Doctoral Thesis

Urban mobility pricing with heterogeneous users

Author(s):
Chakirov, Artem

Publication Date:
2016

Permanent Link:
https://doi.org/10.3929/ethz-a-010638579

Rights / License:
In Copyright - Non-Commercial Use Permitted
DISS. ETH NO. 23174

URBAN MOBILITY PRICING WITH HETEROGENEOUS USERS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ARTEM CHAKIROV

Diplom-Ingenieur, Technische Universität München

born on 19.03.1983

citizen of

Germany

accepted on the recommendation of

Prof. Dr. Kay W. Axhausen, examiner
Prof. Dr. Erik Verhoef, co-examiner
Dr. Alexander Erath, co-examiner

2016
Contents

Abstract xi

Zusammenfassung xiii

Acknowledgments 1

1 Introduction 3

2 Literature Review and Motivation 11

2.1 Urban mobility pricing 12

2.1.1 Pricing differentiation and discrimination 13

2.1.2 Road and congestion pricing 14

2.1.3 Public transport pricing 19

2.2 Heterogeneous values of time 24

2.3 Social equity - Accessibility, Activities, Alternatives 27

3 Methodology 31

3.1 Multi Agent Transport Simulation (MATSim) 33

3.2 Public Transport in MATSim 35

3.3 Congestion pricing 36

3.4 Economic Evaluation 39

3.4.1 Expected Maximum Utility 40
Urban Mobility Pricing with Heterogeneous Users 2015

3.4.2 Generalized cost and realised utility in an agent-based SUE . . . . . . 43
3.4.3 Public Transport Operation Cost . . . . . . . . . . . . . . . . . . . . . 43
3.4.4 Social welfare and consumer surplus . . . . . . . . . . . . . . . . . . . 44
3.5 Simulation methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45

4 Values of time and schedule delay 47

4.1 Utility maximisation in MATSim context . . . . . . . . . . . . . . . . . . . . 47
4.2 Including heterogeneity into the activity-based model . . . . . . . . . . . . 50
4.3 Schedule Delay Modelling – Being Early or Late . . . . . . . . . . . . . . 55
4.3.1 Schedule delay in an activity – based context . . . . . . . . . . . . . . 57
4.3.2 Three forms of schedule delay heterogeneity . . . . . . . . . . . . . . 60

5 Experimental Scenario Set-up 65

5.1 Corridor scenario . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65
5.1.1 Supply . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66
5.1.2 Travel demand and behavioural parameters . . . . . . . . . . . . . . . 67
5.1.3 Socio-economic characteristics . . . . . . . . . . . . . . . . . . . . . . . 69
5.2 Enriched Sioux Falls scenario . . . . . . . . . . . . . . . . . . . . . . . . . . 77
5.2.1 Supply . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 77
5.2.2 Travel demand and behavioural parameters . . . . . . . . . . . . . . . 79
5.2.3 Socio-economic characteristics . . . . . . . . . . . . . . . . . . . . . . . 81

6 Congestion Pricing: Simulation Results and Discussion 83

6.1 System convergence and stability . . . . . . . . . . . . . . . . . . . . . . . . 85
6.2 Social and Consumer Welfare . . . . . . . . . . . . . . . . . . . . . . . . . . 87
6.2.1 Unimodal scenario - car only . . . . . . . . . . . . . . . . . . . . . . . . 88
6.2.2 Multi-modal scenario with varying levels of bus service . . . . . . . . 96
6.3 Distributional Effects ............................................. 101

6.3.1 Second heterogeneity dimension - varying $\mu$ and $\eta$ ratios ... 104

6.4 Fare and toll revenues ........................................... 106

6.5 Scheduling preferences ........................................... 107

6.6 Travel times and travel speeds ................................. 110

6.7 Effects of trip distance and travel mode ...................... 111

7 Public Transport Pricing: Simulation Results and Discussion 115

7.1 Optimal flat fare .................................................. 116

7.2 In-vehicle valuation of time under crowding ................. 122

7.2.1 Crowding in an agent-based context ...................... 122

7.3 Dynamic pricing of public transport externalities .......... 124

7.3.1 Agent-based cost internalisation approach ............... 125

7.3.2 Challenges in the welfare evaluation ..................... 128

7.4 Social and consumer welfare .................................. 128

7.4.1 Impact of crowding disutility .............................. 128

7.4.2 Benefits of dynamic, marginal cost pricing ............. 129

7.4.3 Monetary revenues .......................................... 132

7.5 Distributional effects of dynamic fares ....................... 134

7.6 Status quo and the importance of modal split .............. 135

7.6.1 Flat public transport fare .................................. 136

7.6.2 Dynamic public transport and congestion pricing ....... 137

7.7 Large scale scenarios and alternative activity chains ....... 140

8 Conclusion and Policy Implications 147

8.1 Multi-modal modelling approach and user heterogeneity .... 148
8.2 Application to a large-scale scenario .............................................. 151
8.3 Planning practice and outlook ....................................................... 152
Bibliography .................................................................................. 169
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Overview and context of principal topics in this work</td>
<td>8</td>
</tr>
<tr>
<td>3.1</td>
<td>Iterative MATSim cycle</td>
<td>33</td>
</tr>
<tr>
<td>3.2</td>
<td>Discrete, step-based link charge vs. continuous link charge, averaged over the 5 minutes period, following the time at which an agent entered a tolled link</td>
<td>59</td>
</tr>
<tr>
<td>4.1</td>
<td>Marginal utilities function in MATSim context for the morning commute</td>
<td>59</td>
</tr>
<tr>
<td>5.1</td>
<td>Corridor scenario set-up, with bus stops located every 600 meters and home and work locations normally distributed on the either end of the corridor</td>
<td>66</td>
</tr>
<tr>
<td>5.2</td>
<td>Histogram of incomes with dashed line indicating the mean of 73,320 USD</td>
<td>70</td>
</tr>
<tr>
<td>5.3</td>
<td>Lorenz curves and Gini coefficients for different heterogeneity parameters</td>
<td>72</td>
</tr>
<tr>
<td>5.4</td>
<td>Distribution of Value of Time for different heterogeneity factors (car mode)</td>
<td>72</td>
</tr>
<tr>
<td>5.5</td>
<td>Fitted density distributions of schedule delay cost for different types and degrees of heterogeneity</td>
<td>75</td>
</tr>
<tr>
<td>6.1</td>
<td>Convergence of average score (utility) with linear annealing of replanning rates in MATSim for runs with 1000 and 2000 iterations (proportional heterogeneity, 5 min bus service headway)</td>
<td>87</td>
</tr>
<tr>
<td>6.2</td>
<td>Effect of congestion pricing on social welfare and consumer surplus for different degrees of heterogeneity (n) and different types of schedule-delay heterogeneity, with only car mode available. Base case ((n = 0)) is the scenario with homogeneous travellers</td>
<td>89</td>
</tr>
<tr>
<td>6.3</td>
<td>Effect of congestion pricing on social and consumer welfare for different degrees of proportional heterogeneity, defined by factor (n) and additional ratio heterogeneities of (\mu) and (\eta)</td>
<td>98</td>
</tr>
<tr>
<td>6.4</td>
<td>Welfare effects of congestion pricing, dependent on availability of public transport with proportional heterogeneity</td>
<td>99</td>
</tr>
</tbody>
</table>
6.5 Changes in consumer welfare dependent on income after introduction of congestion pricing for different levels of proportional heterogeneity.

6.6 Boxplot of consumer welfare changes as a result of congestion pricing for two different levels of bus service.

6.7 Mode shares of car trips before and after congestion pricing (proportional heterogeneity, 2 min bus headway, last iteration).

6.8 Boardings, alighting and occupancy for the east bound bus service starting at 07:36 a.m. (proportional heterogeneity, 2 min bus headway, n = 3).

6.9 Consumer welfare changes from congestion pricing in $, for the unimodal scenario.

6.10 Revenues from congestion pricing and bus fares for the scenario with \( \gamma \) heterogeneity.

6.11 Departure rates and average values of time for each departure time, aggregated over 10 min bins (proportional heterogeneity, no bus service).

6.12 Average congestion charge paid per trip dependent on the trip departure time, aggregated over 10 min bins (proportional heterogeneity, no bus service).

6.13 Mean travel times for different degrees of proportional heterogeneity n and levels of bus service, before and after congestion pricing. Base case (n = 0) is the scenario with homogeneous travellers.

6.14 Mean travel speeds for different degrees of proportional heterogeneity n and levels of bus service, before and after congestion pricing. Base case (n = 0) is the scenario with homogeneous travellers.

6.15 Departure time vs. avg. in vehicle distance, aggregated over 10 min bin for morning and evening rush hours (proportional heterogeneity, n = 1, no bus service).

6.16 Travel distance dependent consumer benefits of congestion pricing: homogeneous case vs. proportional heterogeneity (n = 5).

7.1 Social welfare dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

7.2 Mode share dependent on the bus fare level for proportional heterogeneity with \( n = 1, 5 \) min bus headway.

7.3 Consumer welfare dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

7.4 Operator revenues dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

7.5 Congestion pricing revenue dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

7.7 Social and consumer welfare with crowding disutility and internalisation of public transport externalities for proportional heterogeneity (5 min bus headway).

7.8 Revenues from internalisation of public and private transport externalities for proportional heterogeneity (5 min bus headway).

7.9 Bus fare distribution for proportional heterogeneity with \( n = 1 \) (5 min bus headway).
7.10 Bus fare distribution for proportional heterogeneity with $n = 5$ (5 min bus headway). ........................................... 134
7.11 Changes in consumer welfare dependent on income after introduction of marginal, dynamic bus fares for different levels of proportional heterogeneity. ........................................... 135
7.12 Social welfare dependent on bus fare level and degree of proportional heterogeneity with fixed car cost of $30$ $\$/$ per day (5 min bus headway). ........................................... 137
7.13 Operator revenues dependent on bus fare level and degree of proportional heterogeneity with fixed car cost of $30$ $\$/$ per day (5 min bus headway). ........................................... 137
7.14 Social and Consumer welfare with crowding disutility and internalisation of public transport externalities for proportional heterogeneity and fixed car cost of $30$ $\$/$ per day (5 min bus headway). ........................................... 138
7.15 Social and consumer welfare with crowding disutility and internalisation of public transport externalities for proportional heterogeneity and fixed car cost of $30$ $\$/$ per day (5 min bus headway). ........................................... 139
7.16 Changes in consumer welfare dependent on income after introduction of marginal, dynamic bus fare for different levels of proportional heterogeneity and fixed car cost of $30$ $\$/$ per day (5 min bus headway). ........................................... 139
7.17 Revenues from internalisation of public transport and congestion externalities for proportional heterogeneity (5 min bus headway). ........................................... 140
7.18 Effect of congestion pricing on social welfare and consumer surplus for different degrees of heterogeneity in the Sioux Falls scenario. ........................................... 142
7.19 Effect of congestion pricing on social welfare and consumer surplus for different activity chains and varying degrees of heterogeneity in the Sioux Falls scenario. ........................................... 144
7.20 Sioux Falls mode share depended on the degree of heterogeneity and individual’s activity chain (NCP). ........................................... 144
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Bus operation cost according to Australian Transport Council (2006)</td>
<td>44</td>
</tr>
<tr>
<td>5.1</td>
<td>Bus operation cost</td>
<td>67</td>
</tr>
<tr>
<td>5.2</td>
<td>Activity constraints</td>
<td>68</td>
</tr>
<tr>
<td>5.3</td>
<td>Behavioural and monetary simulation parameters for the Corridor scenario</td>
<td>69</td>
</tr>
<tr>
<td>5.4</td>
<td>Overview of heterogeneity types</td>
<td>71</td>
</tr>
<tr>
<td>5.5</td>
<td>Value of time ($/hr) distributions for different heterogeneity factors</td>
<td>73</td>
</tr>
<tr>
<td>5.6</td>
<td>Schedule delay early values for different heterogeneity factors in ($/hr)</td>
<td>75</td>
</tr>
<tr>
<td>5.7</td>
<td>Schedule delay late values for different heterogeneity factors in ($/hr)</td>
<td>76</td>
</tr>
<tr>
<td>5.8</td>
<td>Sioux Falls scenario: activity constraints</td>
<td>81</td>
</tr>
<tr>
<td>5.9</td>
<td>Behavioural and monetary simulation parameters for the Sioux Falls scenario</td>
<td>81</td>
</tr>
<tr>
<td>6.1</td>
<td>Parameters of simulation sets</td>
<td>84</td>
</tr>
<tr>
<td>6.2</td>
<td>Overview of abbreviations and terminology</td>
<td>86</td>
</tr>
<tr>
<td>6.3</td>
<td>Effects of congestion charge on social welfare (EMU) for proportional, $\alpha$ and $\gamma$ heterogeneity in the unimodal &quot;car only&quot; scenario</td>
<td>91</td>
</tr>
<tr>
<td>6.4</td>
<td>Effects of congestion charge on consumer welfare (EMU) for proportional, $\alpha$ and $\gamma$ heterogeneity in the unimodal &quot;car only&quot; scenario</td>
<td>91</td>
</tr>
<tr>
<td>6.5</td>
<td>Effects of congestion charge on realised social welfare and consumer welfare for proportional heterogeneity and &quot;car only&quot; scenario, averaged over last 100 iterations</td>
<td>93</td>
</tr>
<tr>
<td>6.6</td>
<td>Effects of congestion charge on consumer welfare social welfare and consumer welfare for $\alpha$ heterogeneity and &quot;car only&quot; scenario, averaged over last 100 iterations</td>
<td>93</td>
</tr>
<tr>
<td>6.7</td>
<td>Effects of congestion charge on realised travel and schedule delay cost for proportional heterogeneity and &quot;car only&quot; scenario, averaged over last 100 iterations</td>
<td>94</td>
</tr>
<tr>
<td>7.1</td>
<td>Parameters for the scenario with crowding disutility</td>
<td>124</td>
</tr>
</tbody>
</table>
Abstract

Efficient allocation of transport supply and demand is a key element of any well-functioning modern city. Use of mobility pricing as a transport demand management tool can assist to achieve this goal by helping to provide for an economically and socially optimal use of the available infrastructure. However, pricing of public goods can affect equity and have unintended redistribution effects, making its application challenging and unpopular with the public. This is primarily due to the complexity of optimal pricing problem and the limitations of travel behaviour and transport demand allocation models. Modelling and predicting impact of pricing policies requires capturing user’s heterogeneity, understanding the distribution of diverse transport needs and assessing the available alternatives, in all of their dimensions and interdependencies.

This thesis applies an activity- and agent based simulation framework (MATSim) to study the effects of optimal congestion and public transport pricing on social and consumer welfare in a multi-modal context and in the presence of user heterogeneity in travel time valuations and trip characteristics. Here, MATSim’s ability to model interaction between public and private transport, including congestion delays and boarding denials due to capacity constraints, proves to be essential.

In presenting a new approach for the inclusion of value of time and schedule delay heterogeneity as well as values of in-vehicle crowding into the MATSim framework, this thesis focuses on two major aspects: the impact of the degree of heterogeneity on the benefits from congestion and dynamic public transport pricing policies and the impact of availability of alternative modes of transport on equity. Using multi-modal Corridor and extended Sioux Falls scenarios, the results of this study confirm the potential for dynamic mobility pricing schemes based on internalisation of external costs, to create significant gains in social and consumer welfare. The scale of these gains depends on the scale of road congestion, public transport crowding and demand distribution between the available modes; however, all of these factors are affected by the degree of user heterogeneity. Increasing heterogeneity in the presence of multiple modes with different travel costs can lead to self-organization effects, thereby diminishing the gains of congestion pricing policies. Furthermore, effects on consumer welfare for individuals with low values of time are highly dependent on the availability and service
level of an alternative mode. These findings can have significant impact on urban transport policy design when applied to a city-specific transport system and a population with a higher degree of economic and social inequality.
Zusammenfassung


Diese Arbeit verwendet eine aktivitäts- und agentenbasierte Simulation (MATSim), um die Auswirkungen der optimalen Straßennutzungsgebühren, sowie der dynamischen Tarife im öffentlichen Verkehr in einem multimodal kontext und unter Berücksichtigung der heterogenen Zeitwerte und differenzierten Präferenzen der Reisenden, zu untersuchen. Dabei erweist sich die Fähigkeit von MATSim, die Interaktion zwischen dem öffentlichen und privaten Verkehr, einschließlich Stauverzögerungen und Kapazitätsbeschränkungen der Fahrzeuge, detailliert zu modellieren, als unerlässlich. Darüber hinaus integriert diese Arbeit die negativen Effekte überbesetzter Busse in die agentenbasierte Simulation und nutzt diese, um anhand der Grenzkosten festgelegte dynamische Bustarife zu evaluieren.

Um die Heterogenität der Nutzer zu berücksichtigen, präsentiert diese Arbeit einen neuen Ansatz zur Integration der individuellen Zeitwerte in die Nutzenfunktion der einzelnen Aktivitäten. Zwei Hauptaspekte stehen dabei im Mittelpunkt: die Auswirkungen des Heterogenitätsgrades auf das ökonomische Nutzen der dynamischen Bepreisungspolitik, sowie die Auswirkungen der Verfügbarkeit von alternativen Verkehrsmittel auf die Wohlfahrt des Verbrauchers und deren Verteilung. Die Ergebnisse dieser Studie bestätigen das erhebliche Potenzial der, auf der Internalisierung externer Kosten basierenden dynamischen
Acknowledgments

This work was carried out during the years 2011-2015 at the Future Cities Laboratory (FCL), Singapore-ETH Centre and at the Institute for Transport Planning and Systems (IVT), Swiss Federal Institute of Technology (ETH Zurich).

I would like to express my special appreciation and deepest gratitude to Prof. Dr. Kay W. Axhausen, who gave me his continuous support, guidance and freedom in the development of my research. His insightful advice and vast knowledge in the interdisciplinary field of transport planning and economics have been an invaluable source of inspiration and motivation all throughout my work.

I am also very much indebted to my co-examiners: Prof. Dr. Erik Verhoef and Dr. Alexander Erath, who supported me during this thesis with their expertise and provided many valuable and insightful comments. Their guidance, detailed and constructive feedback as well as strong encouragement helped me tremendously to sharpen my existing ideas and to develop many new ones.

I would also like to thank all my colleagues and friends at the Future Cities Laboratory, who helped to create a truly interdisciplinary, dynamic, stimulating and simply extraordinary working environment. The time I have spent at FCL will remain in my memory as full of inspirational moments and unforgettable experiences.

My special appreciation goes to the members of Transportation and Mobility team at the Future Cities Laboratory and Transport Planning Group at ETH Zurich whose support has been instrumental for this thesis. Without their help this work could have never been accomplished. I am very thankful for their advice, multitude of ideas and valuable inputs. A special thanks goes to Pieter Fourie, Sergio Ordoñez and Michael van Eggermond for their generous help, friendly assistance, fun company and countless great moments.

I also wish to express my sincere gratitude to the whole MATSim community and all developers, who contributed countless hours of research and development to this innovative project. In particular, generous support by Dr. Christoph Dobler, Dr. Michael Balmer and Dr. Marcel Rieser, who readily shared their expertise, experiences and code, has proven to
be tremendously helpful in understanding, applying and developing MATSim. I am also very thankful Prof. Kai Nagel and members of his group at the Institute of Transport Systems Planning and Transport Telematics at TU Berlin for fruitful discussions, informative input and numerous contributions to MATSim project, which served as a valuable base and inspiration for this work.

A special thanks goes to Adrianne Joergensen, Alina Klaus, Pieter Fourie and Alexander Erath for their tremendous help in the proofreading and editing of the manuscript. Last but not least, I would like to thank my family and all my friends, near and far, for the continuous encouragement, understanding and motivation. Their love, patience and support made this work possible.
Chapter 1

Introduction

"The difference between planning for automobility versus accessibility is the difference between planning for movement versus planning for people and places."

Robert Cervero (Cervero, 1997)

Transport planning is traditionally concerned with design, evaluation, adoption and optimisation of transport infrastructure and policy to meet existing and future transport demand. This demand is driven by travel decisions of individuals who need to move from one location to another in order to be able to engage in activities, which could not have been performed otherwise. By enabling economic activity and social interaction, transport infrastructure provides benefits to society as a whole and should be designed and operated in a socially optimal way. Individual travel or mobility needs usually vary according to lifestyle, income level, age or personal taste, and can significantly change over time. Balancing satisfaction of the diverse mobility needs and cost of transport provision is a delicate and challenging task. Access to transport on the one hand and its externalities and cost on the other have a major impact on quality of life, economic welfare, well-being and equality of opportunities in both, developing and developed societies across the globe.

Today cities are already housing more than a half of worlds population and proportion of urban dwellers continues to rise (UN, 2013). Rapid urbanisation concentrates different lifestyles and their likewise diverse mobility needs into dense urban agglomerations. Accounting for these needs and providing reliable and efficient urban transport infrastructure for all members of society is essential to economic welfare, productivity (Venables, 2007), social and economic equality (Lucas, 2012) and well-being (De Vos et al., 2013).

Addressing the divergent transport needs of urban dwellers within the constraints of modern cities and metropolitan areas requires innovative solutions that go beyond mere construction,
road space extension, and provision of more extensive public transport. To that end, advances of digital technology and information systems enable and simplify application of new, innovative demand management techniques. Next to personalised information provision, pricing has been identified as a powerful demand management tool, capable of addressing market inefficiencies and enabling more optimal use of the existing infrastructure.

Pricing of goods and services related to transport can take many different forms, such as road pricing, fuel taxes, parking costs or public transport charges. With mobility referring to the physical movement of people (Litman [2011]), mobility pricing in general characterises the combination and integration of all types of transport related pricing (Vrtic et al. [2007]). However, finding an optimal integrated mobility pricing policy is a highly complex and challenging task. Pricing different aspects of mobility results in a number of varying long- and short-term effects with often broader economic and social implications, such as e.g. interaction with labour or real-estate markets. Hence, from the transport demand management perspective more flexible and differentiated pricing forms represent an attractive alternative to flat taxation mechanisms. In particular, dynamic road pricing and public transport charging schemes enable targeted and differentiated application and as the digital technology evolves, are gaining popularity among transport economists and planners.

Yet, the design, communication, and implementation of differentiated mobility pricing schemes for use of transport infrastructure and services, has proven to be a challenging task. Pricing of public infrastructure, such as roads, is often regarded as a form of regressive taxation. However, informed pricing scheme design with regard to local transport system, land-use, activity patterns and other spatial and economic characteristics of the place and the population, enables the planner to come up with progressive, neutral or regressive pricing schemes (Santos and Rojey [2004]). This requires evaluation and assessment of policy impact, going beyond aggregate measures of potential efficiency gains. A disaggregate approach to distributional effects, accessibility and policy impacts on individual mobility becomes even more important in consideration of transport-related inequality and social exclusion.

One of the major challenges in designing an economic instrument for influencing and changing individual behaviour is the assessing the heterogeneity of economic agents and predicting their reactions to provided policy incentives. Heterogeneous preferences among travellers are based on a variety of factors: age, income, sex, family situation, work schedule, and attitudes and lifestyles. The figure that commonly combines heterogeneous user preferences in one measure, is the monetary value placed on the time gains resulting from reduction in travel times. It is also referred to as value of time (VoT) or value of travel time savings (VTTS) (Small and Verhoef [2007]). Discussed in more detail in section 4, values of time are often linked to personal or household incomes, but can also vary with travel conditions, trip characteristics and personal preferences (Axhausen et al., 2008). A recent overview and discussion of the valuation of travel...
time was given by Small (2012). For sake of mathematical convenience and simplification, most transport models apply the same average values of time to all travellers. Though appropriate in some cases, such an aggregated approach fails to provide detailed results for study and evaluation of transport pricing policies. As Hensher and Goodwin (2004) point out, next to the average value of VTTS, accounting for its distribution is decisive for unbiased analysis and result evaluation, e.g. when forecasting market share. Informed by empirical impact measurements of heterogeneous user characteristics and travel behaviours, transport models used for planning and evaluation must be able to assess equity implications and identify the winners and losers of the policy measures in question. This is not only important for a sensible, efficient and fair transport system development and policy-making, but also for its clear public communication. Engaging the general public and visualising benefits of proposed measures on disaggregate level is decisive for gaining public support.

Most approaches to understanding the interplay between pricing policies and heterogeneous user preferences employ dynamic economic models of traffic congestion, often building upon and extending the classic bottleneck model, as initially presented by Vickrey (1969). Taking a single-trip perspective, the bottleneck model connects the distribution of departure and arrival rates with corresponding travel times given the capacity constrains of the transport infrastructure. Marginal utilities of activities, which play a crucial part in trip timing decisions, are only considered indirectly. Full day scheduling dynamics are usually omitted. The policy impact on travellers’ well-being is measured in terms of consumer surplus, or the difference between total trip cost (given the aggregated transport demand) and the maximum trip cost travellers would be willing to accept.

Based on Vickrey’s bottleneck model and its dynamic extensions, various subsequent publications have shown the major effects of heterogeneous values of time on outcomes of economic policy and project appraisal. Among others, Bates (1996), Arnott et al. (1988) and van den Berg and Verhoef (2011a) highlight that the type and extent of heterogeneity strongly influence the welfare effects of congestion tolling and different tolling schemes. They forcefully argue for the need to include heterogeneity in demand models with trip-timing choice. Ignoring it in design of transport pricing schemes and policies can lead to biased prediction and unintended redistribution effects with significant social and economic implications. These risks are especially prevalent in urban environments where a multitude of activity choices and alternative modes with varying travel cost can be considered as imperfect substitutes. In these scenarios, the evaluation of pricing policies with regard to heterogeneous traveller preferences becomes even more vital, but also more complex. Only a few publications address the impact of heterogeneity in a multi-dimensional, multi-modal context (e.g. van den Berg and Verhoef 2013). The complexity of the problem makes
the application of deterministic, analytical frameworks challenging, as these are not well suited to deal with typically skewed distributions of user preferences with multiple levels of heterogeneity and struggle to incorporate all relevant choice dimensions. In particular, for cases of real world scenarios involving dense metropolitan regions, an analytical approach is often unable to provide the level of detail and resolution required for designing differentiated transport demand management policies.

An alternative, more holistic approach for modelling of heterogeneous travel behaviour on the one side and complex interaction between transport demand and supply on the other, is based on individual activity patterns and disaggregated simulation of interactions between all relevant economic agents. In contrast to the trip-based approach, the activity-based perspective on travel demand modelling focuses on individuals’ activity schedules as drivers of transport demand. Such an approach enables an explicit representation of diverse mobility needs and associated constraints. Based on the theory of consumer choice, activity scheduling decisions and associated travel behaviour are usually assumed to follow the economic principle of individual utility maximisation (Axhausen and Gärling, 1992). Here, discrete choice modelling theory represents an essential and established tool for modelling of individual behaviour (Ben-Akiva and Lerman, 1985; Train, 2003). Together with a multi-agent simulation, which operates on the level of individuals with specific socio-economic and demographic characteristics, this approach enables the representation of otherwise intractable states and outcomes. It also allows to achieve these results without imposing the unrealistic demanding conditions, often required for an analytical solution.

