Microeconomic Interpretation of MATSim for Benefit-Cost Analysis

Benjamin Kickhöfer and Kai Nagel

This chapter explains how MATSim's agent-based framework can be interpreted from a micro-economic perspective and how it can be used for the economic evaluation of transport policies, e.g., for BCA (Benfit-Cost Analysis). The text of this chapter is partly taken from Kickhöfer (2014, Section 2.3).

Typically, the process of economic policy evaluation consists of three steps: First, forecasting changes in the system by modeling users' reactions to a policy (Section 51.1). Second, assigning some (potentially monetary) valuation to these changes (Section 51.2). And third, applying an appropriate aggregation rule (Section 51.3). As will be shown in the next sections, these steps are neither completely independent nor completely dependent on each other.

51.1 Revisiting MATSim's Behavioral Simulation

Estimating policy intervention benefits relies on a sound descriptive model able to predict individuals' related behavioral changes. As explained in Section 1.2, agents in MATSim optimize their mobility behavior over several iterations by reacting to the behavior of other agents. Even if one assumes homogeneous individual preferences in the behavioral parameters of their utility functions (see Section 3.4), activity locations and activity patterns of agents typically differ, meaning that the simulation deals with heterogeneous decision makers. It thus seems reasonable to interpret the simulation from a discrete choice modeling perspective (see Chapter 49). Another attractive reason to use this interpretation lies in the well-established approaches to estimate user benefits and system welfare changes in discrete choice models.

As shown by Nagel and Flötteröd (2012, also see Chapter 47 and Section 49.1.1), the MATSim choice model is equivalent to a standard MNL model under the following two conditions: first, valid choice sets have been found for all individuals; second, the score of each plan has converged to its expectation value (self-consistent state). An approximation of this can be reached by switching innovation off (Section 4.5.3) and forcing scores to convergence (Section 3.3.4, also see Section 49.1.1). Still, the following methodological issues remain:

1. **Choice set incomplete**: The maximum number of plans J in each agent's choice set is limited by memory constraints; the choice set for decision making is, hence, unlikely to be complete.

2. **Plans correlation from innovation**: Plans might be correlated. This is very likely if they are modified or replaced by best-response re-planning modules (e.g., the route choice module), since they always have a tendency to generate the same answer. However, random mutations, in general, also tend to result in correlated plans, since the concept of a mutation implies only a small move away from the parent. This violates the required IIA (Independence from Irrelevant Alternatives) property of the choice set necessary for a MNL model.

3. **Plans correlation from plans removal**: The current MATSim implementation has a tendency to retain similar, i.e., correlated, plans when the number of plans has grown beyond J, because the current default plans remover deletes the plan with the lowest score, which is also typically most different from other plans. As a result, normally only very similar plans—with very similar scores—remain in the choice set.

These three issues can lead to biased behavior, which would have consequences for economic evaluation. Possible solutions for these shortcomings are discussed in Section 49.2, and again, from a different angle, in Section 97.3. For the rest of this chapter, it will be assumed that the above issues are solved, and that a consistent solution has been found for the system states before and after the policy change. However, the following text briefly discusses possible impacts of the above issues on policy appraisal results, to facilitate better understanding.

51.2   Valuing Human Behavior at the Individual Level

Following de Jong et al. (2007), a major advantage of the agent-based approach is a seamless integration of (i) forecasting behavioral changes as a reaction to changes in the system, and (ii) the subsequent economic evaluation. In this section, it is shown how estimated agent-specific preferences, which determine behavior, can directly be used for deriving individual VTTSs and how they need to be modified for running a MATSim simulation to obtain individual utility differences resulting from a policy change. The next Section 51.3 will then focus on how these individual utility changes can be used to derive an indicator of overall welfare change for the considered population.

