Optimal Public Transport Pricing

Towards an Agent-based Marginal Social Cost Approach

Ihab Kaddoura, Benjamin Kickhöfer, Andreas Neumann, and Alejandro Tirachini

Address for correspondence: Ihab Kaddoura, Department of Transportation System Planning and Telematics, Berlin Institute of Technology (Technische Universität Berlin), Berlin, Germany (kaddoura@vsp.tu-berlin.de). Benjamin Kickhöfer and Andreas Neumann are at the Department of Transportation System Planning and Telematics, Berlin Institute of Technology (Technische Universität Berlin), Berlin, Germany. Alejandro Tirachini is at the Civil Engineering Department, Transport Engineering Division, University of Chile (Universidad de Chile), Santiago, Chile.

Abstract

In this paper, marginal cost-pricing rules are applied to an agent-based model. User-specific optimal bus fares are estimated by simulating user interactions at a microscopic level. We consider external effects of prolonged in-vehicle and waiting times (including the effect of full buses on boarding denials) induced by passengers boarding vehicles. We give a new-look to the relationship between optimal fare and travel distance, which is found to depend crucially on the type of external cost that is considered when calculating fares. Conditions for optimal bus fares that increase or decrease as a function of trip length are numerically found.

Date of final version: July 2014

1.0 Introduction

In the transport sector congestion, air pollution, noise, and accidents are sources of inter- or intra-sectoral external costs, leading to welfare losses and an inefficient market equilibrium (van Essen *et al.*, 2008). External effects arise if marginal private user costs deviate from marginal social costs. Marginal social cost pricing means setting generalised prices equal to the sum of marginal producer costs, marginal private user costs, and marginal external costs. Therefore, in the public transport market, the optimal fare amounts to the difference between marginal private and marginal social costs. By charging the optimal fare, external effects are internalised and taken into account by users. Thus, the right incentives are given to achieve market efficiency and maximise social welfare.

Since theoretical first-best conditions do not exist in reality (for example, due to under-priced competing modes, difficult computation, and unfeasible application), second-best solutions are required (Proost and van Dender, 2001; Verhoef, 2001; Small and Verhoef, 2007). Insights from the first-best solution may help to develop a second-best pricing strategy. Furthermore, the first-best solution can be used as a theoretical benchmark for the evaluation of other measures.

In this study, general public transport marginal cost-pricing rules are introduced in an agent-based simulation (MATSim), in which passenger-specific first-best fares can be calculated even in large-scale scenarios. To our knowledge, this is the first model that considers microscopic user-by-user bus fares, calculated with the objective to maximise social welfare. Accounting for time-dependency and queue formation allows for simulating the interaction of activity-scheduling decisions and public transport pricing. Three sources of external delays are accounted for when calculating the effect of an extra passenger taking the bus: (i) delay for passengers on-board; (ii) delay for passengers waiting at downstream stops; and (iii) extra waiting time for passengers unable to board a full bus. That is, depending on these three effects — for example, the total delay imposed on other passengers during a trip — a specific fare is computed for each individual. In this study, other externalities such as crowding costs and road congestion are excluded. The microscopic price differentiation is compared against the classical flat fare rule in order to analyse differences in terms of revenue, optimal frequency of service, and social welfare. In reality, charging an individual-specific fare for each trip would be difficult to implement and could even be considered unethical. Therefore, we analyse temporal and spatial effects by calculating average fares for each user-specific pricing scenario, which may be used as a starting point for more practical (second-best) pricing schemes — for instance, charging users anonymously according to time of day and location of boarding (and perhaps also location of disembarking). Importantly, our approach is useful to analyse the relationship between optimal pricing and trip length at a microscopic level, an issue previously addressed with analytical models by Mohring (1972), Turvey and Mohring (1975), and Kraus (1991), as discussed in the next section. A numerical example provides insights on the sensitivity of the user-specific optimal bus fare to travel distance and to the headway between vehicles, in which the relevance of introducing user-specific bus fares — as opposed to optimal flat fares — is evident.

The rest of the paper is organised as follows. Section 2 provides a discussion of the relevant literature on public transport economics that serves as a background for the analysis undertaken in this paper, Sections 3 and 4 present the simulation approach and

scenario used to obtain results, Section 5 describes the marginal cost-pricing effects that are included in our framework, Section 6 provides a discussion of results, and Section 7 summarises the main conclusions of the study.

2.0 Literature Review

Most studies on transport externalities focus on the private transport sector. Congestion is usually the largest part of all external costs in peak periods, whereas in off-peak periods, other externalities such as pollution and accidents have been found to be at comparable levels with congestion (de Borger *et al.*, 1996; Parry and Small, 2009). In the context of urban public transport modes, external effects among public transport users are investigated to a lesser extent (Maibach *et al.*, 2008; Prud'homme *et al.*, 2012). Nash (2003) addresses the urban rail sector and relates delays to the traffic volume by regression, finding delay effects to be the most significant external effect. A number of recent survey-based studies find that crowding externalities can be a substantial part of total travel costs (Wardman and Whelan, 2011; Prud'homme *et al.*, 2012; Haywood and Koning, 2013). While most discomfort studies address in-vehicle crowding, some studies also include crowding in access-ways and platforms (Lam *et al.*, 1999; Douglas and Karpouzis, 2005). Several authors have incorporated external effects in analytical models that usually account for delays due to capacity constraints, boarding and alighting passengers, the disutility of crowding, and congestion effects among vehicles at bus stops or train stations (see, for example, Mohring, 1972; Turvey and Mohring, 1975; Kraus, 1991; Parry and Small, 2009; de Palma *et al.*, 2015; Tirachini, 2014; Tirachini *et al.*, 2014). Turvey and Mohring (1975) analyse delay effects induced by boarding and alighting passengers to: (i) passengers inside the bus; and (ii) passengers waiting at a transit stop who will enter the bus either further along the route or at the same stop behind agents who get on and off the bus. However, the authors explain, in case of random passenger arrival patterns at transit stops, the second effect disappears on average. That is because agents with shorter waiting times compensate for agents with a longer waiting time. The only external effect considered in Mohring (1972) are delays that boarding and alighting passengers impose on passengers who are on-board vehicles. In this case, optimal fare increases as a function of the number of passengers on-board vehicles, because those passengers experience an extra delay when a new passenger gets on or off a vehicle. For the case of a feeder route, Mohring (1972) finds marginal cost to be inversely related to the trip length, because when buses approach the Central Business District (CBD), they increase in passenger number. However, Turvey and Mohring (1975) point out that this rule does not hold for the case of limited vehicle capacities and the probability of fully loaded buses, causing passengers not to be able to board the first arriving bus. That is, short-distance passengers may not be able to board a bus full of longer-distance passengers. Along the same lines, Kraus (1991) found that when including an external discomfort cost, long-distance sitting passengers impose a discomfort cost on short-distance passengers who are unable to find a seat, and therefore are forced to stand.