Compared with traditional, aggregate models, also known as 4-step models, an agent- and activity-based approach prioritizes a more user-centred perspective on transport demand. Focusing on individuals’ characteristics and activity schedules, it inherently operates at a higher resolution of spatial and temporal characteristics, and is therefore highly suitable for the design and optimisation of transport demand management policies, such as dynamic mobility pricing.

Comprehensive models of individual travel decisions coupled with detailed models of interaction between transport supply and demand enable new insights into collective behaviour and emergent phenomena. Hence, the disaggregated approach of an agent-based simulation framework simultaneously allows for choices of mode, departure time, or location with high degrees of spatial and temporal resolution, thereby facilitating the inclusion of heterogeneous user preferences based on agent-specific attributes. Due to the exponential growth of computational power in recent decades, the simulation of such systems at a large scale has become increasingly feasible and affordable. This has boosted the development of integrated agent- and activity-based simulation models as promising alternative approaches to the diverse urban challenges of transport needs and preference heterogeneity.

However, despite being advocated as one of major advantages and strengths of an agent-based
Modelling approach, so far only limited effort has been made to include truly heterogeneous user preferences and values of time into current modelling and simulation frameworks (e.g. Horni 2013; Kickhöfer 2014). This is partially due to the lack of disaggregated, empirical data, such as estimates for inter-personal diversity of behavioural characteristics, and individual utilities for performing work and leisure activities according to time of day, activity type, or personal attributes. Furthermore, there are still a number of existing challenges in an adequate software design capable of incorporating multi-dimensional user heterogeneity.

Use of an agent- and activity based simulation framework to study effects of heterogeneous user preferences on pricing policies in a uni- and multi-modal context is the main topic of this work. By developing a concept for inclusion of income-dependent values of time and schedule delay into an agent- and activity-based transport modelling paradigm, this thesis investigates their impact on economic welfare assessment and distributional effects of congestion pricing policy for private and public transport. The availability and level of public transport services proves to be decisive factors in determining the impact of heterogeneous preferences, either with or without congestion pricing policies in place. Using a multi-modal Corridor and Sioux Falls scenarios implemented in a multi-agent transport simulation framework MATSim (MATSim 2015), this work engages its ability to capture the physical interactions between private and public transport while incorporating the capacity constrains of public transport vehicles and detailed simulation of boarding and alighting processes. Building upon these capabilities, the existing model is further extended by exploring the impact of in-vehicle crowding, with passenger discomfort and operational service delays associated with it. Such details, which are so often neglected in macro- and mesoscopic simulation models, play a decisive role in the distribution of gains and losses resulting from the introduction of congestion and dynamic public transport pricing policies.

Another, very important aspect of policy and project assessment is the modelling of all major degrees of freedom, available to an individual to react to changes in transport policy and supply. As e.g. Zöllig and Axhausen (2012) demonstrate, evaluating travel times savings without departure time and location choice dimensions in an accessibility-based assessment of infrastructure investments, does not capture all utility gains. The ability to include these additional yet relevant choice dimensions within a co-evolutionary agent and activity-based framework is an important aspect of this approach.

Figure 1.1 visualises the main components of transport demand and supply, highlighting the focal points of this work. Factors, dimensions and interdependencies targeted in detail are shown in black; the role and impact of grey-coloured elements is not explicitly considered here, but will be incorporated the future research. The thick, grey connection arrows between transport infrastructure, demand management and operations, and transport demand summarise a number of reciprocal correlations, which are not explicitly discussed in the scope of this
work. Furthermore, while considered interdependencies have mainly short-term effects, other relations such as land-use and activity-location choice represent long-term developments and are not considered in this work.

The graph emphasises the focus of this work on the use of mobility pricing as a demand management tool for a multi-modal urban transport networks. Aiming to induce behavioural changes in mode choice and activity scheduling, pricing effects have to be evaluated in an integrated approach and under consideration of all relevant degrees of freedom. Applying an agent-based simulation enables identification of the market equilibrium conditions, given transport demand, supply, pricing policy and operational strategies. The ultimate objective is to understand the forces and dynamics between mobility pricing, public transport service provision, and socio-economic population diversity in order to obtain efficient and fair transport system optimisation.

This work is structured as follows: Chapter 2 places the presented approach into the context of an existing body of literature by touching on topics of transport and mobility pricing, the relevance of heterogeneous preferences, and role of transport planning and policy in social equity and welfare. Subsequently, Chapter 3 describes the applied methodology, focusing on
agent- and activity-based simulation and economic evaluation techniques. The concepts of heterogeneous values of time and schedule delay cost as well as their inclusion in an agent- and activity-based transport simulation framework are discussed in Chapter 4. Chapter 5 presents the experimental simulation set-up as well as the relation between degree of heterogeneity and inequality measures, which is followed by the discussion of simulation results and their implications in Chapter 6. While Chapter 6 focuses exclusively on congestion pricing, Chapter 7 presents the analysis of dynamic public transport fares based on internalisation of external public transport cost. Chapter 7 also presents analysis of mobility pricing effects and heterogeneous users when studied within a more complex Sioux-Falls scenario. In conclusion, Chapter 8 summarises the findings of this thesis and highlights their policy implications.
Chapter 2

Literature Review and Motivation

"Traffic congestion is caused by vehicles, not by people in themselves."

Jane Jacobs (Jacobs, 1961)

As pointed out in the introduction, rapid growth of densely populated urban areas and megacities raised demand for design of innovative transport management solutions and policies in order to ensure an efficient personal mobility provision. Understanding and modelling transport demand and travel behaviour in an urban context in all its diversity, is a key criteria for fulfilling this goal. From the planning perspective this requires focusing on personal preferences, activities and needs as fundamental drivers of individual travel decisions. Such paradigm shift from a flow- to an individual- and from a trip- to an activity - centred approach fosters development of new, disaggregated modelling concepts, simulation tools and data collection techniques. Modelling individual activity chains as the underlying layer of travel demand, coupled with an agent-based simulation approach, allows to account for multiple dimension of heterogeneous user preferences in a behaviourally consistent manner. This enables design, optimisation and evaluation of advanced transport policies scenarios, such as dynamic and differentiated infrastructure pricing, from the disaggregate perspective. Winners, losers and policy redistribution effects can be identified on the level of individuals and households, with shifts in social equity and accessibility becoming apparent.

In this context, leading up to the methodological framework and experimental set-up of this work, this chapter sets a wider frame and establishes the foundation and motivation of this thesis. Based on the existing body of literature, three major topics are discussed: urban mobility pricing, role of traveller’s heterogeneous preferences and the notion of equity in conjunction with mobility provision and pricing policies.
2.1 Urban mobility pricing

Free markets governed by the unregulated forces of supply and demand do not necessarily lead to the socially optimal outcome. Setting appropriate regulatory framework and providing incentives in order to correct for market failures and inefficiencies, represents one of the central tasks of public policy and administration. In this context, passenger transport has traditionally been a heavily regulated market with a major involvement of a public sector in planning, financing, development and operations of transport infrastructure. Over the past decades, deregulation policies gained popularity in political and academic circles and reforms and privatisation of transport operators were carried out in many parts of the industrialised world. However, optimal scale of public involvement and the most efficient balance between public and private transport provision remain controversial issues. Economies of scale and number of negative externalities such as congestion, pollution or accidents provide a good argument for not leaving transport markets entirely to themselves. Furthermore, given the crucial role of transport in social and economic activity and its importance for social equity and economic development, private and public interests do not always coincide. Development and operating strategies aiming for the increase in private profits can often contradict policies beneficial to social welfare maximisation.

In free markets price of goods or a services emerges from a balance of demand and supply, created by Adam Smith’s "invisible hand of the market". At the same time pricing can also be used as a powerful instrument for market regulation and correction of market distortions. However, the large degree of economic complexity makes market interventions a challenging and mostly an inefficient measure with potentially disastrous consequences for economy and welfare. From the economic theory perspective, a justification for market intervention with pricing mechanism can be provided by the need for internalisation of external costs in order to achieve a socially optimal market equilibrium and increase social welfare.

Market inefficiencies or market failures often arise from presence of externalities - cost imposed on third parties, who are not involved in the decision to consume the particular good or service. In the context of transport infrastructure as a public good with limited capacity, a various sorts of externalities on other users and non-users can occur. Most prominent negative externalities in transport are travel delays, pollution, changes in land value, discomfort during travelling and safety hazards (Anas and Lindsey, 2011). In economics, inclusion of external costs into the total cost born by the consumer, also referred to as internalisation of externalities, is considered to be a powerful policy for achieving a better market outcome and increasing social welfare.

In transport policy, various pricing mechanisms and policies can be applied in order to account for the full cost of personal mobility and achieve a better allocation of transport infrastructure and capacity to the existing travel demand. Such pricing policies include, but are not limited to,
road pricing, public transport fares, parking charges or taxation of fuel and car ownership. Often applied simultaneously and given their different temporal, spatial and social scales, interdependencies between various pricing policies lead to a number of reinforcing as well as opposing effects. Often advocated but rarely implemented, an integrated pricing strategy and urban transport policy design is essential in order to explore synergies and achieve a socially optimal market outcome with minimal resources and limited market intervention (May et al., 2006).

As this work examines pricing policy as a transport demand management tool, operating on short- and medium time scales, the subsequent discussion focuses on road and public transport pricing, highlighting some of the most relevant publications in these fields.

### 2.1.1 Pricing differentiation and discrimination

Transport pricing strategies can pursue different, mostly conflicting objectives: profit maximisation, transport system optimisation for the good of society, social justice and egalitarian access to transport or taxation and income redistribution in a wider economic context. Finding an optimal pricing strategy according to either of these objectives is a highly challenging and complex task. It requires understanding, modelling and predicting human behaviour on the one hand and capturing multi-dimensional supply characteristics on the other. In order to trigger rational reactions yielding a desired outcome, pricing schemes have to be transparent and simple enough to be understood by consumers. However, even in presence of cognitive restrictions, design of mobility pricing strategies offers a great potential and various degrees of freedom for optimisation and innovation. In particular, the advance of personal digital information technology and development of personal advice applications, new ways of comprehensive trip planning and pricing scheme communication are expected to empower travellers and change their travel behaviour (Batty et al., 2012).

Differentiated and dynamic pricing schemes, facilitated by digital technology, enable a highly targeted and precise application of pricing as a well-defined demand management tool. This provides more flexibility and facilitates accurate composition of incentives, yielding better results at lower cost and with reduced side effects. Today, divergent pricing strategies for distinct imperfect mobility substitutes, such as different transport modes are common. So is differentiated pricing of specific services directly related to transport provision, such as parking. The introduction of spatial and temporal dynamics in pricing of public transport or road space, however, often struggle to win public acceptance. In case of road pricing, one of the most advanced implementation of time and location specific pricing schemes is to be found in Singapore (Olszewski and Litman, 2005). As of August 2015, it incorporates 77 gantries around the city perimeter and along all major expressways, which differentiate between
different vehicle types and change prices in steps of up to 5 minutes during the peak hours (Land Transport Authority, 2015b). Singapore has also adopted a distance-based public transport fare system and experiments with incentives and discounts for off-peak travel in public transport. A recent trial run included a waiver of train fares for passengers travelling towards the city centre before the morning peak (Land Transport Authority, 2015c).

On the demand side, prices can be differentiated between population groups according to specific social and economic characteristics. However, price discrimination as it is widely accepted in aviation, is a lot less common in urban mobility. Beyond a wide popularity of flat discounts and special season cards for young or elderly population groups, any discrimination dependent on other socio-economic characteristics is uncommon. Alternative ideas, such as trip-based pricing or tradable mobility credit schemes exist (Yang and Wang, 2011; Wu et al., 2012), but are rather only theoretical constructs and thought experiments at this stage. An increasing popularity of car and bicycle sharing schemes and integrated mobility planning approaches, however, can help to drive innovation in this area and bring such schemes to real-world implementation.

2.1.2 Road and congestion pricing

As discussed above, the way road pricing schemes are designed and implemented as well as the objectives used for the price setting and optimisation can be very diverse. Road pricing applied to eliminate road congestion follows the objective of social welfare maximisation and is usually referred to as congestion pricing.

Traffic congestion, a common phenomena in many urban areas around the world, is a result of travel demand exceeding the supply of road capacity during a certain time period. In economic terms, congestion is a consequence of market distortions arising from heavily used public roads becoming a congested public good with a rival nature of private goods (Hau, 2005). With speed on the roads dropping below the free flow speed, the travel costs perceived and borne by an individual road user do not fully reflect the total cost imposed on the society. In modern economies the cost of congestion adds up to significant amounts and is estimated to reach up to 1.5% of GDP in certain countries (Nash et al., 2003; Schrank et al., 2010; de Palma and Lindsey, 2011). Next to the obvious time and productivity losses, a number of side effects such as accident rate, pollution, noise, personal stress and well-being contribute to negative effects and external cost associated with traffic congestion (Mayeres et al., 1996; Parry et al., 2007).

The idea to use road pricing as a tool for regulating travel demand and alleviating congestion was introduced by Pigou (1920) and Knight (1924) and later developed by Vickrey (1963). Since, road and in particular congestion pricing has been a topic of many publications, addressing its economic fundamental principles (e.g. Hau, 2005), public acceptability
(e.g. Schade and Schlag, 2003), equity effects (e.g. Levinson, 2010), spatial impact (e.g. Condeço-Melhorado et al., 2011) or evaluation of effects after its real-world implementation (e.g. Santos and Bhakar, 2006; Börjesson et al., 2012). Detailed overviews and comprehensive summaries of road pricing evolution, state-of-the-art and latest methodological advances are provided by de Palma and Lindsey (2011), Tsekeris and Vöß (2009) and Lindsey (2006). Subsequent paragraphs outline the methodological approaches to model, simulate and evaluate road pricing as well as highlight the practical challenges and related key publications.

From a theoretical stand point, there exists a broad consensus among urban economists and transport planners on the benefits of congestion pricing and its capability to efficiently regulate demand and reduce congestion cost and its externalities (Lindsey, 2006). However, translating theoretical findings derived based on simplified models under a particular set of assumptions into real-world implementations is challenging and implications of pricing policies in a wider economic and social context as well as appropriate counter measures are subjects of heavy debates.

As a consequence of public controversy around road and congestion pricing policies, the number of cities and regions, which have ventured to implement road pricing schemes and succeeded in gaining public support for it, remains rather limited. Along with dynamic pricing schemes limited to a single facility (e.g. I-15 in California, Sydney Harbour Bridge) or nearly fixed zone-based pricing (e.g. London, Stockholm, Norway), so far only Singapore adopted some form of a comprehensive congestion related pricing scheme covering the city centre as well as all major expressways (Olszewski and Litian, 2005). Comprehensive comparison and discussion of road pricing implementations around the world can be found in Small and Verhoef (2007), Tsekeris and Vöß (2009) and de Palma and Lindsey (2011). Many of the implemented road pricing schemes have proven to be successful in mitigating congestion and also found broad public acceptance and support once their benefits became evident (Eliasson and Jonsson, 2011; Santos et al., 2008). However, drawing universal conclusions and learning from these experiences is hard, as given a unique transport infrastructure, economic conditions and spatial configurations, each city requires a tailored pricing design and equity assessment a priori.

Allocation of revenues from congestion charging is considered to be vital for the overall effect of a pricing policy with regards to consumer welfare and equity. From the stand point of the general population, distrust to the authorities, lack of transparency as well as time delay between policy implementation and noticeable benefits from its realisation, can promote perception of congestion charge policy being another form of taxation used to finance bureaucracy and overblown administration expenses. Among economist and transport planners, however, the allocation of road toll revenues is also a subject to debate. Earmarking of congestion charge revenues for a public good as e.g. extension of public transport, is often considered as one of the most effective measures beneficial for social welfare and public
policy acceptance (Eliasson and Mattsson 2006). However, as interactions between congestion charges and other markets such as e.g. labour market can offset welfare benefits resulting from internalisation of congestion externalities, alternative revenue distributions and more flexible ways of returning the revenue to the taxpayer can be even more beneficial and politically more popular (Litman 1996). Such revenue distribution can take many forms and does not need to be exclusively related to transport. As for example Parry and Bento (2001) point out, congestion charges and associated increase in commuting cost can have a discouraging effect on labour force participation and lead to the overall welfare loss, if toll revenues are not allocated for the reduction of labour taxes.

Regardless which policy prevails, the communication of congestion pricing benefits to the general public by the decision-making authorities is crucial for gaining its support and acceptance. It should also be noted, that even at optimal pricing point, congestion is not fully eliminated. Hence, traffic speeds significantly below the free flow speed in presence of pricing may play another important role in the perception of congestion pricing efficiency by the road users.

Effects and mechanisms of action of road pricing measures as demand management tool can heavily depend on particular transport network, alternative modes of transport and other policies such as e.g. flexibility of working hours. Major efficiency gains commonly result from changes in trip timings and peak spreading. Changes in route choice and more equal spatial distribution of demand through the network is another important dimension. Further, in presence of public transport, congestion charges can also be used as a tool to control mode share and induce mode shifts from private to public transport.

There also exists a number of secondary, often neglected effects of congestion pricing. Parry and Bento (2002) point out the critical importance of interplay between congestion charge and other distortions within the transport systems, such as public transport subsidies, accident and pollution externalities or congestion on routes not subject to road pricing. Such distortions can have substantial impact and either increase welfare gains of congestion pricing or turn them into welfare losses.

Robustness of the transport system and travel time reliability is another important issue, potentially affected by congestion pricing policies. Unexpected events such as accidents can affect systems performance and have lasting effects hours after the actual incident occurrence. Frequency and severity of unexpected events can have major impacts on travel choices, in particular for travellers with low degree of scheduling flexibility. Attempts to capture and quantify value of reliability have been undertaken in a number of publications, e.g. Brownstone and Small (2005), Fosgerau and Karlström (2010), Ehreke et al. (2015), Carrion and Levinson (2012) provide one of the most recent, broad overviews of this topic. Demand regulating pricing policy can help to absorb shocks introduced into the system through unexpected events.
An efficient transport demand allocation and its temporal and spatial distribution across the network, prevent dense aggregation of people and vehicles. Thus, it can help to ensure the availability of excess capacity in the system at any given point in time and therefore minimise consequences of unexpected events.

Long-term effects of congestion pricing require extensive studies and are also often difficult to assess. Impact on urban development, residential and commercial location choice, activity spaces and vibrancy of urban areas can be as vital as short-term mitigation of traffic congestion. As one of the few publications, Anas and Hiramatsu (2013) incorporate changes in travel, housing and labour markets into the assessment of road pricing effects for the Chicago metropolitan area. However, in most cases, a comprehensive framework relating short-term changes in travel costs to long-term changes on a wider urban scale and in other markets, is still to be developed and offers a vast potential for research. Projects such as e.g. SustainCity (2011) and SimMobility (2015) represent such attempts.

In practice, addressing the challenges associated with the transition from theory to real-world implementations of efficient congestion pricing schemes, requires the definition and development of methods and tools, which are capable to capture the full complexity of the problem. The challenges to be addressed can be roughly grouped into the following categories:

- adequate modelling approach,
- suitable evaluation techniques at the different spatial and temporal scales across various socio-economic strata,
- multi-objective optimisation of a second-best pricing scheme
- effective public communication.

Modelling of road pricing involves understanding, representing and predicting human behaviour in a complex socio-economic urban transport system. As discussed in the section 2.2, one of the major complications in capturing redistribution, equity and wider socio-economic effects of road pricing is associated with modelling of travellers heterogeneity. Development and application of an adequate model, that is capable of accounting for all relevant factors and hence accurately explain and predict the behavioural response of travellers to mobility pricing policies ultimately requires trade-offs between realism and restrictions in time and monetary budgets of the modeller. Data availability together with conceptual realisation and feasibility often represent the most critical issues.

---

1 Second-best pricing usually refers to a simplified implementations of road pricing schemes guided by the principle of internalisation of marginal external costs, but where the idealised assumptions underlying the standard economic analysis are not fully met (Small and Verhoef, 2007).
Existing modelling paradigms include different approaches to this problem such as analytical static models, analytical dynamic models and stochastic simulation approach. Major differences between these approaches are related to the level of aggregation, unit of analysis, temporal resolution, behaviour representation and level of detail when modelling interaction between relevant actors and infrastructure. Some of the relevant aspects and characteristics of trip- vs. activity based approaches are discussed in the section 2.2 below.

Given a suitable modelling approach, the challenge of optimal design of a charging scheme with its spatial characteristics, technical implementation and price differentiation across time, road and vehicle type still remains. Decisions on the technical realisation as e.g. placement of cordon locations, definition of the spatial extent of the tolling area or the time dynamics of the pricing regime, are often based only on expert knowledge and iterative learning procedures from monitoring the effects of the current pricing scheme. This is partly due to the fact, that the majority of transport models simply are not designed for providing the resolutions required for the design and assessment of dynamic pricing schemes.

The feasibility of different evaluation methodologies is strongly dependent on the modelling approach. Assessment and evaluation of pricing policies should capture multi-faceted behavioural responses to road pricing in order to determine its benefits and shortcoming for all parts of society and economy. Quantifying different social and economic effects of road pricing in urban networks across various population groups, geographical regions or economic sectors on varying timescales is equally important for model prediction as for the impact evaluation after the actual policy implementation. Such data serves not only as an indicator for the performance of the road pricing scheme, but should also be used to asses the quality of the applied model. Furthermore, in the evaluation process also secondary effects of congestion mitigation, which are usually not directly considered by the model during the planning phase, such as reliability, pollution, long-term climate effects, urban design and city planning, lifestyle, should not be neglected.

Public opinion plays a crucial role for the actual implementation of a congestion pricing policy. Reasons for public rejection of road pricing measures are diverse. Often perceived as another form of taxation, time delay for materialisation of its benefits, subjective perception of its effects and scepticism towards government as administrator of revenues, make it an unpopular measure. Most prominent arguments used to explain the public rejection of road pricing policies focus on the individual’s self-interest. Personal disadvantages, or in economic terms loss in consumer welfare, experienced by the majority of travellers after introduction of congestion pricing are commonly seen as drivers of public opinion. According to Hau (2005), the public rejection of congestion pricing results from losses experienced by the vast majority of travellers. The only exception thereby is the case of hypercongestion, in which all drivers benefit from the imposed congestion tolls. Another argument, commonly brought up in
connection with objections to road pricing policies, is the question of equity and perception of fairness of such measures.  Ison and Rye (2005) argue, minimising inequity is key for acceptance of congestion pricing. More recently, Eliasson (2014) provides an alternative view on the opinion formation towards congestion pricing, suggesting that individual attitudes and preference in case of congestion pricing are often unstable due to the lack of direct experience. Hence, how the policy is framed and communicated and which terminology is used, is crucial for gaining support for the proposed pricing policy. However, to ensure effective, direct and generally insightful communication, requires modelling tools and evaluation techniques capable to provide disaggregated information on policy impact relevant to an individual citizen.

### 2.1.3 Public transport pricing

As cities all over the world continue to attract new people and maintain strong population growth, public transport plays a major role in providing mobility and ensuring smooth flow of people within ever expanding urban areas. As a form of mass transport, it facilitates urban growth and enables densification. In particular, given the still dominant mono-centric city structures and separation between commercial and residential areas, bus and rail service lines became irreplaceable parts of urban transport networks, providing sufficient capacities to move millions of commuters, tourists and other travellers on a daily basis.

At the same time, development, maintenance and operations of public transport infrastructure has its cost. Distribution of these substantial cost between users, service providers and the general public is subject to the political decision process. Pricing of public transport and the objectives of pricing scheme design are closely related to the broader question of public transport financing on the one hand and optimal travel demand allocation, on the other. Given the number of effects on overall transport system and interdependencies between transport and economic activity, the problem of finding a socially optimal set of bus or train fares can not be considered in isolation and requires a holistic economic modelling approach.

Design of public transport fare schemes requires consideration of its role from two angles: financing of operations and transport demand management. Much attention has been traditionally given to the question of financing. Pursuing economic efficiency and profitability of public transport operations, however, may stand in conflict with its social role and second order benefits derived from it. Given complex links between fare revenue, network size, quality of service, total ridership and egalitarian provision of mobility and accessibility, profit maximisation strategies, as they are desirable from the operators point of view, tend to be much less beneficial for society than e.g. a ridership maximising policy alternative (van Goeverden et al., 2006). Such policies, however, often require subsidies of public transport operations, which can introduce a new set of inefficiencies and distortions into the system.
Planned and built to move large amounts of people in an efficient manner, public transport systems provide a number of external benefits. These include incentives to economic activity and labour force participation provided by lower transport and commuting cost and increase in accessibility for all population groups. Additionally, the number of benefits associated with reduction in car travel represents another important contribution to the positive effects of increased usage of public transport services. As argued by Mohring (1972), growing number of passengers is commonly addressed with an extension of public transport capacity, achieved by increase of operational service frequencies and introduction of new service lines. At the same time, such capacity extending measures also result in a better service provision, lowering the user’s travel cost and making public transport more attractive mode of transport. This creates a positive feedback loop, also known as "Mohring effect" (Mohring, 1972). This effect especially applies to bus operations, which in comparison to rail networks, are much cheaper and faster to extend.

The interdependency between transport demand and decreasing average cost of service provision translates into economies of scale, making a strong case in favour of subsidies to public transport operations and infrastructure. The self-financing constraint to public transport service provider appears to result in a socially sub-optimal service provision with low operating frequencies and large vehicle capacities (Jara-Díaz and Gschwender, 2008). Hence, the positive effects of scale economy, combined with other external benefits of public transport make a strong case for subsidising its operations and infrastructure. More specifically, two lines of arguments are commonly used to make a case for public transport subsidies (Parry and Small, 2009). The first targets the direct impact of increased service frequency on waiting times and therewith the asymmetry between more passengers given the same number of services and less services given the same number of passengers. Or in other words, while the private cost of scaling bus operations are constant, the social cost are decreasing. Hence, from an economic perspective, marginal social cost being lower than average social cost provides a typical case for subsidy, required to achieve a better resource allocation (Turvey and Mohring, 1975). The second rationale addresses external benefits of restraining car use by increasing public transport service quality. The induced mode shift from private to public transport eases road congestion and reduces number of negative external effects associated with it, such as pollution, noise, accident rates etc. In line with this argumentation, Parry and Small (2009) observe with an example of three metropolitan areas, that fare subsidies of 50 percent or more of operating costs are welfare improving at the margin. These findings seem to still hold, even after incorporating burden from subsidies on the broader tax system. Additionally, equity and re-distributional effects of public transport subsidies as well as its central role in accessibility provision should not be neglected and are discussed in detail in the section 2.3 below. However, also such effects require detailed evaluation in a broader context and as Nelson et al. (2007) show on example of Washington metropolitan area, location specific features of population, land-use,
transport network and policy can also lead to high income groups benefiting disproportionately from public transport service provision.