51.2.1   The Utility of Time

The MATSim scoring function of plan (= alternative) i consisting of $q = 0..N-1$ activities and trips has been introduced in Chapter 3 in the following form:

$$U = \sum_q U_{act,q}(t_{dur,q}, ...) + \sum_q U_{trav,q}(t_{trav,q}, ...) , \tag{51.1}$$

where monetary payments (e.g., tolls) are included in $U_{trav,q}$ and the index i was dropped for notational convenience.[1]

An approximate argument about optimal time allocation can be made as follows: Assume the constraint $T - \sum_q t_{dur,q} - \sum_q t_{trav,q} = 0$, i.e., that the time per day is limited by $T = 24h$, during which all trips and activities need to be completed. Let us now also assume that all travel times are fixed; i.e., we ignore the possible optimization from departure time or mode switches and concentrate on the activity time allocation problem. Optimizing under this constraint leads to the Lagrangian

$$L = \sum_q U_{act,q}(t_{dur,q}, ...) + \sum_q U_{trav,q}(t_{trav,q}, ...) + \mu \cdot (T - \sum_q t_{dur,q} - \sum_q t_{trav,q}) , \qquad (51.2)$$

where μ is the Lagrangian multiplier corresponding to the time constraint.[2]

Solving the optimization problem leads to

$$0 \;\overset{!}{=}\; \frac{\partial L}{\partial t_{dur,q}} = U'_{act,q}(t_{dur,q}, ...) - \mu \qquad (\forall q) \qquad (51.3)$$

and the time constraint equation from above, where $U'_{act,q} := \partial U_{act,q}/\partial t_{dur,q}$. Equation (51.3) states that, at the optimum and without further constraints, the $t_{dur,q}$ need be selected for all activities q such that all $U'_{dur,q}(t_{dur,q}, ...)$ are the same and equal to μ.

Equation (51.2) can also be seen as a linearized version of the indirect utility function; for example, reducing travel duration by Δt_q affects not only $U_{trav,q}$, but will also lead to a utility change of $\mu \cdot \Delta t_q$ from the constraint, which can be interpreted as the linearized utility effect of spending that time otherwise.[3] In consequence, the **marginal utility of time spent traveling** reads

$$\frac{\partial L}{\partial t_{trav,q}} = U'_{trav,q}(t_{trav,q}, ...) - \mu . \qquad (51.4)$$

μ is the **marginal utility of time as a resource**—the marginal utility generated by increasing T, i.e., by making the day longer than 24 hours. The marginal utility of time spent traveling is thus determined by μ, modified by "any enjoyment or dislike of the travel itself" (Small, 2012).

To get a handle on the MATSim utility function in Equation (51.1), μ and $U'_{trav,q}$ need to be obtained separately: μ in order to calibrate $U'_{act,q}$ as in Equation (51.3) and $U'_{trav,q}$ to calibrate the direct utility of time spent traveling, the offset to the marginal utility of time as a resource. This will be further discussed in Section 51.2.4.

51.2.2   The Utility of Money

Time allocation theory (DeSerpa, 1971; Jara-Díaz and Guevara, 2003) makes a similar argument for money, with a budget constraint similar to the time constraint. Just as the time constraint leads to a marginal utility of time as a resource, the budget constraint leads to a marginal utility of money as a resource.

---

[1] Strictly speaking, at this point, it would make more sense to stay with the scores S that MATSim generates. Section 51.2.5 discusses the relation between MATSim scores S, systematic utility V and total utility U in more detail. However, since the following text uses terms like "marginal utility of time" or "marginal utility of money", equations are also noted using U instead of S.

[2] This should not be confused with the scale parameter from discrete choice theory; here, to be consistent with time allocation theory, μ represents the marginal utility of time as a resource and corresponds to β_{dur} in Chapter 3.

[3] A reminder: the indirect utility function describes utility as a function of the value of the constraint that emerges when, for each value of the constraint, utility is maximized.

