Optimal Road Pricing:

Towards an Agent-based Marginal Social Cost Approach

Ihab Kaddoura*, Benjamin Kickhöfer

Technische Universität Berlin, Germany

* Corresponding author (e-mail: kaddoura@vsp.tu-berlin.de)

January 2, 2014

Abstract

In this paper, an new approach is developed to calculate optimal time-dependent user-specific congestion tolls in an agent-based simulation framework. First, the ideas behind this agent-based marginal social cost pricing approach are presented by means of illustrative examples. An agent database keeps track of the agents' interferences on each road segment of the network. It calculates each agent's contribution to two distinct sources of delays: (i) the capacity utilization in the link's point queue, and (ii) the capacity utilization of the link's road space. The resulting delay effects are converted into monetary terms and charged from the causing agent. Thus, the external congestion costs are internalized since they now influence the agent's evaluation of different travel alternatives. To the best knowledge of the authors, this approach is unique in calculating and pricing dynamic congestion effects among car users on a microscopic, truly agent-based level. The implementation is then applied to the well-known Sioux Falls scenario. The simulation experiments show that the user-specific marginal social cost pricing approach results in higher social welfare than the reference scenario without road pricing. Additionally, the average external congestion costs are observed to be higher during peak periods and on urban road segments. The approach proofs therefore to be applicable in large-scale real-world scenarios, potentially with heterogeneous users, and to be combined within the same simulation framework with optimal public transport pricing or optimal exhaust emission pricing.

Keywords: road pricing, optimal toll, user-specific road charges, agent-based simulation, congestion, external effects, internalization, queue model

1 Introduction

The idea to internalize marginal external costs which are equal to the difference between marginal social costs and generalized user prices goes back to Pigou (1920). The author claimed that it is possible to change people's behavior towards a more efficient use of limited road capacities by imposing a toll that is equal to the marginal external costs. The assumption is that the toll would make users to take into account their contribution to congestion costs, which emerge from additional travel times that they impose on other users by the use of road infrastructure.

Current estimates indicate that congestion causes the largest part of the transport related externalities (see, e.g., Maibach et al., 2008, p.103). Similar to that study, Small and Verhoef (2007) find for car travel in the US, that marginal congestion costs cause roughly 65% of the total variable external costs. For peak hours, other studies come to similar findings (see, e.g., de Borger et al., 1996; Parry and Small, 2009). For this reason, most studies in the literature focused on congestion pricing.

For car travel, the most prominent example for finding optimal toll levels is the bottleneck scenario, originally introduced by Vickrey (1969) for homogeneous users. The idea is that a queue and the resulting time losses at a bottleneck can be eliminated by introducing a time-dependent toll: this optimal pricing scheme induces a flow of vehicles that is equal to the capacity of the bottleneck. Since this scenario can be solved analytically, it has been studied widely in the transport economic literature. Several extensions were applied: Vickrey (1973) introduced proportional heterogeneity in the Values of Travel Time Savings (VTTS) and Values of Schedule Delay (VSD). Arnott et al. (1993) incorporated price-sensitive demand into the model, and van den Berg (2011) combines continuous heterogeneity and price-sensitive demand.

For public transport travel, Mohring (1972) introduced the concept of total cost minimization (sum of user and operator costs) in a single bus route example. Many researchers extended this basic model by considering congestion costs, namely crowding (Oldfield and Bly, 1988; Kraus, 1991; Jara-Díaz and Gschwender, 2003), and bus congestion (Tirachini and Hensher, 2011).

By deriving general principles that are valid for different scenarios, all of the above studies laid the foundations for a better understanding of optimal tolling. They are well suited to understand the economic principles behind price setting. However, due to their simplified nature, these models are less appropriate to handle large-scale scenarios, non-deterministic demand, or changes in capacities.

For large-scale agent-based models, there have been attempts to approximate the optimal toll: Nagel et al. (2008) charge travelers on each road segment with respect to the time spent traveling. The authors find that this toll results in time gains and, thus, in an increase in social welfare. For an evacuation scenario, Lämmel and Flötteröd (2009) compute an approximation to the system optimum by imposing marginal social costs on each traveler. The authors develop a formula which approximates these marginal social costs for stationary flow conditions. Also with this model, the total travel time is reduced compared to a Nash-equilibrium approach. However, in an agent-based model, there is no need to calculate marginal social costs based on aggregated flows. As Kaddoura et al. (2013) show for a multi-modal corridor, marginal social costs can be calculated on an agent-by-agent basis, i.e. maintaining the truly agent-based perspective. One advantage of the latter approach is that heterogeneity of travelers can directly be considered, which is not true for the calculation via aggregated flows. In the same mindset, Kickhöfer and Nagel (2013) prove that it is possible to calculate first-best, agent-specific, time-dependent air pollution tolls within the same framework.

Therefore, the present study aims at calculating optimal road tolls by investigating the marginal congestion cost imposed by users on other users at a microscopic, truly agent-based level. The approach is tested in a framework that accounts for dynamic congestion and possibly allows to be applied in scenarios with complex networks and heterogeneous demand. In that sense, the present study can be seen as a step towards an agent-based integration of optimal congestion pricing (this study), optimal public transport pricing (Kaddoura et al., 2013), and optimal air pollution pricing (Kickhöfer and Nagel, 2013).

The paper is structured as follows: Sec. 2 describes the simulation framework which is used for marginal social cost approach presented in Sec. 3. The newly developed pricing approach is then applied to a simple test scenario (Sec. 4) and a more complex test scenario which is based on the city of Sioux Falls (Sec. 5). Finally, Sec. 6 summarizes the findings and provides an outlook on future studies.

2 Methodology

In this study, the agent-based microsimulation MATSim[1] is used for the investigation of optimal road pricing. Sec. 2.1 gives a general overview of MATSim, Sec. 2.2 describes the underlying traffic flow simulation which is of substantial significance for this paper. For further information on the simulation framework MATSim, see Raney and Nagel (2006).