Most cities around the world make use of arguments in favour of public transport subsidies and use them to cover up to 89% of operational costs of public transport (Parry and Small 2009). Hong Kong and Singapore represent the few exceptions, where service providers are able to fully cover the operating expenditure through revenue (Kenworthy and Laube 2001). However, as Vivier and Pourbaix (2006) note, given the variety of factors influencing mode share, ridership and operational costs at a specific location, no clear relationship between coverage rate of operating expenditure thorough fare revenue and operators efficiency can be observed. Counterarguments against public transport subsidies usually include non negligible cost of public funds, inefficient use of capital and labour and low cross price-elasticities between car usage and public transport fares (Proost and Dender 2008; Winston and Shirley 2010). Furthermore, Proost and Dender (2008) stress the important role of the shadow values of transport tax revenues and demonstrate on examples of Brussels and London case studies, that welfare gains can also be achieved by removing existing subsidies to public transport and parking, internalising transport externalities and providing optimal frequencies of public transport service.

Being a highly political issue, transparent communication of subsidies, cost and benefits of financing structures and fare schemes is key to gain public support for any public transport related policy. Highlighting of subsidies by transport operators may even produce crowding-in effects and help to maximize public valuation and therewith willingness to pay for public transport services (Drevs et al. 2014).

A meaningful approach to optimal public transport pricing design is based on the internalization of external cost of public transport. Similar to road congestion, public transport is prone to crowding, with imbalance of demand and supply leading to a number of negative externalities. Once a critical vehicle occupancy rate or waiting passenger density at a station is reached, each passenger contributes to the level of discomfort as well as extended boarding and alighting times. As travellers, however, do not bare the cost imposed on fellow travellers, inefficiencies in allocation of supply and demand can arise.

A substantial body of literature exists on externalities of public transport associated with crowding. A comprehensive summary of effects of passenger density and crowding was recently presented by Tirachini (2013). In this paper authors subdivide effects of crowding into 7 categories, addressing prior research publications related to each of them: effect on in-vehicle time, effect on waiting time, effect on travel time reliability, effect on well-being, effect on the valuation of travel time savings and effect on optimal public transport supply and fare.
This thesis focuses specifically on two major crowding externalities: (i) increasing personal in-vehicle travel time valuation due to growing feeling of discomfort, lack of seating and psychological stress associated with crowded environment and (ii) direct effects of crowding on travel times due to rising friction among passengers leading to prolonged dwelling processes.

Similar to congestion pricing, internalisation of crowding externalities provides a promising avenue for addressing inefficiencies in operations of public transport systems. To this end, dynamic public transport fares, based on external costs can be used as an efficient demand management tool and help to control temporal and spatial distribution of mobility demand.

In the following section, the two major negative external effects, as considered in this thesis and used for design of dynamic public transport pricing, are discussed.

**Travel time delay and passenger friction**

Boarding and alighting processes of passengers at stops and stations are central element of public transport operations and represents a significant share of total travel times. Each passenger requires certain amount of time to board and alight a public transport vehicle. This time contributes to the total passenger service time (boarding and alighting) at a stop or a station. Passenger service time and the time required to open and close doors determines the total time, which the vehicle spends at a stop and is commonly referred to as *dwell time* (Lin and Wilson, 1992; Transportation Research Board, 2010; Sun et al., 2013). As various studies find, dwell time makes up on average about 20-26% of the total travel time and plays therefore a significant role in public transport operations (Levinson, 1983; Tirachini, 2013). Dwell times depend on a number of factors, such as demographic characteristics of passengers, fare payment system, platform design, bus configurations and level of crowding (Tirachini, 2013; Sun et al., 2013). As each boarding and alighting passenger contributes to the total dwell duration and therewith travel time of all other passengers, this time delay can be considered as a negative externality.

Another effect, which can have severe impact on individual’s travel time and travel experience, is the time delay experienced due to the inability to board a full vehicle. Having to wait for the next service not only significantly increases waiting time and impacts travel time reliability, but can also be associated with experience of stress and frustration.

Turvey and Mohring (1975) were one of the first to discuss travel time and waiting time delays, as negative public transport externalities. In their work, Turvey and Mohring (1975) address the question of optimal bus fares under first-best condition and factors contributing to marginal social cost pricing. The authors identify the following contributions to excess marginal social cost imposed by an individual passenger on other travellers: (i) delay of passengers on the bus during boarding and alighting and passengers boarding behind an individual at a particular bus
stop, (ii) delay imposed on passengers who could not board a full vehicle and (iii) additional cost to the operator in order to be able to provide a constant service quality for all following passengers.

Recently, Kaddoura et al. (2015) presented one of the first attempts to implement marginal cost pricing for bus operations within an agent-based simulation approach. Partly following findings of Turvey and Mohring (1975), the authors implement a calculation of user-specific optimal bus fares based on internalization of cost associated with prolonged in-vehicle and waiting times due to additional passengers. Chapter 7 of this thesis draws partly on the methodology and technical implementation presented by Kaddoura et al. (2015).

A crucial aspect for modelling, evaluation and internalisation of travel time delays in public transport operations, is the actual model of boarding and alighting operations. In this context, findings on variability of dwell times by Sun et al. (2013) are of the particular interest. As authors observe based on smart card data records from Singapore’s bus system, dwell time per passenger is not constant, but increases after a certain vehicle occupancy is reached. In other words, crowding leads to disproportional delay at the stops and therefore longer travel times for all passengers. Details and implementation of this model in an agent-based simulation framework are discussed in section 3.2 below.

The role of user heterogeneity in effects of dynamic public transport pricing is a less studied area. Turvey and Mohring (1975) note the importance of heterogeneous valuations of time among travellers for design of optimal fare schemes. In particular, the increase in service frequency or travel speeds is assumed to attract a group of passengers with higher values of time as compared to the existing passengers. The nature of analytical model applied, however, leads to Turvey and Mohring (1975) ignoring the user heterogeneity in their analysis.

**Value of time under crowded conditions**

As a result of thousand years of evolution, humans placed in a densely crowded, enclosed environments with number of strangers naturally experience number of behavioural implications. These can include feeling of discomfort, increased alertness, perception of threat to personal safety and in extreme cases even anxiety. Caused by perceived lack of control, restricted movement space, limited behavioural freedom as well as stimulus overload, crowding can be defined a ‘psychological state characterised by stress and having motivational properties’ (Bell et al., 2001).

A number of studies confirms, what for a regular public transport user might seem to be an intuitive assumption: travelling in a crowded bus or train is a significantly less enjoyable experience than being comfortably seated in a sparsely occupied public transport vehicle. Among other, Whelan and Crockett (2009) and Hensher et al. (2011) use stated choice
experiments to demonstrate an increase in average values of time and therewith in willingness to pay to reduce travel time, after a certain vehicle occupancy has been reached. Li and Hensher (2011) present a review of evidence for willingness to pay for avoiding crowded conditions during the journey, using studies conducted in the UK, USA, Australia and Israel. Also Wardman and Whelan (2011) reach similar conclusions and provide one of the most extensive meta-analysis of British rail crowding valuation studies.

There are commonly three different ways to quantify crowding valuations: time multipliers, monetary value per trip or monetary value per time unit (Li and Hensher 2011). Though not utterly independent, the majority of recent publications uses time multipliers due to better and easier transferability and comparability between cities or countries. The time multiplier is thereby commonly defined as a ratio between value of travel time savings in crowded conditions and value of travel time savings in uncrowded conditions.

One of the challenges in measuring value of crowding, is its heterogeneity across individual travellers. From meta regression by Wardman and Whelan (2011), a systematic variations in time multipliers including seated or standing, load factor and trip purpose can be observed. Personal characteristics and preferences, however, seem to be the most dominant factor influencing individual’s value of crowding. This is also confirmed by Theler and Axhausen (2013), who show, that perceptions of a vehicle being full, crowded or overcrowded diverge and depend on individual’s socio-demographic characteristics.

In face of overwhelming evidence for increase in values of time in crowded conditions and its impact on optimal public transport operations, a number of studies have included the crowding phenomenon into the transport demand models (e.g. Kraus, 1991; Jara-Díaz and Gschwender 2010; Tirachini et al., 2010; Fröhlich et al., 2012). More recently, Prud’homme et al. (2012) quantified severe welfare loss from crowding on Paris subway, also highlighting its effect on modal and route choice.

In summary, it can be said, that though the importance of crowding effects on public transport has been recognized and quantified, dynamic public transport fares based on the internalisation of crowding externalities remain a much less studied field as compared to the dynamic road congestion pricing. This work sets to address this shortcoming and investigate interaction dynamics between optimal pricing of private and public transport.

2.2 Heterogeneous values of time

As pointed out in the introduction, valuation of travel time and its distribution across the population, trips purposes or times of day can play a determining role in the economic project and policy appraisal and evaluation. However, a variety of factors affecting inter- as well as
intra personal differences of values of time pose a considerable challenge in understanding heterogeneity and adequately accounting for it.

Theoretical frameworks for valuation of time build on models of time and income allocation, as initially presented by Becker (1965) and later refined and extended by i.a. DeSerpa (1971), and Jara-Díaz (2003). Traditionally, the values of time are estimated using travel diaries, micro-census data or stated preference (SP) surveys (e.g. Jara-Díaz and Guevara, 2003). Alternatively few studies were performed based on revealed preference (RP) observations or the combination and comparison of SP and RP sources (e.g. Brownstone and Small, 2005), combining the strengths of both approaches. It is worth pointing out, that techniques of data collection and estimation of values of travel time require careful design and analysis as this can have major impact on the estimation and the mean of VTTS distribution (Fosgerau, 2006). A detailed overview of the theory and state of the art methodologies for value of time modelling and estimation can be found in Small and Verhoef (2007), Small (2012) and Börjesson and Eliasson (2014).

Users heterogeneity matters in particular, when designing, both, static and dynamic pricing policies. It affects welfare gains and redistribution effects and can become a decisive factor in favour or against the implementation of certain projects and policies. The significance of heterogeneity in values of time and relevance of their spread and distribution among travellers was demonstrated and discussed by i.a. Small and Yan (2001), Hensher and Goodwin (2004), Verhoef and Small (2004), Fosgerau and van Dender (2010). The case for elaborate inclusion of heterogeneity into the models becomes even stronger within a dynamic modelling context with inclusion of time choice and schedule delay cost, as shown by van den Berg (2014), van den Berg and Verhoef (2013, 2011a,b).

Different valuations of travel time and heterogeneous scheduling preferences can arise from variety of factors, such as interpersonal variations due to socio-economic and demographic characteristics, personal tastes, nature and type of activities as well as time or social context in which an activity is performed. Furthermore, trip characteristics and travel conditions such as trip length, congestion or crowding can have significant impact on the travel time valuation (Axhausen et al., 2008). Additionally, intra-personal variations such as benefits gained from an activity or a preference for its timing and duration can occur from day to day. Such variety of impact factors requires extensive data collection efforts and identification of most crucial factors in regard to the policies and projects under study.

The focus of the research presented in this work, lies in studying the impact of heterogeneous values of time on pricing policies in a multi-modal context, using an agent- and activity based simulation approach. Thus, heterogeneity in values of time is modelled based only on the household income. Other impact factors are not considered. For the purpose of this work, a direct representation of income elasticities based on estimates from the Swiss value
of time study by Axhausen et al. (2008), is applied and implemented in the activity-based simulation framework. Axhausen et al. (2008) estimate income elasticities of values of time considering number of heterogeneity dimensions and demonstrate high variation in time values based on the trip purpose, personal income and trip distance. To this end the authors use two stated preference surveys conducted in Switzerland. The detailed methodological approach is presented in the Chapter 3.

The trip-based approach to transport demand modelling, as part of the conventional four-step model, has been a well-established method for transport planning on city, regional or countrywide levels for decades. Using the trip as a basic unit of analyses, the four step model follows four basic steps: trip generation, trip distribution, mode choice and traffic assignment. It commonly applies spatial and temporal aggregation of single trips into flows in order to represent and forecast travel demand. Based in the conventional bottleneck model by Vickrey (1969) and Small (1982), the trip based approach focuses on trip characteristics, considering trip causes and drivers behind travel decisions as exogenous factors. Thus, it neglects the underlying structure of individual activity schedules, making inclusion of multi-dimensional heterogeneity in user preferences and travel decisions a cumbersome task.

Given a wide spectrum of heterogeneity sources, approaches based on individual travel behaviour models and daily activities appear to be more suitable for capturing the diversity in the preferences of travellers and commuters. Hence, driven by the necessity of more complete understanding and modelling of travel decisions, activity-based approach emerged as an alternative modelling paradigm (Axhausen and Gärling, 1992; Bowman and Ben-Akiva 2001; Ashiru et al., 2004). Focusing on individuals activity schedules, it enables explicit and intuitive representation of temporal and spatial travel dimensions, linking cause and effect of travel behaviour. In addition, the activity-schedule based approaches are more suitable for time shift and location choice modelling and analysis with a high level of resolution.

Attempts to extend the tradition single - trip approach and enable modelling of trip chains and flexible scheduling have been the topic of research in number of publications. Recently Jenelius et al. (2011) presented a two trip model with focus on analytical derivation of traveller delay cost and value of time in the presence of flexible activity scheduling. Building on top of previous work by Ettema and Timmermans (2003), Jenelius et al. (2011) provide a generalisation of a single-trip model using marginal activity utility functions for preference representation. However, to stay analytically tractable, the simple model remains restricted to two departure times as decision variables. Inclusion of other dimensions as mode, route and destination choice is considered as desirable but challenging. Following a similar idea, Li et al. (2014) use an activity-based approach to incorporate commuters day-long activity schedule and time decisions into the Vickrey’s bottleneck model for the morning and evening trips. The modelling approach based on utility maximisation, is analytically explored for a special case
of constant marginal utilities.

However, the degree of complexity arising with interdependencies of individual activity- and trip-chains and accurate modelling of individuals choice dimensions as well as feedback loops from interaction between travellers and infrastructure pushes analytical approaches to their limits. Agent-based simulation enables to overcome these limitations and incorporate behavioural complexity, diversity and interaction on an individual level. Exploiting these advantages, this work applies agent- and activity-based simulation to investigate heterogeneous values of time and different types of schedule delay heterogeneity on trip characteristics as well as user and social welfare using non-linear activity utility functions. A multi-modal approach with direct interaction between private and public transport on the network as well as spatial distribution of home and work activity locations add new dimensions and degrees of realism to the model. This enables the analysis of heterogeneity impact on design and effects of mobility pricing schemes. Chapter 3 presents the applied methodology.

2.3 Social equity - Accessibility, Activities, Alternatives

With transport acting as glue between economic and social processes, wider impact of transport policy on equity and its role in economic welfare distribution seems self-evident. Profound impact of any mobility pricing scheme and the multidimensional interlacing with human behaviour, social structure and economic benefits and disadvantages, makes the design and implementation of optimal and fair transport policy under a variety of physical, economic and social constrains a demanding task. In contrast to the general consensus among economists over the basic principles and aggregated social benefits of congestion pricing, its equity impact and redistribution effects are contested. As described above, according to economic theory imposing congestion charges results in welfare gain for the society due to the increases of the social surplus. The economically efficient state though does not comprise any statement on the equity or the fairness of resource distribution.

Equity, as relative measure of fairness or freedom from bias or favouritism (Merriam–Webster 2013), can be considered from various and divergent stand points. It is common to distinguish between vertical equity, which refers to the distribution of costs and benefits between groups and horizontal equity, which relates to the distribution of costs and benefits within a certain group (Ecola and Light 2009). Vertical equity is usually defined in terms of personal or household income, and horizontal equity is often associated with spatial effects and is defined in terms of mobility or accessibility (Anas and Lindsey 2011). As was shown by Arnott et al. (1994) and Hau (2005), low income groups often suffer disproportionately from road pricing, making it often a regressive measure prior to recycling of revenues from it (Basso and Jara-Diaz 2012). An overview of studies on equity and revenue recycling from road
pricing are given by Ecola and Light (2009) and Levinson (2010). More recently Anas and Lindsey (2011) discussed the equity issue and distributional impact of road pricing from a more general perspective of urban transport externalities. In another recent work, based on cordon toll scenarios of Chicago central area, Anas and Hiramatsu (2011) point out the strong dependency of the welfare distribution on the size of the cordon toll area, the spatial income distribution as well as the need of considering the effects of road tolls on the real estate market.

As pointed out earlier, redistribution of revenues plays a crucial role with regards to the impact of congestion charging policies on consumer welfare and therefore is central to the formation of public opinion towards the policy. From a transport planners perspective, understanding both, horizontal and vertical equity changes as a result the pricing policy is crucial for effective revenue redistribution. Such distribution policy, as it could be the commitment of toll revenues to the improvement of public transport as in case of Stockholm, can lead to an overall progressive effect of road pricing (Eliasson and Mattsson 2006). However, direct compensation of losers might be not the only way to achieve such outcomes and alternative distribution models with financial incentives for a broader population can be equally efficient and feasible (Litman 1996).

Independent of revenue allocation, a detailed understanding of disaggregated effects on different user groups and the assessment of winners and losers based on socio-economic characteristics as well as the spatial distribution of home, work and activity locations, is indispensable for the design of progressive and fair transport policies. Such disaggregated analysis of road pricing effects enables correction of inequalities and establishment of a Pareto improving policy, with everybody being left equal or better off than before. A disaggregated modelling approach is also necessary for evaluation of consequences of redistribution measures in the form of discounts and exemptions. Such direct interventions into the operation of a pricing scheme can impact and compromise the efficiency of the whole system and its performance, holding a risk of triggering undesirable second order effects.

Most transport projects and policies entail redistribution effects, which can be either regressive of progressive. Direct perceptible pricing impacts on consumer, before materialisation a potential return and redistribution of revenues, is eventually the most crucial determinant for public support and therefore successful implementation of congestion pricing. However, beyond the apparent monetary transfers, availability of means of transport can have a much deeper impact on equity and social exclusion. Access to transport correlates with individual’s job opportunities as well as activity locations available to fulfil her or his needs. In this context, disaggregated approaches as the concept of activity spaces (Schönfelder and Axhausen 2003) as measures of social exclusion, and accessibility (Geurs and van Wee 2004), gained increasing attention and popularity. Evaluation of accessibility can follow different approaches and be either related to an individual as a measure of available opportunities or to a place, as a measure
of its attainability.

From a public point of view, pricing policies and price discrimination in particular, are often perceived as unfair and unjustified measures, especially before the actual implementation of the policy in question (Eliasson, 2014). At the same time well designed charging schemes provide an opportunity to combine efficient demand management with funding streams for improving the situation of people with transport disadvantages. To make a differentiated, dynamic and integrated pricing scheme fair and efficient and convince the public benefits, requires spatially and socially differentiated design and evaluation measures (Preston and Rajé, 2007).

A number of publications address the development of accessibility measures and highlight their benefits for project and policy evaluation as a more comprehensive welfare indicator (e.g. Zöllig and Axhausen, 2012; El-Geneidy and Levinson, 2006; Levinson, 1998), stressing a need for a disaggregated approach with inclusion of individual’s spatiotemporal constraints and feedback mechanisms between accessibility and travel behaviour (Geurs and van Wee, 2004). More recently, Cascetta et al. (2013) presented a behavioural definition of accessibility, derived from number of opportunities as perceived by an individual, for satisfying his or her needs. With number of perceived opportunities being dependent on the availability of transport alternatives and their cost, this approach to accessibility aligns with the focus of this work on integrated, multi-modal approach to pricing scheme design and evaluation. Focusing on the response to pricing policies and associated welfare gains and losses on the availability and service quality of alternatives using an activity- and agent-based simulation, this work responds to the need for integrated techniques for project and policy assessment.
Chapter 3

Methodology

"Agent-based modeling is a way of doing a thought experiment. Although the assumptions may be simple, the consequences may not be at all obvious."

Robert Axelrod (Axelrod, 1997)

Transportation demand arises from travel decisions of individuals, who use the available transport supply to move from one location to another, mainly for sake of performing activities at different locations, which could not have been performed otherwise. From a modelling perspective, properties and behaviour of a transport system as a whole represent an emergent phenomena. Combination of travel choices made on an individual level given the predefined constraints, results in interactions among individuals and physical transport network, and leads to the emergence of a complex collective behaviour. The global system behaviour and system properties are not apparent from either the behavioural rules of individuals or from physical properties of the transport network - a characteristic often used to define a complex system as such (Darley, 1994). Agent-based modelling and simulation provides a framework for modelling and studying of such complex transport systems composed of autonomous, interacting agents in a given supply scenario.

Agent-based simulation is considered as third way of doing research, combining deductive and inductive approaches for scientific discovery (Axelrod, 1997). Following a deductive methodology, a specific set of rules and assumptions is defined in order to derive consequences of these assumptions. However, in contrast to deduction, simulation usually does not present a rigorous proof, but provides provides "an orderly formal framework and explanatory apparatus" (Foster, 2006). Analysis of the newly generated data is performed based on an inductive reasoning approach, as it is also used for empirically obtained data sets.

Agent-based models set themselves apart from a more general definition of a complex system
by two common properties: the heterogeneity of agents and the individual’s capability to adopt various states in the course of a simulation. Hence, the agent-based simulation approach enables modelling of social self-organization processes and emergent socio-economic behaviour on a scale and with level of detail unattainable by analytical approach (e.g., Helbing, 2012).

In the context of activity-based transport simulation, agent heterogeneity relates to a specific set of individual activity locations, daily schedules and, as discussed in this work, also values of time. Activity types and trip stages using a particular mode represent a dynamic state of an individual. While some of the agent characteristics and state parameters are constant, others, such as the activity schedule or activity locations, are subject to learning and evolutionary processes, evolving over the course of the simulation. In this work an agent-based modelling and simulation is applied with three objectives in mind: it aims to demonstrate the importance of user heterogeneity for mobility pricing design in a multi-modal urban environment, to present a proof of concept for the capability of an agent- and activity based simulation to capture heterogeneity in such an application and to examine gained insights against the intuition, experience and literature.

In transport modelling, agent-based simulation is often closely associated with the activity-based modelling approach, where activity patterns and the spatial distribution of their locations are considered as driving force behind transport demand. Thus, travel behaviour and associated choices can only be modelled in the context of an activity agenda, commonly incorporating one weekday. McNally and Rindt (2008) give an extensive overview of emergence and evolution of activity-based models and their substantial advantages over a conventional trip-based approach.

The inherent complexity of human activity behaviour requires extensive data and computational power for its modelling and simulation, in particular when performed on the level of individuals and with high degree of spatial resolution. Dependent on scale, number of agents and level of detail based on which the interaction among agents is modelled, conducting agent- and activity-based simulations can require substantial computational resources. The exponential growth of computing power coupled with new data collection methods helped to foster research in this field, leading to development of various agent- as well as activity-based simulation models (Balmer et al., 2008; Arentze and Timmermans, 2000; Mahmassani and Abdelghany, 2003; TRANSIMS Open Source, 2013). Given the vast range of personal socio-economic characteristics and individual preferences influencing travel behaviour, agent-based simulation posses essential properties for study of user heterogeneity and associated emergent phenomena arising from collective behaviour. This work adopts and extends the Multi-Agent Transport Simulation framework MATSim (MATSim, 2015; Horni et al., forthcoming), which is described in detail below.
3.1 Multi Agent Transport Simulation (MATSim)

The MATSim framework integrates travel demand based on individual activity schedules with simulation-based dynamic traffic assignment. One of its major strength, is its capability of detailed modelling and simulation of multi-modal networks. Joint simulation of private and public transport based on its queuing model allows time-dependent calculation of travel times accounting for spill-over effects and direct interaction of private and public transport. To this end, MATSim includes a model of interaction dynamics based on the actual physical properties of vehicles and network links, which represent road infrastructure. By capturing the queueing process of vehicles at the end of each network link, it enables accurate modelling of the congestion dynamics, affecting travel times of cars and buses alike.

Based on a co-evolutionary algorithm, agents alter their behaviour from iteration to iteration, evaluating new routes, alternative transport modes, departure times or secondary activity location choices in the process. Each agent tries to find an optimal daily schedule, which maximizes its utility function. Following each iteration of the queue-based network assignment, the activity scheduling and travel choices of each agent are evaluated and scored, leading to the generation of an individual choice sets. The selection of travel alternatives from the choice set of each agent is performed using a random utility model. As described in detail by Nagel and Flötteröd (2009), after a number of iterations individual utilities collectively converge and the system reaches a stable agent-based Stochastic User Equilibrium (SUE) (Sheffi 1985).

The major stages of MATSim’s iterative cycle are visualised in Figure 3.1. Initial demand represents the population of individuals with pre-assigned, initial daily schedules and activity locations. During the iterative cycle, commonly executed dozens or hundreds of times, agents search for better schedule alternatives along the available choice dimensions, evaluating these
in the execution and scoring phases. Once the equilibrium state has been reached, the travel times and individual plans are analysed on aggregate and disaggregate scales.

The full compatibility with the discrete choice theory enables the inclusion of new determinants of individual decisions and the application of established choice modelling concepts for the economic welfare evaluation. Featuring a modular architecture, MATSim allows for flexible management, adoption and extension of behavioural features and individual choice dimensions. In the context of this work, the departure and mode choice models are of particular relevance and route choice plays a subordinate role. Focusing on commuting trips with simple home-work-home activity chains, this work does not take the activity location choice into account, therefore not further discussing it below.

The departure time choice enables agents to alter their departure times from activities. Selected agents modify their departure times and activity durations of a daily plan randomly within a pre-defined time window, keeping other choice dimensions unchanged. Analogous, if an agent is selected for mode choice, it can change the mode of its journeys in the next iteration. As the mode choice has to be consistent (taking bus in the morning to work and car back home is improbable, due to non-availability of the vehicle at the work place), mode choice is altered at a sub-tour level (round-trip). The routing of individual trips in MATSim can be performed based on one of the common routing algorithms, such as Dijkstra or Landmarks A-Star routing (Lefebvre and Balmer, 2007). In this work, the route for each trip is generated based on the pre-defined trip mode and either the time-dependent Dijkstra fastest path algorithm for private transport or multi-node Dijkstra routing algorithm of public transport (Rieser, 2010; Ordóñez Medina and Erath, 2013). More details on the architecture and functionality of modules handling these choice dimensions within the MATSim framework can be found in Balmer et al. (2008), Horni et al. (forthcoming) or MATSim (2015).