Beyond analytical models, marginal social cost pricing on public transport has not received much attention in activity-based simulations, which are less simplifying and more applicable to real-world scenarios. Analytical models usually ignore more complex

user reactions; that is, time adaptation, and the situation of coordinated arrival patterns in which passengers schedule their arrival according to a published timetable of services (see Jansson (1993) for a model with coordinated arrival patterns). While the cost of prolonged in-vehicle times as well as the disutility of crowding and discomfort are addressed by several authors, external waiting costs resulting from boarding and alighting passengers is less intensively discussed. In Kaddoura *et al.* (2014), flat bus fares are optimised applying a simulation-based grid search approach, finding delay effects due to passenger transfers and capacity constraints to have a major impact on the social welfare and optimal pricing structure. The present study takes up these findings and addresses the setting of optimal public transport pricing by investigating the marginal cost imposed by public transport users on other users, at a microscopic user-by-user level. Results in terms of social welfare, distribution of optimal fares, and optimal frequency of service can be used as a benchmark for the analysis of other pricing policies such as flat fares.

3.0 Methodology

In this research, the open-source agent-based micro-simulation MATSim[1] is used for the optimisation of public transport fares. In this section a general overview of MATSim is given, including the special characteristics of simulating public transport. For further information of the simulation framework MATSim, see Raney and Nagel (2006).

3.1 MATSim overview
Private and public transport users are modelled as individual agents that have a mental and physical behaviour. The iterative simulation approach consists of the following steps:

1. **Plans generation:** All agents independently generate daily plans that contain the planned activities, departure times, and transport modes for intervening trips.
2. **Traffic flow simulation:** All agents simultaneously execute their selected plans and interact in the physical layer.
3. **Evaluating plans:** All executed plans are evaluated with respect to individual utility functions, taking into account both the activities and trips.
4. **Mutation/selection:** For the next iteration, some agents generate new plans by mutating copies of existing plans. The remaining agents select plans depending on their differences in utility.

The iterative process of coupling the physical and mental behaviour enables agents to improve their plans and generate plausible alternatives. After disabling the mutation (see step 4) for the final iterations, the agents choose their plans with respect to a multinomial logit model. The system's inherent randomness determines which plan is individually chosen and, as a consequence, how the traffic behaviour changes from iteration to iteration. Once the demand is relaxed, meaning the average utility of executed plans remains stable over several iterations, and assuming that the travel alternatives form valid choice sets,

---

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

the system state is in an approximate stochastic user equilibrium (Nagel and Flötteröd, 2012).

3.2 Public transport in MATSim

A transit schedule contains all transit lines, routes, stops, and departures. It describes all planned transit vehicle operations in the system. Each vehicle is separately simulated in the traffic flow simulation. Depending on the vehicle type and the number of boarding and alighting passengers, transit vehicles can be delayed and actual departures may differ from the planned schedule. A parallel door operation mode allows simultaneous boarding and alighting. A serial mode gives alighting passengers priority. In the case where a vehicle is fully loaded, additional boardings are denied and passengers will have to wait for the next vehicle. For a detailed description of MATSim's public transport dynamics, see Neumann and Nagel (2010), and Rieser (2010).

4.0 Scenario: Multi-modal Corridor

The scenario that is used for the simulation experiments in this study is based on the setup and simulation approach used in Kaddoura *et al.* (2014).

4.1 Setup

4.1.1 Supply

For the simulation experiments, a multi-modal corridor with a total length of 20 km is considered, along which all activities are located. Between 4 a.m. and midnight the corridor is served in both directions by a constant number of buses that are operated by a single company. Capacities are set to 100 passengers per vehicle. Transit stops are placed every 500 m along the corridor. Assuming a walking speed of 4 km/h, access and egress times depend on the beeline between transit stop and activity location. The minimum dwell time at each transit stop is set to 0 sec; thus the bus only stops at transit stops if passengers intend to board or alight. Buses are assumed to decelerate when approaching bus stops, regardless of whether passengers will board/alight or not. That is, the deceleration and acceleration time at bus stops is constant for all transit stops, and is therefore included in the average bus speed. Considering an average bus speed of 30 km/h, a slack time of 20 min when reaching a corridor's endpoint and ignoring passenger transfers amounts to a cycle time of 2 h. As transit vehicles are delayed by passengers, actual cycle times and headways differ from the planned schedule. Bus door operation mode is serial so that alighting precedes boarding. For each vehicle, the average boarding time is set to 1 sec per person (assuming two doors per vehicle and 2 sec for each person and door, obtained with a boarding system with contactless card fare payment (Wright and Hook, 2007)) and alighting times to 0.75 sec per person. Delayed transit vehicles will try to follow the schedule by shortening slack times. For the alternative car mode, the free-flow speed is set to 50 km/h. In order to focus on dynamic delay effects within the public transport mode, interferences between cars and buses are excluded, and roads are not affected by congestion. Therefore, car travel times only result from the distance between two activity locations and the free-flow speed.

Figure 1
Initial Departure Time Distribution for Bus Users; Car Mode is Similar
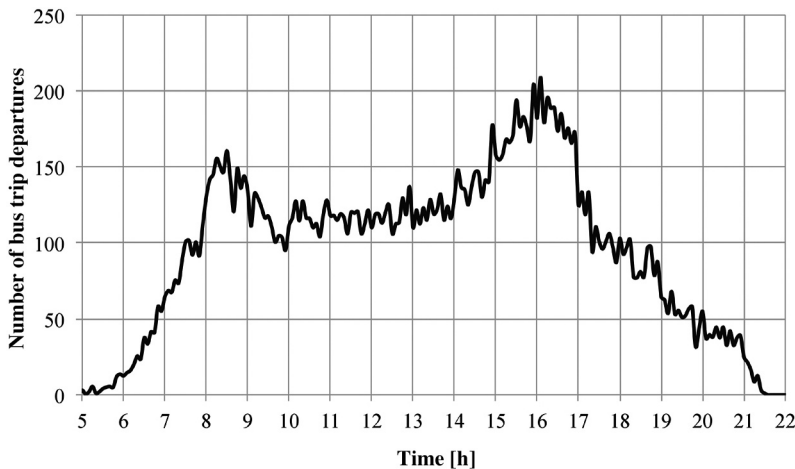
4.1.2 Demand
On the demand side, 20,000 travellers are considered with randomly distributed activity locations along the corridor. The activity patterns are split into two types: 'Home–Work–Home' (35 per cent of all travellers) and 'Home–Other–Home' (65 per cent of all travellers). Initial departure times from activity 'Home' to 'Work' follow a normal distribution, with mean at 8 a.m. and a standard deviation of 1 h. Agents are assumed to head back home 8 h after starting work. The activity type 'Other' has an initial duration of 2 h and is uniformly distributed from 8 a.m. to 8 p.m. The initial modal split for all trips is 50 per cent car and 50 per cent bus. The overlay of peak and off-peak demand (both activity patterns) is shown in Figure 1.