However, MATSim does currently not include such a monetary budget constraint. It is also questionable whether it should be introduced: the typical theoretical argument assumes the possibility of increasing one's income by working more hours. It is questionable if this functions in European countries, where work contracts typically include a fixed number of working hours, which cannot easily be changed. Hence, an alternative derivation of the marginal utility of money is necessary.

Assume that $U_{trav,q}$ includes a change in the monetary budget, $\lambda \cdot \Delta m$, e.g., invoked by fares or tolls. Then

$$\frac{\partial U}{\partial m} = \frac{\partial U_{trav,q}}{\partial m} = \lambda, \tag{51.5}$$

that is, reducing the monetary budget by Δm reduces the utility by $\lambda \cdot \Delta m$. We will therefore interpret λ as the **marginal utility of money**.[4] Taking the first derivative of L with respect to m would lead to the same result.

In contrast to the marginal utility of time above, we do *not* break down the marginal utility of spending money for travel into a marginal utility of money as a resource, and an offset for spending money on a particular purpose (for an example of this decomposition, see, e.g. Munizaga et al., 2008). Because there is no monetary budget constraint, there is also no neutral Lagrange multiplier that would give the marginal utility of money as a resource.

This, however, leads to the problem that if there are multiple monetary channels, they may have different marginal utilities of money. For example, the marginal utility of toll payments is larger than the marginal utility of payments for fuel—i.e., people find it less irritating to pay for fuel than to pay tolls (see, e.g., Vrtic et al., 2008). That is, each monetary channel, such as fuel cost, toll, public transport fare, or a toll refund, may lead to different preference estimates.

To our knowledge, there is no best solution to this problem in the literature. For the time being, we work with forcing all alternatives' cost-related parameters to a uniform value in preference estimation. However, choice modelers typically avoid limiting the model's degrees of freedom in this way, since it suppresses some information contained in the data.[5] It is therefore often impossible to obtain necessary parameter estimates from the literature. Where raw data is available, the same model can be re-estimated with a uniform marginal utility of money across alternatives (see, e.g., Kickhöfer et al., 2011; Tirachini et al., 2014).

Also, Small (2012) points out that the "neutral" marginal utility of money as a resource is difficult to estimate; for example, it is *not* the marginal utility of income. As an alternative research avenue, we could hypothesize that a measure's monetary channels are included in the choice experiment. For example, a travel time improvement in a value-of-time study could come together with a hypothetical income tax increase, *or* with a hypothetical toll. A rudimentary version of this actually takes place in Switzerland, where large infrastructure investments are bundled with tax increases that pay for them before they are put to public vote (see, e.g., BAV, 2013).

51.2.3   Value of Time

The **VTTS of trip** q is now defined as the marginal utility of time spent traveling (Equation (51.4)), divided by the marginal utility of money (Equation (51.5)), i.e.,

$$VTTS_q = \frac{\partial L / \partial t_{trav,q}}{\partial L / \partial m}, \tag{51.6}$$

---

[4] This constant, potentially person-specific, implies that income effects (Herriges and Kling, 1999; Daly et al., 2008; Dagsvik and Karlström, 2005; Jara-Díaz and Videla, 1989) do not play a role, i.e., that changes in expenses resulting from transport policies are not strong enough to change λ. In microeconomic theory, λ is the usual variable for the marginal utility of money and corresponds to β_m in Chapter 3.

[5] J. de Dios Ortúzar, personal communication.

where we are using the indirect utility function since we assume that the traveler compares optimal allocations before and after the change.

With $\partial L / \partial t_{trav,q} = U'_{trav,q} - \mu$ from Equation (51.2) one obtains

$$VTTS_q = -\frac{U'_{trav,q}}{\lambda} + \frac{\mu}{\lambda}\ ,\tag{51.7}$$

μ/λ is sometimes called the value of time as a resource.