Incorporating heterogeneous value of time preferences in MATSim framework requires specification individual parameters for each agent’s utility function, or as it is referred to in MATSim’s terminology a scoring function, which evaluates daily schedule performance based on individual attributes. As this work focuses on the income related variations in values of travel time, the household income of each individual is attached to the representing agent as an agent-specific property. This enables its direct incorporation into the value of time and schedule delay calculation within the individual’s scoring function. Section 4.2 describes the concept behind the value of time variations in an agent-based context in detail. The same agent-specific values of time have also to be taken into account in the routing module during the shortest path calculations in order to ensure consistency between routing and scoring.
3.2 Public Transport in MATSim

MATSim provides a fully integrated simulation of public transport operations, based on a detailed model of interactions between passengers, buses, trains and private transport vehicles. Each public transport vehicle is modelled and specified according to its physical characteristics such as size, capacity and number of seats. Buses and trains move on the road or rail network, respectively, following the queue-based traffic dynamics. Sharing the same road network, buses interact with cars and therefore, in absence of dedicated bus lanes, are also subject to delays caused by car congestion. Vice versa, a bus stopping at a bus stop without a bus bay on a one-lane street, will delay cars following it, for the duration of passenger alighting and boarding process. In case a public transport vehicle reaches its maximum capacity, further boardings are denied, leaving travellers at the station to wait for the next bus or train.

The duration of the dwell process itself depends not only on the number of passengers boarding and alighting, but also on vehicle characteristics, such as the number of doors and the total occupancy. Here, the work by Sun et al. (2013) provides valuable insights into the dwell time dynamics and is applied in an agent-based simulation context in this thesis. Based on the electronic smart card fare collection system in Singapore, the authors estimate a bus dwell time model, which incorporates the friction between boarding, alighting and on-board passengers. Sun et al. (2013) show, that a non-linear approach, as presented in Equation 3.1, performs best in relating the physical characteristics of vehicles to the boarding and alighting dynamics observed in the field.

\[
D_w = \max \{(B - 1) \cdot b + (\max(On - \gamma \text{Cap}, 0)) \cdot a, (A - 1) \cdot a\},
\]

Here, \(D_w\) indicates the total dwell time without door opening or closing time, and \(B\), \(A\) and \(On\) is the number of boarding, alighting and on-board passengers respectively. \(a\) and \(b\) are the estimated parameters and \(\gamma \text{Cap}\) is the critical vehicle occupancy as in proportion of the total capacity. As Sun et al. (2013) note, the passenger friction becomes relevant once the occupancy reaches 50% ∼ 60% of the vehicle capacity.

Another important aspect, is the stochastic variability of dwell times due to factors not captured by the model. Sun et al. (2013) show, that there exists a linear relation between the standard deviation of dwell times as predicted by Equation 3.1 and dwell times as observed in the data set for all bus types. Following this observation, the total bus dwell time is modelled in MATSim as the normal distribution, with the mean \(D_w\) and standard deviation \(Std = 2.22 + 0.12 \cdot D_w\). In order to prevent unrealistically short dwell times, the minimum dwell time is capped at \(\frac{1}{2}D_w\).

Furthermore, the presented dwell time model does not consider additional delays resulting from the bus deceleration and acceleration as well as door operations. Hence, these delays are
included as constants. The door opening and closing time is assumed to add 1 second each, and acceleration and deceleration add 4 seconds each, to the total bus stop time. This results in additional delay of 10 seconds for each dwell process. This delay will become in particular relevant, once external costs occurring during public transport operations are discussed in Chapter 7.

Concluding, the high level of detail in modelling of public transport is crucial for capturing the interaction between public and private transport in case of shared road space as well as crowding effects and dynamic phenomena such as bus bunching.

### 3.3 Congestion pricing

Following the utility maximisation theory, every economic agent pursues the goal of maximising its total utility. It does so by adjusting its own behaviour and decisions based on the individual experience. This optimisation is carried out according to the individual’s utility function and up to the point, where unilateral change in behaviour does not lead to any additional utility gains. This eventually results in a stable overall state of the system, commonly referred to as user equilibrium in case of transport networks. Put forward by Wardrop (1952) specifically for route choice and known as Wardrop’s first principle, it is closely related to the concept of Nash equilibrium in the game theory (Nash, 1950), where no individual decision maker can be better off by changing its strategy without cooperation. One of major differences between the two equilibrium definitions, is that in contrast to Wardrop’s user equilibrium, the Nash equilibrium does not require cost of all routes being equal, given a finite number of decision makers (Altman and Wynter, 2004). Furthermore, Wardrop (1952) postulates a second principle, which implies existence of an equilibrium with minimal average travel time, also known as system optimum, given a cooperative behaviour among individual users.

Pricing, as discussed above, can be applied as a tool to eliminate arising inefficiencies and influence travel decisions of individuals in order to trigger a shift from an otherwise stable user equilibrium towards a more socially optimal state - the system optimum. The concept of first-best congestion pricing is based on the idea of internalisation of the congestion delay externality. Charging road user a toll equal to the cost she imposes on all other travellers on the same route by adding to the congestion delay, leads to an efficient allocation of network capacity among users. Adding this cost, commonly referred to as marginal external congestion cost (mecc), to each travellers trip cost, minimises the total travel time in the network and lets the system converge to a state equivalent to the system optimum, as defined by Wardrop (1952). The sum of private and external cost is often referred to as the total social cost.

Congestion pricing refers to an implementation of road pricing scheme with road charges
set based on the principle of marginal external congestion costs. It is also common to
differentiate between first-best and second-best pricing policies. First-best road pricing
represents an idealised, optimal case, where the road charges paid by each driver exactly
correspond to the marginal external congestion costs, she or he is responsible for. In practice,
however, implementation of such accurate pricing schemes is impossible due to a number
of existing constraints, such as e.g. limitations in users ability to understand the complexity
of pricing scheme and rationally adjust its behaviour, or technical feasibility of first-best
policy implementation. Second-best pricing refers to an optimal pricing scheme under the
existing constraints, resulting in an approximation of first-best pricing and deviation from
direct marginal external costs. The challenge in design of second-best pricing instruments
is to satisfy the existing constraints and at the same time to minimise the distortion of the
theoretically optimal first-best case, representing a hypothetical benchmark. Examples and
further discussion on the design of the second-best congestion charging policies is provided by
e.g. Verhoef et al. (1996); Verhoef (2000); Zhang and Yang (2004); Verhoef (2005).

Congestion pricing based om marginal external costs also represents an approach highly
suitable for an implementation in an agent-based modelling paradigm and traffic dynamics
simulation based on a queuing model. Here, mecc can either be implemented as a trip or
link based congestion charge. Kaddoura and Kickhöfer (2014) attempt to conceptualise and
implement a trip-based congestion charging approach in an agent-based simulation context.
Encountering some conceptual challenges on the way, Kaddoura and Kickhöfer (2014) present
a proof-of-concept for time-dependent user-specific congestion tolls, which increase realised
user and social benefits. However, authors do not explicitly conduct an economic welfare
evaluation, addressing the question of how close the achieved results approximate first-best
pricing policy outcomes.

In this work a link-based congestion charging approach is adopted. Thereby mecc can be
computed on a link-by-link basis through out the network without the need to directly account
for prices on one link affecting pricing policy on all the other links in the network (Yang and
Huang, 1998; Small and Verhoef, 2007; Safirova et al. 2007). This property significantly
simplifies the implementation of an optimal congestion pricing strategy within the large-scale
agent-based simulation framework (Quinet and Vickerman, 2004). Using cost aggregation
over a specified time bin, a constant toll is attributed to each link and paid by all travellers
entering the link during the duration of the time bin. Such implementation of the mecc
pricing approximation for an agent-based queuing model was initially presented by Lämmel
and Flötteröd (2009) and later refined by Lämmel (2011).

A simplified approximation of mecc, based on the assumption of stationary flow condition
being satisfied as long as the queue at a link’s end persists, leads to the following definition:

\[ mecc_l(t_0) \approx t_l^{end}(t_0) - t_l^{lv}(t_0) \]  

(3.2)
with \( mecc_l(t_0) \) denoting the external cost that one additional agent causes by entering a link \( l \) at the time \( t_0 \). \( t^{end}_l(t_0) \) denotes the time at which the congestion, that the "causative" agent contributed to by entering a link at \( t_0 \), dissolves. And \( t^{in}_l(t_0) \) is the time at which the "causative" agent enters the bottleneck at the end of the link \( l \) (for detailed derivation see Lämmel [2011]). In other words, under the assumption of a constant, maximal outflow rate at the link \( l \), \( mecc \) equals to the time the link \( l \) remains congested after the "causative" agent passed through it. Testing this computationally inexpensive method with continuous evaluation of social cost for each agent on every link in the network, Lämmel and Flötteröd (2009) and Lämmel (2011) present simulation results for an optimization of routing in an evacuation scenario supporting the efficiency of this approach.

It is also important to note, that while the approach in Equation 3.2 represents an approximation of the first-best congestion pricing, the assumption of stationary flow is valid only as long as there is no congestion spillback at intersections. A spillback from congested link to links further upstream can lead to over- as well as undercharging of some agents, with its overall effects depending on network topology and traffic flows.

In this thesis, the implementation of the presented marginal social cost pricing approximation in MATSim framework is performed according to the algorithm presented in Lämmel (2011, chap. 3.1) and under consideration of practical implementation issues discussed in Lämmel (2011, chap. 4.1.3). Furthermore, the implementation is also based on the source code, initially developed by Christoph Dobler and obtained through personal communication and public MATSim Sourceforge code repository (July 2013). In the course of the agent-based congestion pricing implementation, number of practical aspects, as also observed by Christoph Dobler, emerged and are addressed as discussed in following.

First, in order to facilitate a smooth convergence of the system to a stable equilibrium state, changes in travel cost induced by pricing policies should be introduced gradually, allowing agents to adjust their behaviour to a new situation. Hence, the evolution of link-based congestion charge in MATSim is not only based on the last iteration, but is implemented as a moving average of the last 10 iterations.

Second, time bins of 5 minutes were chosen for aggregation of tolls, representing a reasonable time resolution for a stable convergence of congestion charges. However, this aggregation can result in overcharging some travellers during the first or last time bin of the congestion period. Setting congestion charges for the whole length of a time bin, during which a vehicle queue emerges and disappears, also affects drivers who travel on the tolled link just before or after the actual congestion period. Additionally, aggregation of congestion charges into time bins, can result in an abrupt transition between enforcement of a congestion charge and free road usage. Due to the approximation of congestion charge as presented in Equation 3.2, this is especially the case for the time at the beginning of the charging period. In order to prevent
an abrupt price change from free road use to a high congestion charge, drivers are charged an average charge which prevails on the entered link during the next 5 minutes. This simulates a smooth transition at the beginning of the congestion charge period and prevents a hike of the link travel cost. Visualised in Figure 3.2, the applied congestion charge averaging facilitates a convergence to a stable equilibrium state.

The convergence behaviour of the system in MATSim, given the marginal external congestion charge implementation as described above, is discussed in Chapter 6.

3.4 Economic Evaluation

As pointed out in the introduction, the central focus of this work lies in the assessment of social and economic impact of urban transport polices as in particular congestion pricing given a population of users with heterogeneous preferences. Benefits and losses of policy in question have to be quantified and assessed from societal as well as individual points of view. Hence, defining an adequate economic evaluation methodology, capable to comprehensively capture policy impact on different levels is crucial for interpreting of simulation results and the robustness of conclusions derived.

The Expected Maximum Utility (EMU) approach allows straightforward calculation of consumer surplus and is consistent with the discrete choice theory, the basis of agents’ daily schedule selection in MATSim. Challenges associated with EMU calculation, alternative welfare indicators as well as transition to social welfare calculation in presence of public transport operations are discussed in following.
3.4.1 Expected Maximum Utility

Mostly used in simulation, the discrete choice modelling approach provides a realistic way of modelling behavioural responses on an individual level to changes in supply and demand, and enables the emergence of complex system behaviour as a result of traveller’s personal preferences. A rich body of literature provides a natural and consistent way for applied welfare evaluation based on random-utility maximizing discrete choice models (e.g. de Jong et al. (2005), Train (2003), Ben-Akiva and Lerman (1985), Small and Rosen (1981), McFadden (1981) to name just a few major contributors). Referred to as Expected Maximum Utility (EMU), its calculation considers utilities of all alternatives available to the individual traveller and applies a logarithmic term to reflect the principle of diminishing marginal utility. EMU can be considered as a straightforward way for calculation of consumer welfare, which at the same time can be interpreted as a general measure of accessibility (Ben-Akiva and Lerman (1985). For a logit model, which is used as standard SUE model in MATSim, EMU formulation can be derived as follows:

A choice probability for an alternative $i$ with a deterministic utility $V_i$ is given by equation 3.3, with $J$ as the total number of available alternatives.

$$P(i) = \frac{e^{\mu V_i}}{\sum_{j=1}^{J} e^{\mu V_j}}.$$  

In an activity-based simulation context each alternative is a daily plan and $J$ is the total number of plans for an agent to choose from. Consequently, the logsum term for a choice set $J$ is defined as the natural logarithm of the denominator in equation (3.3):

$$V_J = \frac{1}{\mu} \cdot \ln \sum_{j=1}^{J} e^{\mu V_j}.$$  

Here, $\mu$ represents the scale parameter of the disturbance term and $\epsilon$ is to be understood as the degree of decision makers rationality or the ability of the user to distinguish between the utilities of different alternatives (see section 2.3 in Kickhöfer (2014) for detailed explanation). Being dependent on the size of the choice set, the logarithmic formulation reflects the idea of decreasing marginal utility of additional alternatives.

Under the assumption of marginal utility of income staying constant over changes from the
particular policy, the expected change in the consumer surplus as result of policy introduction for a traveller $n$ is formulated as in equation [3.5] with $\alpha_n$ indicating marginal utility of income and superscripts 0 and 1 referring to the states before and after the change (de Jong et al. (2005)).

$$\Delta E(CS_n) = \frac{1}{\alpha_n \mu} \left[ \ln \sum_{j=1}^{J_1} e^{\mu V_1^j} - \ln \sum_{j=1}^{J_0} e^{\mu V_0^j} \right].$$

(3.5)

Extending the EMU formulation from the change in individual welfare of an agent $n$ to a population of $N$ individuals, the total change in the consumer welfare is computed as a sum of individual changes (Equation [3.6]).

$$\Delta W = \sum_{n=1}^{N} \Delta E(CS_n).$$

(3.6)

**Challenges in Choice-Set Generation**

One of the main premisses in discrete choice theory is the assumption of independence of irrelevant alternatives (IIA), where adding additional alternatives should not change the decision makers choice for any existing alternative. This is a strong condition to be satisfied, in particular in cases when the set of potential alternatives is very large. Hall (2003, Chapter 2) discusses this intricate problem in detail for departure time and route choice. The main challenge is containing the number of possible choice alternatives while establishing independence between relevant potential choices.

MATSim can be considered as a multidimensional choice set generator, which generates and evaluates a number of alternative daily plans by varying prior plans in predefined dimensions (e.g. route, mode, departure time or location choice) with every iteration. With a commonly used "best score" criteria for selection of daily plans, only a limited number (typically five) of best performing plans are kept as the agent’s choice set in its memory, discarding other, worse performing alternatives. Though this guarantees a smooth and stable conversion of the overall system, the plans saved for each agent in its memory tend to become very similar as the number of iterations increases. Hence, these similar plans become unsuitable for the welfare calculation, because they violate the IIA condition (see also Kickhöfer, 2014).

Oliveros (2013) and Nagel et al. (2014) attempt to address this problem within an agent-based
modelling context, focusing on randomized components in route choice for public transport and car modes respectively. Alternative attempts focus on strategies for selection and removal of already existing plans. Grether (2014) investigates a strategy based on pathsize logit approach, aiming to provide a more realistic model of the decision process for selection and storage of daily plans in agents memory and in its individual choice set. However, a cohesive multi-modal and time-dynamic approach to this problem requires further research and in-depth evaluation, as alternative choice sets may alter the stability and convergence process of the simulation.

To circumvent this problem and ensure compliance with the IIA condition, an alternative solution for generation of a choice set suitable for welfare evaluation is proposed and applied in this work. For each agent, the chosen plan of the final iteration of the simulation run is picked and used as a base for a rule-based definition of alternatives; thereby, departure time alternatives are defined following the approach by Antoniou et al. (1997), where authors used a set of five alternatives for each departure.

In an activity-based context, trip departure times are essentially activity end times. Hence, a choice alternative is determined not by a single departure time, but by the set of all departure times through the day, resulting in exponential growth of the choice set with increasing number of activities during the day. For the home - work - home activity chain considered in this study, departure times with +1h and -1h relative to the observed chosen alternative are considered. Applied to morning and evening commute, this results in a set of 9 possible departure time combinations. Here, combinations, which alter work durations by two hours are discarded as being too substantial and therefore improbable schedule variations (earlier departure in the morning (-1h) and later in evening (+1h), or later departure in the morning (+1h) and earlier in the evening (-1h)). Generating these alternatives for both, car and public transport modes, results in a total of 14 possible daily schedules. In case walking appears to be a realistic alternative (travel duration less than 1h), it is as well added to the choice set. As walking is considered to be unaffected by traffic and crowding conditions, only one utility maximizing departure time is chosen. This results in a choice set of 14 to 15 alternatives for each agent, with the size of the choice set of an individual agent remaining constant in all scenarios. Though this appears to be a fairly large number, taking into account that it is a full day activity plan and looking at it from a single trip perspective, leaves us with maximum 7 choices for the morning commute (3 departure time for car and bus, and one for walking) and dependent on the first choice between 3 and 1 alternatives for the evening commute. These appear to be appropriate numbers, in line with number of choices considered by other studies.

An exogenous generation of alternative daily plans requires to evaluate the total utility of each plan, or in MATSim terminology plan’s score. However, simulating every non chosen alternative for each agent and keeping the behaviour of all the other agents constant at the same time, would result in a very large number of simulation runs, equal up to 14 times the size of 42.
the scenario’s agent population. Hence, executing these full scale simulation runs is infeasible, even for small scale scenarios.

To this end, the utility of every non-chosen alternative is evaluated based on the travel times of the last iteration of the initial simulation run, using a pseudo-simulation approach presented by [Fourie et al. (2013)]. This is basically equivalent to the simulation, where only the agent of interest would change to a different daily plan, without affecting the system state as a whole.

### 3.4.2 Generalized cost and realised utility in an agent-based SUE

In an agent-based simulation framework each agent follows the goal of maximising its utility given an individually parametrised objective function, also referred to as utility function in discrete choice theory and scoring function in a MATSim context. In activity based models, the utility function commonly incorporates utilities gained from activity performance and (dis)utilities associated with travelling. In context of MATSim’s co-evolutionary optimization algorithm, each agent’s experienced utility of the simulated day is calculated at the end of each iteration. Sum of all utilities from the chosen alternatives across agents, can be interpreted as a form of generalised cost given the overall system state in this iteration and is refereed to as $\sum \text{Realised utility}$ [Zöllig and Axhausen (2012)]. In MATSim framework, average of Realised utilities ($RU$), also called simulation score, is often used as an indicator of convergence of the overall system and stability of the stochastic user equilibrium. As $RU$ directly measures the actual utilities of chosen alternatives, it also can be adopted as qualitative indicator of overall economic performance. Zöllig and Axhausen (2012), for example, use it next to the EMU calculation for the assessment of infrastructure investments with an agent-based accessibility approach. A major advantage of $RU$ indicator is the ease of averaging across multiple iterations. As discussed in detail in Chapter 6, this allows to account for variations from iteration to iteration due to the stochastic nature of the equilibrium.

### 3.4.3 Public Transport Operation Cost

Providing a high service level of public transport comes at a cost. Operating a bus line with higher headway requires more buses, more drivers and results in more vehicle kilometres. At the same time, as demonstrated in the following sections, the level of public transport service and availability of alternatives to the car mode has a decisive impact the on economic benefits of congestion pricing in a multi-modal urban environment, especially in the presence of heterogeneous values of time. However, conducting welfare evaluations and analysing sensitivity of pricing policy effects in the presence of varying bus service frequencies, requires to account for operational as well as capital cost of service provision.
Matching the travel behaviour parameters, which are based on the survey data from Sydney 
Tirachini et al. (2014), formula and parameters for cost estimates of bus operations in the 
corridor scenario are borrowed from national guidelines for transport management in Australia 
Australian Transport Council (2006) and were previously used for simulation-based public 
transport fare and frequency optimization by Kaddoura et al. (2015). The total cost $C$ of bus 
operations for one day are calculated according to equation (3.7)

$$C = (d_{vkm} \cdot c_{vkm} + t_{vh} \cdot c_{vh}) \cdot O + N_v \cdot c_{vday},$$ (3.7)

with $d_{vkm}$ as total vehicle kilometres per day, $c_{vkm}$ monetary cost per km, $t_{vh}$ total operational 
vehicle hours, $c_{vh}$ monetary cost of vehicle operation per hour, $O$ factor for overhead cost, $N_v$ total number of vehicles and $c_{vday}$ daily capital cost. Here, the first part of the equation accounts 
for variable operational cost and the second part for the fixed cost. The daily unit cost $c_{vday}$ and 
the cost per vehicle kilometre $c_{vkm}$ are dependent on the vehicle capacity, with cost functions 
derived from linear regression and shown together with other parameters in Table 3.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Formula</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{vkm}$</td>
<td>$0.006 \cdot \text{capacity} + 0.513 ,$$/,\text{vkm}$$</td>
<td>$$$/,\text{km}$</td>
</tr>
<tr>
<td>$c_{vday}$</td>
<td>$1.6064 \cdot \text{capacity} + 22.622 ,$$/,\text{vday}$$</td>
<td>$$$/,\text{day}$</td>
</tr>
<tr>
<td>$c_{vh}$</td>
<td>$33 ,$$/,\text{vh}$</td>
<td></td>
</tr>
<tr>
<td>$O$</td>
<td>$1.21$</td>
<td></td>
</tr>
</tbody>
</table>

The operator profit can be calculated as the difference between the sum of collected fares and 
operational cost $C$. With $N_{pt}$ indicating a total number of bus trips and $f_{bus}[\,$$/\,\text{trip}]$ the bus 
fare, for operators profit $\Pi_{PT}$ follows:

$$\Pi_{PT} = N_{pt} \cdot f_{bus} - C$$ (3.8)

In the welfare calculation, as described in the following section, operators profits (or losses) 
are considered as part of the total social welfare.

### 3.4.4 Social welfare and consumer surplus

In this work, social welfare is defined as a sum of consumer benefits, monetary payments made 
by travellers for the usage of transport infrastructure and cost of bus operations, as presented in 
equation (3.9). The public transport operation cost, as discussed above, also incorporate overhead 
cost. This definition is based on an idealized assumption of zero transaction cost and no losses
as a result of monetary transfers from travellers to bus and road operators. Money collected from congestion charging and public transport is assumed to be returned to society in one form or another, incurring no additional toll collection or administration cost in the process.

\[
Social\ \text{Welfare} = \text{Consumer\ Welfare} + \text{Fare\ Revenue} + \text{Toll\ Revenue} - \text{PT\ Operation\ Cost} \quad (3.9)
\]

### 3.5 Simulation methodology

The simulation methodology used in this work draws on a wide body of experience from previous MATSim based studies and related research. The initial demand for all scenarios is based on a set of relaxed plans with car as a single mode and one plan per agent. After introduction of a new bus service frequency or congestion pricing policy, 1000 iterations are performed to allow the system to reach the state of stochastic user equilibrium (SUE). In the first 800 iterations, agents generate and evaluate new plans with the two major replanning strategies: mode choice and departure time choice. During this plan generation process the share of agents trying out new plans is linearly decreasing, with a total of 40% in the first iteration and no new plans generated after iteration 800. This process is referred to as linear annealing and prevents extreme system state changes or "jumps" after disabling of the plan generation process. In the last 200 iterations, agents choose from the existing 5 plans in their memory with probabilities given by the logit model.

In particular, for small networks the introduction of linear annealing plays an important role for systems performance during the transition of the simulation from the choice generation to the choice selection phase. Without linear annealing approach, the expectation of certain fraction of travellers switching to a different, non-optimal route, mode or departure-time, gets incorporated during the iterative process of agents schedule optimisation, into the choice set of daily plans of each agent. The system virtually learns about non-optimal, random behaviour of a consistent share of the population. In large-scale simulation scenarios and with low replanning rates to begin with, this effect is often hardly visible and can be neglected. However, in small scale scenarios it becomes significant, as given a limited network size, the equilibrium state tends to be sensitive to perturbations.

The performance of linear annealing and its effect on convergence of the overall system to equilibrium are discussed in Chapter [5]
Chapter 4

Values of time and schedule delay

"A good sustainability and quality of life indicator: The average amount of time spent in a car."

Paul Bedford (Nozzi, 2003)

This chapter introduces the theoretical framework for accounting for heterogeneity of values of time and schedule delay within an agent- and activity based context and relates it to the traditional, analytical approach of transport economics. Building upon the established MATSim framework, the new approach to include heterogeneity in (dis)utilities of activity performance and models of different types of schedule delay heterogeneity are presented and discussed.

4.1 Utility maximisation in MATSim context

An activity- and agent-based approach to transport modelling diverges from the concept of trips and flows and focuses on the modelling of the decision-making process of individual agents based on static or dynamic activity schedules. In this context, an agent is commonly defined as an individual pursuing the goal of utility maximisation within a certain time frame - often one day. In case of the MATSim framework, individual agents follow a trial and error approach, based on a co-evolutionary algorithm. Thus, using an iterative learning technique, agents optimise their travel choices and daily schedules from iteration to iteration. This leads to the generation of choice sets of alternative daily schedules and results in the convergence to a Stochastic User Equilibrium (SUE). Detailed discussion of this approach and the transition from trips to individual, behavioural travellers was previously thoroughly treated by Nagel and Flötteröd (2012). This work tries to closely follow the notation and structure in description of individual utility functions applied by Nagel and Flötteröd (2012), and presents an extension of this approach to an agent population with heterogeneous values of time.
A traditional, trip-based approach usually focuses on the cost of each trip, whereby the total cost of travel consists of cost of travel time, schedule-delay and additional monetary expenses. In contrast to the trip-based approach, that pursues the objective of single trip cost minimisation, activity-based models operate based on the maximisation of total utility gained by each economic agent during a certain time period. Hence, in the discourse of the activity-based modelling it is more common to speak about utilities and (dis)utilities instead of cost.

Given a personal utility function for every economic agent, MATSim is well suited for the incorporation of heterogeneity on individual and trip levels into the simulation model.