4.2 Simulation approach

4.2.1 Choice dimensions
The iterative learning mechanism of MATSim employs a utility-based approach. Each iteration agents choose from an existing set of daily plans with respect to a multinomial logit model. In order to generate plausible travel alternatives, agents are enabled to modify their daily plans according to the dimension's mode and time. That is, agents are enabled to switch between public transport and car. In this study, cars are assumed to be available for all agents. Additionally, agents adapt their activities' scheduling decisions. Activities can be shifted, extended, or shortened with respect to activity specific opening/closing times and typical durations (see Table 1).

4.2.2 Utility functions
Executed daily plans are evaluated taking into account the activity and trip-related part of the utility:

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

205

Table 1
Activity Attributes

Type	$t*$	Opening time	Closing time
Home	12 h	Undefined	Undefined
Work	8 h	6 a.m.	8 p.m.
Other	2 h	8 a.m.	8 p.m.

where V_{plan} is the total utility of a daily plan; n is the total number of activities or trips; $V_{act,i}$ is the utility for performing activity i; and $V_{trip,i}$ is the utility of the trip to activity i. The first and the last activity are wrapped around the day, and handled as one activity. Thus, the number of activities and trips is the same. The trip-related utility is calculated as follows:

$$V_{pt,i,j} = \beta_{v,pt} \cdot t_{i,v,pt} + \beta_{w,pt} \cdot t_{i,w,pt} + \beta_{a,pt} \cdot t_{i,a,pt} + \beta_{e,pt} \cdot t_{i,e,pt} + \beta_c \cdot c_{i,pt},$$
$$V_{car,i,j} = \beta_0 + \beta_{tr,car} \cdot t_{i,tr,car} + \beta_c \cdot c_{i,car}, \tag{2}$$

where V is the utility for person j on their trip to activity i. Attributes considered for car trips, indicated by *car*, are the travel time $t_{i,tr,car}$ and monetary costs $c_{i,car}$ that depend on the travel distance. For public transport trips, indicated by *pt*, different time components are considered: the in-vehicle time $t_{i,v,pt}$, the waiting time $t_{i,w,pt}$, the access time $t_{i,a,pt}$, the egress time $t_{i,e,pt}$, as well as monetary costs $c_{i,pt}$. To calculate the positive utility gained by performing an activity, a logarithmic form is applied (Charypar and Nagel, 2005; Kickhöfer *et al.*, 2011):

$$V_{act,i}(t_{act,i}) = \beta_{act} \cdot t*_{,i} \cdot \ln\left(\frac{t_{act,i}}{t_{0,i}}\right), \tag{3}$$

where t_{act} is the actual duration of performing an activity (when the activity is open), $t*$ is an activity's 'typical' duration, and β_{act} is the marginal utility of performing an activity at its typical duration. In the equilibrium, all activities at their typical duration are required to have the same marginal utility; therefore, β_{act} applies to all activities. $t_{0,i}$ is a scaling parameter linked to an activity's priority and minimum duration. In this study, $t_{0,i}$ is not relevant, since activities cannot be dropped from daily plans. Activities' opening and closing times determine in which time slots activities can be performed (see Table 1). Outside these time slots, agents do not earn a positive utility; thus, they are penalised by the opportunity costs of time $-\beta_{act}$.

4.2.3 Parameters

Behavioural parameters for the utility functions are based on estimations for Sydney (Tirachini *et al.*, 2014). Table 2 depicts the estimated parameters (flagged by a hat) and Values of Travel Time Savings (VTTS). Table 3 shows the adjusted parameters that are used in the present study and fit the activity-based approach of MATSim. As described in Kickhöfer *et al.* (2011, 2013), time-related parameters are split into opportunity costs of time and a disutility of travelling. Since the VTTS for egress based on the estimations is implausible high and egress times are constant in the present paper, the egress time parameter $\beta_{e,pt}$ is set equal to the parameter for access $\beta_{a,pt}$. The simulation approach does not account for parking and walking times of car users; agents can enter and leave

Table 2
Parameters and VTTS Based on Tirachini et al. *(2014)*

$\hat{\beta}_{tr,car}$	−0.96	[utils/h]
$\hat{\beta}_{v,pt}$	−1.14	[utils/h]
$\hat{\beta}_{w,pt}$	−1.056	[utils/h]
$\hat{\beta}_{a,pt}$	−0.96	[utils/h]
$\hat{\beta}_{e,pt}$	−3.3	[utils/h]
$\hat{\beta}_{c}$	−0.062	[utils/$]
$\hat{\beta}_{act}$	n.a.	
$VTTS_{tr,car}$	15.48	[$/h]
$VTTS_{v,pt}$	18.39	[$/h]
$VTTS_{w,pt}$	17.03	[$/h]
$VTTS_{a,pt}$	15.48	[$/h]
$VTTS_{e,pt}$	53.23	[$/h]

their private vehicles directly at an activity location and start an activity straight away. In order to compensate for that, the alternative specific constant β_0 was re-calibrated. The result is an alternative specific constant for car of $\beta_0 = -0.15$, leading to 50 per cent public transport trips for a headway of 9 min and a fare of $1.50. $c_{i,car}$ is the monetary cost component for car trips depending on the travel distance and a cost rate of 0.40 $/km (Australian Dollar [AUD], $1.00 = EUR 0.70 [July 2013]). $c_{i,pt}$ is the fare which is either a flat fee that has to be paid every time an agent boards a bus or the sum of all user-specific prices that are paid during a trip.

4.2.4 Operator's profit and social welfare
The operator cost model that is applied in this study follows an approach for urban regions in Australia (ATC, 2006):

$$C = (vkm \cdot c_{vkm} + vh \cdot c_{vh}) \cdot O + vNr \cdot c_{vday}, \tag{4}$$

Table 3
Parameters and VTTS Used in the Present Paper

$\beta_{tr,car}$	0	[utils/h]
$\beta_{v,pt}$	−0.18	[utils/h]
$\beta_{w,pt}$	−0.096	[utils/h]
$\beta_{a,pt}$	0	[utils/h]
$\beta_{e,pt}$	0	[utils/h]
β_{c}	−0.062	[utils/$]
β_{act}	+0.96	[utils/h]
$VTTS_{tr,car}$	15.48	[$/h]
$VTTS_{v,pt}$	18.39	[$/h]
$VTTS_{w,pt}$	17.03	[$/h]
$VTTS_{a,pt}$	15.48	[$/h]
$VTTS_{e,pt}$	15.48	[$/h]

Table 4
Operator Costs Based on ATC (2006)

C_{vkm}	0.006 capacity + 0.513 [\$/vkm]
C_{vday}	1.6064 capacity + 22.622 [\$/vday]
C_{vh}	33 [\$/vh]
O	1.21

where C is the operator cost per day; *vkm* is the vehicle kilometres per day; c_{vkm} is the cost per vehicle kilometre; *vh* is the vehicle hours per day; c_{vh} is the cost per vehicle hour; O is the overhead; *vNr* is the total number of operating vehicles per day; and c_{vday} is the daily capital cost per vehicle. c_{vkm} and c_{vday} are obtained by linear regression, whereas c_{vh} and O are constant for different vehicle types (see Table 4).