2.1 MATSim Overview

In MATSim, the transport demand is modeled as individual agents who have a mental and physical behavior. At first, so-called 'daily plans' are generated for each agent. A plan describes activity types, locations, departure times and transport modes for the trips placed in between the activities. The adaptation of demand to supply follows an iterative approach that involves the ensuing three steps:

1. **Traffic flow simulation:** All agents execute their selected plans in the physical layer. For a detailed description of the traffic flow simulation, including the interactions among agents, see Sec. 2.2.

2. **Evaluating plans:** The agents evaluate their executed plans considering both the activities and the trips.

3. **Learning:** A certain share of agents generates new plans for the next iteration. An existing plan is copied and then mutated. The remaining agents select a plan for the next iteration by choosing among their existing plans according to a multinomial logit model.

Repeating these steps over many iterations couples the mental and physical behavior. This enables the simulation outcome to stabilize, the agents improve and generate plausible travel alternatives. Assuming the travel alternatives to represent valid choice sets, the system state converges to a stochastic user equilibrium (Nagel and Flötteröd, 2012).

2.2 Traffic Flow Simulation: Queue Model

As all agents simultaneously execute their selected plans, they interact in the physical environment. The traffic flow simulation is based on a queue model developed by Gawron

[1] Multi-Agent Transport Simulation, see www.matsim.org

(1998). In MATSim, a road segment (link) is the smallest spatial entity. Each link is modeled as a *First In First Out* queue that has three attributes: free speed travel time t_{free}, flow capacity c_{flow}, and storage capacity $c_{storage}$. Every vehicle has to spend at least the free speed travel time on a link before leaving it. The flow capacity may cause congestion by defining the maximum number of vehicles that can leave a link within a given time span. This is in the literature often referred to as 'bottleneck congestion' (see, e.g., van den Berg, 2011). The storage capacity restricts the maximum number of vehicles on a link and may cause spill-back effects on upstream links. For each time step (typically 1 second), the state of every link's queue is updated. A vehicle is moved from link l_1 to link l_2 if (i) the free speed travel time has passed, that is, the vehicle has arrived at the end of the link, (ii) the inverse of the flow capacity has passed since the last vehicle left link l_1 and (iii) the storage capacity on link l_2 is not reached.

In this study, an implementation of the traffic flow simulation called *qsim* is used. This is, at the time of writing, the default setting in MATSim. The simulated traffic flow which is obtained from the queue model is consistent with the flow dynamics described by the macroscopic fundamental diagram (see Agarwal et al. (2013) who use the *qsim*; and see Zheng et al. (2011) for a modification of the *qsim*).

3 Agent-based Marginal Social Cost Pricing

Of all existing external effects, this study focuses on congestion effects among users within the car mode. External delay effects are computed based on the simulated traffic flow. An agent who leaves a link prevents all immediately following agents from leaving that link for the time of $\frac{1}{c_{flow}}$. A following agent that arrives at the end of the same link during that time interval will be queued at the end of the link. In order to trace back the delays to their origin, all links keep track of the agents' movements. Each time an agent leaves a link, the agent ID and the time are saved as a temporary status of the congestion that results from the flow capcacity constraint. Once the queue on a link dissolves, e.g. if for a time interval of $\frac{1}{c_{flow}}$ no agent is moved to the next link, that information is deleted and the tracking of delays starts again.

For each agent that moves from one link to the next, a total delay d_{total} is calculated as the difference of the actual travel time (t_{act}) and the free speed travel time (t_{free}). If an agent moves from one road segment to the next with $d_{total} > 0$ this agent is referred to as the affected agent. At the same time the affected agent can also be a causing agent imposing delays on other agents. As described in Sec. 2.2, the applied queue model accounts for two sources of delay effects among agents: the flow capacity and the storage capacity. Thus, the total delay either results from one of these constraints or a combination of both:

$$d_{total} = d_{flow} + d_{storage}, \tag{1}$$

where d_{flow} is the delay caused by the flow capacity constraint; and $d_{storage}$ is the delay resulting from the storage capacity constraint on the downstream link. d_{flow} and $d_{storage}$ may be caused by different agents.

3.1 Internalization of Congestion Effects Without Spill-back ($d_{storage} = 0$)

Each time step an affected agent moves from one road segment to the next one, the temporary status of the congestion resulting from the flow constraint is called, e.g. the agents that have left the link from the moment the congestion has dissolved for the last time. In case the congestion has already disappeared on that link, d_{flow} is zero and d_{total} results from spill-back effects only. Otherwise, if the queue caused by the flow constraint has not yet dissolved, beginning with the last agent, this approach iterates through all agents that have left the link since the congestion has last dissolved.

In case $d_{total} \leq \frac{1}{c_{flow}}$, the last agent previously leaving the link is the only causing agent imposing an effect of d_{total} on the affected agent. In case $d_{total} > \frac{1}{c_{flow}}$, there are several causing agents imposing a delay on the affected agent. Each agent is considered to cause a delay of $\frac{1}{c_{flow}}$ and is therefore charged with the equivalent monetary amount. The latter is calculated by multiplying the delay time by the (possibly individual) Value of Travel Time Savings (VTTS) of the affected agent. Each time a portion of the total delay effect is internalized, $\frac{1}{c_{flow}}$ is deducted from the total delay. Once the remaining delay is zero, all causing agents are identified and the total delay effect is internalized.