The utility functions of each agent consist of the sum of (dis)utilities from all activities and trips performed through the day:

\[ U_{total,a} = \sum_{i=1}^{m} U_{act}^{(i)}(a) + \sum_{i=1}^{n} U_{trv}^{(i)}(a) \]  

(4.1)

where \( U_{total,a} \) is the total utility from a given daily activity schedule of an agent \( a \). Here \( m \) is the total number of activities and \( n \) the total number of trips between the activities within one day. Nagel and Flötteröd (2012) consider the number of trips being equal to the number of activities \( (m = n) \), as the first and the last activity of the day are both the same type and take place at the same location, and therefore can be counted as one. However, as discussed below in section 4.3 within an activity-based framework the schedule-delay cost actually depends on the marginal utility of an activity and its typical duration. With home being the most common overnight activity, counting its evening and morning parts as one activity implies the same schedule delay cost in the evening and early departure cost in the morning. This does not necessary hold, as value of time and schedule delay cost vary substantially by time of day, as shown for the morning commute by Tseng and Verhoef (2008).

The (dis)utility of travelling in an activity-based context includes the (dis)utility of travel time and (dis)utility of monetary travel cost:

\[ U_{trv}^{(i)}(a) = \beta_{trv}^{(i)}(a) \cdot t_{trv}^{i} + \beta_{mon}^{(i)}(a) \cdot c \]  

(4.2)

with \( t_{i} \) being the travel time of the trip \( i \) performed by an agent \( a \) and \( c \) the monetary cost of the trip. Furthermore, \( \beta_{trv}^{(i)}(a) \) is the agent’s specific sensitivity to the travel time, which in the most of the previous publications differed only between trip modes. However, as seen i.a. in Axhausen et al. (2008) and Hess et al. (2008) \( \beta_{trv}^{(i)}(a) \) also heavily depends on the socio-demographic characteristics of the traveller, length and duration of the trip, trip purpose and the congestion.

\(^1\)In the context of MATSim a more technical term "score" is used in some publications. It is interchangeable with (dis)utility and refers to the same value.
level on the road or within the public transport vehicles. The same applies for $\beta_{\text{mon}(i)}^a$ whose variation relates not only to income differences, but also depends on the origin of the cost and can be further differentiated between fuel cost, parking fees, public transport fare, road tolls etc.

It is important to note that though the travel (dis)utility in the context of an activity-based schedule contains the same components as the travel (dis)utility of the trip-based approach, the two are not exactly the same. Within a trip-based approach, travel (dis)utility is the only cost considered and the opportunity cost of time or as it is also called, the marginal utility of time as a resource is included into this cost. In contrast, an activity-based approach considers (dis)utilities of travelling and activity performance within a certain time period, and the opportunity cost of time is included implicitly, as during travelling no utility from activity performance can be earned. This raises the issue of estimation of separate values of travel (dis)utilities and opportunity cost of time, which was also discussed by Kickhöfer (2014).

The utility activity performance is a logarithmic function of activity duration with a linear parameters controlling for activities’ typical activity duration as introduced by Charypar and Nagel (2005). The major benefit of the logarithmic form is the decreasing marginal utility with always positive values above the specified scaling parameter. The activity utility function for an activity $i$ in MATSim is defined as

$$U^{\text{act}}_a(t_i) = \beta^{\text{act}}_a(i) \cdot t_{\text{typ}}(i) \cdot \ln \frac{t_i}{t_0(i)}, \quad (4.3)$$

where $t_i$ is the duration of an activity $i$ performed by an agent $a$, $t_{\text{typ}}(i)$ the typical duration of this activity and $\beta^{\text{act}}_a(i)$ the marginal utility of the activity $i$ for an agent $a$ at its typical duration.

At this point it is assumed that activities of the same type (e.g., home, work, shopping) have same typical durations for all agents. This does not have to be necessary the case, as e.g. typical duration of work activity can vary among agent dependent on the individual’s profession and type of job. $t_0(i)$ is a scaling parameter, which determines after which period an activity starts to have a positive utility and does not influence marginal utility of an activity. For the marginal activity utility at its typical duration follows:

$$U^{\text{act}}'_a(t) = \beta^{\text{act}}_a(i) \cdot t_{\text{typ}}(i) \cdot \left. \frac{1}{t} \frac{d}{dt} \right|_{t=t_{\text{typ}}(i)} = \beta^{\text{act}}_a(i). \quad (4.4)$$

Theoretically, in equilibrium the marginal utilities of all activities need to be the same. Hence, the typical activity durations $t_{\text{typ}}(i)$ determine the ratios of different activities with the same utility factor $\beta^{\text{act}}_a$. However, in reality, with time-dependent travel cost this does not hold.
The departure time and therefore the marginal utility of the preceding activity at the time of departure is chosen to optimise the total utility from activity performance and travel cost. e.g. the utility loss and penalty for leaving work early but getting home without congestion might outweigh the higher marginal utility of work than home.

If a particular trip of an agent \(a\) takes a shorter amount of time, the time savings can be distributed between any activities of the day with due regard to additional constraints as opening or closing time of locations, where the activities are performed, in case such restriction are in place. This results from the fact, that under the equilibrium conditions the marginal utilities of all activities have to be the equal.

It is also worth noting, that in the event of time pressure between two activities with same utility factor \(\beta_{a}^{act}\), the agent will prefer to allocate the scare time resources to the activity with a shorter typical duration as it follows from Equation (4.4):

\[
\frac{\beta_{a}^{act} \cdot t_{typ}(1)}{(t_{typ}(1) - \epsilon)} > \frac{\beta_{a}^{act} \cdot t_{typ}(2)}{(t_{typ}(2) - \epsilon)} \quad \text{for} \quad t_{typ}(1) < t_{typ}(2).
\]

For the choice of the trip start time, which equals the departure time from an activity, however, this effect is mostly superimposed by the time dynamic travel cost.

### 4.2 Including heterogeneity into the activity-based model

In the econometric discrete choice modelling theory, the utility function consists of two parts: the deterministic part and the random error term. The random error term represents the unobserved heterogeneity and individual’s decision variables not known to the modeller (McFadden, 1978). While MATSim’s transport demand and behaviour model is based on the discrete choice modelling framework, the utility function does not explicitly include the error term for route, mode and departure time choice dimensions. Instead, the stochasticity inherently present in co-evolutionary choice-generation algorithm implicitly introduces randomness into the simulation. Two agents with exactly the same characteristics may end up with different departure times, travel modes and routes, depending on the mutation order, which their plans underwent during the iterative replanning process (Horni, 2013).

Evaluation of utility functions for daily activities and trips on individual level, as it is performed in an agent- and activity-based simulation, represents a well suited framework for inclusion of heterogeneous user preferences. As indicated in the previous section, the modular architecture of MATSim provides an option for extension of the "scoring function" module, allowing to define custom marginal (dis)utilities of travelling, activity performance or monetary expenses
on an individual agent level.

A number of previous studies approached the subject of explicit inclusion of heterogeneous user preferences in the MATSim context. Incorporating destination choice in microsimulation, Horni (2013) addresses the large significance of unobserved preferences in traveller’s destination choice by explicitly adding a random error term to each individual’s utility function. In his work, Horni (2013) discusses computational and methodological challenges, focusing on shopping and leisure activities and efficient application to large-scale scenarios.

Individual’s personal socio-demographic and economic attributes, however, are one of the most decisive sources of heterogeneity. So far, personal attributes in MATSim have been mainly used for activity chain assignment and car ownership models during the synthetic population generation process (Horni et al., 2011; Erath et al., 2012b).

Some of the earlier studies of heterogeneity in the MATSim context, presented by Kickhöfer (2014), Kickhöfer et al. (2011), Kickhöfer et al. (2010), focused on heterogeneity in user perception of monetary travel expenses, such as road tolls. Hence the marginal utility of money was multiplied by an income-dependent term to reflect differences in perception of monetary travel expenses. More recently Nagel et al. (2014) used heterogeneous values of time and varying sensitivities to road tolls to demonstrate the benefits of adding randomness to routing, when faced complex interactions between toll levels and values of time. In all previous studies, however, (dis)utilities of travel time and activity performance were left homogeneous in the population. This work takes a different approach by incorporating heterogeneous values of time directly on the level of individuals’ (dis)utilities of activity performance and travelling.

To account for deterministic taste heterogeneity, this work implements a continuous interaction formulation approach, applied among others by Axhausen et al. (2008) and Hess et al. (2008) for the Swiss value of time study. Following the notation by Axhausen et al. (2008), impact of $y$ on utility of $x$ can be defined as

$$f(y, x) = \beta_x \left( \frac{y}{\hat{y}} \right)^{\lambda_{y,x}} x,$$

with $y$ being the value of the variable influencing the sensitivity of the alternative’s attribute $x$ and $\hat{y}$ being a reference value of this variable. For example, $y$ could be income, age or trip distance and thereby influence the sensitivity to monetary cost, travel time or crowding conditions by altering the utility parameter $\beta_x$. The reference value $\hat{y}$ can be chosen arbitrary, but it usually make sense to use the sample mean value, guaranteeing that $\beta_x$ represents the sensitivity to $x$ at the average value of $y$. Axhausen et al. (2008) and Hess et al. (2008) focus on heterogeneity resulting from variations in income, trip distance and trip purpose. For
the purpose of this study, however, only the formulation for income dependent variability of monetary travel cost (dis)utility ($TC$), as presented in (4.7), is adopted.

$$TC(inc, \tau) = \beta_{\text{mon}} \left( \frac{inc}{\hat{inc}} \right) ^{\lambda_{\text{inc,mon}}} \cdot \tau,$$

(4.7)

with $inc$ being the income variable influencing the sensitivity to monetary expenses $\beta_{\text{mon}}$ and $\hat{inc}$ being a sample mean value used as a reference. The income dependent factor with parameter $\lambda_{\text{inc,mon}}$ in the formulation above will be referred to as the income-sensitivity factor.

How the marginal utility of travel time savings (mUTTS) and the value of time can be derived in MATSim framework, was recently discussed by Nagel et al. (2014). As the implementation of heterogeneity in an activity-based context builds upon these definitions and for the sake of completeness, parts of discussion from Nagel et al. (2014) are repeated below.

In order to derive the marginal utility of travel time savings (mUTTS), for sake of simplicity, a case where travel time savings are distributed between two activities is considered. However, the following argumentation can be easily transferred to any number of activities $m$ within one day.

As pointed out above, in MATSim the total (dis)utility of travelling combines the travel time dependent penalty and the opportunity cost of the time spent on the trip. Hence, the marginal utility of travel time savings (mUTTS) can be expressed as the utility change resulting from the reduction of travel time and the gain in activity duration by the time period $\epsilon$. Under the assumption that the duration trip $i$ can be shortened by $\epsilon$ and the time gain is dedicated to either the subsequent activity, the activity preceding the trip or any combination of the two, it follows:

$$mUTTS_a = \frac{\partial}{\partial \epsilon} \left( \beta^{\text{trv}}_a \left( t^{\text{trv}}_i - \epsilon \right) + \beta^{\text{act}}_a \cdot l^{\text{typ}}_i \cdot \ln \frac{t_i + \epsilon \cdot \alpha}{t_{0(i)}} + \beta^{\text{act}}_{i+1} \cdot l^{\text{typ}}_{i+1} \cdot l^{\text{int}}_{i+1} \cdot (1 - \alpha) \right)$$

(4.8)

$$= -\beta^{\text{trv}}_a + \beta^{\text{act}}_a \cdot t^{\text{typ}}_i + \beta^{\text{act}}_{i+1} \cdot t^{\text{typ}}_{i+1} \cdot (1 - \alpha),$$

with $0 \leq \alpha \leq 1$. The last step of the equation (4.8) takes advantage of the equality of mUTTS for all activities in the equilibrium state. Hence, with $\beta^{\text{act}}_a \cdot t^{\text{typ}}_i = \beta^{\text{act}}_{i+1} \cdot t^{\text{typ}}_{i+1}$, the mUTTS can be expressed using only the opportunity cost of either preceding or following activity.
The two terms contributing to the marginal utility of travel time savings are:

- The marginal (dis)utility of travelling, which can be dependent not only on mode as described by [Nagel et al. (2014)](#), but also affected by individuals’ socio-demographic characteristics, congestion and crowding conditions during the trip as well as trip distance and duration.

- The marginal utility of time as a resource for the activity following the trip, as defined in the equation (4.3).

Marginal utility and marginal value of travel time savings are linked by marginal value of money $\beta_{a}^{\text{mon}}$:

$$mVTTS_a = \frac{mUTTS_a}{\beta_{a}^{\text{mon}}} = -\frac{\beta_{a}^{\text{trv}(i)}}{\beta_{a}^{\text{mon}}} + \frac{\beta_{a}^{\text{act}(i+1)}}{t_{i+1}} \cdot \frac{t_{\text{typ}(i+1)}}{t_{i+1}} \beta_{a}^{\text{mon}},$$  \hspace{1cm} (4.9)

with $t$ being the actual activity durations and $t_{\text{typ}}$ the typical or ideal activity duration. From here [Nagel et al. (2014)](#) derive heterogeneous utilities of money $\beta_{\text{class}}^{\text{mon}}$, using estimated mVTTS values for three different classes of vehicles/drivers and assume all other utility parameters to be homogeneous across the population. Incorporating the heterogeneous perception of tolls in routing and scoring and using a real world case study, [Nagel et al. (2014)](#) obtain improved results by adding additional randomness to the router in form of the randomised road tolls.

Following the logic of activity based modelling approach, it can be argued that the variation in values of time across the population is the consequence of varying utilities, which individuals gain from performing a particular activity. Especially, the connection between value of time, marginal utility of money and utility gained from working is apparent. Translating this argumentation into an activity-based modelling framework, the heterogeneous values of time should be reflected in variations of $\beta^{\text{act}}$ among agents. Then the higher willingness to pay for the reduction of travel times emerges intrinsically from the higher activity utilities and travel (dis)utilities of agents with a high personal or household income.

However, as the factor $\lambda^{\text{inc,mon}}$ is estimated for the heterogeneous perception of travel cost, as defined in Equation (4.7), the heterogeneity then is first added to the marginal utility of money $\beta_{a}^{\text{mon}}$. Thus the heterogeneous utility of money for an agent $a$ becomes:

$$\beta_{a}^{\text{mon}} = \beta_{a}^{\text{mon}} \left( \frac{\text{inc}}{\text{inc}} \right)^{\lambda^{\text{inc,mon}}},$$  \hspace{1cm} (4.10)

After substituting (4.10) into (4.9) and rearranging the heterogeneity multiplier of the monetary cost into the nominator in order to reflect activity utility as the source of heterogeneity in value
of time, marginal value of travel time savings can be formulated as

\[
mVTTS = \frac{-\beta_{\text{trv}}^{(i)} + \beta_{\text{act}}^{(i+1)} \cdot \frac{t_{\text{typ}}^{(i+1)}}{t_{i+1}}}{\beta_{\text{mon}}}.
\]

The notation \(\beta_{TT\text{mode},p}\) highlights the dependency of the marginal travel (dis)utility factor on travel mode and the purpose of the trip, which as [Axhausen et al. (2008)] show, have a significant impact on valuations of travel time. However, as this thesis focuses only on income induced heterogeneity and only commuting trips in the simulation scenario below, the superscript \(p\) is omitted.

Under the assumption of an activity duration being in the neighbourhood of its typical duration, linearisation around \(t_{i+1} = t_{\text{typ}}^{(i+1)}\) makes the mVTTS independent of the duration of the activity following the trip. As [Nagel et al. (2014)] point out, using this approximation for routing substantially simplifies robust software design.

Combining the homogeneous utility parameters for travel and activities with the heterogeneity factors, leads to:

\[
mVTTS = \frac{-\hat{\beta}_{\text{trv}}^{(i)} + \hat{\beta}_{\text{act}}^{(i+1)}}{\beta_{\text{mon}}} .
\]

with

\[
\hat{\beta}_{\text{trv}}^{(i)} = \beta_{TT\text{mode}} \left( \frac{\text{inc}}{\text{inc}} \right)^{-\lambda_{\text{inc},\text{mon}}}
\]

\[
\hat{\beta}_{\text{act}}^{(i)} = \beta_{\text{act}} \left( \frac{\text{inc}}{\text{inc}} \right)^{-\lambda_{\text{inc},\text{mon}}}
\]

As equation 4.12 shows, the above formulation adjust the marginal utilities of activity performance and travelling for each individual, keeping the marginal utility of monetary travel expenses constant across the population. This transformation does not change individual’s mVTTS, but has an important behavioural effect. Becoming intrinsic property of the model,
heterogeneity in VTTS does not anymore directly depend on level of monetary travel expenses. It also paves the way towards activity specific marginal utilities and consistent inclusion of other heterogeneity factors as well as potential addition of budgeting concept into the modelling framework. Moreover, as it is demonstrated in the following section, this approach facilitates inclusion of schedule delay heterogeneity in an activity-based context as it intrinsically emerges from income dependent variation of the marginal activity utility.

4.3 Schedule Delay Modelling – Being Early or Late

In the dynamic models of congestion the term schedule delay commonly refers to the time difference between preferred arrival and actual arrival time at an activity or preferred and actual departure time from an activity. The cost associated with these, for travellers undesired time deviations is referred to as the schedule delay cost. Affected by a broad range of factors, schedule delay cost display a wide inter- and intra-personal variability, leading to challenges in study and quantification of its distributions (Small et al., 2005). However, as discussed in the Chapter 2 a number of recent studies (e.g. van den Berg, 2014; van den Berg and Verhoef, 2013, 2011a,b) show, that heterogeneity in schedule delay cost matters. It can have a significant influence on welfare and distributional effects in transport planning and policy design, and therefore requires more attention from modellers and practitioners.

Schedule delay cost are commonly divided into schedule delay early ($\beta$) – arriving at an activity location before the desired activity can be started and the schedule delay late ($\gamma$) – arriving later than planned.

Though schedule delay early and schedule delay late both represent an indicator of individual’s flexibility to breach external timing constraints, variations of these values among individuals can have very different roots. Schedule delay early relates to the ability to use idle time in a productive or enjoyable manner. While one person might be able to gain some value from it, others can perceive it as a wasted waiting time. Schedule delay late on other side is the flexibility of being somewhere later than preferred or the gravity of consequences of being late. For example, for a high-income financial industry employee who uses quite, early morning hours in the office to answer emails, the loss from being at work earlier than planned is rather limited. However, coming in after stock market has opened is a no go and can have severe consequences.

Timing and schedule delay

Congestion and traffic delays and thereby schedule delay cost arise from large share of population having the same or very similar preferred arrival or departure time, as it is the case for rigid working hours across many professional groups. Hence, the definition of preferred
arrival times and their distribution across the population is a strong determining factor of travel and schedule delay. In transport economic literature in general and in the classic bottleneck model in particular, preferred arrival time $t^*$ is modelled as one fixed value for all travellers. de Palma and Lindsey (2002) study travellers heterogeneity in preferred arrival and departure time, focussing on its effect on differences between morning and evening commuting peaks.

From a perspective of an individual, schedule delay early cost $\beta$ and the preferred arrival time are the two sides of the same coin. Schedule delay cost can essentially be interpreted as the degree of rigidity of the preferred arrival and activity start timings. An extreme case with schedule delay cost being close to zero implies individual’s indifference to the preferred arrival timings.

In an agent- and activity based model, the idealised concept of single preferred arrival time $t^*$ can be extended to a preferred arrival time window. Within this time window, an individual is completely insensitive to the arrival time and could arrive at any time without gaining or loosing. In case of the time invariant travel cost, any arrival within the preferred time window leads to the same benefits for each individual. Activity-based approach allows to account for the variability of preferred arrival time not only by definition of opening and closing times, but also with additional penalties for arrival and departures outside of pre-defined time windows.

**Commuting**

Work activity has traditionally been the most prominent example for the illustration and study of schedule delay cost. The often rigid working hours lead to the high penalisation of arriving at a workplace too late or leaving it too early. Though, high schedule delay cost can also occur in the context of other activities as e.g. being late for picking up children, a dinner reservation or a date, cost of delays for leisure and other activities are hard to study and quantify. This is mainly due to the high variability of such costs for a given individual based on a trip purpose, time of day or type of social interaction or joint activity planning.

Schedule delay heterogeneity for commuting work trips can be linked to several causes. Individual’s type of job and associated working hours constraints are naturally considered as the most relevant factor. A common assumption is the direct proportional relation of schedule delay heterogeneity with the income dependent values of time. However, correlation between values of time and schedule delay does not exclude other factors, such as type of jobs, which might alter the ratio of marginal values of time and schedule delay for each individual. Within the dynamic bottleneck models of congestion these individual ratios are considered as decisive. Different types of heterogeneity, such as proportional, $\alpha$ and $\gamma$ heterogeneity capture different types of the relationship between value of time and schedule delay and are discussed in detail below, in section 4.3.2.
Dynamic propagation

A number of publications studied schedule delay in the context of the morning commute problem, using the same time preference for the arrival at the workplace among all travellers. The starting point is commonly the bottleneck model presented by Vickrey (1969), which is then extended with heterogeneous preferences or additional choice dimensions (see Small (2012) for an overview). However, comparatively little has been published about the evening peak, where traveller’s deviation from preferred departure-time lead to an additional cost. Vickrey (1973), Fargier (1983) and later de Palma and Lindsey (2002) compared the morning and evening commutes and demonstrated that the convenient symmetry of the two peaks in case of identical travellers breaks down when heterogeneous trip timing preferences, values of time and schedule delay cost are considered. Yet, only little attention was given to the interdependency of morning and evening peaks and the whole day schedule dynamic. As activities require a minimal duration for their performance, delays tend to propagate through the day and often have implications for all consecutive activities.

4.3.1 Schedule delay in an activity – based context

Accurately capturing different levels and dimensions of user’s preference heterogeneity among individuals and trips in a single model is a challenging task. From the transport economics perceptive, the ratios of individual’s value of time, schedule delay early and schedule delay late play a determining role in trip scheduling decisions and are therefore of major relevance for the welfare evaluation. Hence, it is common to derive absolute values of schedule delay from this relationship as it is observed in empirical studies. The values commonly used are based on assumptions by Arnott et al. (1990): $\mu = \frac{\alpha}{\beta} = 2$ and $\eta = \frac{\gamma}{\beta} = 3.9$. This basically implies, that willingness to pay to save a minute of being at some activity location earlier than intended is only a half of willingness to pay to save a minute of travel time and approximately one forth of saving a minute of being late.

In an activity-based approach as it is used in this work, the value of time can be considered as being composed from two components: opportunity cost of time and (dis)utility of travelling. Here the schedule delay cost is intuitively related to the first.

In case of the agent- and activity-based framework MATSim, (dis)utility of schedule delay late is defined as an extra penalty term, which is added to the utility function in case the predefined activity start timing constraints have been breached (Nagel and Flötteröd, 2009). Schedule
delay cost for being late in MATSim is defined as follows:

$$U^{late(i+1)} = \beta^{late} \cdot t_{late(i+1)},$$

(4.15)

where $$\beta^{late}$$ is the marginal (dis)utility of being late and $$t_{late(i+1)}$$ is the time duration by which the start of the activity $$i + 1$$ got delayed relative to the preferred activity start time. The definitions of activity and travel utility functions within an activity based framework as discussed in section 4 lead to the opportunity cost of time being intrinsically included into the travel cost. As no utility from activity performance can be earned during the trip, the foregone utility is accounted for within the cost of delay and as [Nagel and Flötteröd, 2009] argue correctly, does not need to be added to the late arrivals.

It is straightforward to extend equation (4.15) from homogeneous to the heterogeneous case. Substituting $$\beta^{late}$$ with $$\hat{\beta}^{late}_a = f(inc, act(i))$$ leads to schedule delay cost late being dependent on income and activity type of a particular agent $$a$$.

The schedule delay early occurs when a traveller receives an additional penalty for arriving at an activity location before the desired arrival time. This is usually the case if an activity can not yet be started due to external constraints such as e.g. business opening hours. Transferring this notion into an activity-based context, an early arrival corresponds to the cost of doing nothing and therefore to the opportunity cost of time, if no additional extra cost is added [Nagel and Flötteröd, 2009]. However, with time dynamic values of time, the opportunity cost of time and therewith the cost of schedule delay early can vary in the course of the day. Following the notion that being early at an activity location can be interpreted as a consequence of an early departure from the previous activity, the schedule delay early is determined by the utility of the activity $$i$$ preceding the trip $$i$$. Under the assumptions that this activity $$i$$ could have been performed longer by $$t = t_{early(i+1)}$$ and that its duration is close to $$t_{typ(i)}$$, the cost of schedule delay early can be expressed as follows:

$$CSD_{early}(t_{early(i+1)}) = \hat{\beta}^{act(i)}_a \cdot t_{typ(i)} \cdot \left( \ln \frac{t_i + t_{early(i+1)}}{t_0(i)} - \ln \frac{t_i}{t_0(i)} \right)$$

$$= \hat{\beta}^{act(i)}_a \cdot t_{typ(i)} \cdot \ln \left( \frac{t_i + t_{early(i+1)}}{t_i} \right)_{t_i = t_{typ(i)}}$$

$$= \hat{\beta}^{act(i)}_a \left( \frac{inc}{inc_{mean}} \right)^{-\lambda_{inc,mean}} \cdot t_{typ(i)} \cdot \ln \left( 1 + \frac{t_{early(i+1)}}{t_{typ(i)}} \right).$$

(4.16)
Linearising around $t_{\text{early}} = 0$ results in:

$$CSD_{\text{early}}(t_{\text{early}}(i+1)) \approx CSD_{\text{early}}(0) + CSD'_{\text{early}}(0) \left( t_{\text{early}}(i+1) - 0 \right)$$

$$= \beta_{\text{act}}(i) \left( \frac{inc}{\hat{inc}} \right) \cdot \lambda_{\text{inc,mon}} \cdot t_{\text{early}}(i+1).$$

As it can be seen from (4.17), $CSD_{\text{early}}$ corresponds to the marginal utility of the previous activity at its typical duration.

Using the notation common in transport economics for linear factors of value of time ($\alpha$), schedule delay early ($\beta$) and schedule delay late ($\gamma$), these parameters can be summarized for an activity-based framework as follows:

$$\alpha = mVTTS \cdot \beta_{\text{mon}} = -\hat{\beta}_{\text{trv}}(i) + \hat{\beta}_{\text{act}}(i+1);$$
$$\beta = \hat{\beta}_{\text{act}}(i);$$
$$\gamma = \hat{\beta}_{\text{late}}.$$

In order to illustrate the definitions of value of time and schedule schedule delay in MATSim’s activity-based framework, Figure 4.1 shows marginal utility functions for a morning commute. Under the assumption of fixed work end time and an instantaneous switch from home to work activity, the agents departure time would be naturally equal to its preferred arrival time $t^*$, with $t^*$ being defined by the equality of marginal home and work utilities. However, given a constant
marginal (dis)utility of travel time, an agent will have to balance between marginal utilities and time dependent travel cost in order to find its optimal schedule. Values of time $\alpha$, $\beta$ and $\gamma$, as visualised in Figure 4.1 are defined in equation 4.18 and were discussed above. Cost of schedule delay late $\gamma$ occur when agent arrives at work after 9am. Given the assumptions of travel time $t = 0$, schedule delay can be basically interpreted as the (dis)utility of staying at home after 9am and is therefore visualised in Figure 4.1 as a part of marginal utility of the home activity.

