IHOP3 – Economically consistent simulation of travel behavior with MATSim

Olivier Canella (WSP), Gunnar Flötteröd (KTH), Joacim Thelin (Sweco)

May 2, 2017

Abstract

The IHOP project series aims at building Sweden's next generation strategic transportation model system. IHOP1 investigated the feasibility of deploying a dynamic and disaggregate network simulation package. IHOP2 developed a technical framework for integrating travel demand models and network assignment packages through the MATSim technology. IHOP3 moves on to ensure an economically consistent analysis of the travel behavior simulated in such a system.

The concrete challenge addressed by IHOP3 is as follows. Sweden's national travel demand model Sampers is static (i.e. it does not model time-of-day) and aggregate (i.e. it models representative person groups but no individual travelers). The person/network simulation system MATSim is, on the other hand, dynamic (full days are simulated second-by-second) and disaggregate (individual synthetic travelers interact in a simulated network environment).

Given these different resolutions of time and travel demand, different utility functions are used in Sampers and MATSim, which in turn leads to different models of travel experience, leading ultimately to different cost-benefit analysis results. The objective of IHOP3 is to devise a simulation method that allows for the economically consistent integration of Sampers and MATSim, resulting in the specification of a common, fully dynamic and person-centric, utility function in both Sampers and MATSim. The proposed solution, which is already partially implemented in Sampers/MATSim, is demonstrated in a small simulation setting, with the objective to indicate its scalability and readiness for implementation in a production version of the IHOP system.

This project has been funded by Trafikverket under contract TRV 2016/12446.

1 Introduction

This document describes how the static and aggregate travel demand model Sampers and the dynamic and disaggregate person/network simulation system MATSim can be combined into a model system that allows for economically consistent cost-benefit analysis (CBA). This report refers to Sampers version 4 throughout. The two essential steps undertaken in this work are

1. the implementation of a Sampers micro-simulator, in which synthetic travelers make individual travel choices, and
2. the development of a simulation logic in which both Sampers and MATSim use the same fully dynamic utility function.

Step 1 has been realized by Sweco and WSP; it is document in Section 2. Step 2 has been realized by KTH; it is documented in Section 3. Either section comprises an experimental validation. The remainder of this introduction outlines the respective approaches.

1.1 Sampers micro-simulation

The so far existing Sampers model is aggregate in that it predicts, for each population subgroup and for each considered sequence of travel tours, comprising mode(s) and destination(s), the probability that the respective subgroup travels according to the respective tour sequence. MATSim, on the other hand, relies on a synthetic population of individual travelers, which are representative for the real study region population. The two key advantages of the synthetic population representation are (i) ability to represent arbitrary population heterogeneity and (ii) full compatibility with a dynamic network flow micro-simulation.

A micro-simulation version of Sampers has hence been developed. This version aims at following the same model specification as the original Sampers. (It will turn out that this does not impair the Sampers/MATSim utility consistency.) The difference is merely in how the model outputs are generated. The previously existing Sampers computes choice probabilities per subgroup and travel alternative. Multiplying the size of a subgroup with its choice probabilities yields the (fractional) number of subgroup members that use each travel alternative. The micro-simulation, on the other hand, (i) creates as many synthetic persons per subgroup as this group is large and (ii) draws one travel alternative for each synthetic person according to the choice probability distribution in the corresponding subgroup. The result is a synthetic traveler population, containing for each individual a chosen tour sequence with mode(s) and destination(s), that can be directly fed into MATSim.

1.2 Consistent utility representation in Sampers and MATSim

Running the Sampers micro-simulation as explained above uses Sampers' original utility function to simulate the travel choices (tour sequences comprising mode and destination information) of a synthetic traveler population. MATSim then simulates (i) the choice of routes and departure times for every individual in this population and (ii) the resulting network-wide vehicle flows and travel times. MATSim is fully dynamic in the sense that (i) network conditions (flows, travel times) change continuously throughout a simulated day, and (ii) the simulated route and departure time choices account for these dynamics. In other words, MATSim adds dynamic route and departure time information to the mode and destination choices made in Sampers. The overall result are detailed travel plans for every synthetic individual that comprise a sequence of tours with destination and travel mode information, the routes along which these tours are executed, and the departure time for every single route.

To leverage this level of detail in a credible CBA, a consistent perception of time-of-day in Sampers and in MATSim is established. The solution is *not* to add dynamics to Sampers because this would (i) require to develop and estimate a new, dynamic Sampers version and (ii) detach this dynamic Sampers version from the further development of the nation-wide static version. Instead, a Sampers post-processing logic is developed that takes as input multiple travel choices per individual from the static Sampers micro-simulation version, plus dynamic network information from MATSim. Using statistical techniques, this post-processor then selects one Sampers tour sequence per synthetic traveler such that this selection becomes statistically equivalent to having been made in a fully dynamic version of Sampers. This design is robust to changes to Sampers, in that it continues to function even when the Sampers logic is changed (for instance, re-estimated), as long as the Sampers outputs required by the post-processor can still be provided.

2 Sampers micro-simulation

This section describes how Sampers is enriched with functionality (i) to simulate (draw) individual travelers consistently with the population information contained in the SAMS database, and (ii) to simulate (draw) for each traveler one or more travel plans, consisting of a sequence of tours that are annotated with destination and mode information. The programming work has been undertaken in the not yet released development version of Sampers 4. The resulting code is located in the repository https://bitbucket.org/KTH-TLA/sampers-2015/, branch IHOP3_JT2. Given that Sampers 4 is itself under development, the results presented here are tentative and to be understood as a proof of technical feasibility.

2.1 Simulation of synthetic travelers

The representative traveler groups defined in Sampers' SAMS database are split into simulated travelers (individuals, agents) for each individual zone as follows.

Consider a roulette wheel consisting of an inner wheel that can spin within a fixed outer wheel. Put, with uniform distances, as many marks on the inner wheel as there are travelers in the considered zone. Then, mark for each representative traveler group a region on the outer wheel such that the size of this region is proportional to the share of that group in the considered zone. Spin the inner wheel and observe its final position. (The numeric simulation of this process consists in uniformly randomizing the angle of the inner wheel.) The number of inner marks pointing at an outer region indicates the number of individual travelers generated within the corresponding representative population group. In cases where representative groups are annotated with value ranges (for instance "age between 30 and 40 years"), one concrete realization is generated by sampling from the respective value range.

Figure 1 shows an example for eight travelers and three population subgroups with weights 0.3, 0.2 and 0.5. In the displayed configuration, three travelers (number 2, 3 and 4) are created with the properties of representative group 1 (weight $w = 0.3$), one traveler (number 5) with the properties of group 2 (weight $w = 0.2$), and four travelers (number 6, 7, 8 and 1) with the properties of group 3 (weight $w = 0.5$).