A simple example is given in Fig. 1: three agents are considered to move from link l_a to link l_b. All agents enter l_a at 1 sec intervals: the first agent a_1 at time step $t = 0$, the second agent a_2 at time step $t = 1$ and the third agent a_3 at time step $t = 2$. l_a is set as follows: $t_{free} = 10$ *sec*; $c_{flow} = 1200$ veh/h ($\frac{1}{c_{flow}} = 3$ *sec*; one agent every 3 sec). The first agent leaves link l_a after 10 sec (at time step $t = 10$) and blocks the link for 3 time

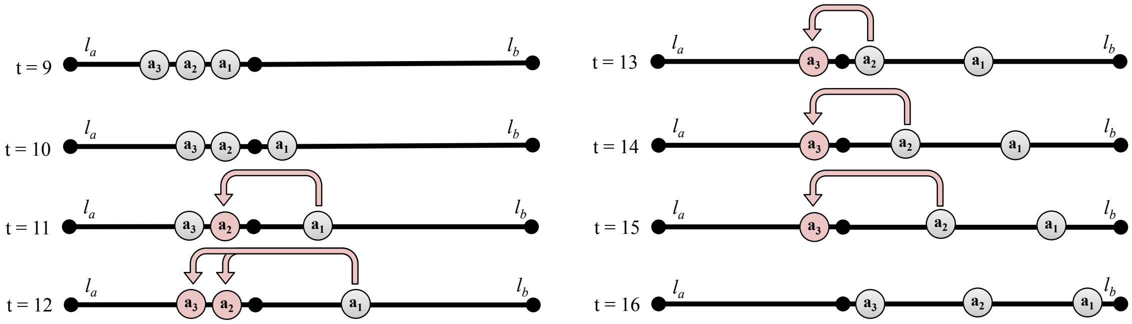


Figure 1: Three agents moving from l_a to l_b; a red arrow indicates an external delay effect.

steps (including the current time step). That is, the earliest link leave time for the next agent on that link is $t = 13$. The second agent arrives at the end of l_0 at time step $t = 11$ and is therefore delayed for 2 sec ($t = 11$ and $t = 12$). As $d_{total} \leq \frac{1}{c_{flow}}$ is true, the last agent that has previously left the link is considered as the causing agent. Hence, a_1 has to pay for the delay effect of 2 sec imposed on a_2. a_2 leaves l_a at time step $t = 13$. Again, the link is blocked for 3 time steps including the current time step. That is, the earliest link leave time for the next agent is $t = 16$. a_3 arrives at the end of l_a at $t = 12$, is then queued for 4 time steps and leaves the link at $t = 16$. As $d_{total} > \frac{1}{c_{flow}}$ is true, there are several agents causing the delay. Beginning with the last agent, the delay is allocated on the agents that left l_a before. a_2 is considered to cause the delay of $\frac{1}{c_{flow}} = 3\ sec$ and is therefore charged the equivalent monetary amount. Deducting the internalized delay from the total delay that a_3 was delayed, 1 sec remains. The next agent that previously left l_a is a_1. Thus, a_1 is charged again, this time for a delay effect of 1 sec that is imposed on a_3.

3.2 Internalization of Congesiton Effects With Spill-back ($d_{storage} \neq 0$)

In case an agent is delayed when leaving a link even though the queue resulting from the flow capacity on this link has dissolved, the delay can be ascribed to the storage capacity of a downstream link only ($d_{total} = d_{storage}$). If an agent enters a link and thereby occupies the last space, the link is blocked. Other agents on upstream links are prevented from entering that link. Once an agent leaves the congested link and the number of agents drops below the storage capacity, the next agent can follow. Four interpretations of this delay effect come to mind:

1. The agents at the queue's origin, e.g. the bottleneck link, are responsible for the

spill-back and charged with the delay.

2. The agent who is blocking the last space on the downstream link is responsible and charged with the delay.

3. All agents in front are equally responsible and *share* the resulting delay. The sum of tolls is equal to the monetized delay effect.

4. All agents in front are equally responsible and each agent pays the full sum of the resulting delay. Following this interpretation, the sum of tolls may exceed the monetized delay effect.

In this study, the agents that induce delays at the bottleneck link are considered as the causing agents for spill-back related delays that are imposed on agents on upstream links (interpretation 1).

To give an example, Fig. 2 depicts three agents that move along link l_a, l_b and l_c. l_b is the bottleneck link with $\frac{1}{c_{flow}} = 5\ sec$; $\frac{1}{c_{flow}}$ on l_a and l_c is 1 *sec*. $c_{storage}$ on l_b is set to 1 vehicle; $c_{storage}$ on l_a and l_c is 2 vehicles. t_{free} on l_b is 1 sec; t_{free} on l_a and l_c is set to 2 sec. The agent a_1 enters l_a at $t = 1$, a_2 at $t = 2$, and a_3 at $t = 3$. At $t = 4$ agent a_1 leaves l_b and blocks the link for 5 time steps (including the current time step). That is, the earliest leave time for the next agent on l_b is $t = 9$. At $t = 5$ agent a_2 reaches the end of l_b but is not allowed to leave l_b because of the flow capacity of l_b. At the same time step a_3 arrives at the end of l_a and is delayed by the storage capacity of l_b. At $t = 9$ agent a_2 is moved from l_b to l_c with a delay of $d_{total} = d_{flow} = 4\ sec$. As described in Sec. 3.1 agent a_1 is the causing agent and has to pay for this delay. When a_2 leaves l_b the storage capacity is released. Thus, in the same time step a_3 is able to move from l_a to l_b with a delay of $d_{total} = 4\ sec$. As the queue resulting from the flow capacity on l_a has already dissolved, the delay results from $c_{storage}$ only ($d_{total} = d_{storage}$). According to interpretation 2, a_2 is the causing agent (a_2 was blocking the next link). This interpretation may be implemented by assuming the last agent that has left the upstream link to be the causing agent that occupies the last space on the downstream link.[2] However, according to interpretation 1, a_1 is considered as the causing agent imposing a delay on a_3 (a_1 caused the spill-back at the bottleneck link). In this study, this is implemented by saving the delays that are not internalized and passing them on from link to link. Once the bottleneck link is reached

[2] For intersections and agents that turn left or right, this implementation may lead to the wrong agent.

all causing agents are identified and the total delay is internalized.[3] In this example, a_3 leaves l_b at $t = 14$ with $d_{total} = d_{flow} = 4$ *sec*. This delay results from the flow capacity on l_b and a_2 is the causing agent. To internalize the delays resulting from spill-back effects, d_{total} is increased by the remaining (and not internalized) delay from the previous link. Hence, the total delay which is used for the internalization procedure described in Sec. 3.1 amounts to 8 sec and now includes d_{total} from the current link (4 sec) plus $d_{storage}$ from the previous link (4 sec). As the queue on l_b has not yet dissolved, both, a_1 and a_2 are identified as the causing agents that have to pay for delay imposed on a_3.