Up to this point, the definitions of value of time and schedule delay involved a trip-dependent value of time based on the type of activities $i$ preceding and $i + 1$ following the trip. As highlighted above, this represents an important aspect and source of heterogeneity. However, as this work mainly focuses on income dependent valuations of time during the commuting trips, from this stage on identical value $\beta^{\text{act}}$ is used for the marginal utilities of all activities. Sub- and superscripts $i$ are omitted in the future notation.

For a homogeneous user case, values of time in Equation (4.18) become constant and do not depend on the socio-economic characteristics of an individual agent, as it is shown in Equation 4.19.

\[
\begin{align*}
\alpha &= mVTTS \cdot \beta^{\text{mon}} = -\beta^{\text{trv}} + \beta^{\text{act}} \\
\beta &= \beta^{\text{act}} \\
\gamma &= \beta^{\text{late}}
\end{align*}
\]

(4.19)

### 4.3.2 Three forms of schedule delay heterogeneity

Three forms of user’s scheduling preference heterogeneity are commonly considered in the literature: proportional heterogeneity, $\alpha$-heterogeneity and $\gamma$-heterogeneity (van den Berg (2014)). Being defined by the ratios of value of time and schedule delay cost $\mu = \frac{\alpha}{\beta}$, $\eta = \frac{\gamma}{\beta}$ and $\lambda = \frac{\alpha}{\gamma}$, different model forms try to capture varying trade-offs and user preferences between travel and arriving before or after the preferred arrival time at a given activity location. It is worth to note, that the ratios $\mu$ and $\lambda$ have also a behavioural interpretation, reflecting individual’s willingness to accept greater schedule delays, early or late, in order to reduce travel time (van den Berg and Verhoef 2011b). The greater $\mu$ or $\lambda$ are, the more important is travel time in comparison to schedule delay.

**Proportional heterogeneity**

Initially addressed by Vickrey (1973), proportional heterogeneity is based on the assumption of a direct link between value of time and schedule delay values. A person with high value of
travel time savings, is expected to have higher absolute values of schedule delay. However, the ratios of value of time and schedule delay $\mu$, $\eta$ and $\lambda$ remain constant for all travellers. Values of schedule delay are rigidly linked to individuals value of time and vary proportionally over travellers, following the same distribution as values of time.

By defining $\beta_{late}$ in the equation (4.18) as $\beta_{late} = \beta_{const} \cdot \beta_{act}$ with $\beta_{const} = \text{const.}$, the heterogeneity in an activity-based framework, as presented above corresponds to the proportional heterogeneity type. For the definition of value of time and schedule delay in the activity-based MATSim context follows:

$$\alpha = -\beta_{trv} + \beta_{act} = -\beta_{trv} \cdot \left(\frac{inc}{inc}\right)^{-\lambda_{inc,mon}} + \beta_{act} \cdot \left(\frac{inc}{inc}\right)^{-\lambda_{inc,mon}}$$

$$\beta = \beta_{act} = \beta_{const} \cdot \left(\frac{inc}{inc}\right)^{-\lambda_{inc,mon}}$$

$$\gamma = \beta_{const} \cdot \beta_{act} = \beta_{const} \cdot \beta_{const} \cdot \left(\frac{inc}{inc}\right)^{-\lambda_{inc,mon}}$$

(4.20)

From the equations (4.20), it is easy to see, that ratios $\mu = \frac{\alpha}{\beta}$ and $\eta = \frac{\gamma}{\beta}$ stay constant for all agents, with $\alpha$, $\beta$ and $\gamma$ varying proportionally with the income factor. This corresponds to the definition of proportional heterogeneity.

A major drawback of this heterogeneity model is that it does not account for different scheduling flexibilities as e.g. based on profession or workplace. In fact, travellers with higher income may have more flexible jobs and more power to adjust their personal schedule, while workers with lower income, often employed in the service, healthcare, education or manufacturing industries, tend to be bound by strict opening and shift schedules. Hence, commuters with the high value of time due to their high income might have lower relative cost of arriving late or early at their work place, as observed by Koster and Koster (2013) and also discussed by van den Berg (2014). A potential way to address this problem is to model values of schedule delay as an inverse function of income. This, however, would alter the ratio $\mu$ among individuals and lead to the presence of the so called $\alpha$ – or ratio heterogeneity, which is discussed in following paragraph.

$\alpha$ – heterogeneity / ratio heterogeneity

The second type of the schedule delay heterogeneity attempts to capture the variable perception of being early or late among individuals. Hence, it relates to the varying ratios of value of time and schedule delay, as it is not considered in case of the proportional heterogeneity. It is referred to as $\alpha$-heterogeneity, as for convenience in modelling practice the value of time $\alpha_a$ usually varies among individuals while other values stay fixed. However, the term ratio or $\mu$
heterogeneity appears to be more appropriate, as being characterised by constant \( \eta = \frac{\gamma}{\beta} \) and varying \( \mu = \frac{\alpha}{\beta} \) and \( \lambda = \frac{\alpha}{\gamma} \), this model of heterogeneity aims to capture variability of individual flexibility for arriving late or early, which might be either independent or somehow correlated to the individual’s value of time.

The agent-based simulation approach allows to easily investigate scenarios closer to reality, with proportional heterogeneity and heterogeneity \( \alpha \) overlapping and therefore creating a 2-dimensional distribution. Similar approach was previously studied by (van den Berg and Verhoef, 2011b), who used a symmetric triangular distribution for \( \beta \) and \( \mu \). One of the major issues hereby is the lack of data on the distribution of schedule delay early values within the population and its correlation with individual or household incomes and value of time. This issue is discussed in detail in the Chapter 5.

Under the assumption of \( \beta \) being linearly dependent on \( \alpha \) by a randomly distributed factor \( \zeta_{\beta} \), the joint definition of proportional and \( \alpha \)-heterogeneity is as follows:

\[
\begin{align*}
\alpha &= -\beta^{trv} + \beta^{act} = -\beta^{trv}_{\text{cost}} \cdot \left( \frac{\text{inc}}{\text{inc}_{\text{mon}}} \right)^{-\lambda_{\text{inc,mon}}} + \beta^{act}_{\text{cost}} \cdot \left( \frac{\text{inc}}{\text{inc}_{\text{mon}}} \right)^{-\lambda_{\text{inc,mon}}} \\
\beta &= \zeta_{\beta} \cdot \beta^{act} = \zeta_{\beta} \cdot \beta^{act}_{\text{cost}} \cdot \left( \frac{\text{inc}}{\text{inc}_{\text{mon}}} \right)^{-\lambda_{\text{inc,mon}}} \\
\gamma &= \eta \cdot \beta
\end{align*}
\]  

(4.21)

With \( \eta = \frac{\gamma}{\beta} = \text{const.} \) and \( \mu \) and \( \lambda \) varying across the population, conditions of \( \alpha \) heterogeneity are fulfilled.

\( \gamma \) – heterogeneity

The third type of heterogeneity is usually referred to as \( \gamma \)-heterogeneity, where the value of schedule delay late \( \gamma \) varies among individuals, but other values stay constant. What really matters here, however, is the variation of the ratio \( \eta = \frac{\gamma}{\beta} \), which expresses the willingness of individuals to arrive at an activity rather early than late.

Similar to the case of \( \alpha \) heterogeneity, simultaneously varying values of \( \gamma \) and \( \alpha \) across the population represents a more realistic situation. Hence, the term \( \gamma \)-heterogeneity is slightly misleading as it basically refers to the \( \eta \)-heterogeneity.

With the schedule delay early \( \beta \) varying proportionally with \( \alpha \) and individual’s \( \gamma \) value being linked to \( \alpha \) by a randomly distributed factor \( \zeta_{\gamma} \), values of time and schedule delay can be
expressed as follows:

\[
\alpha = -\beta^{trv} + \beta^{act} = -\beta^{trv}_{cost} \cdot \left( \frac{inc}{inc} \right)^{-\lambda_{inc,mon}} + \beta^{act}_{const} \cdot \left( \frac{inc}{inc} \right)^{-\lambda_{inc,mon}}
\]

\[
\beta = \beta^{act} = \beta^{act}_{const} \cdot \left( \frac{inc}{inc} \right)^{-\lambda_{inc,mon}}
\]

\[
\gamma = \frac{\gamma}{\beta} \cdot \beta^{act} = \frac{\gamma}{\beta} \cdot \beta^{act}_{const} \cdot \left( \frac{inc}{inc} \right)^{-\lambda_{inc,mon}}
\]

Following from equation (4.22), \( \eta = \frac{\gamma}{\beta} \) and \( \lambda = \frac{\alpha}{\gamma} \) become heterogeneous ratios, while \( \mu = \frac{\alpha}{\beta} \) stays constant. This corresponds to the common definition of \( \gamma \) heterogeneity as discussed above.

As this work assumes only income dependent value of time heterogeneity, the subscripts in \( \lambda_{inc,mon} \) will be omitted, substituting \( \lambda_{inc,mon} \) with \( \lambda \) for a better readability in the following chapters.
Chapter 5

Experimental Scenario Set-up

"One of the fundamental aims of scenarios is to create multiple options; so they must be flexible, in order to incorporate both quantitative and qualitative variables; they attempt to single out the interrelationships among these variables in backgrounds characterised by both swift changes and high complexity."

Antonio Martelli (Martelli, 2014)

This chapter presents the experimental scenario set-up, designed to evaluate effects of heterogeneous values of time in a multi-modal context. This includes description of transport supply, transport demand as well as heterogeneous behavioural parameters based on individuals’ socio-economic characteristics.

5.1 Corridor scenario

For the initial evaluation of the presented methodology and effects of heterogeneous travellers, a rather simple, multi-modal Corridor scenario is chosen. Essentially limiting the agents’ choice dimensions to departure time and mode choice facilitates the identification of impact from modelling users’ heterogeneous values of time as well as effects of congestion and dynamic public transport pricing policies. At the same time, limiting agents to two degrees of freedom enables a comparison between analytical economic models and agent-based simulation methodology.

Figure 5.1 shows a sketch of the scenario set-up. The 20 km long corridor, with home locations distributed on one and work locations on the other side, consist of three lanes in each direction with flow capacity of 800 vehicles per hour and lane. The multi-modal scenario also includes a
Figure 5.1: Corridor scenario set-up, with bus stops located every 600 meters and home and work locations normally distributed on the either end of the corridor.

bus line, that is operating in each direction along the full corridor length, with bus stops located every 600 meters. The scenario can be interpreted as an arterial road of a mono-centric city with housing located on the outskirts and work places in the city centre.

5.1.1 Supply

In the multi-modal Corridor scenario, cars and buses share the same road space. As highlighted above, this implies that bus operations and travel times are affected by road congestion. It is also assumed that every bus stop is equipped with a bus bay and therefore buses do not block following traffic while stopping for boarding and alighting passengers. Furthermore, bus headways are varied between 10 and 2 min in order to investigate the impact of varying service levels and public transport capacities on effects of pricing policies with heterogeneous user preferences. For accurate assessment of welfare effects of varying bus service levels, it is important to account for both, operation and capital cost of the service provision. To this end, the model and cost estimates presented by the [Australian Transport Council (2006)] and discussed in section 3.4 are applied. As only morning and evening peak hours and commuting trips are considered in the analysis, cost of bus operations from 6am - 10am and 5pm - 21pm are taken into account. The capital cost are accounted for in full scale, neglecting the potential use of the same public transport vehicles through the day and serving a wider ridership of non-commuting passengers. Though this may lead to overestimating cost of bus operations during peak-hours, the assumption of peak-hour operations subsidising the off-peak service is reasonable. Accounting for potential delays along the route and providing a time buffer for
Table 5.1: Bus operation cost

<table>
<thead>
<tr>
<th>Headway (min)</th>
<th>Number of vehicles</th>
<th>Fixed cost</th>
<th>Operating cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>12</td>
<td>$2'006</td>
<td>$6'280</td>
<td>$8'286</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>$4'013</td>
<td>$12'559</td>
<td>$16'572</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>$10'032</td>
<td>$31'398</td>
<td>$41'430</td>
</tr>
</tbody>
</table>

on-time service in the opposite direction, a total time of 1h is assumed for a vehicle and driver to serve the 20 km corridor in one direction. The resulting cost estimates are presented in Table 5.1.

5.1.2 Travel demand and behavioural parameters

The agent population consist of 8'000 agents, all with the same daily home - work - home activity chain. The home locations of agents follow a normal distribution along the west side of the corridor, with $\mu = 6.67$ km and standard deviation $\sigma = 3.33$ km. On the other side, the work locations are distributed around $\mu = 13.33$ km with the same standard deviation $\sigma = 3.33$ km. Furthermore, both home and work locations are uniformly distributed on the north and south side of the corridor, with maximal distance of 1 km to each side. This results in a maximal crow fly distance to the closet but stop of 1.04 km.

For the single corridor scenario two choice dimension are relevant: departure time and mode choice. Given the simple network configuration in a corridor, route choice does not feature as an independent choice dimension at this stage. In the simulation set-up presented here, a time window of 120 minutes is used for each single modification of departure time. A new departure time, which is the same as the activity end time, is picked randomly according to a uniform distribution in a range of $+/- 60$ minutes of the initial departure time, as evaluated in one of the preceding iterations.

In contrast to the sharp preferred arrival time for all travellers commonly used in the models of transport economics, defining a preferred activity start window in context of an agent-based simulation approach enables evolution of departure and arrival time distributions during peak hours as observed in reality. Consequently, the width of the time window for an activity start as well as its relation to activity’s typical duration affect how peaked or spread out the demand during peak-hours will be and therefore the degree of congestion in the network.

In the agent based context, the preferred arrival and departure times to and from activities are determined by typical activity duration parameters as well as additional constraints for the time intervals during which each activity can actually be performed (also called facility opening times). Based on a logarithmic form of the utility function for all activities, as presented in Equation 4.3, the typical activity duration and the marginal utility parameters essentially
Table 5.2: Activity constraints

<table>
<thead>
<tr>
<th>Activity</th>
<th>Typical duration</th>
<th>Opening time</th>
<th>Latest start time</th>
<th>Earliest end time</th>
<th>Closing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>14h</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Work</td>
<td>9.5h</td>
<td>8.00 a.m.</td>
<td>9.00 a.m.</td>
<td>6.00 p.m.</td>
<td>7.00 p.m.</td>
</tr>
</tbody>
</table>

shift the function response along the utility axis. An additional set of parameters: "latest start time" and "earliest end time" helps to translate the definition of schedule delay cost to the activity-based framework. Arrival at an activity location after "latest arrival time" or departure from an activity before the "earliest departure time" induces an additional schedule delay penalty. Disutility of the schedule delay depends on the model of scheduling preferences, as discussed in section 4.3.2. Table 5.2 presents activity timing constraints as defined in the Corridor scenario.

The behavioural parameters for the Corridor scenario are borrowed from the enriched, agent- and activity-based Sioux Falls model, as presented in Chakirov and Fourie (2014) and initially estimated by Tirachini et al. (2014) from results of a stated choice survey conducted in Sydney in 2009 (Hensher et al., 2011). Table 5.3 summarizes the parameters as they are used in the Corridor scenario.

Inherently, the marginal travel time related disutility coefficients estimated in traditional discrete-choice models combines the opportunity cost time and the additional disutility caused by the travel time with the corresponding mode. Applying this behavioural parameters in an activity-based model requires to split the estimated utility parameters into its components and assign a separate utility for activity performance and a disutility for travel time. The value of time equals to the estimates from the survey. This approach is consistent with economic approaches which use the inherent opportunity cost of time and additional utilities or (dis)utilities of time, depending how it is spent. Jara-Diaz et al. (2008) present a choice-experiment with a separate estimation of the marginal utility of time and (dis)utility of travelling. However, in most cases obtaining such disaggregated values from stated or revealed preference data, which is not designed for it, is difficult or even impossible. Previous studies (e.g. Kickhöfer, 2014; Kickhöfer et al., 2011; Kaddoura et al., 2015) worked with the assumption of travelling by mode with a smallest disutility being as good or bad as doing nothing and therefore being only equal to the opportunity cost of time. The "doing nothing" situation occurs in case of early arrival at an activity location or departure after the closing time and therefore corresponds to the disutility of schedule delay early \( \beta \) or in MATSim context the (dis)utility of waiting. The empirical evidence, however, suggests that waiting is commonly perceived two to three times as bad as the time spent travelling. Hence, in this work the split of marginal travel time related disutility for car mode from Tirachini et al. (2014) is defined by the ratio \( \alpha : \beta = 2 \) (Arnott et al., 1990).
Table 5.3: Behavioural and monetary simulation parameters for the Corridor scenario

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{act}$</td>
<td>+ 0.48 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,car}$</td>
<td>- 0.48 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,pt}$</td>
<td>-0.66 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,walk}$</td>
<td>-1.401 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{wait,pt}$</td>
<td>-1.458 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{cost}$</td>
<td>-0.062 [utlis/$]</td>
</tr>
<tr>
<td>$\beta_{0,car}$</td>
<td>-0.562 [utlis]</td>
</tr>
<tr>
<td>$\beta_{0,pt}$</td>
<td>-0.124 [utlis]</td>
</tr>
<tr>
<td>$\beta_{0,walk}$</td>
<td>0.0 [utlis]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT Fare</td>
<td>2 $ / trip</td>
</tr>
<tr>
<td>Car cost per km</td>
<td>0.2 $ / km</td>
</tr>
<tr>
<td>Parking cost</td>
<td>12 $ / day</td>
</tr>
</tbody>
</table>

5.1.3 Socio-economic characteristics

Socio-economic characteristics are a major source of heterogeneity in any urban population. A wide body of literature investigates their influence on activity types and locations and the associated travel behaviour, as highlighted and discussed in Chapters 1 and 4.

In the scope of this study, income is the single socio-economic characteristic attached to the agent population. As previously noted in section 4.3 while income is often considered to be the major driver between proportional schedule delay heterogeneity, it is definitely not the only factor affecting VOT variations. This is particularly the case for intra-personal differences in schedule delay cost and thus for $\alpha$ and $\gamma$ heterogeneity. When employing income variations as the source of heterogeneity in VOT, it is also important to question the level on which the monetary income matters the most. For the larger parts of the developed world with traditional families as the dominant household form, it is fare to assume that the total household income has a dominant impact on personal valuation of time for the majority of adult family members.

In this work household income represents the single socio-economic characteristic attached to the agent population. The household incomes in the Corridor scenario are generated using the income distribution from the synthetic population of the Sioux Falls scenario presented by Chakirov and Fourie (2014). In case of Sioux Falls scenario the distribution emerges intrinsically during the iterative proportional fitting inflation procedure. In order to recreate the same incomes distribution, a log-normal probability density function was fitted to the Sioux Falls income distribution and used to randomly draw incomes for the 8’000 agents of the Corridor scenario. Figure 5.2 shows the resulting histogram of incomes. No correlation between personal income and home or work locations exists and the daily commute distance is therefore independent of the income. Furthermore, car ownership is not specifically modelled in the Corridor scenario and is equal to 100%.

As mentioned before, one of the main objectives of this study is to provide a better understanding for the role of heterogeneous values of time for scenario and policy evaluation.
in an activity- and agent-based context. Advancing from a homogeneous to a heterogeneous user perspective, two factors play a major role: the shape and parameters of the distribution of a particular socio-economic independent variable and the sensitivity of user behaviour or user decisions to the variable. Both factors are strongly dependent on the economic strength, level of equality and culture in a particular region (e.g. the share of transport cost from total living expenses plays an important role).

Following the formulation for heterogeneous preferences presented in Equation 4.7, the degree of heterogeneity and the sensitivity of the utility function to income can be controlled by adjusting the parameter $\lambda$. Using a scale factor $n$ for variation of $\lambda$, the contribution of travel cost to the utility term becomes:

$$TC(inc, \tau) = \frac{\beta_{mon}}{m} \left( \frac{inc}{\hat{inc}} \right)^{n \cdot \lambda} \cdot \tau, \quad with \quad m = \frac{1}{N} \sum_{p=1}^{N} \left( \frac{inc_p}{\hat{inc}} \right)^{n \cdot \lambda}.$$  

(5.1)

Here $\tau$ is the amount of monetary expenses and $m$ the normalization factor equal to the average of the income dependent correction term. As the mean value of the income sensitivity factor is $\neq 1$, introducing it with a constant $\beta_{mon}$, which was estimated separately, leads to an increase in the average value of time. Hence, to ensure comparability between the homogeneous reference case and various heterogeneity scenarios, the mean value of the travel cost factor and thereby the average value of time across the population is kept constant. This can be achieved by introducing the normalization term. In practice, if estimation of marginal utility of travel cost is performed directly using continuous formulation, the normalisation term becomes redundant as it would be intrinsically incorporated into the $\beta_{mon}$ value.

For heterogeneous (dis)utilities of activity performance and travelling in the MATSim
framework accordingly follows:

\[
\hat{\beta}_{trv} = \beta_{trv} \cdot m \cdot \left( \frac{inc}{\hat{inc}} \right)^{-n \cdot \lambda}, \quad \hat{\beta}_{act} = \beta_{act} \cdot m \cdot \left( \frac{inc}{\hat{inc}} \right)^{-n \cdot \lambda}.
\] (5.2)

The role of the factor \( n \) used to scale the parameter \( \lambda \) can be interpreted in two ways. On the one hand, scaling \( \lambda \) can be seen as a change in perception sensitivity of the monetary travel cost expenses relative to the household income. This sensitivity can vary based on a number of economic and cultural characteristics of a particular geographic region or even for the same person dependent on a trip purpose. As Axhausen et al. (2008) observe, the \( \lambda \) for business trips is significantly higher than for commuting trips.

On the other hand, slightly rewriting equation 5.1 as indicated in equation 5.3 allows to interpret different \( n \) - factors as variations in spread of the underlying income distribution and therefore varying inequality.

\[
TC(inc, \tau) = \beta_{mon} \frac{m \left( \frac{inc}{\hat{inc}} \right)^{\lambda}}{\tau} = \beta_{mon} \frac{m \left( \frac{inc_{new}}{\hat{inc}_{new}} \right)^{\lambda}}{\tau}.
\] (5.3)

A common measure of inequality is the Gini-coefficient. It indicates the deviation of income distribution from a perfectly equal distribution and is 0 in case of perfect equality of incomes and 1 for a population where one person is receiving all the income (Gini, 1921; Cowell, 2011). Geometrically, it can be visualized as the area difference between the integral of the real cumulative income percentage curve and the integral of the perfectly equal cumulative income percentage curve. Such plots are refereed to as Lorenz curves and are presented in Figure 5.3 for different \( n \)-factors and the underlying income distribution of the Corridor scenario. It is interesting to note that Gini-coefficient of 0.20 (\( n=0.5 \)) is rather close to a more equal country such as as Sweden, 0.39 (\( n=1 \)) is slightly under the US average and close to the Gini estimates for UK and 0.69 (\( n=2 \)) is only slightly above the Gini-coefficient estimates for South Africa (World Bank, 2015; Central Intelligence Agency, 2013). The scenarios of \( n=0 \) is the homogeneous user case and \( n=3, n=5 \) can be considered as extreme reference cases, or as scenarios with not only high income inequality, but also higher sensitivity to travel cost. Such situation could be expected in places, where the share of transport cost of the total living expenses is especially high.

Table 5.4: Overview of heterogeneity types.

<table>
<thead>
<tr>
<th>References in text / literature</th>
<th>( \alpha )</th>
<th>( \mu = \alpha / \beta )</th>
<th>( \eta = \gamma / \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>proportional</td>
<td>variable</td>
<td>const.</td>
<td>const.</td>
</tr>
<tr>
<td>( \alpha ) heterogeneity, ( \mu ) ratio heterogeneity</td>
<td>variable</td>
<td>variable</td>
<td>const.</td>
</tr>
<tr>
<td>( \gamma ) heterogeneity, ( \eta ) ratio heterogeneity</td>
<td>variable</td>
<td>const.</td>
<td>variable</td>
</tr>
</tbody>
</table>
Figure 5.3: Lorenz curves and Gini coefficients for different heterogeneity parameters.

![Lorenz curves and Gini coefficients](image)

Figure 5.4: Distribution of Value of Time for different heterogeneity factors (car mode).

![Distribution of Value of Time](image)

Proportional heterogeneity

The bean plot in Figure 5.4 shows the effective value of time distributions dependent on the heterogeneity factor \( n \) and based on behavioural parameters and the income distribution as presented in section 5.1.2. Table 5.5 provides the summary of these distributions, highlighting the growing gap between minimum and maximum values of \( \alpha \) and dropping median for constant mean values. It is important to note, that though the distribution for heterogeneity factors \( n = 3 \) and \( n = 5 \) can be considered as extreme cases of income inequality (Gini = 0.86 and 0.98 respectively), research indicates that a wider spread and long-tail of value of time distribution might be more common than assumed. Fosgerau (2006) investigates the distribution of the value of travel time savings from a Danish value of time study (Burge et al., 2004) and finds a large spread between \( \alpha \) at lower and higher end of the spectrum. For conditioned parametric distributions, estimates of 20th and 80th percentiles of value of
time distribution differ by more than a factor 50. Based on this observations, de Palma et al. (2011) expect the preferences for schedule delay early and late to vary on the same scale. This is reflected in the proportional heterogeneity scenario with constant ratios of $\mu = \frac{\alpha}{\beta} = 2$ and $\eta = \frac{\gamma}{\beta} = 3.9$, as used by Arnott et al. (1990) and number of subsequent publications (e.g. van den Berg and Verhoef 2011a, 2013).

In the remainder of this work, expressions such as degree of heterogeneity, spread in values of time or increasing n-factor are used interchangeably.

Table 5.5: Value of time ($/hr) distributions for different heterogeneity factors.

<table>
<thead>
<tr>
<th>n factor</th>
<th>Mean</th>
<th>Median</th>
<th>Max.</th>
<th>Min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>15.5</td>
<td>15.5</td>
<td>19.0</td>
<td>12.4</td>
</tr>
<tr>
<td>1</td>
<td>15.5</td>
<td>15.6</td>
<td>23.2</td>
<td>9.9</td>
</tr>
<tr>
<td>2</td>
<td>15.5</td>
<td>15.1</td>
<td>34.3</td>
<td>6.2</td>
</tr>
<tr>
<td>3</td>
<td>15.5</td>
<td>14.5</td>
<td>49.8</td>
<td>3.9</td>
</tr>
<tr>
<td>5</td>
<td>15.5</td>
<td>12.9</td>
<td>100.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Adding $\mu$ heterogeneity

Data on the real distribution of the schedule delay early $\beta$ and $\gamma$ is scarce. The lack of studies on this topic is not at least due to the large number of factors on which the schedule delay cost depend. Next to personal socio-economic characteristics, particular trip purpose, timing and social interactions determine the degree of loss from being at desired location too early or too late. Focusing only on commuter trips, the ability to make use of the time before the compulsory working hours is highly dependent on the individual’s profession and type of job, allowing the assumption for relative stability of individual preferences from day to day. Given the lack of clear evidence, inferring direct correlation of schedule delay with value of time, particularly in case of commuting trips, is a reasonable assumption.

