(6)}$$

where:

$\widetilde{D}_k$ = raw outcome of iteration k,

D_k = smoothed outcome of iteration k that is fed back to the next iteration,

There are two standard ways in which MSA can be applied to ensure convergence of a demand model combined with network simulation:

- Averaging demand trip tables
- Averaging LOS variables

As shown in Figure 1, these methods also can be effectively combined for AB models, although it is not necessary for simple 4-step models

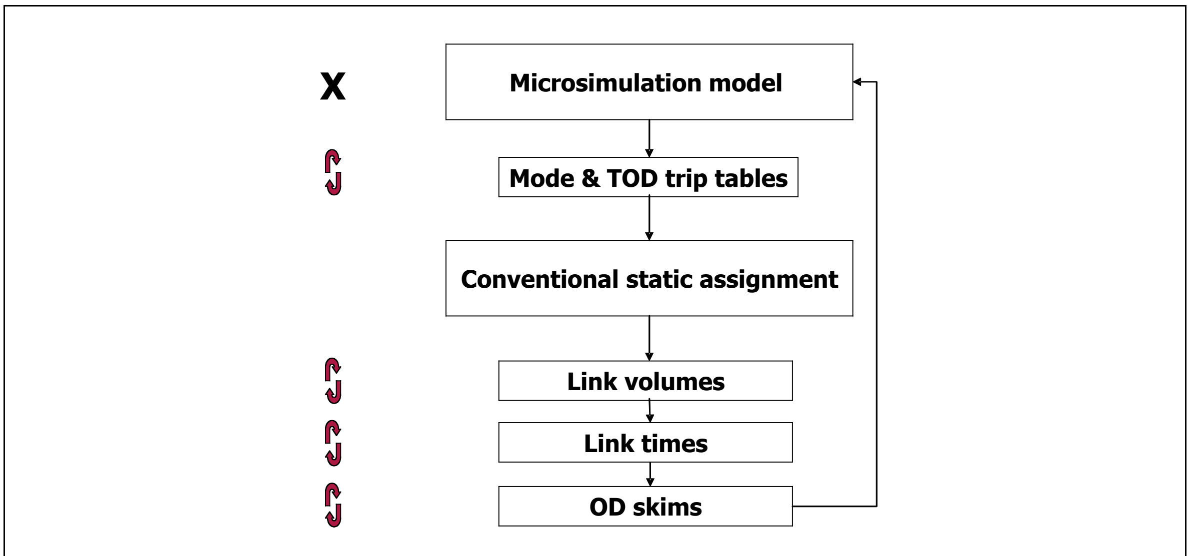

Figure 1. Equilibrium / feedback options

In many applications, a MSM demand model can be considered as a trip table generator followed by conventional assignment and skimming procedures. In practical planning situations, the model users are interested in aggregate outcomes of the MSM and do not usually track individual record details. With aggregation of MSM outcomes to OD demand flows, , equilibrium strategies with a MSM model are not principally different from equilibrium strategies applied with conventional models, with the exception of the different and more sophisticated way for the generation of the raw trip table in each iteration. In terms of methods for averaging model outputs/inputs, the following should be noted:

- Original output of MSM procedure (individual household / person / tour / trip characteristics) cannot be meaningfully averaged between iterations since it represents a unique set of discrete values associated with a different list of agents at each iteration.
- Trip tables can be averaged in the same way as for conventional models.
- LOS skims can be averaged in the same way as for conventional models. There are three different technical ways for averaging LOS skims:
    - Directly average origin-destination (OD) skim matrices.
    - Average link times and then skim OD LOS matrices.
    - Average link volumes, calculate corresponding link times, and then skim OD LOS matrices (preferred method).

Averaging link volumes is a better strategy that results in a faster convergence in over-congested networks, since link volumes are more stable than link travel times which are derived from an exponential function of link volumes.

Application Experience

The New York region represents an extreme example and challenge for modeling demand-network equilibrium for the following reasons:

- Very high levels of congestion that constitute a good example of a setting where a perfect convergence would be difficult to achieve even with an aggregate model.
- Huge number of persons (20,000,000), size of the regional network (4,000 zones), and multi-class trip tables (7_4,000_4,000) that result in significant model run time even for a single global iteration.
- Full variety of possible behavioral responses of travelers to changing LOS variables (switching modes, destinations, and/o time-of-day) that objectively contributes to instability / and non-convergence.

In this paper we have only presented some results of the application of some averaging strategies, with no enforcement methods applied yet. The numerous tests that comprise this research implemented with the New York model can be summarized in the following way:

- The most effective convergence strategy has been found to be a parallel MSA applied for both trip tables and link volumes producing synthetic LOS skims based on these link volumes for each subsequent iteration.
- There is a good level of convergence achieved with respect to network link volumes and aggregate county-to-county trip tables (28_28).
- In practical terms, the first 3-4 global iterations result in a reasonable equilibrium state while implementing additional 5-6 iterations brings only a marginal improvement.
- It can be clearly seen that the further improvement in convergence cannot be achieved without overcoming average Monte-Carlo error, and that further refinement of the procedure is bound to an effective handling of Monte-Carlo variability through enforcement.