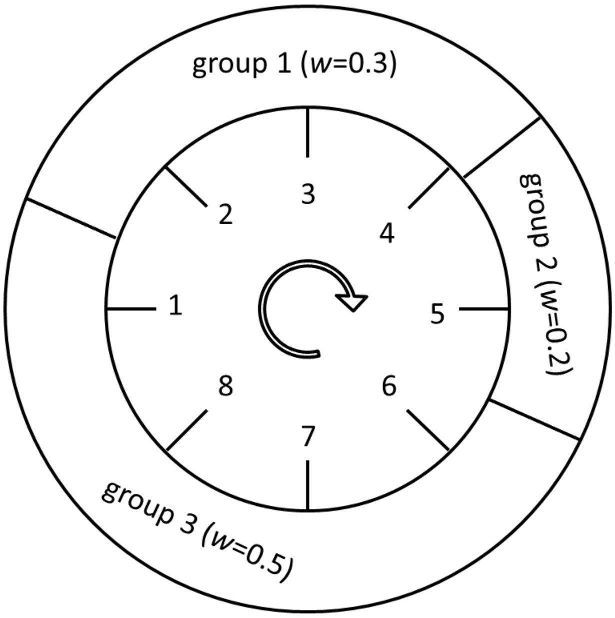

Figure 1: Roulette-wheel population simulation

2.2 Simulation of tour sequence choices

This is again simulated per zone. Sampers defines the probability of drawing T tours, with the (purpose, destination, mode) of tour t being denote by(p_t, d_t, m_t), as

$$\underbrace{\Pr(T)}_{\text{number of tours}} \cdot \underbrace{\Pr(p_1, \ldots, p_T \mid T)}_{\text{tour purposes}} \cdot \prod_{t=1}^{T} \underbrace{\Pr(m_t \mid p_t)}_{\text{mode choice}} \underbrace{\Pr(d_t \mid p_t, m_t)}_{\text{destination choice}} . \tag{1}$$

Each agent is independently assigned a tour sequence pattern according to this distribution by first drawing a number of (zero to four) tours and then drawing the tour purposes given that number. Then, a mode and a destination are drawn independently for each tour: First, a mode is drawn given the tour purpose, and then a destination is drawn given the tour purpose and the mode. The deployed sampling logic can again be imagined as a roulette process, now with a single mark on the inner wheel, the regions on the outer wheel representing the choice alternatives, and the their sizes being proportional to the respective choice probabilities.

2.3 Test results

The following results aim at demonstrating the technical feasibility of the Sampers micro-simulation approach. Note that the development version of Sampers used here has not yet been calibrated against a household survey. Simulation experiments have been performed for the Stockholm region.

Synthetic individuals

The number of individuals in the representative groups of the new Sampers version are integers. As a consequence, the roulette randomization would not even be needed here. It anyway perfectly reproduces the frequencies for sex, age group and income group. In the case of decimal weights per representative group (as it is for instance the case in the household survey), the roulette selection makes a sampling error of at most one individual, but is by design unbiased (meaning that it makes on average an error of zero).

Tour generation

The first considered indicator is the number of home-based tours per individual. The result is shown in Table 1. The displayed frequencies are computed from overall 1'898'069 tours in Stockholm county. One observes a sampling error of at most 0.1 %.

The second considered indicator is the number of tours per activity type, shown in Table 2. Here, the deviations are larger and cannot be explained by finite sample sizes; the average error over all activity types is 7 %. The reason

Table 1: Validation: number of tours per day

Tours per day	Sampers probabilities	Simulated frequencies
0	26.8 %	26.7 %
1	50.3 %	50.3 %
2	18.3 %	18.3 %
3	3.8 %	3.8 %
4	0.9 %	0.9 %
Total	100 %	100 %

Table 2: Validation: number of tours per activity

Activity	Sampers frequencies	Simulated frequencies	Difference	(%)
Work	418'476	422'446	3'970	1 %
SchoolE	135'722	147'533	11'811	9 %
SchoolS	28'178	29'032	854	3 %
SchoolA	33'811	29'467	-4'344	-13 %
Recreation	294'061	430'509	136'449	46 %
ShoppingDaily	251'261	310'459	59'198	24 %
ShoppingRarely	172'127	144'551	-27'576	-16 %
Visit	162'268	155'667	-6'601	-4 %
Escort	62'864	40'096	-22'768	-36 %
SHCC	88'342	77'738	-10'604	-12 %
Other	112'724	94'736	-17'988	-16 %
Business	17'858	15'835	-2'023	-11 %
Total	1'777'692	1'898'069	120'377	7 %

for this is that, due to time budget constraints, a simplified model specification has been used when implementing the sampling logic, meaning that the implemented and desired distribution are systematically different. Clearly, these simplifications will need to be replaced by a proper implementation in a production version of the code.

Mode choice

The last considered indicator is the modal share for work trips. Results are shown in Table 3. The biggest difference is 0.4 %.

2.4 Generation of output data (input to MATSim)

Overall, the simulation processes described here lead to the following Sampers micro-simulation outputs.

1. Synthetic population, including person-specific attributes (age, income etc.) and home locations.
2. A configurable number of simulated travel plans (tour sequence with modes and destinations) per synthetic person.
3. For each travel plan,
   (a) its systematic utility as evaluated in Sampers, split into (i) travel-time independent utility and (ii) the remaining travel-time dependent utility;
   (b) its choice probability according to (1);
   (c) the destination zone identifiers for all tours in the plan.

Table 3: Validation: modal share for work trips

Work	Car Driver	Car Passenger	Public Transport	Walk	Bicycle
Sampers probabilities	53.9 %	2.1 %	20.5 %	10.2 %	13.2 %
simulated frequencies	53.5 %	2.1 %	20.7 %	10.4 %	13.3 %
Difference	-0.4 %	0.0 %	0.2 %	0.2 %	0.1 %

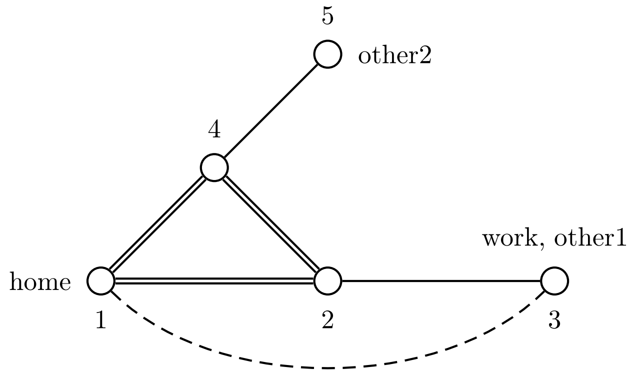

Figure 2: Example network

Items 1 and 2 are written jointly into a MATSim population file in compliance with MATSim's standard specification (population_v5.dtd). A computation of the data elements in item 3 has also been implemented, but this data is currently not written to file because a suitable standard file format has been made available only very recently by the MATSim development team.

3 Consistent utility representation in Sampers and MATSim

Since the methods described in this part of the report are somewhat more abstract, the main ideas are expressed in the context of a small example, which is introduced upfront.

3.1 Example

Consider the network shown in Figure 2. All solid (road) links in this network are bi-directional and have a maximum speed of 50 km/h. The double-lined solid links have a capacity of 1000 veh/h, and the single-lined solid links have a capacity of 750 veh/h. The dashed line represents a public transport (PT) alternative that operates congestion-independently at half the car speed limit. The geometrical scale is such that the distance between node 1 and 2 is one kilometer.

5'000 persons[1] have their *home* location at node 1 and their compulsory *work* location at node 3. An optional *other* activity is possible either at node 3 (*other1*) or at node 5 (*other2*). Every person selects one of the following eight tour sequences:

1. *work* by car
2. *work* by car, then *other1* by car
3. *work* by car, then *other1* by PT
4. *work* by car, then *other2* by car
5. *work* by PT
6. *work* by PT, then *other1* by car
7. *work* by PT, then *other1* by PT
8. *work* by PT, then *other2* by car

A tour sequence is subsequently also denoted as a (travel) plan.