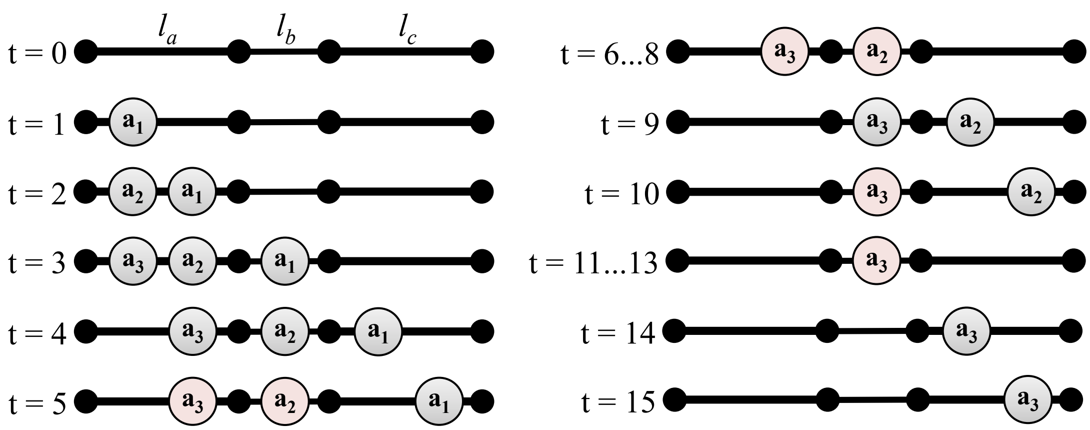


Figure 2: Three agents moving along l_a, l_b and l_c; bottleneck: l_b ($\frac{1}{c_{flow}} = 5$ *sec*)

More complex is the case in which the queue that results from the flow capacity has not yet dissolved and d_{total} results from both capacity constraints. If the remaining delay is not zero after iterating through all agents previously leaving the link, the remaining delay is considered as $d_{storage}$, e.g. the affected agent is delayed resulting from spill-back effects caused by storage capacity constraints on downstream links.

3.3 Tolls as Incentives

The user-specific toll calculated by the presented approach is the equivalent monetary amount of the total delay effects that an agent imposes on other agents. By charging these tolls, the delay effects are internalized and considered in the utility of an agent's plan. This, in turn, influences the choice probabilities of the underlying logit model or, in

[3] For spill-back effects along several links, in the current implementation, $d_{storage}$ may not be completely internalized if the affected agent does not pass by the bottleneck link.

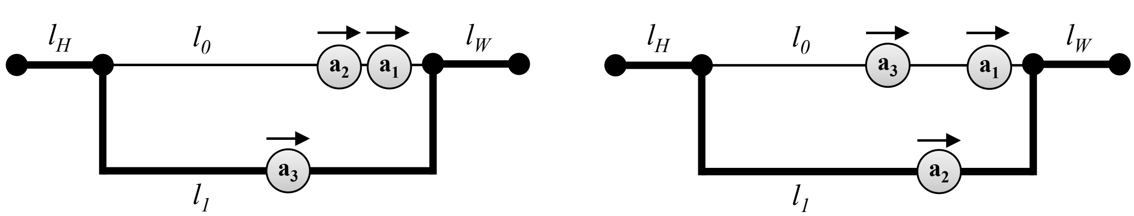
other words, influences the decision making process of agents.

At the end of each iteration, average travel times and average toll levels are calculated for every link and every time bin. In this study, a time bin always covers 15 min. Based on this information about travel times and toll levels, the router proposes possible routes in the next iteration.

Combined with the iterative approach, the above two steps enable agents to adjust their behavior in order to maximize their utility. As each individual's utility now also includes congestion costs, the system state converges towards an approximation of the social welfare maximum (or approximate system optimum).

4 Test Scenario

The approach described in Sec. 3 is verified using a simple test scenario. As depicted in Fig 4 the network consists of two alternative routes from link l_H to link l_W, either along link l_0 or along link l_1. Link l_0 is set as follows: $t_{free} = 10$ *sec*; $c_{flow} = 1200$ vehicles per hour ($\frac{1}{c_{flow}} = 3$ *sec*; one vehicle every 3 sec). Link l_1 is considered not to be congested (c_{flow} is very large), but t_{free} is set to 13 sec. For both l_H and l_W, t_{free} is zero and c_{flow} is sufficient high not to cause congestion. Furthermore, for all four links $c_{storage}$ is very large, so there are no spill-back effects. On the demand side, three agents are considered to travel from l_H to link l_W. They start at 1 sec intervals: the first agent a_1 at time step 0, the second agent a_2 at time step 1 and the third agent a_3 at time step 2. The agents are only allowed to choose between the two possible routes, either along l_0 or l_1. The scenario is used for two simulation experiments: No pricing (Fig. 3a) and the marginal social cost pricing approach (Fig. 3b).



(a) Experiment 1: No pricing

(b) Experiment 2: Marginal social cost pricing

Figure 3: Test scenario: Different routing approaches

4.1 Simulation Experiment 1: No Pricing

In this simulation experiment, we allow the agents to improve over several iterations only taking into account their own travel time (these cost components are called marginal generalized private costs). Hence, a user equilibrium is obtained: The first and the second agent use link l_0. The first agent leaves the link without any delay at time step 10. The second agent arrives at the end of link l_0 1 sec later (at time step 11) and is therefore delayed for 2 sec. Nevertheless, that agent is still 1 sec faster than on link l_1. After agent a_2 leaves link l_0 at time step 13, again the link is blocked for 3 time steps (including the current time step). That is, the earliest link leave time for the next agent on link l_0 is time step 16. That is, the third agent is better off taking link l_1. Instead of being queued for 4 sec on link l_0, that agent prefers the additional 3 sec free speed travel time on link l_1. The total delay effect amounts to 2 sec and the total travel time of all agents amounts to 35 sec.