51.2.4   From Estimated to MATSim Parameters

As stated above, most value of time studies do not separately estimate μ, λ, and $U'_{trav,q}$ ($\forall q$). Assume that an MNL estimation of behavioral parameters from a mode choice survey between car and PT uses the following utility functions:

$$\begin{aligned} U_{car,q} &= & \hat{\beta}_{trav,car} \cdot t_{car,q} &+& \hat{\beta}_m \cdot \Delta m_{car,q} \\ U_{pt,q} &= & \hat{\beta}_0 &+& \hat{\beta}_{trav,pt} \cdot t_{pt,q} &+& \hat{\beta}_m \cdot \Delta m_{pt,q}\ , \end{aligned}\tag{51.8}$$

where $t_{car,q}$, $t_{pt,q}$, $\Delta m_{car,q}$ and $\Delta m_{pt,q}$ are, respectively, travel times and monetary costs in the different modes, and $\hat{\beta}_x$ are the corresponding parameter estimates As explained in Section 51.2.2, $\hat{\beta}_m$ (the same as λ above) is assumed to be the same for all modes, or more precisely, for all types of expenditure.

According to Equation (51.4), the marginal utility of time spent traveling needs to be split into two components:

1. The marginal utility of time as resource, which needs to be used for $U'_{act}(t_{dur,q}, ...)$ ($\forall q$) in Equation (51.3).
2. The direct marginal utility of time spent traveling, which needs to be used for $U'_{trav,q}(t_{trav,q}, ...)$.

We do not know of any good way to perform this split; Kickhöfer et al. (2011) and Kickhöfer (2014) use the least negative $\hat{\beta}_{trav,mode}$ for μ (i.e., β_{dur}) and then re-calculate all other direct marginal utilities of travel time relative to that. As indicated in Section 3.4 of this book, this is currently the preferred procedure.

51.2.5   From Simulation Output to Evaluation

At the end of the simulation run, each agent n has a number of plans $i = 1..J$, each of them associated with a score $S_{n,i}$, computed according to Equation (51.1). For economic evaluation, the question arises how to aggregate these $S_{n,i}$ into an agent-value S_n, which can then be interpreted as a utility U_n. Possibilities include using:

• the logsum of the agent's plans scores, i.e., $\ln \sum_i e^{S_i}$
• the score of the agent's last executed plan,
• the average of the agent's plans scores, or
• the highest score of the agent's plans.

51.2.5.1   Using the Logsum of the Agent's Plans Scores

In literature, the logsum term

$$logsum_n = \ln \sum_i e^{V_i}$$

has been proposed for applied welfare analysis with Discrete Choice Models (Small and Rosen, 1981; de Jong et al., 2006; Kohli and Daly, 2006; de Jong et al., 2007). Under the assumption of a correctly specified model and choice set, the logsum term represents the EMU (Expected Maximum Utility) for a user with several options $i = 1..J$ in her choice set and the systematic utility of each option i is V_i. It is the expectation value, given that a random (Gumbel-distributed) ε_i is added to each V_i, and that the individual chooses the alternative with the highest $U_i = V_i + \varepsilon_i$.[6] In this interpretation, the (expected/average) MATSim score S_i is equated with the systematic part of the utility V_i.

However, as described in the previous Section 51.1, the use of MATSim as choice set generator yields issues with incompleteness of the choice set and with similarity of daily plans. In the current MATSim implementation, the maximum error occurs when all plans are copies of the best plan, rather than a diversity of plans. An upper bound of this error can be approximated as follows. Without loss of generality, assume that $i = 1$ is the plan with the largest systematic utility. Then

$$logsum_n = EMU_n = \ln \sum_{i=1}^{J} e^{V_i} \leq \ln \sum_{i=1}^{J} e^{V_1} = \ln(J \cdot e^{V_1}) = \ln J + \ln e^{V_1} = V_1 + \ln J .$$

At the same time, obviously

$$logsum_n = EMU_n = \ln \sum_{i=1}^{J} e^{V_i} \geq \ln e^{V_1} = V_1 .$$

Overall,
$$V_1 \leq logsum_n \leq V_1 + \ln J .$$

That is, for a choice set with I alternatives, the true logsum value lies between the systematic utility of the best option, V_1, and $V_1 + \ln J$.