Daily operators' revenues are calculated by multiplying the number of public transport trips per day (T_{pt}) by the fare (f) that is paid during the trip. Operators' profit per day ($\Pi_{operator}$) is calculated as equation (5):

$$\Pi_{operator} = T_{pt} \cdot f - C. \tag{5}$$

The benefits on the demand side are calculated as the Expected Maximum Utility (EMU), taking into account all users' choices. The sum of operator profit and user benefits determines the social welfare W:

$$W = \Pi_{operator} + \sum_{j=1}^{J} \left(\frac{1}{|\beta_c|} \ln \sum_{p=1}^{P} e^{V_{plan}} \right), \tag{6}$$

where β_0 is the negative marginal utility of monetary cost; 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 choice set.

4.2.5 Simulation procedure

For the first 250 iterations of demand relaxation (see Section 3.1), plans are modified by each choice dimension (mode/time) with a probability of 10 per cent. The size of the agents' choice sets is set to six daily plans each. After 250 iterations, the agents are assumed to have a plausible set of plans; thus, choice set generation is switched off. For an additional fifty iterations, the agents choose among their existing daily plans according to a multinomial logit model. To allow for a fast computational performance and assuming the impact of randomness — that is, fluctuations from iteration to iteration — to be negligible, the last iteration is considered as a representative day, and the outcome of the last iteration is used for analysing travel behaviour and social welfare.

## 5.0 Agent-based Marginal Social Cost Pricing

This study enhances MATSim by implementing a marginal social cost-pricing approach. On an agent-based level, external effects (marginal delays and number of affected users) are traced back to their origin. These external effects are internalised by charging the equivalent monetary amount from the agent who is causing the delay effects. Of all existing

external effects (see Section 2), this study focuses on delay effects among users within the public transport mode. That is, road congestion and interferences between buses and cars are excluded (see Section 4.1), and the disutility of travelling (in-vehicle time) is independent from crowding effects inside the transit vehicles or the seat availability (see Table 3). Slack times at the corridors' endpoints are considered as operating hours. As slack times are used to catch up on delays, operating costs are constant and only depend on the number of buses. Hence, marginal operating cost per passenger is $0 and the optimal user-specific fare is equal to the external cost only. External effects that are considered in this study either prolong in-vehicle times or waiting times. Sources for externalities are capacity constraints and delays induced by passengers boarding/alighting.

5.1 Effect 1: in-vehicle time delays due to boarding and alighting passengers
Assuming a scheduled minimum dwell time of 0 sec, transit vehicles will be delayed when agents are boarding or alighting. Therefore, travel times — that is, in-vehicle-times — increase with the number of agents who get on and off the bus. An agent who is boarding imposes a delay of their average boarding time (1 sec) on each agent inside the transit vehicle. To give an example, for two passengers inside a vehicle, the total external effect amounts to 2 sec (see Figure 2). Thus, the equivalent monetary amount that is charged to the causing agent amounts to $VTTS_{v,pt} \times (2/3{,}600)$ h. In this setup, bus drivers are not able to catch up on delays except during slack times at the route endpoints, when there is no passenger inside the vehicle.
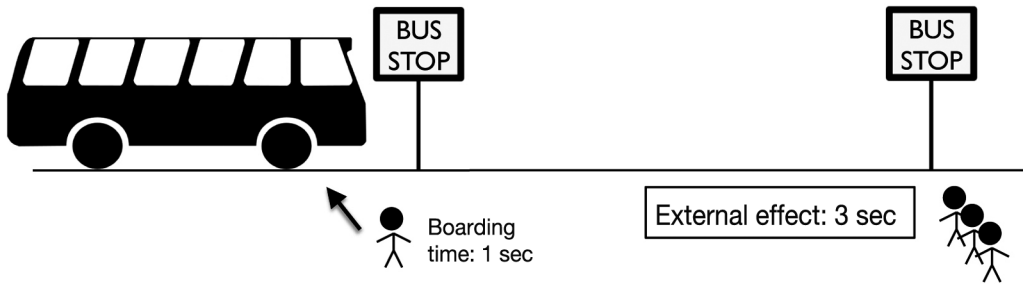
Figure 2
External Effect 1
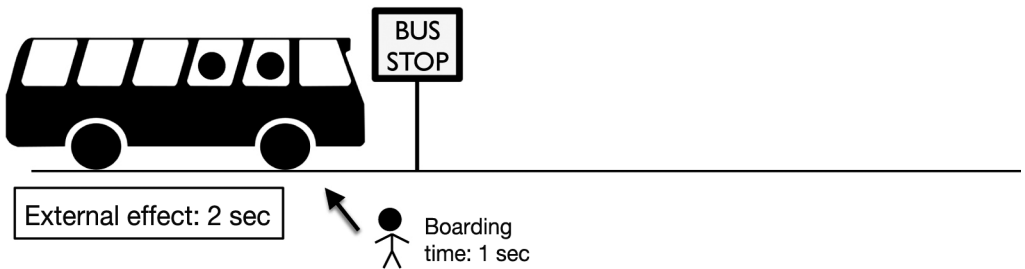
5.2 Effect 2: waiting time delays due to boarding and alighting passengers
As referred to in Section 2, another external effect is discussed in the context of passengers who get on and off buses. In cases where travellers' arrival patterns at transit stops are coordinated, meaning the users try to reach the stop shortly before a bus arrives, there may be users with prolonged waiting times due to other passengers who are boarding or alighting. As we assume agents adapt to the schedule, this effect does not cancel out on average, as is the case for random passenger arrivals in Turvey and Mohring (1975). Nevertheless, the agents' adaptation to bus departures is not perfect and some agents will also benefit from delayed buses. These positive external effects are ignored in this approach. Therefore, the calculation of waiting time delays due to boardings/alightings is an upper limit of the actual delays. If an agent delays the bus by 1 sec and, for instance, three

Figure 3
External Effect 2

agents are waiting at a transit stop further along the route, the total external effect amounts to 3 sec (see Figure 3). Hence, the external cost that is charged to the causing agent amounts to $\text{VTTS}_{w,pt} \times (3/3,600)$ h. However, agents are able to adapt their arrival times at transit stops according to the delay of transit vehicles. To give an example, 600 boarding agents delay a vehicle by 10 min. The waiting time of a passenger further along the route may be less than the delay of the arriving bus; for instance, 1 min. In that case, the external effect imposed on that passenger amounts to the waiting time only. The causing agents do not have to pay for the total delay of the transit vehicle, only for the external effect. Therefore, the actual waiting time is divided by the vehicle delay to derive a factor (1 min/10 min $= 0.1$) to calculate the delay effect that each causing agent has to pay for.