4.2 Simulation Experiment 2: Marginal Social Cost Pricing

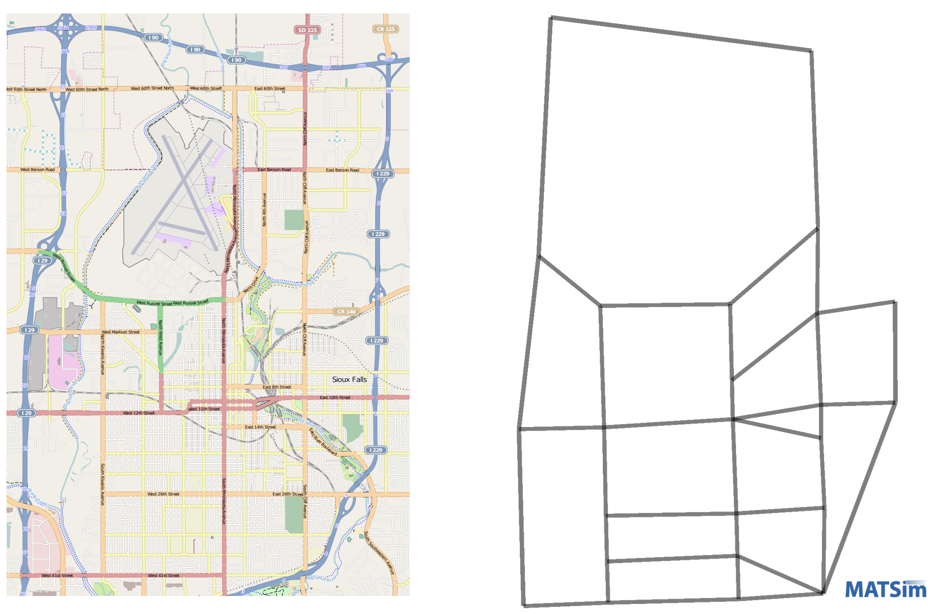
In this simulation experiment, the marginal social cost pricing approach described in Sec. 3 is applied. Agents improve over several iterations, taking into account their own travel time *and* the external delay effects they impose on other agents (together, these cost components are called marginal generalized social costs). Hence, the system optimum in terms of average travel time is obtained: The first agent is still using link l_0. However, the second agent moves along link l_1 and the third agent uses link l_0. Because the time interval between agent a_1 and a_3 when leaving link l_0 is larger, a_3 is only 1 sec delayed by a_1. The total delay effect amounts to 1 sec only and the total travel time of all agents amounts to 34 sec, which is 1 sec less compared to the first simulation experiment without any pricing.

5 Sioux Falls Scenario

In this section, the marginal social cost pricing approach described in Sec. 3 is applied to a more sophisticated test scenario which is based on the city of Sioux Falls, South Dakota, United States. The scenario which is used in this study was generated by Chakirov and Fourie (forthcoming) based on a simplified road network of Sioux Falls which was introduced by LeBlanc et al. (1975) and from then on widely used and modified for numerical illustrations in the field of transport planning.

5.1 Setup

Supply Fig. 4 depicts the Sioux Falls map section and the simplified road network which is used for the simulation. The spatial network configuration is taken from LeBlanc et al. (1975), whereas the physical link parameters that are essential for the dynamic traffic assignment are set according to the real-world road types, assuming 2 lanes, a free speed of 50 km/h, and a flow capacity of 800–1000 cars per hour and lane for urban roads; and 3 lanes, a free speed of 90 km/h, and a flow capacity of 1700–1900 cars per hour and lane for highways (Chakirov and Fourie, forthcoming). In order to focus on external delay effects within the car mode, in this study, the public transport mode is emulated. Ignoring the transit schedule and vehicle capacities, public transport users are teleported between their activity locations assuming a free speed travel time twice as much as for the car mode. For the walk mode we assume the travel time to result from the beeline distance between two activity locations times a factor of 1.3 and a free speed of 3 km/h.



(a) Map section (openstreetmap.org)

(b) MATSim network

Figure 4: Sioux Falls road network

Initial Demand and Scenario Calibration The initial demand was generated by Chakirov and Fourie (forthcoming) using real-world census data, information on residential buildings in the study area and LeBlanc et al. (1975)'s demand matrix. The provided scenario accounts for two activity patterns. 56,904 agents have the activity pattern "Home-Work-Home" and form the peak demand. To account for the non-commuters, in this study, 50,508 agents are considered to have the travel pattern "Home-Secondary-Home". The mode of transportation for the initial daily plans is car. In this study, external effects within the public transport mode are ignored, we only account for congestion effects within the car mode. Therefore, the scenario is calibrated in order to obtain a number of car users which yields a plausible level of road congestion. The simulation is run for a total of 1000 iterations. During the first 700 iterations, choice sets are generated by allowing the agents to change their routes, their departure times, and their mode of transportation. Every iteration, plans are modified by each choice dimension with a probability of 10%; each agent's choice set is limited to 6 daily plans. For the last 300 iterations, the agents only choose from their acquired choice sets according to a multinomial logit model. The resulting modal share is 60.38% car, 30.21% public transport, and 9.41% walking. The output plans of the final iteration are used as input demand for the simulation experiments in Sec. 5.2

Utility Functions Each executed daily plan is evaluated taking into account both, the activity and trip related (dis-)utility:

$$V_{plan} = \sum_{i=1}^{n} \Big(V_{act,i} + V_{trip,i} \Big) \ , \tag{2}$$

where V_{plan} is the utility of an executed plan; n is the total number of activities or trips; $V_{act,i}$ is the utility for performing an activity i; and $V_{trip,i}$ is the utility of the trip to activity i. The first and the last activity are handled as one activity, so the number of activities and trips is the same. The trip related utility is calculated as follows, depending on the chosen mode of transportation.