51.2.5.2   Using the Score of the Agent's Last Executed Plan

Using, for each agent, the logsum over the scores of all plans implies that all these plans are valid behavioral choices. An alternative would be to simply use the score of the last executed plan. The behavioral interpretation consistent with this procedure is that there is no additional relevant randomness beyond what MATSim generates intrinsically. There has been no systematic work in this direction in the MATSim context, but such an approach might be justified in conjunction with the idea of explicitly generating the missing $\varepsilon_{n,i}$ for each person-alternative-pair n, i, then always selecting the best plan, as described in Section 97.4.6.

51.2.5.3   Using the Average of the Agent's Plans Scores

In principle, it is also possible to use

$$S_n = \frac{1}{J} \sum_{i=1}^{J} S_{n,i} \cdot P_{n,i} , \tag{51.9}$$

where $P_{n,i}$ is the probability of plan i for agent n. This can, however, only be justified when the choice probabilities, $P_{n,i}$, are interpreted like mixed strategies from game theory, i.e., that sampling from these probabilities is the true agent behavior. In principle, we cannot see why such an

---

[6] At this point, we assume that V_i is absorbing the scale parameter.

interpretation should be plausible—except that it is statistically the same as Section 51.2.5.2 with the advantage of having less variance. Note, however, that the approach is intertwined with the choice model. If, e.g., $P_{n,i}$ is one for the plan with the highest score and zero for all other plans, then Section 51.2.5.2 and Equation (51.9) are identical.

51.2.5.4   Using the Highest Score of the Agent's Plans

Alternatively, one could simply use the highest score that the agent has in its plan. This would only make sense if true behavior is assumed to always select the plan with the highest score. Again, this should then also be expressed by the choice model, i.e., using the highest score only makes sense when the agent always selects the plan with the highest score, in which case the result becomes the same as Section 51.2.5.2 and 51.2.5.3.

51.2.5.5   More Complicated Variants

Section 49.1.2 discusses the idea that MATSim's typical choice model might be described by a mixture-of-logit model. In that model, ϵ_{ni} remains fixed per agent n and alternative i, but other attributes such as the network conditions vary from one iteration to the next. In Equation (49.5), η_{ni} denotes these random, but simulation-generated, deviations from the average conditions; let us add an index k for the iteration number, i.e., write η_{ni}^k. That is, it is postulated that a real person would know both ϵ_{ni} and η_{ni}^k, but the simulation only knows the latter (through the MATSim score). Equation (49.5) then just describes the resulting choice distribution from what MATSim often does, i.e., apply a logit model to scores that are not averaged.

   At least for η_{ni}^k that are uncorrelated from one iteration to the next it is, however, clear that this will not result in optimal *average* agent behavior – the agent may be pushed towards some choice by a random fluctuation of the η_{ni}^k, but obtaining a much lower score from that choice in the average. Overall, the agent would be better off by first averaging the score of each alternative over many iterations, and then basing her choice on those scores. This goes back to the converged scores of Section 49.1.1.

   Calculating benefits from a mixture-of-logit interpretation becomes thus rather involved: we postulate that the agent sees the full MATSim score, plus some private ϵs; that she optimizes based on the sum of these two; that the MATSim simulation, however, does not know the ϵs and thus has to sample from the logit model; but that the economic utility has to include the effect of the ϵs although we do not know them, as in Section 51.2.5.1. Overall, thus, assigning utility values to such behavior as described by Equation (49.5) requires a better understanding of underlying behavioral rationality. Section 97.4.6 discusses this further.

51.2.5.6   Summary

Overall, there seem to be two consistent strategies to aggregate various plan scores of an agent n into one value:

- If the choice model is a logit model, then using the logsum term over all plan scores as the agent's utility U_i is consistent with the choice model.
- If the choice model is such that the plan with the highest score is selected, then using that score as the agent's utility U_i is consistent with the choice model.