One wishes to implement a Sampers/MATSim model system of this scenario where travel behavior is based on the following, fully dynamic utility ${V_{ni}}^{\text{dynamic}}$ assigned to travel plan i of traveler n:

$${V_{ni}}^{\text{dynamic}} = V_{ni}^{\text{dest,mode}} + V_{ni}^{\text{schedule}} \tag{2}$$

where $V_{ni}^{\text{dest,mode}}$ only evaluates travel time independent properties of the destination(s) and mode(s) contained in the plan, and V_{ni}^{schedule} only evaluates the remaining travel-time dependent terms of that plan. These terms are specified as follows.

[1]A scaling trick is used here, in that only 1000 persons are simulated on a network in which flow and space capacities are scaled down by a factor of 0.2 (Flötteröd, 2016).

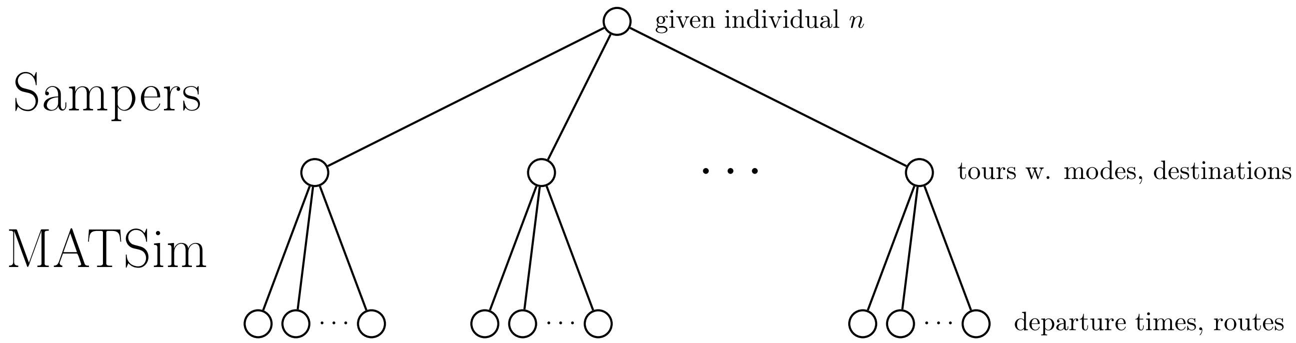

Figure 3: Hierarchical model system structure

- $V_{ni}^{\text{dest,mode}}$ is for simplicity set to a flat zero for all alternatives. This means that in the absence of scheduling costs, the simulated travelers in Sampers choose uniformly between all travel plan alternatives. This specification is clearly simplistic, yet it suffices to demonstrate all model properties of interest.
- V_{ni}^{schedule} evaluates the time structure of the given plan. It is defined through

$$V_{ni}^{\text{schedule}} = \text{ASC}_i + \beta_{\text{travel}} \cdot \text{travel-duration}_i \ldots + \sum_{\text{activities } a \text{ in plan } i} \beta_{\text{act}} \cdot \text{desired-duration}_a \cdot \log\left(\text{realized-duration}_a\right) \quad (3)$$

where ASC_i is an alternative-specific constant for plan i, $\beta_{\text{travel}} = -6\,\text{h}^{-1}$ is a negative (penalty) coefficient for the total (summed over all trips) travel-duration$_i$ in plan i, $\beta_{\text{act}} = 6\,\text{h}^{-1}$ is a positive (reward) coefficient for activity implementation, and desired-duration$_a$ and realized-duration$_a$ are activity a's desired and realized duration, respectively. The logarithmic form of the activity utility terms has the effect that travelers prefer realized activity durations that are close to the respective desired values. This constitutes a basic instance of MATSim's built-in scheduling utility function. It can be enriched with further parameters (such as activity priorities, public transport ticket prices, ...) and even be replaced by a systematically different specification (for instance one with individual-specific parameters, non-linear travel time costs, ...). Further details can be found in Kickhöfer and Nagel (2016).

Another consequence of the logarithmic valuation of activity time is that splitting one activity into two activities with in total the same length increases utility. Plans comprising only a single tour are hence given a positive ASC (of value 102) in order to avoid such plans becoming completely unattractive attractive alternatives in the given, synthetic scenario. All other plans receive an ASC of zero.

The desired duration of *work* is 8 hours and that of *other* activities is 2 h. The remaining duration of the day is allocated to the *home* activity. The realized duration of an activity depends on when a traveler arrives at and departs from an activity location: An activity can only be performed during its opening time, which is 7:00-18:00 for *work* and 9:00-22:00 for *other*; the *home* activity is always available.

Overall, this setting can be interpreted as a hierarchical equilibrium model system.[2] The lower-level network model (MATSim) takes as input the tour sequence (including modes and destinations) for every traveler. It then computes, for each traveler, subjectively optimal routes and departure times. The upper-level demand model (Sampers) takes as input the resulting performance of these tour sequences (from MATSim) and selects, for every traveler, the subjectively optimal tour sequence (which in turn is input to MATSim). Figure 3 illustrates this structure.

3.2 Consistency problems with a naive model system setup

In a naive integration of Sampers and MATSim, the time-dependent link travel times computed by MATSim would be flattened into static inter-zonal travel time matrices per mode. (The same could be done with travel distances and monetary costs; this report only refers to travel times for presentational simplicity.) Based on this data, Sampers would select a tour sequence for every simulated traveler. The resulting mode/destination travel plans would then be inserted into MATSim, where they would be enriched with approximately equilibrated routes and departure times. The resulting network travel times would again be aggregated and returned to Sampers, etc. This is roughly the approach taken in the IHOP2 project, which focused on technical model interoperability (Canella et al., 2016).

The fundamental problem one faces in this setting is that Sampers is a static model that treats time merely as a tour attribute, whereas MATSim is a dynamic model that treats time as a degree of freedom along which a travel schedule is laid out.

[2] MATSim computes, like every other stochastic network simulator that moves discrete vehicles, only approximate equilibria.

More specifically, Sampers is not designed around all-day travel plans but around tours (travel from home to an activity location and back). The utility of a tour accounts for the travel time of that tour as an attribute but is independent of the travel times of all other tours. In consequence, time pressure cannot be modeled, such that it becomes possible that Sampers predicts travel plans where the sum of travel and sensible activity durations exceeds 24 hours. Neither is the time at which a trip is made accounted for, even though traveling at the onset of the peak hour, during the peak hour or off-peak can lead to very different travel times.

MATSim's scheduling utility function (3) also evaluates time as an attribute (in its first addend), but it additionally values the time spent performing activities (in the sum constituting its second addend) subject to a 24-hours time budget. This specification accounts for time pressure in the sense that increasing travel time leaves less time for activity participation, and adding an activity to a plan requires to take the time for participating in this activity from other activities. Further, the travel time of every single trip is departure-time dependent, accounting for the within-day evolution of congestion and delay.

In consequence, the synthetic travelers in Sampers and MATSim perceive different utilities for the same travel plan; one when choosing their tours, destinations, and modes in Sampers, and another when choosing their routes and departure times in MATSim. This leads to two problems.