$$\begin{aligned} V_{car,i,j} &= \beta_{0,car} + \beta_{tr,car} \cdot t_{i,tr,car} + \beta_c \cdot c_{i,car} \\ V_{pt,i,j} &= \beta_{0,pt} + \beta_{tr,pt} \cdot t_{i,tr,pt} \\ V_{walk,i,j} &= \beta_{tr,walk} \cdot t_{i,tr,walk} \ , \end{aligned} \tag{3}$$

where V is the utility for person j on his/her trip to activity i. Attributes for car trips are indicated by *car*, for public transport trips by *pt*, and for walk trips *walk*. β_0 is the alternative specific constant; $t_{i,tr}$ is the travel time; β_{tr} is the marginal utility of traveling; c_i are all monetary costs payed during a trip; and β_c is the marginal utility of monetary costs. The positive utility gained by performing an activity follows an approach introduced by Charypar and Nagel (2005):

$$V_{act,i}(t_{act,i}) = \beta_{act} \cdot t_{*,i} \cdot \ln\left(\frac{t_{act,i}}{t_{0,i}}\right), \tag{4}$$

where t_{act} is the duration of performing an activity; t_* is an activity's "typical" duration, and β_{act} is the marginal utility of performing an activity at its typical duration; $t_{0,i}$ is a scaling parameter linked to an activity's priority and minimum duration. In this study, activities cannot be dropped from daily plans, thus $t_{0,i}$ is not relevant. Being at an activity location before or after the activities' opening time is penalized by the opportunity costs of time $-\beta_{act}$ (see Tab. 1).

Table 1: Activity attributes

Activity	Typical Duration	Opening Time	Closing Time
Home	12 h	undefined	undefined
Work	8 h	6 a.m.	8 p.m.
Secondary	1 h	8 a.m.	11 p.m.

Parameters Behavioral parameters are based on estimations by Tirachini et al. (2012) for Sydney, depicted in Tab. 2. $\hat{\beta}_{tr,car}$ is the marginal utility of traveling by car, $\hat{\beta}_{v,pt}$ is the marginal utility of the in-vehicle time when traveling by public transport, $\hat{\beta}_{a,pt}$ is the marginal utility of the access time, and $\hat{\beta}_c$ is the marginal utility of monetary costs. Tab. 3 shows the adjusted parameters that are used for the simulation experiments and match the applied activity-based approach. As described in previous studies (for example Kickhöfer et al. (2011, 2013); Kaddoura et al. (2013)) the time related parameters are divided into the opportunity costs of time and the disutility of traveling. Furthermore, for traveling by public transport, the behavioral parameter of the in-vehicle time is used, and for walking as travel alternative, we use the parameter for the access time. The

Value of Travel Time Savings (VTTS) for the car mode and for walking is equal 15.48 \$/h (Australian Dollar [AUD], \$1.00 = EUR 0.65 [Dec 2013]), for the public transport mode the VTTS amounts to 18.39 \$/h. In addition to the possible road charges, car users

Table 2: Estimated parameters from Tirachini et al. (2012)

$\hat{\beta}_{act}$	n.a.	
$\hat{\beta}_{tr,car}$	-0.96	[utils/h]
$\hat{\beta}_{v,pt}$	-1.14	[utils/h]
$\hat{\beta}_{a,pt}$	-0.96	[utils/h]
$\hat{\beta}_{c}$	-0.062	[utils/\$]

Table 3: Adjusted parameters used in the present study

$\beta_{act} = -\hat{\beta}_{tr,car}$	$+0.96$	[utils/h]
$\beta_{tr,car} = \hat{\beta}_{tr,car} - \beta_{act}$	0	[utils/h]
$\beta_{tr,pt} = \hat{\beta}_{v,pt} - \beta_{act}$	-0.18	[utils/h]
$\beta_{walk} = \hat{\beta}_{a,pt} - \beta_{act}$	0	[utils/h]
$\beta_{c} = \hat{\beta}_{c}$	-0.062	[utils/\$]

pay monetary cost depending on the distance traveled, the cost rate ist set to 0.40 \$/km. As agents can enter and leave their cars directly at the activity locations and the applied simulation approach does not account for parking and walking times of car users, $\beta_{0,car}$ is set to -0.16, which is equivalent to 10 min walking. For public transport trips access, egress and waiting times as well as fares are not accounted for. To compensate for that, $\beta_{0,pt}$ was re-calibrated yielding an alternative specific constant for public transport trips of -0.2.

Social Welfare The user benefits are calculated as the Expected Maximum Utility (EMU):

$$EMU = \sum_{j=1}^{J} \left(\frac{1}{|\beta_c|} \ln \sum_{p=1}^{P} e^{V_{plan}} \right) \quad , \tag{5}$$

where j is an individual agent; J is the total number of agents; p is a daily plan; and P is the total number of plans in the agent's choice set. The social welfare is defined as the sum of the user benefits and the toll revenues.

5.2 Simulation Experiments

In a first experiment, we run the simulation **without road pricing**. In a second experiment tolls are set according to the **congestion pricing** approach described in Sec. 3. The simulation experiments are compared in terms of social welfare. In each experiment, the simulation is run for another 1000 iterations. For the first 700 iterations, all agents are allowed to change the mode of transportation, the route (only for car) and the departure time. Each choice dimension modifies a plan with a probability of 10% per iteration. For the last 300 iterations the agents only choose among their acquired choice sets according to a multinomial logit model. We therefore calculate the toll revenues, travel times and trips per mode as the average over the last 300 iterations.

5.3 Results

Social Welfare Tab. 4 depicts the changes in social welfare, including the change in user benefits and toll revenues when setting the tolls according to the approach described ind Sec. 3. In the simulation experiment with user-specific road pricing, the social welfare increases by \$36,753. Even though charging the road users a total of \$12,541 (average over the last 300 iterations), the user benefits increase by \$10,546. That is, the increase in user benefits overcompensates the road charges. The explanation for the increase in welfare is the users' change in behavior which raises the efficiency within the car mode. Applying marginal social cost pricing, users adjust their departure time, route and transport mode by taking into account the delay costs that they impose on others. The number of car trips (average over the last 300 iterations) is reduced by 1,711 (0.8% of the total demand). Hence, the car mode is released and the level of congestion decreases, indicated by the total travel time (average over the last 300 iterations) which is reduced by a total of 2,838 hours compared to the simulation run without pricing. The average travel time per car trip is reduced by more than 1 min, from 6 min 42 sec down to 5 min 28 sec (average over the last 300 iterations). The external congestion costs amount to \$55,489 in the final iteration of the non-pricing experiment. They are much higher, compared to the pricing experiment in which the toll revenues are equal to the internalized external congestion costs (\$12,541).