5.3 Effect 3: waiting time delays due to capacity constraints
If transit vehicles are fully loaded and boardings are denied, the passengers' waiting times increase. In this case, the agents inside the transit vehicle impose a negative external effect on the agent(s) who are not able to enter the bus. The external delay effect depends on when the next transit vehicle is arriving at the transit stop. For example, if the next transit vehicle arrives 600 sec later, the total external effect amounts to 600 sec (see Figure 4). The passengers of the vehicle that denied boarding have to pay $\text{VTTS}_{w,pt} \times (600/3,600)$ h in total. Assuming all agents in the transit vehicle cause this externality, each agent in the bus is charged an equal partial amount. In cases where boarding is denied a second time, the agents in the second bus pay for the interval until the third bus arrives at the transit

Figure 4
External Effect 3
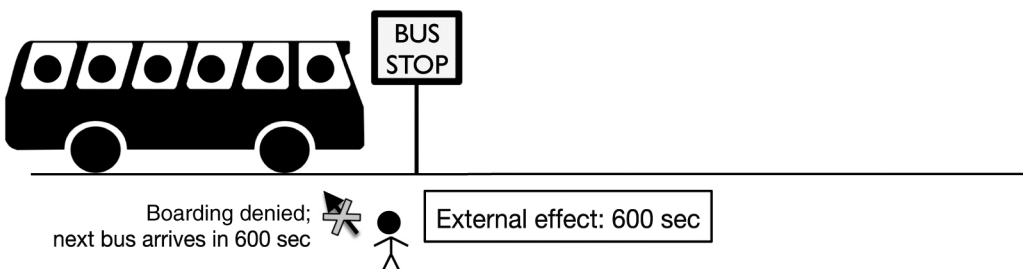
stop, and so forth. In cases of bus bunching, overtakings may cause a positive external effect, since left-behind passengers are able to get on a bus that is faster than the previous bus. In this setup, buses do not overtake each other.

5.4 Simulation experiments

In a first experiment, prices are set according to the cost resulting from prolonged in-vehicle times only (effect 1). The other external effects described above exist in the simulation, but are excluded from calculating marginal cost fares. In a second experiment, prolonged waiting times due to passengers boarding/alighting are included: prices are set according to the external in-vehicle and waiting cost resulting from passengers boarding and alighting (effects 1 and 2). Externalities due to capacity constraints exist, but are not considered when setting prices. In a third experiment, all three external effects are included in calculating the optimal fare (effects 1, 2, and 3). Unless explicitly stated, marginal cost pricing means the inclusion of all three effects. To validate the results in terms of social welfare and external delay effects, a grid search approach is used to identify the optimal fare where all users are charged by the same amount. The grid search is implemented as introduced in Kaddoura *et al.* (2014). The relaxation of the initial demand (see Section 3.1) is embedded into an external loop in which the flat fare is systematically varied between \$0 and \$3 (in steps of \$0.025). The marginal social cost pricing requires a single simulation run, whereas the grid search requires one simulation run for each possible public transport fare. Hence, in terms of performance, the marginal social cost-pricing approach is much faster than the grid search optimisation approach. Both pricing strategies — user-specific and flat fares — are run for different levels of transport supply. Scheduled headways are increased by 2 min between 3 min and 13 min.

6.0 Results

6.1 Social welfare

Figure 5 depicts the absolute welfare level obtained for each headway and pricing rule. As a benchmark to evaluate the effectiveness of different pricing rules, we also consider a free service with public transport fares that amount to \$0 per trip (no pricing), yielding the lowest social welfare for all simulated headways. Charging the optimal flat fares that are found to be greater than \$0 raises social welfare, especially for larger headways and binding vehicle capacity constraints. User-specific pricing leads to an even higher social welfare than charging an optimised flat fare. That is true for all three cases, even though some of the existing externalities are excluded from setting marginal cost fares in experiments 1 and 2. Internalising effects 1 and 2 yields a higher welfare than effect 1 only, especially for larger headways and a higher average waiting time. However, for headways with passengers who are not able to board a bus due to capacity constraints, the inclusion of effect 3 has the most significant effect on social welfare.

Analysing the distribution of bus trips over time reveals that peaks are flattened in the user-specific pricing scheme compared to the flat-pricing strategy. Since user-specific fares are higher during peak times (see later in Section 6.2), and commuters as well as non-commuters can choose their activity start and end times within wide time spans, bus users are able to shift their trip departure times in order to avoid peak periods. That is, by forcing

Figure 5
*Welfare for Each Headway and Pricing Rule*
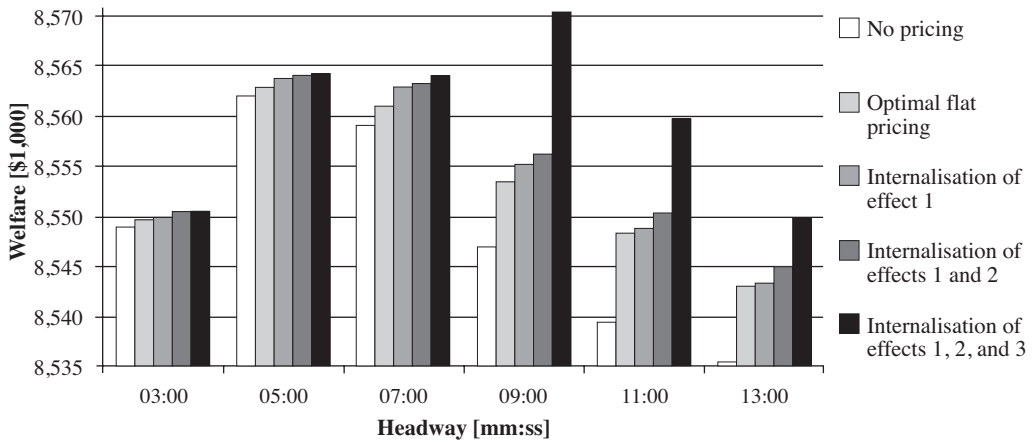
some users to either depart off-peak or to switch to the alternative (uncongested) car mode, the efficiency within the transport system increases. Benefits from shorter in-vehicle and waiting times within the public transport mode overcompensates for lower benefits resulting from users switching to car or arriving earlier or later at the activity location. Therefore, setting user-specific prices raises the social welfare if these external effects are taken into account.

A scheduled headway of 5 min is observed to be the optimal headway for all pricing rules, except the one including external waiting cost due to fully loaded transit vehicles. When setting fares equal to all existing external effects (internalisation of effects 1, 2, and 3), an overall welfare optimum is found for a headway of 9 min. That is, the pricing rule has an effect on the optimal level of public transport supply.

The observations of different welfare levels depending on headway and pricing rule are related to the number of boarding denials. Delay effects due to capacity constraints become significant for headways starting from 7 min; for smaller headways they are not significant. All pricing schemes reduce the number of boarding denials, especially when considering waiting time delays due to capacity constraints (internalisation of effects 1, 2, and 3). The maximum waiting cost related to capacity constraints is found for a headway of 9 min. As more users will switch to the alternative car mode for larger headways, the number of boarding denials decreases. Hence, for pricing strategies that do not explicitly account for capacity constraints, the welfare optimum is the largest headway without a significant number of boarding denials (5 min). However, in cases where the pricing strategy explicitly accounts for capacity constraints, the welfare maximum is found for a headway yielding the maximum occurrence of boarding denials (9 min). A headway of 3 min leads to no occurrence of boarding denials, whereas for a headway of 9 min a maximum number of boarding denials is obtained. Therefore, these two headways are now analysed in more detail.