1. Sampers selects tours/destinations/modes based on *expected* tour travel time utilities that may be systematically different from the *experienced* scheduling utilities in MATSim. This can lead to systematic errors in the predicted travel behavior.

2. A CBA that compares scenarios based on Sampers utilities cannot take into account the dynamic network reality. This prohibits the consistent analysis of measures that either have a time-dimension (such as time-dependent tolls) or have relevant effects along a time dimension (such as peak hour spreading).

The main challenge when resolving this problem is the fact that Sampers represents the choices of all tours within a travel plan as independent events once the tour sequence is given; hence the product form of (1). This renders it impossible to directly insert MATSim's all-day scheduling utility (3) into Sampers.

3.3 Layout of consistent model system

As explained immediately above, Sampers' tour-based design is structurally incompatible with MATSim's all-day travel plan approach. Sampers is, however, able to simulate all-day travel plans, only that the distribution according to which these plans are sampled is different from what MATSim assumes (sequences of independent vs. dependent tours). The chosen approach is hence to consider Sampers' modeling of tour travel time as an attribute as an inexact approximation of the scheduling utility (3) that is actually experienced by the synthetic travelers in MATSim: Even though the real Sampers model is in terms of time representation less accurate than MATSim, it still accounts for time in some way, it has been estimated and validated from real data, and it has shown to make, within its scope, sensible predictions. It is hence assumed that the Sampers is a suitable tool to create a set of travel *alternatives* for every synthetic person. The remaining problem then becomes to extract, for every simulated individual, its actual travel behavior from the set of travel alternatives predicted by Sampers.

3.3.1 Sampers as a choice set sampler

As explained in Section 2, Sampers has been enriched with micro-simulation capabilities, meaning that it is able to draw (i) a synthetic traveler population and (ii) one or more travel plans for each individual in that population.

Let C_n be the full travel plan choice set of individual n (i.e. the set of all plans that can possibly be generated by running the Sampers micro-simulation). Let $P_n^{\text{Sampers}}(i \mid C_n)$ be the probability that Sampers predicts travel plan i for individual n; this is the same probability as written out (1). For each traveler n, Sampers is run $M \geq 1$ times. The resulting plans are put into a sampled plan choice set $D_n \subseteq C_n$, with duplicates being removed. D_n hence contains between 1 and M elements. For each plan $i \in D_n$, its sampling probability $P_n^{\text{Sampers}}(i \mid C_n)$ and all quantities defining its utility in Sampers are known. The probability that alternative i is contained in the resulting choice set D_n turns out to be

$$\Pr(i \in D_n) \quad = \quad 1 - \left[1 - P_n^{\text{Sampers}}(i \mid C_n)\right]^M . \tag{4}$$

Figure 4 illustrates this function for an example where there are only three possible plans 1, 2, 3. The Sampers plan choice probabilities are $P_n^{\text{Sampers}}(1 \mid \{1,2,3\}) = 0.5$, $P_n^{\text{Sampers}}(2 \mid \{1,2,3\}) = 0.4$, and $P_n^{\text{Sampers}}(3 \mid \{1,2,3\}) = 0.1$. The x-axis of the figure shows the number M, i.e. how many times Sampers has independently created a travel plan, and the y-axis shows the probability of a particular plan being included in the choice set. All curves approach

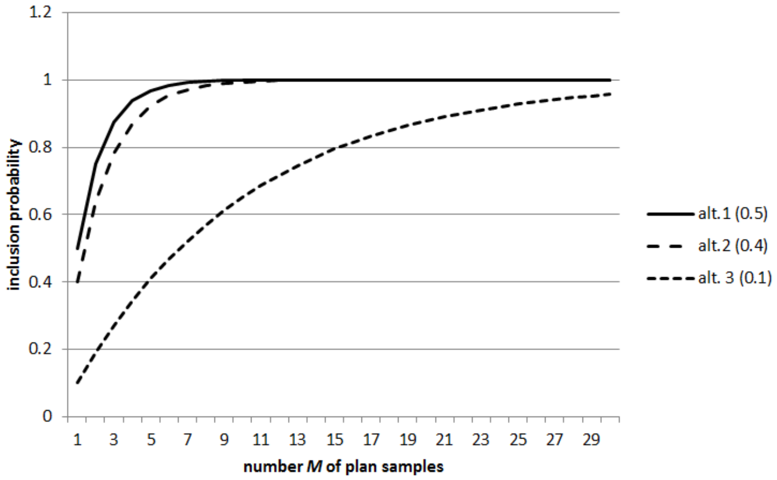

Figure 4: Plan inclusion probabilities $\Pr(i \in D_n)$ over M for three plan alternatives with $P_n^{\text{Sampers}}(1 \mid \{1,2,3\}) = 0.5$, $P_n^{\text{Sampers}}(2 \mid \{1,2,3\}) = 0.4$, and $P_n^{\text{Sampers}}(3 \mid \{1,2,3\}) = 0.1$.

one, meaning that every plan that can be predicted by Sampers will eventually be contained in the choice set, given that one makes M sufficiently large. Further, for any given M, the plan with the highest Sampers choice probability has also the highest probability of being included.

The result of this procedure is, for every traveler n, a choice set D_n of plans that has been generated based on the static Sampers travel plan choice model. One now has to select one element out of this choice set D_n such that this selection resembles the result of executing a version of Sampers that uses the desired utility function (2).

3.3.2 Consistent selection of one travel plan

The objective pursued here here is twofold.

1. Assign to every travel plan in a given traveler's choice set a systematic all-day travel plan utility that takes into account MATSim's scheduling utility (3) instead of Sampers' tour-specific utilities of travel time as an attribute.
2. Select one plan from that choice set according to this MATSim-consistent plan utility, taking additionally into account the random utility structure of Sampers' nested logit choice model.

Sampers' plan choice model (1) represents the choices of all tours within a travel plan as independent events once the tour sequence is given. This is not the same as first defining a random utility for each possible travel plan (comprising a sequence of one or more tours) and then selecting the random utility maximizing plan. To obtain a specification that complies with requirement 1, all travel time independent terms in the systematic Sampers utility of a given travel plan $i \in D_n$ are summarized in a single number $V_{ni}^{\text{dest,mode}}$. For this, the systematic utilities of independently chosen tours, still excluding travel time terms, are summed up.[3] A MATSim-consistent representation of time is then obtained by adding to $V_{ni}^{\text{dest,mode}}$ the scheduling utility V_{ni}^{schedule} assigned by MATSim according to (3) to that all-day travel plan.