Examples of convergence statistics from these tests are shown in **Figure 2**. The following different feedback strategies were tested with 9 global iterations implemented for each strategy:

- **Direct** – full update of trip tables and LOS skims with no averaging.

- **MSA** – full update of trip tables and standard MSA for link volumes.
- **Root MSA** – full update of trip tables and "square root" MSA for link volumes.
- **MSA Trip** – parallel standard MSA applied for trip tables and link volumes.

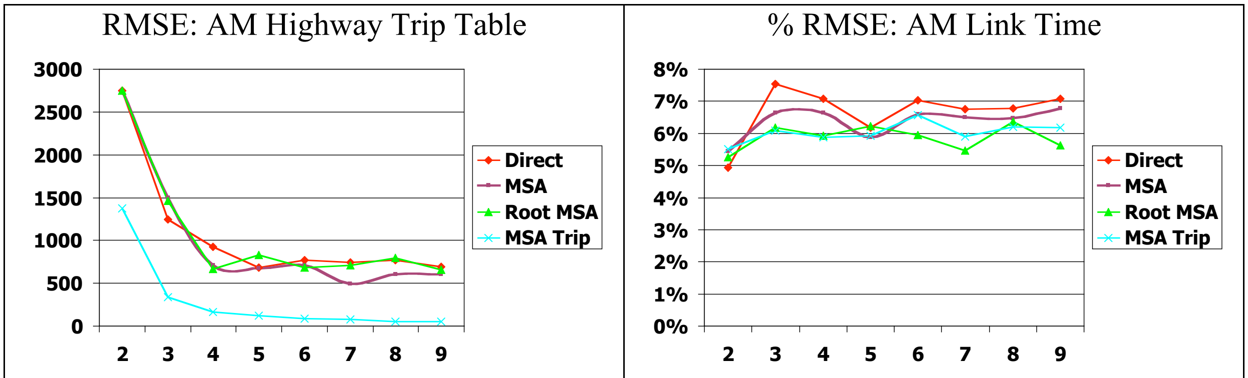

Figure 2. Examples of convergence statistics

The left-hand side of **Figure 2** relates to the Root of Mean Squared Error (RMSE) calculated for AM period highway trip table aggregated to 28_28 county-to-county flows. RMSE relates to the difference between two successive global iterations. It can be seen that the strategy that includes a parallel MSA for trip tables and link volumes achieves the best and absolute convergences. All other strategies that do not include averaging of trip tables result in a certain non-convergence level though it is comparatively small (±500 trips for an aggregate county-to-county flow that is less than 2%).

The right-hand side of **Figure 2** relates to the % RMSE (relative to the average value) for highway link times. It can be stated that there is no significant difference between the strategies with only a relative disadvantage of direct feedback. There are two major reasons for unavoidable fluctuations at the level of 5-6% for any of these strategies: one stems from Monte Carlo error, and the other from the highway assignment results themselves which are not stable for the over-congested AM period and even a small fluctuation of trip table may result in a significant change in route choice. This is, however a typical case with all types of models, and can only be diminished by very large (and often impractical) increases in the number of iterations of the user equilibrium assignment procedures. .

Conclusions

The following conclusions can be made:

- If the purpose of the MSM process is to produce a stable trip table for a Build scenario that can be compared in analysis with a similarly stable Baseline model output, such as for FTA New Starts application, then a very good level of convergence has been observed with an MSA strategy applied over trip tables and link volumes, and in this sense, it is shown that AB MSM models can accomplish this property, just as aggregate 4-step models do.
- Three major practical strategies for implementing feedback for the NY model application have been identified:

- **“Cold” start** that requires 8-9 global iterations. It can start with any reasonable approximation for LOS skims. It is necessary to implement for each Baseline scenario and/or year. It is however, is implemented only for exceptional Build scenarios that are characterized by large-scale regional impacts (for example, Manhattan area pricing).
- **“Warm” start** that requires 3 iterations. Input LOS skims are taken from the last iteration of the Base (Cold start) run. This is a standard procedure for Build scenarios.
- **“Hot” start** that includes a single iteration only starting with the LOS skims from the last iteration of the Base run. This is specific to FTA New Starts applications.

- The biggest source of instability in the current set of tests proved to be associated with stop-frequency and stop-location sub-models as well as the subsequent time-of-day choice. At the same time, the daily tour mode and destination choice part proved to be much more stable. This is an important manifestation of the principle of “structural changes vs. continuous fluctuations” that should be addressed in further research, specifically in an effective combination of averaging and enforcement.

Reference

1. Evans, S.P. (1976) Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment, *Transportation Research*, **10**, 37-57.
2. Florian, M., Nguyen S. (1978) A Combined Trip Distribution, Modal Split, and Trip Assignment Model, *Transportation Research*, **12**, 241-246.
3. Sheffy Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Englewood Cliffs, NJ.
4. Boyce, D.E. (1990) Network Equilibrium Models of Urban Location and Travel Choices: A New Research Agenda. In *New Frontiers in Regional Science*, M. Chatterji and R. E. Kuenne (eds) New York: Macmillan, 238-256
5. Vovsha. P, Schmitt D. (2007) Application of the MORPC Micro-Simulation Model for User Benefit Evaluation for Transit Project. Presented at the *86th Annual Meeting of the Transportation Research Board*, Washington, D.C., CD-ROM.