   In both cases, the choice model needs to be consistent with the behavioral assumption about the agent, i.e., in the first case it needs to be assumed that the model does not know the true agent choice beyond the choice probabilities and the model system thus has to repeatedly sample from these probabilities. In the second case it needs to be assumed that the randomness has already been "frozen" into the score computation (see Section 97.4.6) and the agent thus selects the plan with the highest score.

In both cases, the calculated individual score differences that result from a policy measure can be directly used in order to identify winners and losers.

Some economists claim that the modeler's task of providing information for decision support ends at this point (Ahlheim and Rose, 1989). However, in practice, some (monetary) valuation of the resulting behavioral changes is often required. The next section reviews different possibilities to monetize and aggregate individual utility differences in the MATSim context.

51.3   Aggregating Individual Values

After having obtained the individual changes in terms of utility, it is often necessary to convert these utility changes into monetary terms for economic evaluation, e.g., in BCA. Unfortunately, no "correct" monetization or aggregation approach exists for individual utility differences. This is reflected by the ongoing discussion[7] between transport policy appraisal experts:

1. The first stream argues in favor of a consistency in values used in demand modeling and appraisal (Grant-Muller et al. (2001, p.255), Bickel et al. (2006, p.S4 and p.S8), and Proost[8]). Values from literature should only be used if behavioral model values are not available. These researchers are, however, aware that this procedure potentially limits the comparability of projects in different regions of the same state, or in different member states of the EU (European Union). In consequence, additional indicators such as absolute time savings per income group should also be reported to address equity issues.

2. In contrast to the above, Mackie and Worlsley (2013, p.12) state, that in the United Kingdom, "standard [VTTS] values per minute would be used across incomes, modes and regions. Therefore, their practice is to use behavioral information for modeling but standard values for appraisal." Also Daly (2013) distinguishes between "valuation", i.e., people's willingness-to-pay (or accept) for marginal changes, and "appraisal", i.e., what these changes are worth from a societal point of view.

3. Fowkes (2010), OECD (2006), and Gühnemann[9] argue slightly differently, but in the same direction: modeling and evaluation should be based on the best heterogeneous preferences available; in the evaluation, additional weights should be introduced, e.g., to counter the effect of decreasing marginal utilities of money, or increasing VTTS with income, respectively. These weights would, thus, define the underlying equity concept of the appraisal method.

4. However, as Ahlheim and Rose (1989) point out, no approach to empirically determine these weights is available without assuming some arbitrary a-priori specification. In consequence, every interpersonal comparison of utility changes requires some normative decision and the weights need therefore to be determined on a political level.

One goal of this section 51.3 is to show the impact of a possible integration between behavioral modeling and economic evaluation in the same agent-based framework. First, a conversion into *income equivalents*, and second, a conversion into *time equivalents* (possibly followed by some conversion into money terms).[10] The choice of the procedure depends on a (normative) decision whether one *EUR* or one *h* should be valued equally across individuals. It is, therefore, important

---

[7]   A similar overview on this discussion is given by Börjesson and Eliasson (2014).

[8]   S. Proost, personal communication.

[9]   A. Gühnemann, personal communication.

[10] Kickhöfer (2014) shows that the choice of the monetization and aggregation procedure can have major impact on the results when heterogeneity is assumed in user preferences.

that decision makers and modelers who deal with economic evaluation understand the possible effects of that choice; simply going with the most common approach may not be advisable.

51.3.1   Income Equivalents

Basic Approach   The most common approach used in welfare economics to convert utility changes into money terms is to calculate the monetary amount ΔY_n that one would need to give or take from individual n to offset the impact of the policy on the utility level ΔU_n. According to Equation (51.1), it is calculated as

$$\Delta Y_n = -\frac{\Delta U_n}{\lambda_n} \, . \tag{51.10}$$

Note that the marginal utility of money, λ_n, might be person-specific, e.g., dependent on the person's income.