Further, the random utility in Sampers comprises an error term ε_{ni}, with the concrete distribution chosen for this term leading to Sampers' nested logit choice model structure. To realize objective 2, one needs to draw one element out of D_n such that the probability of obtaining a certain travel plan i is based on the random utility

$$U_{ni}^{\text{dynamic}} = V_{ni}^{\text{dest,mode}} + V_{ni}^{\text{schedule}} + \varepsilon_{ni}. \tag{5}$$

Denote the resulting probability of choosing plan i from D_n for traveler n by $P_n^*(i \mid D_n)$. Algorithm 1 aims at realizing this choice distribution. Intuitively, Steps 1a and 1b draw one plan from D_n, based on random utilities assigned to all elements in D_n. However, the way in which D_n has been composed also has an effect because

[3] At least for models of multinomial logit form, this is equivalent to multiplying the probabilities of independent choices:

$$P(i \mid C) \cdot P(j \mid D) = \frac{e^{V_i}}{\sum_{l \in C} e^{V_l}} \cdot \frac{e^{V_j}}{\sum_{s \in D} e^{V_s}} = \frac{e^{(V_i+V_j)}}{\sum_{l \in C} \sum_{s \in D} e^{(V_l+V_s)}} = P((i,j) \mid (C,D)).$$

Algorithm 1 Sampers post-processing logic

1. `Repeat until a plan has been chosen:`
   (a) `For each` $i \in D_n$, `simulate a realization of` ε_{ni}.
   (b) `Let` $j \in D_n$ `be the plan that has the largest realization`
   $$U_{ni}^{\texttt{dynamic}} = V_{ni}^{\texttt{dest,mode}} + V_{ni}^{\texttt{schedule}} + \varepsilon_{ni}$$
   (c) `Accept` j `as the chosen plan with a probability` $\propto 1/\Pr(j \in D_n)$, `and continue otherwise.`
2. `Return the first accepted alternative.`

alternatives with a higher Sampers choice probability $P_n^{\text{Sampers}}(i \mid C_n)$ also have a higher probability of being included in the choice set D_n and therewith a higher probability of being available for selection. The subsequent Step 1c corrects this by accepting a proposed plan j only with probability that is anti-proportional to its probability of being included in D_n. Overall, the algorithm implements an exact dynamic version of Sampers for $D_n = C_n$ and an approximation thereof for smaller D_n.

3.3.3 Technicalities

Computation of scheduling utility V^{schedule}

Simulated travelers in MATSim select their departure times and routes by computing an approximately utility-maximizing departure time and routing pattern, using the time-dependent travel times they have observed in the previous MATSim iteration(s). Due to simulated congestion, the scheduling utility realized in the subsequent MATSim iteration may deviate from what the agents have expected. This inconsistency is iteratively resolved, in that MATSim runs through many simulated days (iterations), where in every iteration a fraction of the travelers updates their travel plans. These iterations are continued until expected and experienced scheduling utilities are considered sufficiently similar.

When selecting in Algorithm 1, for every simulated traveler n, one travel plan out of its Sampers-generated choice set D_n, a MATSim-consistent scheduling utility V_{ni}^{schedule} must be available for every plan i in that choice set. This means that V_{ni}^{schedule} has to approximate the expected maximum utility which traveler n would receive when selecting subjectively optimal routes and departure times for the fixed tour/mode/destination pattern of plan i.

Since the travel times that would result from choosing a particular plan $i \in D_n$ is a priori unknown, an iterative approach is adopted that estimates travel times respectively scheduling utilities by iterating between Sampers and MATSim. (This in analogy to iterating between Sampers and Emme in a static model system.) Algorithms 2 and 3 present two complementary ways of realizing this.

- Algorithm 2 iteratively updates time-dependent network travel times. These travel times are then used to approximate the scheduling utilities of all plan alternatives available to all agents. An advantage of this approach is that it simultaneously updates the scheduling utilities of all alternatives, even the non-chosen ones. A disadvantage is that it requires to perform the same computation (of scheduling utilities) in two places: once in MATSim, and once in Sampers. This is a more complicated logic than it needs to be, as will become clear immediately below.
- Algorithm 3 iterative updates the scheduling utilities directly, without making the detour of using time-dependent travel times. A disadvantage of this approach is that it updates only the scheduling utilities of the chosen alternative, requiring at least as many outer Sampers/MATSim iterations as there are plans in a sampled choice set. An advantage is that the scheduling utilities used in Sampers are identical to those computed within MATSim, avoiding their approximation in two different places.

In both algorithms, the plan choice set per agent could as well be updated within every iteration.

Simulation of random utility ε

This requires to draw from the random utility distribution assumed in Sampers' nested logit model. This is possible, yet based on rather technical arguments (Devroye, 2012). (There also exists a multitude of general-purpose

Algorithm 2 Iterative estimation of equilibrium travel times

1. `For all agents` $n = 1, \ldots, N$`: Use Sampers to create a choice set` D_n`.`
2. `Assume free-flow travel times` $\mathscr{T}$`.`
3. `For all agents` $n = 1, \ldots, N$`:`
   (a) `Compute scheduling utilities` $V_{ni}^{\texttt{schedule}}$ `for all` $i \in D_n$ `that would be optimal given` $\mathscr{T}$`.`
   (b) `Use Algorithm 1 to select one tour/mode/destination plan out of` D_n `for execution in MATSim.`
4. `Insert all selected plans into MATSim and iterate towards a route/departure time equilibrium.`
5. `Obtain new time-dependent travel times` $\mathscr{T}$ `from MATSim.`
6. `Stop if travel times` $\mathscr{T}$ `have not changed too much, otherwise continue with Step 3.`

Algorithm 3 Iterative estimation of equilibrium scheduling utilities

1. `For all agents` $n = 1, \ldots, N$`:`
   (a) `Use Sampers to create a choice set` D_n`.`
   (b) `Approximate scheduling utilities` $V_{ni}^{\texttt{schedule}}$ `for all` $i \in D_n$ `by Sampers' utility of time as an attribute.`
2. `For all agents` $n = 1, \ldots, N$`: Use Algorithm 1 to select one tour/mode/destination plan out of` D_n `for execution in MATSim.`
3. `Insert all selected plans into MATSim and iterate towards a route/departure time equilibrium.`
4. `For all agents` $n = 1, \ldots, N$`: Set the scheduling utility` $V_{ni}^{\texttt{schedule}}$ `of the executed plan` $i \in D_n$ `to the actual value simulated in MATSim.`
5. `Stop if scheduling utilities have not changed too much, otherwise continue with Step 2.`

techniques for drawing from general distributions, e.g. Ross (2006)) The subsequently presented experiments assume a non-nested multinomial logit structure of the demand model, which means that ε follows a Gumble distribution, which can be simulated using canned statistical software (here, the Apache Commons math library, http://commons.apache.org/proper/commons-math/).

3.4 Example, continued

The scenario introduced in Section 3.1 is subsequently called the “base case” scenario. An alternative “policy scenario” is also considered, in which the capacity of the road between nodes 2 and 3 is increased from 750 veh/h to 1000 veh/h in both directions.

As explained before, Sampers is replaced in this example by a simpler model, which is sufficient to deliver the desired proof-of-concept at a relatively low experimental complexity. This simpler model will in the following still be referred to as “Sampers”. It uses a non-nested multinomial logit choice model. The mode/destination specific part $V_{ni}^{\text{dest,mode}}$ of this utility is set to all-zeros, as explained in Section 3.1. Indeed, even its utility of travel time as an attribute is set to all-zeros. In combination, this means that Sampers uniformly and congestion-independently selects one out of the eight travel plans enumerated in Section 3.1. This setting is relatively simple to implement and analyze, yet it is challenging for the plan choice correction mechanism because all congestion-sensitivity has to be introduced through that mechanism. All behavioral reactions to congestion are hence a consequence of the correction mechanism, and these reactions are clearly identifiable as deviations from a uniform travel plan choice distribution.