Figure 6 depicts the distribution of waiting times for a 3-min and a 9-min headway applying marginal social cost pricing (internalisation of all effects). Waiting times that exceed the scheduled headway due to delayed vehicles or capacity constraints are not

Figure 6
Distribution of Waiting Times (Internalisation of All External Effects)

(a) 3-min headway   (b) 9-min headway
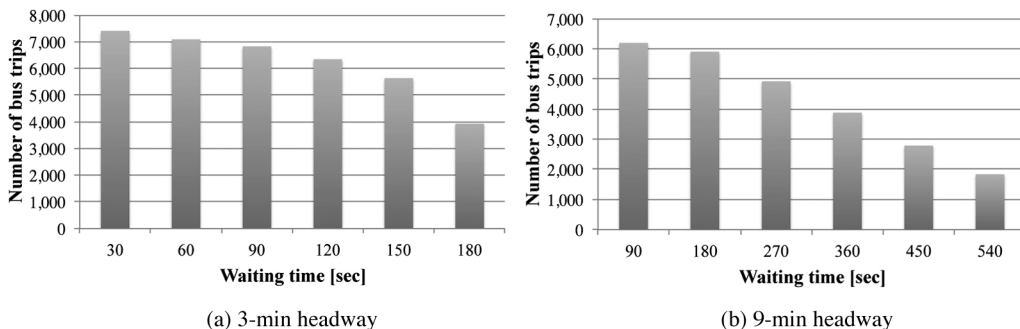
displayed. For both headways, the passengers' arrival patterns at transit stops tend to be coordinated. There are more trips with waiting times shorter than the half of the headway. However, for some trips, waiting times are observed to be longer than the half of the headway and even nearly as long as the scheduled headway. For a 3-min headway, 4,000 trips have a waiting time between 150 sec and 180 sec. For a 9-min headway, almost 2,000 trips are observed with a waiting time between 450 sec and 540 sec. For a detailed investigation of passenger arrival patterns at transit stops depending on the headway and service reliability, see Neumann *et al.* (2013).

That indicates that for both headways, there are travellers who miss a bus by a few seconds. If a bus had arrived later, these passengers would benefit from the delay since they would be able to catch the bus they would have otherwise missed. As described in Section 5, these positive externalities are ignored in the present paper and only the negative effect is accounted for when setting the prices. The distribution of waiting times confirms what was indicated by the positive effect on social welfare when including effect 2. That is, the negative effect of prolonged waiting times due to passengers getting on and off a bus outweighs the positive effect of shortened waiting times.

6.2 External delay cost

In Table 5, the average fare per public transport trip is shown for each pricing rule and a headway of 3 min and 9 min, considering all trips performed during the day. For the optimal flat-pricing scheme, the average fare is the flat amount that has to be paid for each bus trip, while, for the user-specific pricing scheme, the average fare amounts to the average external delay cost which is charged for a public transport trip. As expected, the

Table 5
Optimal Flat Fare and Average User-specific Fare (Per Bus Trip, Considering All Trips Per Day)

	3-min headway	9-min headway
Optimal flat pricing	$0.73	$0.78
Internalisation of effect 1	$0.31	$0.62
Internalisation of effects 1 and 2	$0.53	$1.04
Internalisation of effects 1, 2, and 3	$0.53	$1.69

more effects are included in setting the user-specific fares, the higher the average fare per trip. As there are no occurrences of boarding denials for a 3-min headway, the average internalised external cost remains unaltered when including effect 3. For a 3-min headway, the optimal flat fare amounts to $0.73; setting prices according to the total external cost (internalisation of all existing effects) yields an average fare of only $0.53 (standard deviation: $0.21). Charging all passengers the same optimal amount overprices public transport and results in a lower number of public transport users than applying marginal social cost pricing. However, for a 9-min headway, transit vehicle load factors and delay effects among public transport users are higher, which leads to an optimal flat fare of $0.78 and an average user-specific fare of $1.69, with a standard deviation of $1.79. For this headway, optimal flat pricing seems to under-price public transport services, which result in more public transport users compared to the user-specific pricing scheme.

Total internalised external delay costs for a headway of 3 min amount to $20,245, for a headway of 9 min to $46,443. Excluding the delay cost caused by capacity constraints, the external cost for a 9-min headway amounts to $28,509. Computing the externalities for the flat-pricing strategy shows that, compared to the user-specific pricing strategy, the average external delay costs are higher, in particular during peak times.

Depending on the number of affected users, average internalised external delay costs (all effects) range from $0 to $1.68 per trip for a 3-min headway and from $0 to $8.89 for a 9-min headway. Figure 7a shows the distribution of marginal social cost fares (internalisation of all effects) for a 9-min headway, leading to a total of 27,444 public transport trips, and a 3-min headway with a total of 37,898 bus trips. For a 9-min headway, marginal social cost fares are higher because users have to wait for later vehicles due to buses operating at maximum capacity. For more than 16 per cent of all public transport trips, passengers have to pay more than $2.0. For the distribution of fares between $2.0 and $8.9, see Figure 7b.

During peak times, more users are affected than off-peak. Hence, average peak prices exceed average off-peak fares as well as the obtained averages per day (see Table 5). For trips starting between 2 p.m. and 4 p.m., the average internalised external delay cost is $0.60 for a scheduled headway of 3 min and $2.94 for a headway of 9 min, while trips