The monetary amount $-\Delta Y_n$ from above represents individual Consumer Surplus. Its absolute value is, in the absence of income effects (see Footnote 4 in Section 51.2.1), equal to the Compensating Variation and the Equivalent Variation (Daly et al., 2008). The overall welfare change ΔW for the population with individuals $n = 1..N$ is then calculated by

$$\Delta W = -\sum_{n=1}^{N} \Delta Y_n \, . \tag{51.11}$$

Equity   The above approach is often criticized for equity reasons: if the marginal utility of money is—in the behavioral model—assumed to decrease with income, and these values are directly (without additional weights) used in economic evaluation, rich people will have a stronger impact in the evaluation process than poor people. In turn, this might lead, e.g., to investments in expensive high-speed trains on major corridors rather than affordable train services for everyone. In terms of equity and public acceptance, such specification in the appraisal method might not be desirable. To counter this effect in economic evaluation, the use of standard or equity values is proposed in the literature. In this context, Jara-Díaz (2007, p.106ff) introduces the *social utility of money* and the *social price of time*. For a more general overview of possible solutions how to address equity issues, see Rizzi and Steinmetz (2015).

A rather ad-hoc but simple possibility is to replace the person-specific marginal utility of money, λ_n, with a population average,

$$\overline{\lambda} := \frac{1}{N} \sum_n \lambda_n \, , \tag{51.12}$$

and then

$$\Delta Y_n = -\frac{\Delta U_n}{\overline{\lambda}} \, . \tag{51.13}$$

Following the argument by Fowkes (2010), OECD (2006) and Gühnemann et al. (2011) mentioned above (Item 3), this would be one particular way to introduce the necessary weights. Alternatively, one could think of fixing the social weight of every person to 1.0, and derive the social price of all attributes included in the generalized costs from there (Jara-Díaz, 2007, p.108f).

51.3.2   Time Equivalents

Another option to derive a monetary measure of welfare changes is composed of two steps: First, a conversion of individual utility changes into *equivalent hours of time as a resource* (Jara-Díaz et al., 2008; Mackie et al., 2001). This would be the number of hours ΔT_n that one would need to give or

take from individual n to offset the policy impact on the utility level ΔU_n. Second, a monetization of the resulting numbers through an arbitrary conversion factor, i.e., the monetary value of one hour for the individual or for society.

In the MATSim sense, one could first calculate the corresponding time equivalent by

$$\Delta T_n = -\frac{\Delta U_n}{\mu_n} \ . \tag{51.14}$$

Similar to the marginal utility of money, also the marginal utility of time as a resource, μ_n, might be person-specific.

One option would be simply provide time equivalents, i.e., the BCA would return time equivalents per invested monetary unit. In many situations, however, it is desirable to convert all impacts of a policy into monetary terms, i.e., to compute,

$$\Delta Y_n = \alpha_n \cdot \Delta T_n \ , \tag{51.15}$$

and to compare ΔY_n with investment or changes in external costs. The following options are then possibilities for α_n:

- The obtained time equivalents ΔT_n could be converted in monetary terms using the person-specific resource values of time, i.e.,

$$\alpha_n = \frac{\mu_n}{\lambda_n} \ . \tag{51.16}$$

  This would obviously result in the same monetary amount as the income equivalent approach from Equation (51.10).
- Following Mackie and Worlsley (2013), one could argue that the resource value of time should be the same for every individual, and, thus, use some average value for monetization, e.g.,

$$\alpha_n \equiv \overline{\alpha} = \frac{1}{N} \sum_{n=1}^{N} \frac{\mu_n}{\lambda_n} \ .$$

- As another alternative, one could average over the marginal utility of money only, i.e.,

$$\overline{\lambda} = \frac{1}{N} \sum_{n} \lambda_n$$

  and then

$$\alpha_n = \frac{\mu_n}{\overline{\lambda}} \ . \tag{51.17}$$

  This would highlight that some persons are more pressed for time than others, while, at the same time, using an equal value for the marginal utility of money. Clearly, this gives the same result as Equations (51.12) and (51.13). It does, however, lend itself to a clearer interpretation: first, all utility differences are converted to a comparable scale, i.e., time as a resource (Equation 51.14). Then, these times are converted to a monetary scale, using a conversion factor which includes the pressure for time (i.e., the person-specific μ_n) but assumes an average marginal utility of money.