Despite of its conceptual simplicity, the example scenario is rather pathological from a simulation perspective. The traveler population is completely homogeneous (same home location, same choice set, same utility function), the possibilities of moving through the small network are very limited, and only two different tour destinations are available. The fact that travelers do not spread out according to different preferences or over substantially different alternatives gives the system a tendency to oscillate between the few available alternatives, rendering it very difficult to equilibrate. These difficulties have note been encountered in the preliminary results obtained for the (clearly more heterogeneous) Greater Stockholm scenario (Canella et al., 2016).

To obtain expressive results, fifty outer Sampers/MATSim iterations are evaluated. In every outer Sampers/MATSim iteration, the post-processing logic of Algorithm 1 is used. Algorithm 3 is used to iteratively estimate the equilibrium scheduling utilities.

Different sample sizes $M = 1, 5, 10, 100$ are considered when using Algorithm 1 to draw the choice set D_n per agent n. Given that there are in total eight travel alternatives, the expected number of distinct plans per sampled choice set is approximately $1, 4, 6, 8$, respectively.

- Letting $M = 1$ is equivalent to not using Algorithm 1 because the single plan drawn from Sampers is in this setting the only available alternative for the re-sampling logic of Algorithm 1.
- Letting $M = 100$, it is almost certain that D_n contains all available alternatives. This in turn means that using Algorithm 1 is guaranteed to create a travel behavior that is consistent with the desired choice model stated in Section 3.1.
- Letting $M = 5$ resp. 10 means that for most agents, Algorithm 1 can realize the desired choice distribution only over a true subset of the full set of alternatives.

3.4.1 Network conditions and travel behavior

Results for $M = 100$ are considered, meaning that using Algorithm 1 is guaranteed to create a travel behavior that is consistent with the desired choice model stated in Section 3.1.

To provide an intuition for the physical transport processes in the system, Figure 5 shows realizations of the time-dependent road network link flows for the base case (solid) and the policy case (dashed). These curves will be referred to in the subsequent discussion. Overall, one observes a distinct flow increase between nodes 1,2,3 in the policy case, which is a consequence of the increased flow capacity along this path. Another relevant observation is that there is a substantial morning peak flow on link 4_2 is the base case but not in the policy case. This is so because link 2_3 acts as a bottleneck in the base case, which leads to delay on link 1_2, which in turn motivates some travelers to take the bypass route through nodes 1,4,2 to work. The policy case removes this bottleneck, such that the bypass route becomes unused. The effect of the policy measure on the delay upstream of the bottleneck is illustrated in Figure 6, which compares a realization of the travel time on link 1_2 in the base case and the policy case.

Figure 7a show how the population plan choice shares evolve over outer Sampers/MATSim iterations in the base case, Figure 7b shows the corresponding curves for the policy case. The uniform plan selection during the first eight iterations is the result of every plan being initially selected once for execution in order to obtain a reasonable first estimate of its scheduling utility. This variation of Algorithm 3 may not be necessary in a less artificial scenario where Sampers delivers a more realistic utility than the flat zeros used here. In the following outer Sampers/MATSim iterations, plans choices are simulated according to Algorithm 1. The persistent variability of all curves is a consequence of all travelers re-drawing their travel plans in every outer Sampers/MATSim iteration in the present, basic model system implementation. The policy case converges faster than the base case because removing the road bottleneck between nodes 2 and 3 reduces the system's oscillatory behavior.

Figure 7c shows average (over other Sampers/MATSim iterations) values of the stationary plan shares for all alternatives. Increasing road capacity in the policy case increases the share of travelers going to work by car (first four sets of columns) from 38 % to 50 %, with the share of PT users (second four sets of columns) being reduced correspondingly. Among the four winning alternatives, the share of making a single work tour by car increases less than the shares of plans in which a secondary tour is made. This can be explained by the improved network performance in the policy case, which renders the realization of two tours per day more feasible than in the base case.

The observed plan shares are consistent with the changes in systematic plan utilities shown in Figure 8. The utilities for different types of plans are shown here in monetary units (assuming a value of time of 12 EUR per hour and a currency exchange rate of 9 SEK per EUR, yielding 108 SEK per hour). Travel plans where one goes to work by car become by about 20 SEK more attractive in the policy case, whereas plans where one uses PT to work do not experience a change. This can be explained by PT operating at a congestion-independent speed, cf. Section 3.1. The utility of plans that comprise a secondary tour increases for both destinations (by about 13.5 SEK for *other1* activities performed at node 3 and by about 5 SEK for *other2* activities at node 5). The overall increase in utility can largely be explained by the increased utility of the initial work tour. The utility of plans comprising an activity at node 3 increases more than that of plans with an activity at node 5 because the policy measure consists in increasing the capacity of the road connecting node 3 to the rest of the network.