Figure 7
Distribution of Marginal Social Cost Fares (Internalisation of All Effects)
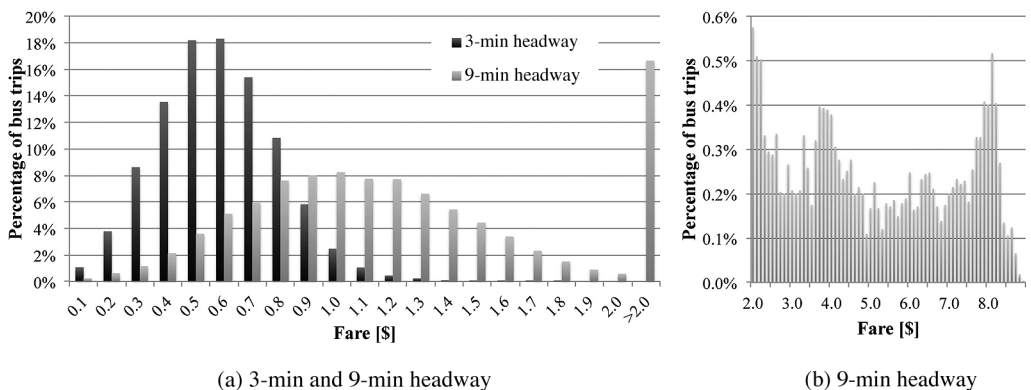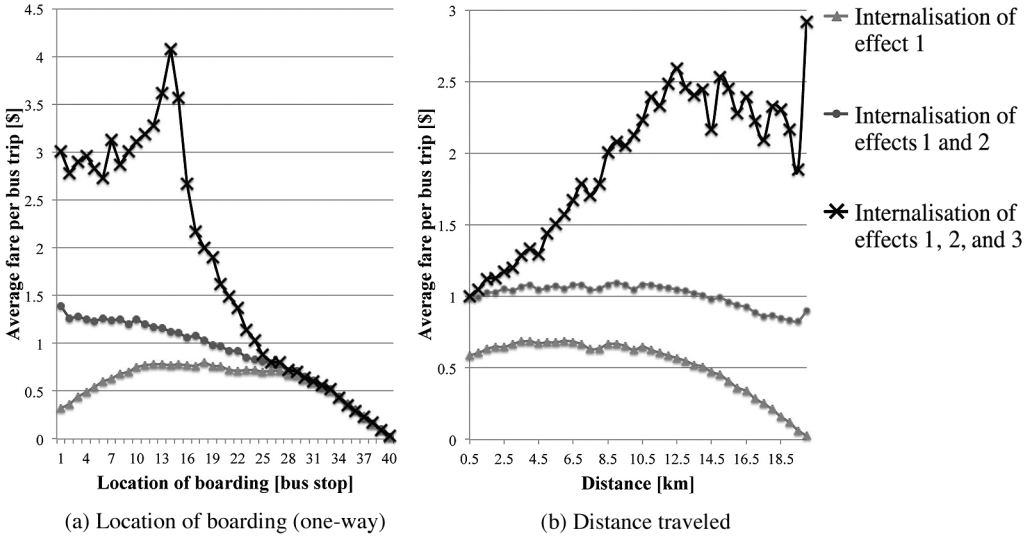

(a) 3-min and 9-min headway   (b) 9-min headway

Figure 8

Average Trip Fare Dependent on Distance and Boarding Location (9-min Headway)

(a) Location of boarding (one-way)   (b) Distance traveled

after 8 p.m. cause an average effect of only $0.21 for a 3-min headway and $0.42 for a 9-min headway. A higher pricing during peak periods increases social welfare by causing users to switch to the alternative car mode or to adapt their activity scheduling in order to avoid peak periods (see Section 6.1).

Beside temporal aspects, externalities are analysed focusing on spatial effects. Figure 8a depicts average bus trip fares dependent on the location of boarding in one direction for a 9-min headway. The average internalised external cost per boarding location depends on which external effects are considered. Beginning with the first experiment, the internalisation just of external effect 1 shows that average trip fares increase for passengers who board at the middle of the corridor. That is because at the beginning and the end of a transit route, transit vehicle load factors are lower, and thus the probability to affect agents inside a vehicle is lower, too. Including effect 2 significantly raises the average internalised external cost per trip for agents boarding at bus stops along the first half of the transit route. That is explained by a higher number of following transit stops and therefore passengers who are affected further along the corridor. The same trend is observed for different headways, also for 3 min where delay effects due to capacity constraints (effect 3) are not present. As clearly shown in Figure 8a, adding effect 3 has a significant impact on the average internalised external cost induced by agents boarding at the first twenty-four transit stops. For agents who board up until transit stop 14, the average external cost-based fare increases. External cost induced by passengers boarding at the following transit stops decreases the later they get on a bus. For transit stops in the second half of the route, there are fewer or no agents that cannot board a bus because of capacity constraints.

Figure 8b shows the average fare per travel distance for all three experiments. Internalised external cost resulting from effect 1 first slightly increases for short distances and then significantly decreases for distances larger than half of the corridor. A similar tendency is observed from Mohring (1972) for feeder routes, as passengers with longer trips board

earlier and delay fewer passengers in the bus. Including effect 2 flattens decreasing average trip fares for long-distance trips, since passengers who board at early transit stops affect all passengers that will be entering the same transit vehicle along the route. Adding effect 3 and computing the marginal social cost fares based on all existing external delays inverts the results obtained without the capacity constraints in place. For trip distances up to 12.5 km, average fares significantly increase with the trip length. For longer distances, average fares fluctuate and slightly decrease, whereas for the maximum trip length of 20 km, the maximum average fare is obtained. Long-distance passengers occupy seats for longer than short-distance passengers. This is in line with findings from Kraus (1991), who observes the same effect when including the disutility of discomfort, and from Turvey and Mohring (1975), who consider fully loaded vehicles. Since the vehicle load factor is the crucial variable, the disutility of prolonged waiting times due to boarding denials and the disutility of discomfort resulting from crowding have a lot in common conceptually. As crowding costs increase for longer distances, the inclusion of crowding would increase the effect observed for the internalisation of boarding denial costs. That is, the average fare would increase in an even steeper way as a function of trip distance than displayed in Figure 8b (internalisation of effects 1, 2, and 3).

7.0 Conclusion

In this study, the agent-based simulation MATSim was successfully used for the optimisation of public transport fares. MATSim was enhanced by implementing a marginal social cost-pricing approach: External delay effects among public transport users are internalised by charging the equivalent monetary amount to the causing agent. As sources for external cost, this study considers the effect of boarding/alighting passengers on in-vehicle time (effect 1), waiting time without an active capacity constraint (effect 2), and waiting time induced by an active capacity constraint (effect 3).

As expected, marginal social cost pricing yields a higher social welfare than charging an optimised flat fare. Starting from effect 1, we find that including more external effects increases the social welfare gain of applying optimised user-specific bus fares. As passengers' waiting times indicate a coordinated arrival at transit stops (following the arrival of buses), it is crucial to include the external effect of prolonged waiting times due to boarding and alighting passengers (effect 2). Adding this effect to the external effect of prolonged in-vehicle times (effect 1) increases the level of social welfare, and also affects the relationship between average optimal fare and travel distance: Since passengers who board at early transit stops affect all passengers that are or will be waiting for the same vehicle along the route, average external costs are higher for early boardings. Thus, for a longer trip length, the average optimal fare decreases to a lesser extent as a function of trip distance. Once passengers are not able to board the first arriving transit vehicle due to limited capacities, social welfare rises most significantly when internalising these delays by charging the extra waiting cost to the passengers inside the vehicle (effect 3). Including that effect increases average external cost-based prices for long-distance trips and early boarding locations along the transit route. In this setup, explicitly accounting for this externality when setting user-specific fares changes the welfare optimal headway from 5 min to 9 min. During peak periods, average external cost-based fares are observed to be higher

than off-peak. As a consequence, users are priced out. They either reschedule their activities and shift their trip departure times to periods with less demand, or they switch to the alternative transport mode (car). Benefits from a higher efficiency within the public transport sector (shorter in-vehicle and waiting times) overcompensate for lower benefits of users who are priced out.