In all cases, the overall welfare change ΔW for the population with individuals $n = 1..N$ is then calculated identically to Equation (51.11), i.e., by $\Delta W = - \sum_{n=1}^{N} \Delta Y_n$.

51.3.3   Income vs Time Equivalents: Discussion

The sections above show how to monetize and aggregate individual utility differences though income equivalents or time equivalents. To summarize:

- Income equivalents put emphasis on the individual willingness-to-pay, whereas time equivalents focus on time pressure.
- The aggregation of income equivalents yields the overall equivalent monetary cash flow that would be generated by the project for the population considered. That is, one *EUR* is valued equally across individuals.
- The aggregation of time equivalents yields the overall equivalent lifetime hours that would be generated by the project for the population considered. That is, one hour of lifetime is valued equally across individuals.
- A monetization of time equivalents using person- and activity-specific resource values of time leads to the same total benefit as directly aggregating income equivalents.
- A monetization of time equivalents using some average value of time as a resource, therefore generally leads to a different total benefit than directly aggregating income equivalents. Such an approach maintains the equal value for one h of lifetime.

51.3.4   Conclusion and Recommendations

Scoring Function   A correct scoring function is central to correct MATSim functioning. The mathematics and understanding of that scoring function need to be derived from time allocation theory in economics. In particular, any marginal utility of travel time needs to be split into the marginal utility of time as a resource (μ in the text above, and β_{dur} in Section 3.4) and an additional direct marginal utility of time spent traveling ($U'_{trav,q}$ in the text above, and $\beta_{trav,mode,q}$ in Section 3.4).

Since most discrete choice models estimate the sum of these two, definition is required about how to split up this sum. A somewhat ad-hoc way to achieve this is to find the mode with the largest ($=$ least negative) marginal utility of time and use that value for the marginal utility of time as a resource. That reference mode's direct marginal utility of time spent traveling is then zero; all other modes' direct marginal utilities of time spent traveling are relative to that of the reference mode.

If one is interested in monetization, i.e., converting utility values into monetary terms, then additionally the marginal utility of money as a resource (λ in the text above, and β_m in Section 3.4) needs to be known. Our current approach to obtain an approximation to λ is to force all monetary preferences in the estimation of a choice model to a unique value. If this is not possible, then one has to make a normative decision which monetary channel is considered most "neutral", i.e., most similar to an "unearned income" channel.

Choice Model and Score Aggregation   MATSim agents normally have more than one plan; each plan has a score. There are two consistent approaches to come up with a utility value from those scores:

- Using a MNL choice model that makes probabilistic draws from those plans using their scores: The correct aggregation is then the logsum of all scores.
- Using a choice model that selects the plan with the highest score: The correct aggregation is then to use the score of that plan.

In both cases, the result is the total utility U of the choice set. In the first case, the logsum term includes an expectation value of the randomness, typically denoted by ε. In the second case, all randomness, if any, needs to be "frozen" into the alternatives, and included into the computation of the score.

Monetization   Individual utility differences resulting from a change in the transport system can be converted into monetary terms by dividing them by λ_n. The result is the change in individual user benefit. Aggregating these individual benefits provides an indicator for the overall welfare change. Since λ_n may vary among agents, e.g., according to their incomes, such approach will put a higher weight on people with small λ_n, typically those with large incomes. An alternative is to use an average $\bar{\lambda}$ for this conversion, even when the behavioral model (= the scoring function) uses person-specific λ_n.

Acknowledgements