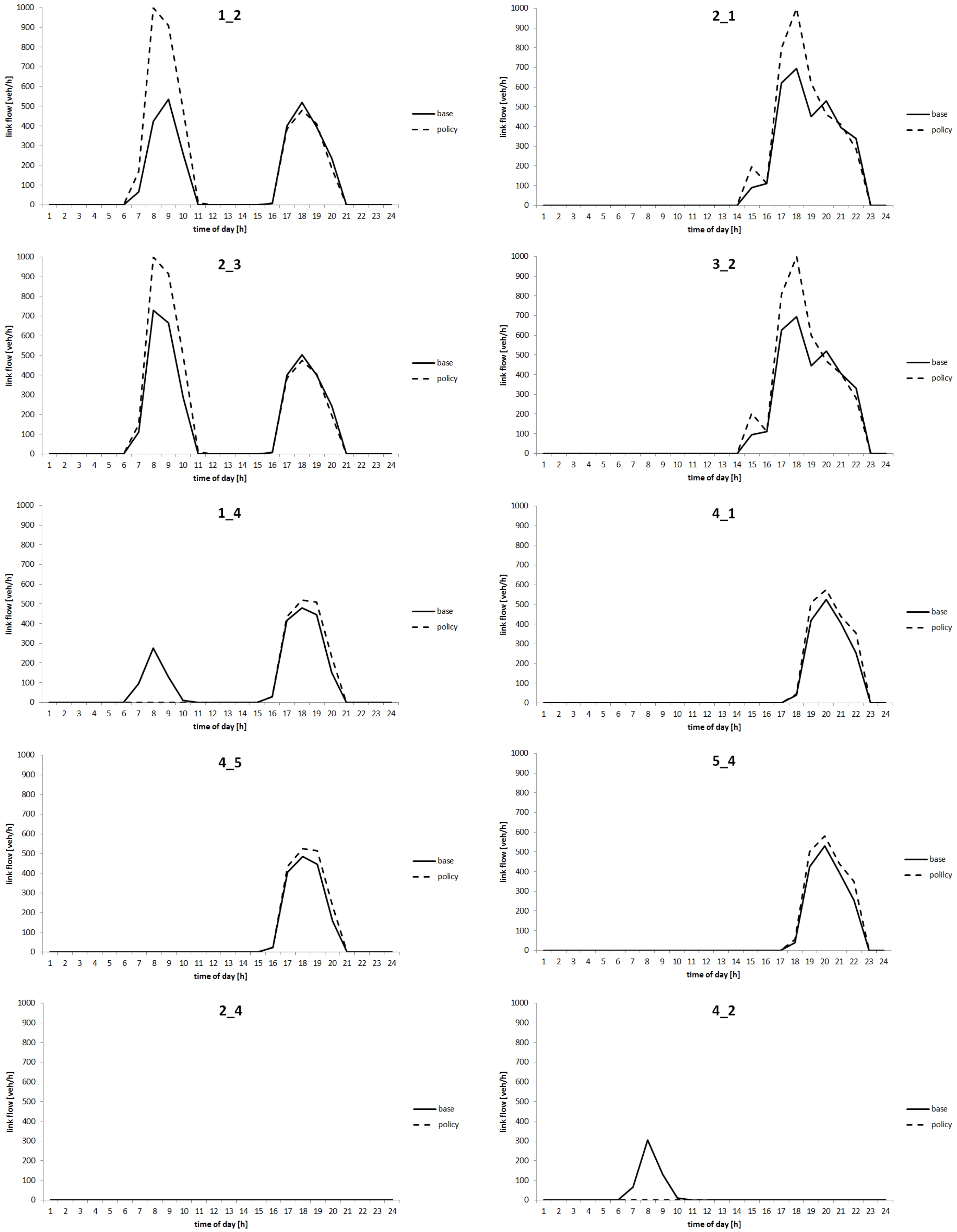

Figure 5: Link flows in base and policy case

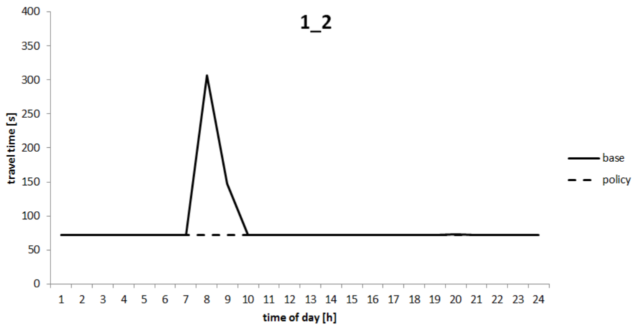

Figure 6: Effect of policy measure on travel time

3.4.2 Expected maximum utilities

The benefit of introducing the policy measure is assessed by computing the expected maximum utility experienced by the traveler population in the policy case minus that experienced in the base case. The expected maximum utility is computed, per agent, by (i) sampling a large number (10'000) of choices according to Algorithm 1 and (ii) averaging the maximum random utility realizations over all choices.

Figure 9 plots the expected maximum utility difference over outer Sampers/MATSim iterations for $M = 1, 5, 10, 100$. The persistent variability of all curves is a consequence of the network simulation stochasticity that could in a real application be averaged out. The predicted benefit of increasing road capacity when computed with $M = 1$ is much larger than the results for $M = 5, 10, 100$. This is plausible because the case $M = 1$ corresponds to a rigid demand in the sense that all travelers have just one plan to choose from, meaning that they cannot react to network congestion by switching to a different plan. The results for $M = 5, 10, 100$ are, apart from fluctuations, visually identical, which indicates that the difference in expected maximum utility is robust with respect to the choice set size. Overall, the benefit of increasing capacity is around 130 kSEK per day; this number could now be further processed in a CBA.

Here, the question may arise how well changes in expected maximum utility can be estimated when the sampled choice set is much smaller than the universal choice set available in Sampers. Two answers can be given.

1. Let $\overline{U}_D$ be the expected maximum utility that results from considering only the choice set D, and let $\overline{U}$ be the true (yet unknown) expected maximum utility that results from considering the full choice set. Only considering D, one hence under-estimates the true expected maximum utility by $\overline{U} - \overline{U}_D$. However, given that performing a CBA only requires to evaluate the *change* in expected maximum utility across scenarios, it is sufficient to ensure that this bias stays constant across scenarios.

   Let $\neg D = C \backslash D$ be the set of all alternatives not in D, and let U_D (respectively $U_{\neg D}$) be the maximum *random* utility over all elements in D (respectively $\neg D$). The maximum *random* utility U of the full model (with choice set $C = D \cup \neg D$) then becomes $U = \max\{U_D, U_{\neg D}\}$, which implies $\overline{U} - \overline{U}_D = \mathrm{E}\{\max\{0, U_{\neg D} - U_D\}\}$. A sufficient condition for the bias $\overline{U} - \overline{U}_D$ to be constant across scenarios is hence that the distribution of the difference in maximum random utility between the choice set D and its complement $\neg D$ does not change across scenarios.

   Practically speaking, this requires to construct D such that a traveler comparing the bundle D to the bundle $\neg D$ does not notice a difference across policy scenarios. The usage of Sampers' static choice probabilities to draw the elements in D, cf. (4), can be instrumented to approximately fulfill this condition by drawing half of D with Sampers' static choice probabilities in the base case and the other half with its choice probabilities in the policy case.[4]

2. If one does not wish to sample alternatives, one can resort to using Algorithm 2 instead of Algorithm 3, meaning that one approximates the systematic utilities even of the non-chosen alternatives by computing best-response time and route choices against expected time-dependent link travel times. Given that these best-response computations also require computing resources, even this approach may reach its limits for very large choice sets; experimentation with the real scenario is necessary to assess how far this approach carries.

Further developments may even succeed to combine the advantages of Algorithm 2 and 3.

[4]This requires to replace (4) by $\Pr(i \in D_n) = 1 - \left[1 - P_n^{\text{Sampers,Base}}(i \mid C_n)\right]^{M/2} \left[1 - P_n^{\text{Sampers,Policy}}(i \mid C_n)\right]^{M/2}$.