External Congestion Effects In the following, the final iteration is analyzed in more detail focusing on the external congestion effects. The average toll per trip amounts to \$0.11 with a standard deviation of \$0.33. Depending on the time of day and route, the congestion costs imposed on other users range between \$0.00 and \$4.08 per trip. During

Table 4: Changes in social welfare when charging marginal social cost prices

	No pricing	Road pricing	**Changes**
User Benefits	\$44,644,110	\$44,654,656	+**\$10,546**
Toll Revenues	\$0	\$12,541	+**\$12,541**
Social Welfare	\$44,644,110	+\$44,680,863	+**\$36,753**

peak times more users are affected than off-peak, hence the average external cost per trip is much higher compared to off-peak periods. For trips starting between 6 p.m. and 8 p.m. (evening peak) an average toll of \$0.23 per car trip is obtained. Whereas, for trips starting between noon and 2 p.m. (off-peak time) the average external delay cost per car trip is only \$0.01.

Fig. 5 shows for each link the average toll of the agents that are driving along this link during the day. The average external congestion cost is higher for urban roads, especially along the intermediate east-west corridor. The average tolls per agent payed on each link range from \$0.00 up until \$0.56. The average tolls per agent are observed to be higher for bottleneck links with a reduced flow capacity compared to the flow capacities of the ingoing links.

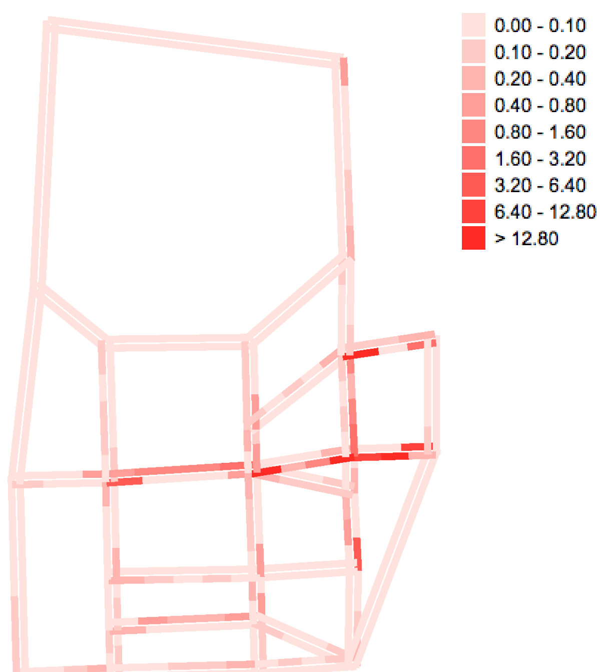


Figure 5: Average tolls payed per link and day [$\$10^{-2}$]

6 Conclusion

In this study, an agent-based approach was developed to calculate optimal user-specific tolls. To the best of the authors' knowledge, this study is unique in calculating dynamic congestion effects among car users at a microscopic, truly agent-based level. This high level of disaggregation allows to consider the heterogeneity in demand, e.g. different VTTS. The presented marginal social cost approach was implemented and tested using the agent-based simulation MATSim which allows the application of large-scale scenarios. For simulating the traffic flows, MATSim applies a queue model in which the flow capacity restricts the number of cars that can pass a link within a certain time interval and the storage capacity defines the maximum number of vehicles per link and effects the spatial propagation of delays (spill-back effects).

The basic idea of calculating marginal social cost is to keep track of the agents' interferences. External congestion effects are then internalized by charging the equivalent monetary amount from the causing agent. Hence, the external congestion costs are included in each agent's utility and thereby taken into account in the decision making process. The calculation of external congestion effects is described for two examples. In the first example, spill-back effects do not occur, the agents in front prevent the following agents on the same link from moving to the next link for a certain amount of time. In the second example, the storage capacity is reduced and the delay effect is propagated on the upstream link.

At first, the user-specific pricing approach was verified using a simple test scenario, which consists of two alternative routes and three agents. The simulation experiments yield a plausible outcome. Next, the marginal social cost approach was applied to a more complex test scenario based on the city of Sioux Falls which is analyzed in more detail. As expected, the simulation experiment with road pricing leads to a higher social welfare compared to the experiment without pricing. By applying marginal social cost prices, the level of congestion is reduced. As the increase in efficiency within the car mode is very large, the gain in user benefits overcompensates for the tolls payed by the users. The external congestion costs are investigated focusing on temporal and spatial effects. First, we find the average tolls per trip to be much higher during peak times compared to off-peak times. Second, the average toll per agent is much higher for urban roads with relative small flow capacities.

The presented first-best approach can be used as a benchmark for second-best pricing strategies. Furthermore, the temporal and spatial differences of the average external congestion costs may be used to develop road pricing strategies that are feasible for real-world application. The results can be used as a starting point for policy making and a manual optimization. Temporal differences of the average external congestion costs may help to implement a peak pricing scheme, e.g. to determine the peak interval or the toll level. Whereas, a spatial analysis of the first-best tolls may reveal in which areas or on which road segments congestion pricing is of particular importance.

The next step is to compare the presented approach with the formula used by Lämmel and Flötteröd (2009) who assume stationary flow conditions. Furthermore, we plan an integrated study which incorporates external congestion costs within the car mode, and additionally delay effects within the public transport mode (Kaddoura et al., 2013), and environmental effects (Kickhöfer and Nagel, 2013).

Acknowledgements

The test scenario in Sec. 4 is based on a talk given by Gregor Lämmel (Forschungszentrum Jülich) and Gunnar Flötteröd (KTH Royal Institute of Technology) at the 32nd Annual Conference on Artificial Intelligence, Paderborn. The authors wish to thank Artem Chakirov (ETH Centre Singapore) for providing the Sioux Falls scenario, Kai Nagel (Technische Univsersität Berlin) and Gregor Lämmel for their helpful comments, and Lars Kröger (Technische Univsersität Berlin) for his helpful assistance in testing and improving the code.