The bottleneck model imposes the restriction $\alpha > \beta > 0$ for the equilibrium to exist. Previous work as van den Berg and Verhoef (2011b) mainly rely on intuitive assumption and choose a triangular distribution for the ratio $\mu = \frac{\alpha}{\beta}$, ranging from 1.01 to 3.

However, as experimental data suggests such restrictions might not always be true. Based on SP data, Tseng and Verhoef (2008) estimate negative utilities for work activity starting 15 to 30 min before the preferred arrival time and resulting in schedule delay early being significantly larger than the value of time: $\beta > \alpha$. This implies that for the questionnaire respondents, time spent travelling in the car appeared to be more valuable than the time spent at work before the desired or required arrival time. This difference appears to be particular strong for workers with lower incomes (Tseng and Verhoef, 2008 Appendix A), indicating flexibility of working hours declining with income.
Agent-based simulation models are not bound to this restrictive assumption, allowing to account for all types of existent preferences. Cases with schedule delay early cost $\beta$ being very small or even going towards zero within a time window which extends beyond the predefined preferred arrival time, indicate high flexibility of working hours and a blurred definition of preferred arrival times. Lower $\beta$ indicate that even in case of being at an activity location before the preferred time, the person can still make use of the "waiting" - time and minimise its cost. Hence, from a perspective of an individual, schedule delay early cost $\beta$ and the preferred arrival time are the two sides of the same coin. Within the preferred arrival time window, an individual is completely insensitive to the arrival time and could arrive at any time without gaining or loosing. In case of time invariant travel cost, any arrival within the preferred time window leads to the same benefits for each individual. The agent-based approach allows to account for the variability of preferred arrival time not only by definition of opening and closing times, but also with a flexible definition of cost for early arrival.

As mentioned in section 4.3, the specification applies in the activity-based MATSim framework intrinsically accounts for the cost of schedule delay early as this time could have been spend earning utility at the preceding activity. Being dependent on logarithmic activity utility function, its marginal costs $\beta$ are not linear but rise logarithmically. However, under the assumption of the ratio between the time deviation $\Delta t$ and typical activity duration $t_{typ}$ being small, the linearisation of schedule delay cost, as presented in section 4.3, is valid.

In line with the objective of a high degree of realism in this study, $\alpha$ - heterogeneity is added on top of proportional heterogeneity, allowing for simultaneous variations of $\alpha$ and $\alpha/\beta$. The ratio $\frac{1}{\mu} = \frac{\beta}{\alpha}$ thereby follows a log-normal distribution with a constant mean $= 0.5$ and the standard deviation of 0.25, for all levels of heterogeneity. As the mean value of $\alpha$ equals 15.48 $$/hr for all degrees of heterogeneity, the mean of $\beta$ remains also constant at $\frac{\alpha}{\mu} = 7.74$$$/hr. The mean of the inverse ratio $\mu = \frac{\alpha}{\beta}$ which is commonly considered in the dynamic bottleneck model, however, is with 2.13 slightly higher than in case of proportional heterogeneity. To ensure comparability between scenarios, however, keeping the mean of absolute $\beta$ values in the activity-based context is considered more crucial.

The distribution of absolute $\beta$ values for different level of heterogeneity is shown in Figure 5.5(a), with red line indicating the constant mean of all distributions. It results from product of two log-normally distributed variables: $\alpha$ and $\beta/\alpha$ and is therefore itself log-normally distributed. Among 8'000 agents in the synthetic population, a few extreme $\beta$ values for $n = 3$ and $n = 5$ lie beyond the range of the x - axis shown in the graph. Table 5.7 summarizes the values of these distributions for each heterogeneity factor $n$.

Figure 5.6 visualizes a 3-dimensional probability density plot of the joint probability distribution of values of time $\alpha$ and the ratio $\mu = \alpha/\beta$ for the two heterogeneity factors $n = 1$ and $n = 3$. The shape of it results from two superimposed distributions. The value of time
Figure 5.5: Fitted density distributions of schedule delay cost for different types and degrees of heterogeneity.

Figure 5.6: Joint probability density distribution for VOT \( = \frac{\alpha}{\beta} \) cost and \( \mu = \frac{\gamma}{\beta} \).

distribution follows from the income distribution and is therefore log-normal. The ratio of \( \frac{\alpha}{\beta} \) is an inverse of log-normally distributed \( \frac{\beta}{\alpha} \). It important to note the difference in range of VOT values between the two scenarios.

Table 5.6: Schedule delay early values for different heterogeneity factors in ($/hr).

<table>
<thead>
<tr>
<th>n factor</th>
<th>Mean</th>
<th>Median</th>
<th>Max.</th>
<th>Min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>7.7</td>
<td>7.5</td>
<td>19.0</td>
<td>3.0</td>
</tr>
<tr>
<td>1</td>
<td>7.7</td>
<td>7.4</td>
<td>21.7</td>
<td>2.8</td>
</tr>
<tr>
<td>2</td>
<td>7.7</td>
<td>7.3</td>
<td>28.1</td>
<td>1.9</td>
</tr>
<tr>
<td>3</td>
<td>7.7</td>
<td>7.0</td>
<td>35.7</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>7.7</td>
<td>6.2</td>
<td>59.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Adding $\eta$ heterogeneity

The cost of lateness is considered to be significantly higher than that of earliness. Again, as in case of earliness it is reasonable to assume a certain correlation between values of time, schedule delay late and household incomes. However, cost of lateness can also be expected to vary with trip purpose and specific characteristics of an activity following a particular trip at a specific time and date. Though number of assumptions can be made based on a common reasoning, empirical evidence is scarce.

In this study, similar to the case of $\alpha$ heterogeneity, $\gamma$ heterogeneity is added on top of proportional heterogeneity. The values of $\gamma$ are now determined by the heterogeneous ratio $\eta = \gamma/\beta$, as the values of beta result from a constant ratio $\mu = \alpha/\beta = 2$. For the distribution of $\eta$ a normal distribution with mean 3.9 and standard deviation 0.975 is assumed. Figure 5.7 visualises the 3-dimensional probability density plot of the joint probability distribution of schedule delay early $\beta$ and the ratio $\eta = \gamma/\beta$ for the two heterogeneity factors $n = 1$ and $n = 3$.

The fitted distribution of absolute $\gamma$ values for different level of heterogeneity is shown in Figure 5.5 (b). The shape of this distribution results from a product of log-normally distributed $\beta = 0.5 \cdot \alpha$ and normally distributed $\eta = \gamma/\beta$, with red line indicating the constant mean of 30.1 $$/hr.

Table 5.7: Schedule delay late values for different heterogeneity factors in ($$/hr)

<table>
<thead>
<tr>
<th>n factor</th>
<th>Mean</th>
<th>Median</th>
<th>Max.</th>
<th>Min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>30.1</td>
<td>30.0</td>
<td>62.1</td>
<td>1.1</td>
</tr>
<tr>
<td>1</td>
<td>30.1</td>
<td>29.8</td>
<td>71.0</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>30.1</td>
<td>28.8</td>
<td>91.7</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>30.1</td>
<td>27.6</td>
<td>120.1</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>30.1</td>
<td>24.2</td>
<td>227.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>
5.2 Enriched Sioux Falls scenario

This section presents an enriched, agent-based small scale scenario with dynamic demand and an integrated public transport system based on the commonly used Sioux Falls test network. Initially presented by Chakirov and Fourie (2014), this scenario aims to provide a realistic test case, with a synthetic population of travellers with attached socio-economic characteristics and a high degree of spatial resolution. This section provides an overview of the scenario, drawing heavily on the working paper by Chakirov and Fourie (2014). In some parts it also adopts the exact wording as used by the authors; however, for the sake of readability, the quotation marks are omitted.

This work applies the enriched Sioux Falls scenario to verify observations based on the rather simple Corridor scenario and demonstrate the effects arising from additional degrees of complexity. In particular, availability of route choice and alternative activity chains distinguish the Sioux Falls scenario from the Corridor scenario presented above. Additionally, spatial distribution of activity locations as well as the implementation of a car ownership model also provide vast potential for further analysis and research.

To generate a diverse synthetic population with spatially distributed activity locations, real world survey and land-use data is used. The socio-demographic and economic characteristics include age and sex on individual and income on household levels. The assignment of home and work locations employs land-use and building information, census data from the City of Sioux Falls, South Dakota as well as commonly used static OD-matrices from LeBlanc et al. (1975). Following sections present transport supply and travel demand with its socio-economic and behavioural characteristics. For more in-depth discussion of the Sioux Falls scenario development process and its properties please refer to Chakirov and Fourie (2014).

5.2.1 Supply

A transport test network should ensure sufficient complexity of travellers’ choice dimensions while limiting computational effort. To this end, the Sioux Falls test network was initially introduced by Morlok et al. (1973) as a starting point. The structure of this network captures the major arterial roads of the City of Sioux Falls, but was never intended to replicate the real city or all characteristics of its transport system, such as travel times and mode share. In the enriched Sioux Falls scenario, the original network is extended with a bus network consisting of 5 bus lines, as previously also presented by Abdulaal and LeBlanc (1979).

In order to match the realistically derived demand, as presented in the next section, properties of the network links are adjusted according to physical properties of the City of Sioux Falls.
road network. Link parameters such as road length, number of lanes and associated maximum car flow rates are set according to the two types of pre-defined road links: highways and urban roads (Figure 5.8(a)), with the following characteristics:

- **Urban road**: 2 lanes with flow capacity 800–1,000 cars per lane per hour;
- **Highway**: 3 lanes with flow capacity 1,700–1,900 cars per lane per hour

The length of all links is set equal to the Euclidian distance between intersections. As the urban links represent only major roads of the city, 2 lanes per direction is a good approximation of real conditions. For a few highway sections, 3 lanes per direction might be a slightly exaggerated value, but for the most part it represents a reasonable approximation. The corresponding flow capacities were chosen according to values indicated in literature such as the Highway Capacity Manual (Transportation Research Board, 2010) or other related research publications (e.g., Ng and Small, 2012).

The public transport network as shown in Figure 5.8(b) is implemented as a bus network consisting out of 5 bus lanes, with the routes as proposed by Abdulaal and LeBlanc (1979). Bus stops are located 600 m apart, and are offset 5 m from the road link. Furthermore, drawing from the feature set of MATSim, each bus stop is assigned to have a bus bay (basically a separate bus-only link, which can hold infinite number of buses), so that no road lanes are
Figure 5.9: Sioux Falls scenario mean household income, aggregated by 100m x 100m cells (Chakirov and Fourie, 2014).

3.2 Home location assignment

The population synthesis, as presented in Section 3.1, provides the total number of households and persons at the census tract level. As the agent-based model allows for spatial resolution at the building level, we are required to assign a specific home location to each household. For this assignment, we used the dataset of buildings, located in close proximity to the Sioux Falls road network. This data set from May, 2013, contains information on all residential buildings, including number of units in each building. It was kindly provided by the City of Sioux Falls GIS division.

In the process, each household is randomly assigned to a residential unit within the tract it belongs to. In case the number of households exceeds the number of units in a specific tract, we allow for multiple occupants per unit after all units are full. This is particularly relevant for the peripheral tracts (e.g., 129, 132, 160), as these tracts extend beyond the rectangular study area around the road network and therefore contain buildings and units which are located outside this area. This results in a high density of home locations on peripheral parts of the study area, right outside the road network. However, as the majority of the road network in these areas represents highways, the increased demand from these regions can be considered as blocked during the boarding and alighting process at the bus stops. Buses on all 5 bus lines are set to operate with a headway of 5 minutes between 6am and 10pm.

5.2.2 Travel demand and behavioural parameters

A realistic socio-demographically heterogeneous demand population is crucial for unlocking the potential of an agent-based simulation. To this end, a synthetic population of households that matches the aggregate distribution of demographic attributes (age, sex and household income) recorded during the 2010 US Census for the 27 census tracts inside and adjoining the city centre of Sioux Falls is produced. The Entropy Maximization approach, applied in each of the census tracts, produces a total of 107'486 persons in 43'936 households. This number, however, also includes non-travelling population as e.g. infants. Figure 5.9 shows the spatial distribution of mean households incomes of the synthetic population. As already pointed out above, a detailed discussion on the generation process of synthetic population can be found in the full paper by Chakirov and Fourie (2014).

In order to keep the scenario transparent, only two trip chain types have been assigned to the travelling population: home – work – home and home – secondary – home. The secondary activities represent various types of shorter activities with more flexible scheduling preferences.
during the day, such as shopping, restaurant visits or other social activities. The survey data used to generate the synthetic population contains the mode of transport to work for persons who do not work at home. Assuming that the employment status of a person correlates with their income, and given that income represents a variable which was controlled for during the population synthesis, all repetitions of a person occurring in the synthetic population are assigned the same employment status as recorded for that person in the survey. This results in a total of 56'904 commuting workers. For secondary activity trip chains, the population without a work trip and older than 21 years is chosen, resulting in a total of 27'206 persons performing secondary activities.

As no information on the real number and distribution of work places within the relevant area was available during the scenario design, the OD-Matrix from [LeBlanc et al. (1975)] is taken as an indicator of the number of work locations in each zone. Subsequently, the assignment of work places to individual workers is performed using a parameter free radiation model presented by [Simini et al. (2012)]. For the assignment of 27'206 secondary activity locations, no capacity restrictions are enforced. All buildings that have commercial use or are marked as community facilities are considered as potential activity locations. The location assignment is again performed with a parameter free radiation model (Simini et al., 2012).

Additional degree of realism is added to the scenario by applying a car ownership model on the household level. The chosen model is an ordered probit model estimated by [Giuliano and Dargay (2006)] based on the US Nationwide Personal Transportation Survey (NPTS) 1995. All parameters of this model are available within the synthetic population and land-use information of the Sioux Falls scenario. Next to socio-demographic characteristics of a household (number of adults, children, pensioners, household income), the model uses attributes of residential location (population density, public transport access and dwelling type). Including public transport accessibility as a variable allows to account for characteristics of the scenario with its added area-wide bus network.

As described in section [5.1.2] activity constraints define the time intervals and the typical activity duration parameter within the activity utility function as it is used in MATSim. Altering these constraints changes the response of the logarithmic utility function along the utility axis. Furthermore, as also pointed out above, additional parameters, such as "latest start time" and "earliest end time", allow for translation of the of schedule delay cost as they are defined for the classical bottleneck-model to the activity-based framework.

For the Sioux Falls scenario applied in this work, a set of constraints, referred to by [Chakirov and Fourie (2014)] as "soft work constrains" is applied (Table [5.8]).

In section [5.1.2] behavioural parameters for the Corridor scenario were presented. As pointed out above, these parameters are borrowed from the same enriched Sioux Falls scenario by
Table 5.8: Sioux Falls scenario: activity constraints

<table>
<thead>
<tr>
<th>Activity</th>
<th>Typical duration</th>
<th>Opening time</th>
<th>Latest start time</th>
<th>Earliest end time</th>
<th>Closing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>13h</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Work</td>
<td>9h</td>
<td>8.00am</td>
<td>9.00am</td>
<td>5.00pm</td>
<td>8.00pm</td>
</tr>
<tr>
<td>Secondary</td>
<td>1h</td>
<td>8.00am</td>
<td>-</td>
<td>-</td>
<td>8.00pm</td>
</tr>
</tbody>
</table>

Table 5.9: Behavioural and monetary simulation parameters for the Sioux Falls scenario

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{act}$</td>
<td>+0.48 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,car}$</td>
<td>-0.48 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,pt}$</td>
<td>-0.66 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{tr,walk}$</td>
<td>-1.62 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{wait,pt}$</td>
<td>-1.458 [utlis/h]</td>
</tr>
<tr>
<td>$\beta_{cost}$</td>
<td>-0.062 [utils/$]</td>
</tr>
<tr>
<td>$\beta_{0,car}$</td>
<td>-0.562 [utils]</td>
</tr>
<tr>
<td>$\beta_{0,pt}$</td>
<td>-0.124 [utils]</td>
</tr>
<tr>
<td>$\beta_{0,walk}$</td>
<td>0.0 [utils]</td>
</tr>
</tbody>
</table>

Chakirov and Fourie (2014) as it is presented here. Initially estimated in the demand model for Sydney by Tirachini et al. (2014), the time-related parameters have to be adjusted for application in the activity-based context. This process was also discussed in section 5.1.2 and is unchanged for the Sioux Falls case. Table 5.9 presents these parameters for the Sioux Falls scenario. The only difference to the Corridor scenario parameters as shown in Table 5.3, is the adjusted cost of walking, which is slightly lower for the Sioux Falls scenario.

5.2.3 Socio-economic characteristics

The importance of socio-economic characteristics for individual’s travel behaviour has been extensively discussed in previous chapters. Section 5.1.3 introduced the income distribution and its impact on individual’s values of time for the Corridor scenario. Modelling of heterogeneous values of time in the enriched Sioux Falls scenario follows exactly the same approach. As the household income distribution for the Corridor scenario is borrowed from the synthetic population of the enriched Sioux Falls scenario, the distributions of values of time among 84’110 active travellers in the Sioux Falls scenario for different degrees of heterogeneity are basically identical with the distributions presented in section 5.1.3 and Figure 5.4. Furthermore, only proportional heterogeneity is studied in the Sioux Falls scenario.
Chapter 6

Congestion Pricing: Simulation Results and Discussion

"An advanced city is not one where even the poor use cars, but rather one where even the rich use public transport."

Enrique Peñalosa (Enrique Peñalosa, 2013)

As highlighted in the Introduction, the impact of any mobility pricing policy on the efficiency of a transport system and its users depends on a variety of physical system properties and behavioural user characteristics. In the process of policy design and evaluation, understanding the sensitivity of results to different parameters is essential for the implementation of a robust and optimal policy leading to the desired outcome. The complexity of an urban transport system, however, leads to a vast number of parameters, variations of which can have a significant impact on the overall performance of the system. Such parameters can include operational variables, individual preferences and behavioural characteristics as well as physical properties of the transport network and the built environment.

Investigating the impact of varying parameters or initial conditions on a transport policy with a simulation-based stochastic user equilibrium approach commonly requires the execution of a new simulation run for each set of relevant parameters. Often applied in the area of numerical weather prediction, such simulation sets, also referred to as ensembles runs, represent a common approach to the evaluation of the dynamic system sensitivity to initial conditions and model parameters. As this work focuses on better understanding the welfare effects of dynamic mobility pricing based on the availability of alternative transport modes and user heterogeneity in values of time as and schedule delay, ensemble runs with all possible combinations of parameters presented in Table 6.1 are conducted and discussed based on the Corridor scenario,
presented in Chapter 5. With heterogeneity degree $n = 0$ being equal to homogeneous users, this results in a total of 128 simulation runs (5 degrees of heterogeneity and 3 heterogeneity types, plus the homogeneous case result in 16 simulation runs. Investigating 4 bus service levels and 2 pricing regimes for each of these demand characteristics, leads to a total of 128 simulation runs).

Another parameter, commonly investigated in the context of the stochastic simulation approach is the initial random seed. In the case of the stochastic user equilibrium in MATSim, previous studies did not observe significant differences in quantitative results when using different random seeds (e.g. Chakirov and Fourie, 2014). Thus, given the general consistency of results obtained with the 128 simulation runs as presented below, the impact of different random seeds is not explicitly discussed.

In order to better understand the effect from the availability of alternatives, or in economic terms substitutes, on the impact from pricing policies, experiments with bus headways of 2 min, 5 min, 10 min and no bus operations at all are conducted. Given a capacity of 90 passengers per bus (34 seating, 56 standing), investigated service headways translate into the overall throughput of the bus line of 2700, 1080, 540 and 0 passengers per hour, respectively.

<table>
<thead>
<tr>
<th>Heterogeneity degree $n$</th>
<th>Bus headway</th>
<th>Heterogeneity type</th>
<th>Pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0)</td>
<td>2 min</td>
<td>homogeneous</td>
<td>no road pricing</td>
</tr>
<tr>
<td>0.5</td>
<td>5 min</td>
<td>proportional</td>
<td>congestion pricing</td>
</tr>
<tr>
<td>1</td>
<td>10 min</td>
<td>alpha</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>no service</td>
<td>gamma</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Parameters of simulation sets

Based on the agent- and activity-based simulation approach, the results presented here are derived from a fundamentally different methodology as applied in the majority of related publications discussed in Chapter 2. Hence, a direct comparisons of the results should be interpreted with care. However, such comparison can also be beneficial for bridging the gap between analytical transport economics approach, which allows to analyse cause and effect in a robust manner, but is often based on a highly simplified set of assumptions, and the agent-based simulation approach, which allows for detailed modelling of travel behaviour and physical interactions, but poses challenges for generalisation of simulation results.

This chapter is structured as follows: first, section 6.1 addresses the methodological question of system convergence and stability of the stochastic user equilibrium. Subsequently, Sections 6.2 and 6.3 evaluate the sensitivity of congestion pricing policies and resulting effects on economic welfare distribution with regards to varying parameters of transport supply and demand, as

---

1 Based on a low-floor single-deck city bus MAN NL323F used by e.g. SMRT in Singapore, EMT Madrid in Madrid, Spain and Dan Bus Company in Tel Aviv, Israel.
listed in Table 6.1. Starting with the unimodal case and extending it to the multi-modal scenario, the importance of integrated transport modelling and policy making is highlighted and discussed. Sections 6.4 - 6.7 provide a different angle, discussing the variations in monetary revenues, impact of scheduling preferences and trip distance effects.

Table 6.2 provides an overview of abbreviations used in figures and tables presented in the course of this discussion.

6.1 System convergence and stability

In section 3.5 the concept of linear annealing, as applied in this work was presented and discussed. Applying this methodology enables a smooth transitions of the simulation from the choice generation phase to the choice selection phase, where each agent chooses from a fixed set of daily plans in its memory. Constantly lowering the replanning rate prevents the system from converging to a state, which is based on the condition of a fixed percentage of agents randomly changing their travel behaviour during each iteration.

Thus, the overall system state and therewith system performance are dependent on the share of agents randomly altering their daily schedule. Given a constant replanning rate, the SUE outcome will depended on the selected replanning rate. Using RU or as referred in MATSim context, the average score, as the performance and convergence measure, the relation between RU and the replanning rate strongly depends on the number of the available choice dimensions. For example, with mode choice as the only choice dimension and mode share of public transport lower than that of car, forcing a certain fraction of agents to try an alternative mode would result in an inflated public transport mode share and a potential increase in RU due to the benefits from congestion reduction outweighing the losses of agents who are 'forced' to take public transport. However, in the case of the Corridor scenario, with departure time choice as another degree of freedom, forcing a part of the agent population to randomly vary their departure times results in the decreasing RU.

To achieve a stable system convergence with the linear annealing approach, the decrease in replanning rate (annealing rate), should be slower than the agent’s average adaptation rate. Or in other words, the change in overall system performance, as measured with RU, should be limited by the replanning rate and be insensitive to the annealing rate. Investigating optimal annealing rates and alternative annealing functions lies outside the scope of this work. However, it is worth to demonstrate systems convergence process, given the scenario set-up and the methodology as described above. Figure 6.1 presents a comparison of two multi-modal Corridor scenario simulation runs with 1000 and 2000 iterations, respectively. It shows, that the RU ( = score) remains insensitive to the changes in the annealing rate or the number of
Table 6.2: Overview of abbreviations and terminology.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Factor determining the level of heterogeneity</td>
</tr>
<tr>
<td>NCP</td>
<td>No Congestion Pricing</td>
</tr>
<tr>
<td>CP</td>
<td>Congestion Pricing</td>
</tr>
<tr>
<td>PT</td>
<td>Public Transport</td>
</tr>
<tr>
<td>$\alpha$ (alpha)</td>
<td>Heterogeneous ratios $\mu = \frac{\alpha}{\beta}$, also referred to as $\alpha$ - heterogeneity. In this scenarios the ratio heterogeneity is added on top of the proportional heterogeneity.</td>
</tr>
<tr>
<td>$\gamma$ (gamma)</td>
<td>Heterogeneous ratios $\eta = \frac{\gamma}{\beta}$, also referred to as $\gamma$ - heterogeneity. In this scenarios the ratio heterogeneity is added on top of the proportional heterogeneity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation measures</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CW</td>
<td>Consumer welfare (also known as consumer surplus or user benefits) calculated as the average EMU per person.</td>
</tr>
<tr>
<td>SW</td>
<td>Social Welfare</td>
</tr>
<tr>
<td>RCW</td>
<td>Realised Consumer Welfare, averaged over last 100 iterations.</td>
</tr>
<tr>
<td>RSW</td>
<td>Realised Social Welfare</td>
</tr>
<tr>
<td>ETTC</td>
<td>Excess Travel Time Cost</td>
</tr>
<tr>
<td>SDC</td>
<td>Schedule Delay Cost</td>
</tr>
<tr>
<td>TC</td>
<td>Travel Cost</td>
</tr>
<tr>
<td>TCSW</td>
<td>Travel Cost Social Welfare</td>
</tr>
<tr>
<td>SW</td>
<td>Sum of consumer welfare (CW), cost of bus operations and monetary revenues from congestion pricing and bus fares.</td>
</tr>
<tr>
<td>RCW</td>
<td>Realised consumer welfare, averaged over last 100 iterations.</td>
</tr>
<tr>
<td>RSW</td>
<td>Sum of Realised Consumer Welfare (RCW), cost of bus operations and monetary revenues from congestion pricing and bus fares, averaged over last 100 iterations.</td>
</tr>
<tr>
<td>ETTC</td>
<td>Cost of travel delay, which come on top the free flow travel cost. Also referred to as queueing cost.</td>
</tr>
<tr>
<td>SDC</td>
<td>Cost of schedule delay early and late.</td>
</tr>
<tr>
<td>TC</td>
<td>Total travel cost or travel disutility (incl. opportunity cost of time).</td>
</tr>
<tr>
<td>TCSW</td>
<td>Sum of total travel cost (TC), cost of bus operations and monetary revenues from congestion pricing and bus fares.</td>
</tr>
</tbody>
</table>
Figure 6.1: Convergence of average score (utility) with linear annealing of replanning rates in MATSim for runs with 1000 and 2000 iterations (proportional heterogeneity, 5 min bus service headway).

iterations. The replanning start and end rate for both simulation runs are the same: at the beginning total of 40% of all agents change their mode (20%) or departure time (20%) at each iteration. This rate gradually drops to 0, with no agent replanning the daily schedules saved in its memory after iteration 800 and 1800, respectively. Though the run with 2000 iterations covers a significantly wider search space, it converges to exactly the same score as the 1000 iterations run.