In future research studies, we plan to isolate each choice dimension's contribution to the overall improvement of social welfare. Further steps are to extend our approach to a real-world transport network (for example, Berlin (Neumann *et al.*, 2014)) and include more effects associated with public transport users (for example, crowding). Furthermore, we will carry out an integrated study in which the methodology presented in this paper is coupled with pricing approaches that account for additional external effects, namely road congestion (Lämmel and Flötteröd, 2009; Kaddoura and Kickhöfer, 2014) and environmental effects, such as air pollution (Kickhöfer and Nagel, 2013).

References

ATC (2006): 'National guidelines for transport system management in Australia', Technical Report, Australian Transport Council.

Charypar, D. and K. Nagel (2005): 'Generating complete all-day activity plans with genetic algorithms', *Transportation*, 32, 369–97.

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

de Palma, A., M. Kilani, and S. Proost (2015): 'Discomfort in mass transit and its implication for scheduling and pricing', *Transportation Research Part B*, 71, 1–18.

Douglas, N. and G. Karpouzis (2005): 'Estimating the cost to passengers of station crowding', 28th Australasian Transport Research Forum (ATRF), Sydney.

Haywood, L. and M. Koning (2013): 'Estimating crowding costs in public transport', Discussion Paper 1293, DIW Berlin.

Jansson, K. (1993): 'Optimal public transport price and service frequency', *Journal of Transport Economics and Policy*, 27, 33–50.

Kaddoura, I. and B. Kickhöfer (2014): 'Optimal road pricing: towards an agent-based marginal social cost approach', Working Paper 14-1, TU Berlin, Transport Systems Planning and Transport Telematics.

Kaddoura, I., B. Kickhöfer, A. Neumann, and A. Tirachini (2014): 'Agent-based optimisation of public transport supply and pricing: impacts of activity scheduling decisions and simulation randomness', *Transportation*, forthcoming.

Kickhöfer, B., D. Grether, and K. Nagel (2011): 'Income-contingent user preferences in policy evaluation: application and discussion based on multi-agent transport simulations', *Transportation*, 38, 849–70.

Kickhöfer, B., F. Hülsmann, R. Gerike, and K. Nagel (2013): 'Rising car user costs: comparing aggregated and geo-spatial impacts on travel demand and air pollutant emissions', in T. Vanoutrive and A. Verhetsel (eds), *Smart Transport Networks: Decision Making, Sustainability and Market Structure*, NECTAR Series on Transportation and Communications Networks Research, Edward Elgar Publishing Ltd, 180–207.

Kickhöfer, B. and K. Nagel (2013): 'Towards high-resolution first-best air pollution tolls', *Networks and Spatial Economics*, 1–24.

Kraus, M. (1991): 'Discomfort externalities and marginal cost transit fares', *Journal of Urban Economics*, 29, 249–59.

Lam, W. H. K., C.-Y. Cheung, and C. F. Lam (1999): 'A study of crowding effects at the Hong Kong light rail transit stations', *Transportation Research Part A*, 33, 401–15.

Lämmel, G. and G. Flötteröd (2009): 'Towards system optimum: finding optimal routing strategies in time-dependent networks for large-scale evacuation problems', in B. Mertsching, M. Hund, and Z. Aziz (eds), *Advances in Artificial Intelligence*, Springer, 532–9.

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

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

Nagel, K. and G. Flötteröd (2012): 'Agent-based traffic assignment: going from trips to behavioural travelers', in R. Pendyala and C. Bhat (eds), *Travel Behaviour Research in an Evolving World — Selected Papers from the 12th International Conference on Travel Behaviour Research*, International Association for Travel Behaviour Research, chapter 12, 261–94.

Nash, C. (2003): 'UNITE (UNIfication of accounts and marginal costs for Transport Efficiency)', Final Report for Publication, funded by 5th Framework RTD Programme, Technical Report.

Neumann, A. and K. Nagel (2010): 'Avoiding bus bunching phenomena from spreading: a dynamic approach using a multi-agent simulation framework', Working Paper 10-08, TU Berlin, Transport Systems Planning and Transport Telematics.

Neumann, A., I. Kaddoura, and K. Nagel (2013): 'Mind the gap — passenger arrival patterns in multi-agent simulations', presented at the Conference on Agent-Based Modeling in Transportation Planning and Operations.

Neumann, A., M. Balmer, and M. Rieser (2014): 'Converting a static macroscopic model into a dynamic activity-based model for analyzing public transport demand in Berlin', in M. Roorda and E. Miller (eds), *Travel Behaviour Research: Current Foundations, Future Prospects*, International Association for Travel Behaviour Research (IATBR), Toronto, Canada, chapter 7, 151–76.

Parry, I. W. H. and K. A. Small (2009): 'Should urban transit subsidies be reduced?', *American Economic Review*, 99, 700–24.

Proost, S. and K. van Dender (2001): 'The welfare impacts of alternative policies to address atmospheric pollution in urban road transport', *Regional Science and Urban Economics*, 31, 383–411.

Prud'homme, R., M. Koning, L. Lenormand, and A. Fehr (2012): 'Public transport congestion costs: the case of the Paris subway', *Transport Policy*, 21, 101–9.

Raney, B. and K. Nagel (2006): 'An improved framework for large-scale multi-agent simulations of travel behaviour', in P. Rietveld, B. Jourquin, and K. Westin (eds), *Towards Better Performing European Transportation Systems*, Routledge, 305–47.

Rieser, M. (2010): 'Adding transit to an agent-based transportation simulation concepts and implementation', PhD thesis, TU Berlin, Transport Systems Planning and Transport Telematics.

Small, K. A. and E. T. Verhoef (2007): *The Economics of Urban Transportation*, Routledge.

Tirachini, A. (2014): 'The economics and engineering of bus stops: spacing, design and congestion', *Transportation Research Part A*, 59, 37–57.

Tirachini, A., D. A. Hensher, and J. M. Rose (2014): 'Multimodal pricing and optimal design of urban public transport: the interplay between traffic congestion and bus crowding', *Transportation Research Part B*, 61, 33–54.

Turvey, R. and H. Mohring (1975): 'Optimal bus fares', *Journal of Transport Economics and Policy*, 9, 280–6.

van Essen, H., B. Boon, A. Schroten, M. Otten, M. Maibach, C. Schreyer, C. Doll, P. Jochem, M. Bak, and B. Pawlowska (2008): 'Internalisation measures and policy for the external cost of transport', Technical Report, CE Delft.

Verhoef, E. (2001): 'Marginal cost based pricing in transport. Key implementation issues from the economic perspective', paper presented at Seminar One of IMPRINT-EUROPE.

Wardman, M. and G. Whelan (2011): 'Twenty years of rail crowding valuation studies: evidence and lessons from British experience', *Transport Reviews*, 31, 379–98.

Wright, L. and W. Hook (2007): 'Bus rapid transit planning guide', Technical Report, Institute for Transportation and Development Policy (ITDP), New York.