References

A. Agarwal, M. Zilske, K. Rao, and K. Nagel. Person-based dynamic traffic assignment for mixed traffic conditions. In *Conference on Agent-Based Modeling in Transportation Planning and Operations 2013*, Blacksburg, Virginia, USA, 2013. Also VSP WP 12-11, see www.vsp.tu-berlin.de/publications.

R. Arnott, A. de Palma, and R. Lindsey. A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 83(1):161–179, 1993. ISSN 00028282. URL `http://www.jstor.org/stable/2117502`.

A. Chakirov and P. Fourie. Enriched sioux falls scenario with dynamic demand. Technical report, Future Cities Laboratory, Singapore - ETH Centre (SEC), forthcoming.

D. Charypar and K. Nagel. Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32(4):369–397, 2005. ISSN 0049-4488. doi: 10.1007/s11116-004-8287-y.

B. de Borger, I. Mayeres, S. Proost, and S. Wouters. Optimal pricing of urban passenger transport: A simulation exercise for Belgium. *Journal of Transport Economics and Policy*, 30(1):31–54, 1996.

C. Gawron. An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model. *International Journal of Modern Physics C*, 9(3):393–407, 1998.

S. Jara-Díaz and A. Gschwender. Towards a general micoreconomic model for the operation of public transport. *Transport Reviews*, 23(4):453–469, 10-12 2003. ISSN 0144-1647. doi: 10.1080/0144164032000048922.

I. Kaddoura, B. Kickhöfer, A. Neumann, and A. Tirachini. Optimal public transport pricing: Towards an agent-based marginal social cost approach. In *hEART 2013 – 2nd Symposium of the European Association for Research in Transportation*, Stockholm, Sweden, 2013. Also VSP WP 13-09, see www.vsp.tu-berlin.de/publications. Awarded as the Best PhD Student Paper at hEART 2013.

B. Kickhöfer and K. Nagel. Towards High-Resolution First-Best Air Pollution Tolls. *Networks and Spatial Economics*, pages 1–24, 2013. doi: 10.1007/s11067-013-9204-8.

B. Kickhöfer, D. Grether, and K. Nagel. Income-contingent user preferences in policy evaluation: application and discussion based on multi-agent transport simulations. *Transportation*, 38:849–870, 2011. ISSN 0049-4488. doi: 10.1007/s11116-011-9357-6.

B. Kickhöfer, F. Hülsmann, R. Gerike, and K. Nagel. Rising car user costs: comparing aggregated and geo-spatial impacts on travel demand and air pollutant emissions. In T. Vanoutrive and A. Verhetsel, editors, *Smart Transport Networks: Decision Making,* *Sustainability and Market structure*, NECTAR Series on Transportation and Communications Networks Research, pages 180–207. Edward Elgar Publishing Ltd, 2013. ISBN 978-1-78254-832-4.

M. Kraus. Discomfort externalities and marginal cost transit fares. *Journal of Urban Economics*, 29:249–259, 1991.

G. Lämmel and G. Flötteröd. Towards system optimum: Finding optimal routing strategies in time-tependent networks for large-scale evacuation problems. pages 532–539, 2009. doi: 10.1007/978-3-642-04617-9_67.

L. LeBlanc, E. Morlok, and W. Pierskalla. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research*, 9:309–318, 1975.

M. Maibach, D. Schreyer, D. Sutter, H. van Essen, B. Boon, R. Smokers, A. Schroten, C. Doll, B. Pawlowska, and M. Bak. Handbook on estimation of external costs in the transport sector. Technical report, CE Delft, 2008.

H. Mohring. Optimization and scale economics in urban bus transportation. *American Economic Review*, 62(4):591–604, 1972. ISSN 0002-8282.

K. Nagel and G. Flötteröd. Agent-based traffic assignment: Going from trips to behavioural travelers. In R. Pendyala and C. Bhat, editors, *Travel Behaviour Research in an Evolving World – Selected papers from the 12th international conference on travel behaviour research*, chapter 12, pages 261–294. International Association for Travel Behaviour Research, 2012. ISBN 978-1-105-47378-4.

K. Nagel, D. Grether, U. Beuck, Y. Chen, M. Rieser, and K. Axhausen. *Multi-agent transport simulations and economic evaluation*, volume 228 of *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, pages 173–194. 2008.

R. H. Oldfield and P. H. Bly. An analytic investigation of optimal bus size. *Transportation Research Part B: Methodological*, 22:319–337, 1988. ISSN 0191-2615. doi: 10.1016/0191-2615(88)90038-0.

I. W. H. Parry and K. A. Small. Should urban transit subsidies be reduced? *American Economic Review*, 99:700–724, 2009.

A. Pigou. *The Economics of Welfare*. MacMillan, New York, 1920.

B. Raney and K. Nagel. An improved framework for large-scale multi-agent simulations of travel behaviour. pages 305–347. 2006.

K. A. Small and E. T. Verhoef. *The economics of urban transportation*. Routledge, 2007. ISBN 9780415285148.

A. Tirachini and D. A. Hensher. Bus congestion, optimal infrastructure investment and the choice of a fare collection system in dedicated bus corridors. *Transportation Research Part B: Methodological*, 45(5):828–844, 6 2011. ISSN 0191-2615. doi: 10.1016/j.trb.2011.02.006.

A. Tirachini, D. Hensher, and J. Rose. Multimodal pricing and optimal design of public transport services: the interplay between traffic congestion and bus crowding. In *Proceedings of the Kuhmo Nectar Conference on Transportation Economics*, 2012.

V. van den Berg. *Congestion Pricing With Heterogeneous Travellers*. PhD thesis, Tinbergen Institute, 2011.

W. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260, 1969.

W. Vickrey. Pricing, metering and the efficient use of urban transportation facilities. *Highway Research Record*, 476:36–48, 1973.

N. Zheng, R. A. Waraich, N. Geroliminis, and K. W. Axhausen. Investigation of macroscopic fundamental diagrams in urban road networks using an agent-based simulation model. In *Proceedings of the 11th Swiss Transport Research Conference (STRC)*, 2011. URL `http://www.strc.ch/conferences/2011/Zheng.pdf`.