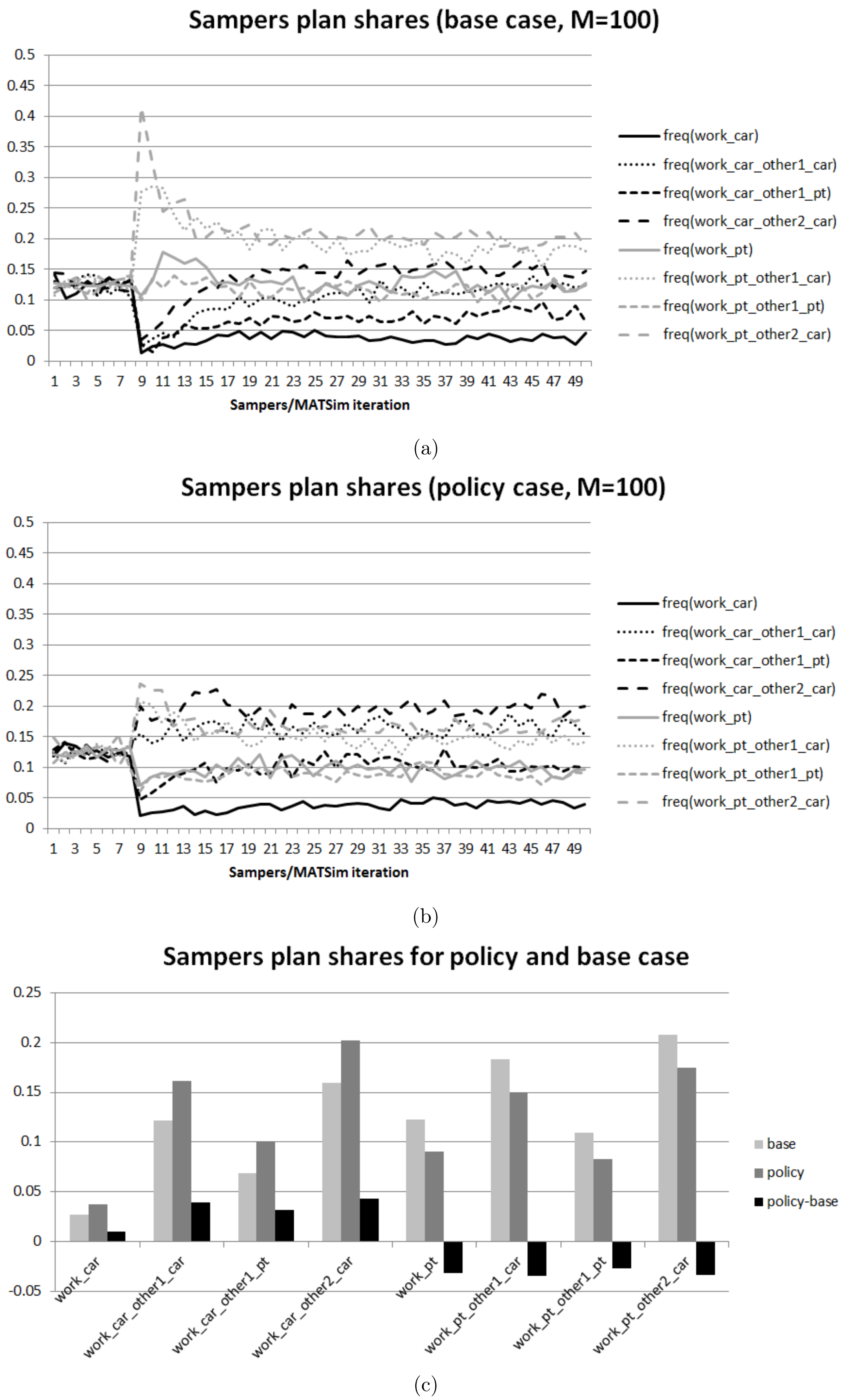

Figure 7: Plan choice shares in different configurations

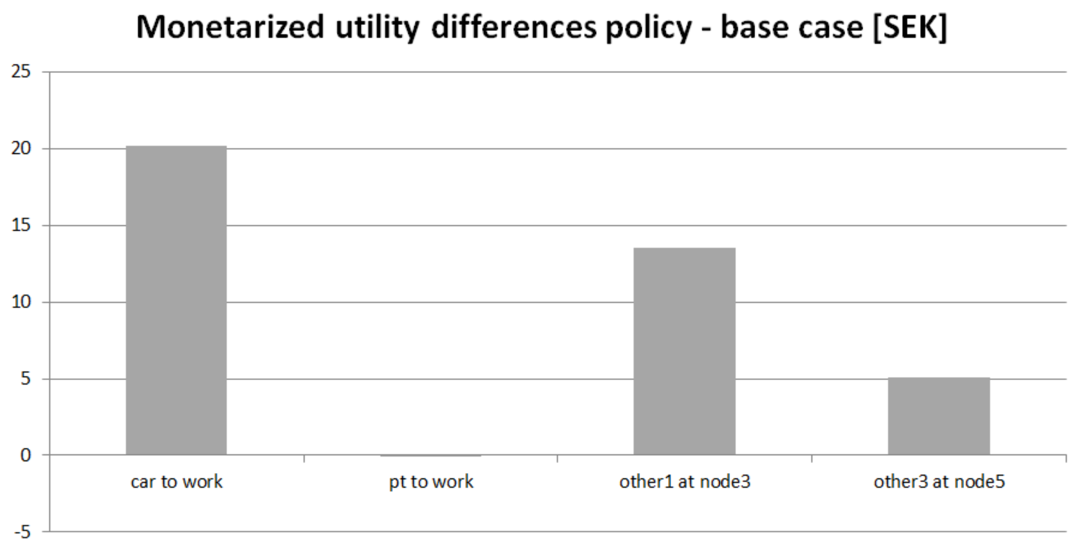

Figure 8: Changes in plan utilities

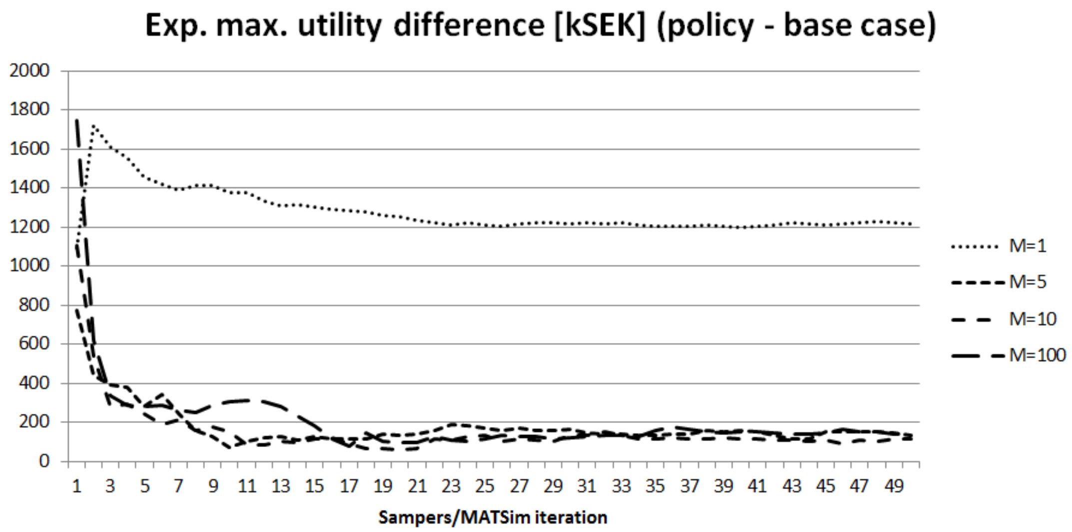

Figure 9: Benefit of policy measure

4 Brief summary

The objective of IHOP3 was to demonstrate that the static and aggregate travel demand model Sampers can be combined with the dynamic and disaggregate network simulation package MATSim in a way that enables credible CBA. As documented in the present report, this objective has been pursued by demonstrating the following two results.

1. Sampers can be run as a person-centric travel demand micro-simulator.
2. Sampers' travel demand predictions can be post-processed such that the resulting travel plan choice distribution resembles the result of replacing Sampers static valuation of time as an attribute by the dynamic scheduling utility used in MATSim.

The first result has been prototypically implemented within the real Sampers system. The second result has been illustrated through a small simulation case study. While the considered scenario is clearly anecdotic, the developed methods are scalable and transferable to systems of metropolitan scale.

References

Canella, O., Flötteröd, G., Johnsson, D., Kristoffersson, I., Larek, P. and Thelin, J. (2016). Flexible coupling of disaggregate travel demand models and network simulation packages (IHOP2), *Technical report*, KTH, Sweco, WSP.

Devroye, L. (2012). A note on the Engelson-Gumbel distribution. School of Computer Science, McGill University.

Flötteröd, G. (2016). Queueing representation of kinematic waves, *in* Horni et al. (2016), chapter 50.

Horni, A., Nagel, K. and Axhausen, K. W. (eds) (2016). *The Multi-Agent Transport Simulation MATSim*, Ubiquity, London.

Kickhöfer, B. and Nagel, K. (2016). Microeconomic interpretation of MATSim for benefit-cost analysis, *in* Horni et al. (2016), chapter 51.

Ross, S. (2006). *Simulation*, fourth edn, Elsevier.