From Figure 6.1 it is also easy to see that once the choice set generation is switched off, the average score (= utility) stays stable. This indicated a stable equilibrium state. However, such stable state does not necessary imply stability on disaggregated level, as e.g. individual welfares or single link tolls.

6.2 Social and Consumer Welfare

Total social and consumer welfare represent fundamental aggregated indicators for policy and project evaluation. As discussed in detail in section 3.4, the calculation of consumer welfare is performed using the Expected Maximum Utility (EMU) approach, based on the chosen daily schedule in the last iterations as well as a choice set of 13 -14 non-chosen alternatives. Total social welfare is calculated as sum of consumer welfare, cost of public transport operations and monetary revenue from the toll and fare collection. It is important to highlight, that the utilities of chosen and non-chosen alternatives used in the calculation of the EMU depend on the system
state of the last iteration of the evaluated simulation run. The same applies to the monetary revenues from pricing and public transport fares. Though, the simulation methodology presented in section 3 ensures stable convergence of the stochastic user equilibrium, minor stochastic variations from iterations to iteration remain. Such variations are more probable to become relevant in case of small scale scenarios with high peak demand and limited degrees of freedom, as e.g. Corridor scenario with its lack of true route choice dimension. Hence, caution should be exercised during comparison of two independent simulation runs, where unfavourable superposition of stochastic variations can hide more subtle changes in social and consumer welfare indicators.

In order to better understand the degree and impact of stochastic variations, average of realised social and user benefits from a set of last 100 iteration runs, is taken as additional measure of evaluation for the unimodal scenario. Averaging over multiple iterations, during which travellers only choose from the existing daily plans in their memory, enables to smooth out effects of stochasticity, providing additional qualitative indicator in support of the EMU - based social and consumer welfare evaluation.

6.2.1 Unimodal scenario - car only

In a unimodal scenario car is the only mode available to commuters. In the absence of a bus service as a viable alternative, departure time choice represents the single degree of freedom along which an agent can optimise its behaviour. It is also important to note, that without congestion pricing and no bus operations in place, social and consumer welfare are identical.

No congestion pricing (NCP)

Figure 6.2 depicts the social and consumer welfare before and after introduction of congestion pricing dependent on degree of heterogeneity $n$, comparing the three types of value of time heterogeneity discussed above: proportional, $\alpha$ and $\gamma$. Solid lines represent scenarios without and dashed lines scenarios with congestion pricing policy. For the base case scenario without congestion pricing (NCP scenario), average social welfare stays almost unaffected by increasing spread and range of values of time. However, the social welfare is slightly higher in the presence of $\alpha$ and $\gamma$ heterogeneity as compared to the proportional heterogeneity case. The higher welfare in $\alpha$ and $\gamma$ heterogeneity scenarios is a result of a discontinuous jump as the ratio heterogeneity is introduced and added on top of the proportional heterogeneity at $n = 0.5$. However, with the increasing $n$ factor welfare graphs of the three different heterogeneity types remain parallel and scenarios with added ratio heterogeneity exhibit the same gradient as the proportional heterogeneity scenario. The cause of the discontinuity at $n = 0.5$ lies in the abrupt introduction of variations in ratios of schedule delay and value of time $\mu = \frac{\alpha}{\beta}$ and
Figure 6.2: Effect of congestion pricing on social welfare and consumer surplus for different degrees of heterogeneity $n$ and different types of schedule-delay heterogeneity, with only car mode available. Base case ($n = 0$) is the scenario with homogeneous travellers.

$\eta = \frac{\beta}{\gamma}$. Though absolute variation in values of schedule delay early and late increase gradually with $n$, a spread in ratios $\mu$ and $\eta$ is introduced abruptly for $n = 0.5$ and remains constant for individual agents over all degrees of heterogeneity ($1 \leq n \leq 5$). These ratios however turn out to be decisive for self-organising effects of $\alpha$ and $\gamma$ heterogeneity and therefore the degree of welfare changes. Variations of individual preferences for trade off between longer travel times and earlier or later arrivals facilitate socially more optimal distribution of departure times, leading to an increase in average social and consumer welfares.

Table 6.3 summarises changes in social and consumer welfare resulting from congestion pricing in the presence of proportional heterogeneity, as visualised in Figure 6.2. It is easy to note the subtle increase in the social welfare with higher $n$ values, resulting in the welfare at $n = 5$ being 0.5% higher as compared to the homogeneous case ($n = 0$). This minor increase is an artefact of generation and evaluation of alternative daily plans with the non-linear activity utility function, using the EMU approach for the consumer welfare calculation. Realised welfare averaged over multiple iterations ($RU$) summarised in Table 6.5 does not exhibit the same increase.

Three conclusion for the effects of heterogeneity in a unimodal NCP scenario can be drawn from these results:

1. In the NCP scenario, the degree of proportional heterogeneity does not affect social and consumer welfare. This independence of welfare from proportional heterogeneity is expected. Without time depended monetary travel expenses, the introduction of proportional heterogeneity does not alter individuals’ ratios of value of time and schedule delay early and late, and therefore provides no incentives for agents to change their travel behaviour compared to the homogeneous values of time scenario. It is also in line with the publications based on
the bottleneck-model, which were already mentioned above. \textsuperscript{van den Berg (2014), van den Berg and Verhoef (2013), van den Berg and Verhoef (2011a) and van den Berg and Verhoef (2011b)} agree, that introducing proportional heterogeneity does not change the departure and travel time patterns.

2. In the NCP scenario, adding $\alpha$ and $\gamma$ heterogeneity has a minor positive effect on social and consumer welfare. Benefits in both scenarios can be interpreted as result of varying rigour of constraints among agents, which are given by the opening time of the work places. Commuters with high $\mu = \alpha/\beta$ ratios encounter low losses from early arrival at work as compared to the time potentially lost in congestion. This encourages behavioural change and shift of departure time away from the peak-hour. Such reduction in travel cost with a mean-preserving increase in heterogeneity of $\mu$ was also shown by \textsuperscript{van den Berg and Verhoef (2011a)} and later confirmed in follow-up publications, as e.g. \textsuperscript{van den Berg and Verhoef (2013)}. Based on the convexity of the equilibrium travel time function, authors demonstrate social benefits from increased $\mu$ ratio of travellers with already high $\mu$, but no social losses from further decrease in $\mu$ ratios for travellers with already low values of $\mu$. Compared to $\alpha$ heterogeneity, $\gamma$ heterogeneity appears to be less studied, with \textsuperscript{van den Berg (2014)} being one of the few publication evaluating effects of heterogeneity in $\eta$ ratios. Also here authors point out the decrease in travel cost with the increasing range of $\gamma$ value among travellers in the NCP case.

3. Effects of proportional heterogeneity and $\alpha$ and $\gamma$ heterogeneities appear to be independent. This was as also observed by \textsuperscript{van den Berg and Verhoef (2011b)}. After adding $\alpha$ and $\gamma$ heterogeneity ($n = 0.5$), welfare curves for all three heterogeneity scenarios run parallel, as degree of proportional heterogeneity $n$ increases. In other words, the effect of adding $\alpha$ and $\gamma$ heterogeneity becomes visible only as the variability in $\mu$ and $\eta$ ratios is introduced and remains unaffected by absolute spread of $\alpha$ and $\gamma$ values.

Concluding, it can be remarked, that for the unimodal case, the agent-based approach to analysis of congestion pricing with heterogeneous user preferences for the most part replicates findings from studies which apply analytical approaches.

**Congestion pricing (CP)**

The introduction of congestion pricing has significant effect on social welfare, increasing it by around 4 % for all degrees of heterogeneity (Figure 6.2\textsuperscript{1}). Tables 6.3 and 6.4 provide a detailed overview of social and consumer welfare gains and losses from congestion pricing for all unimodal scenarios. Here, the $\Delta$ columns indicate the relative welfare changes compared to the NCP scenario.

Though differences between varying degrees of heterogeneity scenarios appear to be of minor
Table 6.3: Effects of congestion charge on social welfare (EMU) for proportional, $\alpha$ and $\gamma$ heterogeneity in the unimodal "car only" scenario.

<table>
<thead>
<tr>
<th>n</th>
<th>SW prop. (p.p)</th>
<th>SW $\alpha$ (p.p)</th>
<th>SW $\gamma$ (p.p)</th>
<th>$\Delta$ SW prop.</th>
<th>$\Delta$ SW $\alpha$</th>
<th>$\Delta$ SW $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>154.2 $</td>
<td>154.2 $</td>
<td>154.2 $</td>
<td>4.12 %</td>
<td>4.12 %</td>
<td>4.12 %</td>
</tr>
<tr>
<td>0.5</td>
<td>154.2 $</td>
<td>154.9 $</td>
<td>154.7 $</td>
<td>4.13 %</td>
<td>3.59 %</td>
<td>3.87 %</td>
</tr>
<tr>
<td>1</td>
<td>154.3 $</td>
<td>154.9 $</td>
<td>154.6 $</td>
<td>4.06 %</td>
<td>3.85 %</td>
<td>3.95 %</td>
</tr>
<tr>
<td>2</td>
<td>154.4 $</td>
<td>155.0 $</td>
<td>154.7 $</td>
<td>4.11 %</td>
<td>3.78 %</td>
<td>3.96 %</td>
</tr>
<tr>
<td>3</td>
<td>154.5 $</td>
<td>155.2 $</td>
<td>154.9 $</td>
<td>4.17 %</td>
<td>3.76 %</td>
<td>3.95 %</td>
</tr>
<tr>
<td>5</td>
<td>154.9 $</td>
<td>155.6 $</td>
<td>155.3 $</td>
<td>4.26 %</td>
<td>3.92 %</td>
<td>4.10 %</td>
</tr>
</tbody>
</table>

Table 6.4: Effects of congestion charge on consumer welfare (EMU) for proportional, $\alpha$ and $\gamma$ heterogeneity in the unimodal "car only" scenario.

<table>
<thead>
<tr>
<th>n</th>
<th>CW prop. (p.p)</th>
<th>CW $\alpha$ (p.p)</th>
<th>CW $\gamma$ (p.p)</th>
<th>$\Delta$ CW prop.</th>
<th>$\Delta$ CW $\alpha$</th>
<th>$\Delta$ CW $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>154.2 $</td>
<td>154.2 $</td>
<td>154.2 $</td>
<td>-3.13 %</td>
<td>-3.13 %</td>
<td>-3.13 %</td>
</tr>
<tr>
<td>0.5</td>
<td>154.2 $</td>
<td>154.9 $</td>
<td>154.7 $</td>
<td>-3.58 %</td>
<td>-2.70 %</td>
<td>-3.42 %</td>
</tr>
<tr>
<td>1</td>
<td>154.3 $</td>
<td>154.9 $</td>
<td>154.6 $</td>
<td>-3.70 %</td>
<td>-2.94 %</td>
<td>-2.58 %</td>
</tr>
<tr>
<td>2</td>
<td>154.4 $</td>
<td>155.0 $</td>
<td>154.7 $</td>
<td>-2.75 %</td>
<td>-2.36 %</td>
<td>-1.83 %</td>
</tr>
<tr>
<td>3</td>
<td>154.5 $</td>
<td>155.2 $</td>
<td>154.9 $</td>
<td>-1.60 %</td>
<td>-1.86 %</td>
<td>-1.73 %</td>
</tr>
<tr>
<td>5</td>
<td>154.9 $</td>
<td>155.6 $</td>
<td>155.3 $</td>
<td>-0.13 %</td>
<td>-0.85 %</td>
<td>-0.56 %</td>
</tr>
</tbody>
</table>

scale compared to total gains of congestion pricing, it is worth to discuss these effects at this point, as they become more pronounced after the introduction of mode choice as an additional choice dimension.

For proportional heterogeneity, a slight increase in welfare gain with increasing degree of heterogeneity can be presumed mainly based on scenarios with wider spread in values of time $n = 3$ and $n = 5$ (Table 6.3, $\Delta$ SW prop.). For lower degrees of proportional heterogeneity, stochastic variations in SUE outweigh this effect. Averaged over last 100 iterations, realised values of social welfare in Table 6.5 (avg. $\Delta$RSW) confirm this assumption. Hence, it can be concluded, that imposing congestion pricing on population of travellers with proportionally heterogeneous values of time, allows for more efficient self-organization of travel and departure times as compared to the scenario with a single, identical value of time for all users. However,
given the activity-based EMU calculation, these effect appears to be of minor scale with departure time choice being traveller’s single choice dimension. The scale of heterogeneity and pricing changes in general will be addressed below.

Adding $\alpha$ or $\gamma$ heterogeneities to the proportional heterogeneity leads to a minor reduction of gains in social welfare from congestion pricing policy. However, the increase in proportional heterogeneity still has the same effect as for homogeneous values of $\mu$ and $\eta$ ratios, leading to increasing gains from congestion pricing as the spread in values of time grows with $n$. But the gains are offset by a noticeable gain drop after the introduction of $\alpha$ and $\gamma$ heterogeneity at $n = 0.5$. Lowering the relative gain from congestion charge by 0.53 % and 0.25 % for $\alpha$ and $\gamma$ heterogeneities respectively, outweighs the gains from increase in proportional heterogeneity, even for the largest spread in values of time ($n = 5$). The lower gains in welfare are the logical consequence of efficiency gains through self-organization processes acting in absence of pricing policies in the presence of heterogeneous $\mu$ and $\eta$ ratios. With departure time choice as a single degree of freedom, social welfare gains from congestion pricing appear to be more sensitive to heterogeneity in $\mu$ and $\eta$ than proportional variations in value of time.

A direct, tangible impact of a mobility pricing policy on travellers before a potential return or redistribution of toll revenues represents on of the main factors affecting individuals’ self-interest. It is eventually the most crucial determinant for the public support for or the opposition to the implementation of congestion pricing. Hence, one of the most interesting results emerges from analysing the effects of congestion pricing under heterogeneous preferences on consumer welfare. For homogeneous travellers, congestion pricing has an adverse effect leading to the average loss of 3.13 %. Though the congestion pricing almost eliminates travel delays, monetary toll payments of a second-best congestion charging scheme exceed gains from eradication of excessive congestion, causing average loss for the consumer. However, the rapidly diminishing average consumer loss with increasing degree of heterogeneity is striking. For extreme case of proportional heterogeneity ($n = 5$), second-best congestion charge has almost no effect on average change in consumer welfare ($\Delta CW = -0.13\%$) even though no alternative routes or modes are available, substantially increasing the social welfare at the same time. This outcome is primary due to the strong distributional effects, with high value of time travellers disproportionally gaining from congestion eliminating pricing policy while schedule delay losses of the commuters with low value of time contribute only little to the overall welfare.

As pointed out above, adding $\alpha$ or $\gamma$ heterogeneities on top of the proportional heterogeneity slightly reduces social welfare gains of congestion charging. However, at the same time it also reduces negative effects of congestion pricing on average consumer welfare. Given certain levels of self-organization without congestion charging due to the varying $\mu$ and $\eta$ ratios, the behavioural intervention through pricing has a less significant effect. With dependency
Urban Mobility Pricing with Heterogeneous Users 2015

Table 6.5: Effects of congestion charge on realised social welfare and consumer welfare for proportional heterogeneity and "car only" scenario, averaged over last 100 iterations

<table>
<thead>
<tr>
<th>n</th>
<th>avg. RSW (p.p)</th>
<th>avg. RCS (p.p)</th>
<th>avg. ΔRSW</th>
<th>avg. ΔRCS</th>
<th>avg. Toll revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>126.8 $</td>
<td>126.8 $</td>
<td>+ 4.7 %</td>
<td>- 4.5 %</td>
<td>92,979 $</td>
</tr>
<tr>
<td>0.5</td>
<td>126.8 $</td>
<td>126.8 $</td>
<td>+ 4.8 %</td>
<td>- 4.1 %</td>
<td>90,419 $</td>
</tr>
<tr>
<td>1</td>
<td>126.8 $</td>
<td>126.8 $</td>
<td>+ 4.9 %</td>
<td>- 4.0 %</td>
<td>90,054 $</td>
</tr>
<tr>
<td>2</td>
<td>126.8 $</td>
<td>126.8 $</td>
<td>+ 5.3 %</td>
<td>- 3.9 %</td>
<td>82,919 $</td>
</tr>
<tr>
<td>3</td>
<td>126.9 $</td>
<td>126.9 $</td>
<td>+ 5.7 %</td>
<td>- 1.6 %</td>
<td>73,653 $</td>
</tr>
<tr>
<td>5</td>
<td>126.9 $</td>
<td>126.9 $</td>
<td>+ 6.5 %</td>
<td>+ 0.8 %</td>
<td>58,056 $</td>
</tr>
</tbody>
</table>

Table 6.6: Effects of congestion charge on consumer welfare social welfare and consumer welfare for \(\alpha\) heterogeneity and "car only" scenario, averaged over last 100 iterations

<table>
<thead>
<tr>
<th>n</th>
<th>avg. RSW (p.p)</th>
<th>avg. RCS (p.p)</th>
<th>avg. ΔRSW</th>
<th>avg. ΔRCS</th>
<th>avg. Toll revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>126.8 $</td>
<td>126.8 $</td>
<td>+ 4.7 %</td>
<td>- 4.5 %</td>
<td>92,979 $</td>
</tr>
<tr>
<td>0.5</td>
<td>127.6 $</td>
<td>127.6 $</td>
<td>+ 4.7 %</td>
<td>- 3.2 %</td>
<td>80,243 $</td>
</tr>
<tr>
<td>1</td>
<td>127.6 $</td>
<td>127.6 $</td>
<td>+ 4.8 %</td>
<td>- 3.2 %</td>
<td>81,065 $</td>
</tr>
<tr>
<td>2</td>
<td>127.6 $</td>
<td>127.6 $</td>
<td>+ 5.0 %</td>
<td>- 2.7 %</td>
<td>78,582 $</td>
</tr>
<tr>
<td>3</td>
<td>127.6 $</td>
<td>127.6 $</td>
<td>+ 5.4 %</td>
<td>- 1.4 %</td>
<td>68,553 $</td>
</tr>
<tr>
<td>5</td>
<td>127.6 $</td>
<td>127.6 $</td>
<td>+ 6.1 %</td>
<td>+ 0.5 %</td>
<td>57,270 $</td>
</tr>
</tbody>
</table>

of individual gains on two heterogeneous values: \(\alpha\) and \(\mu\) or \(\eta\), increase in one of the two heterogeneity dimensions has a less pronounced effect on the average welfare. Resulting from a less extreme distribution of gains and losses across the population, this aspect is returned to and discussed in detail in section 6.3.

As mentioned before, tables 6.5 and 6.6 provide an overview of heterogeneity and congestion pricing policy effects on realised social and consumer welfares, for scenarios with proportional and ratio (\(\alpha\)) heterogeneity respectively. Based only on the chosen alternative, which tends to be the best, the impact of pricing policy is amplified compared to the Logsum-based welfare evaluation. However, the average social and consumer welfare values are notably lower as for Logsum evaluation (Tables 6.3 and 6.4), resulting from neglecting any non-realised utilities of non-chosen alternatives. Yet, the qualitative evidence of an increasing social and consumer welfare gain from congestion pricing in presence of proportional heterogeneity remains true. It is also worth noting, that toll revenues are constantly decreasing as the degree of heterogeneity increases. This can be explained by the fact, that lower prices are needed to motivate users with lower values of time to shift their departures to pre- or post-peak times. Again, decreasing toll revenues with increasing heterogeneity were also noted by researchers in previous work, though to a lesser extent (van den Berg and Verhoef 2013).
Table 6.7: Effects of congestion charge on realised travel and schedule delay cost for proportional heterogeneity and "car only" scenario, averaged over last 100 iterations.

<table>
<thead>
<tr>
<th>n</th>
<th>ETTC</th>
<th>SDC</th>
<th>Δ ETTC</th>
<th>Δ SDC</th>
<th>Δ TCSW</th>
<th>Δ TCCW</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>57,512 $</td>
<td>37,561 $</td>
<td>-86.0 %</td>
<td>+4.5 %</td>
<td>-13.0 %</td>
<td>-12.3 %</td>
</tr>
<tr>
<td>0.5</td>
<td>57,924 $</td>
<td>37,729 $</td>
<td>-86.8 %</td>
<td>+3.1 %</td>
<td>-13.1 %</td>
<td>-11.6 %</td>
</tr>
<tr>
<td>1</td>
<td>57,662 $</td>
<td>37,756 $</td>
<td>-87.2 %</td>
<td>+0.15</td>
<td>-13.7 %</td>
<td>-10.9 %</td>
</tr>
<tr>
<td>2</td>
<td>57,560 $</td>
<td>37,756 $</td>
<td>-86.9 %</td>
<td>-9.5 %</td>
<td>-14.6 %</td>
<td>-8.0 %</td>
</tr>
<tr>
<td>3</td>
<td>57,275 $</td>
<td>37,752 $</td>
<td>-87.4 %</td>
<td>-19.2 %</td>
<td>-15.6 %</td>
<td>-4.5 %</td>
</tr>
<tr>
<td>5</td>
<td>57,076 $</td>
<td>38,029 $</td>
<td>-88.4 %</td>
<td>-38.2 %</td>
<td>-17.7 %</td>
<td>+1.9 %</td>
</tr>
</tbody>
</table>

Scale of heterogeneity and pricing policy impacts

As discussed above, in a unimodal scenario the overall impact of heterogeneity on social welfare changes from congestion pricing appears to be rather minor, while changes in consumer welfare are more significant. Yet, the rather small scale of welfare changes amounting to few percentage points and therewith a rather minor impact of pricing policy is eye-catching. Though highly dependent on the scenario set-up, double digit percentage point gains are commonly seen in bottleneck model based analysis [2, 13, 2011b]. However, the social and consumer welfare in traditional models capture only the time and monetary cost of travel. Welfare gains are expressed as a percentage reduction of the total travel cost. In contrast, in an activity-based context social and consumer welfare are assessed based on the sum of utility gains from activity performance and utility losses from travelling in the course of one day. Hence, changes in welfare from pricing policies are expressed as percentage of not only travel cost but of the total utility earned through the day. This difference becomes more visible, when analysing averages of actual travel cost changes, as presented in Table 6.7. With ETTC being the Excess Travel Time Cost as defined in section 3.4, SDC the sum of Schedule Delay Cost, TCSW the Total Cost based Social Welfare and TCCW the Total Cost based Consumer Welfare, significantly larger percentage changes as a result of congestion pricing policy can be observed. Modeled as an approximation of first-best pricing, the congestion pricing does not fully eliminate travel delays, but reduces them by 86-88%. For homogeneous users changes in departure time patterns lead to an average increase in SDC. This however changes with the increasing degree of heterogeneity, with SDC reduction reaching 38.2% in the most heterogeneous case. The welfare gains based only on travel cost range from 13.0% for homogeneous users and up to 17.7% for highly heterogeneous users. Changes of this magnitude appear more significant and are in line with publications based on the sclassical bottleneck model [2, 13, 2011b].

Parallels to results of the dynamic bottleneck model

As discussed in the review of the relevant literature earlier, conventional static models of
congestion tend to predict losses in the average consumer welfare and for the majority of users as a result of first- and second-best pricing policies, with and without presence of heterogeneity (e.g. Small and Yan 2001, de Palma and Lindsey 2002, Verhoef and Small 2004). Dynamic congestion models, generally using and extending the bottleneck model, present a more differentiated picture. First-best pricing of a single bottleneck does not affect consumer welfare (Small and Verhoef 2007). The toll paid by travellers is equal to the monetary value of time savings resulting from the elimination of congestion delays, leaving the generalised price of travel unchanged. However, for the second-best pricing, which is commonly studied in presence of an untolled alternative such as parallel road or lane without tolls, results diverge. The impact of a second-best pricing policy on the consumer welfare before returning toll revenues to the user can be either negative (van den Berg and Verhoef 2011a) or positive (van den Berg and Verhoef 2011b, 2013). This appears to be highly dependent on dimensions of heterogeneity in value of time and schedule delay taken into account. These finding are also confirmed by van den Berg (2014), who investigates performance of different models of tolling in presence of heterogeneous user preferences.

In essence, the impact of congestion pricing on social welfare identified in agent- and activity based context using the "car only" Corridor scenario is in line with prior publications. Increasing social benefits of congestion pricing with proportional heterogeneity as presented here, are also observed by van den Berg and Verhoef (2011b), van den Berg and Verhoef (2013) and van den Berg (2014). Similarly, these publications also discover decreasing social welfare gains of congestion pricing with increasing spread in $\alpha$ values, or rather $\mu$ ratios. However, due to the second-best nature of congestion pricing presented in this work, effects on consumer welfare are harder to compare. van den Berg and Verhoef (2011b), van den Berg and Verhoef (2013) and van den Berg (2014) observe decreasing general price of travel, which is inversely related to consumer welfare in this work, as a degree of proportional heterogeneity grows. Similar decrease in travel cost and increase in consumer welfare is observed in the MATSim based simulation, but this effect turns out to be more pronounced, eventually changing the average consumer losses to average gains for extremely heterogeneous users.

In contrast, for the $\alpha$ ($\mu$) heterogeneity, van den Berg and Verhoef (2011a), van den Berg and Verhoef (2011b), van den Berg and Verhoef (2013) and van den Berg (2014) find an opposite effect: variability of $\mu$ ratios results in average increase of generalized price under first-best pricing, leading to an average consumer loss. The impact of a second-best pricing appears to be highly dependent on dimensions of heterogeneity, tolling model, scenario set-up and the presence of perfect or imperfect substitutes. For the second-best congestion charge in an agent- and activity based context and without any other alternatives but departure time choice, the effect of heterogeneous $\mu$ ratios is less negative than in the bottleneck model. Given, certain level of self-organization already present in the NCP case, the required pricing intervention to eliminate congestion turns out to be less drastic, weakening its overall effect on the consumer welfare.
welfare. Furthermore, as mentioned above, considering two independently acting dimensions of heterogeneity reduces extreme individual changes, which again affect the average change of consumer welfare in a moderate way and can be particular visible for extreme degree of proportional heterogeneity ($n = 5$). Presence of additional variability in $\mu$ reduces extreme gains of high value of time individuals, who disproportionally contribute to the total welfare and raise its average value.

Almost everything, what was stated about $\alpha$ ($\mu$) heterogeneity, can also be said for the $\gamma$ heterogeneity. In one of the few publications addressing this type of heterogeneity, van den Berg and Verhoef (2013) find overall similar effects as in case of $\alpha$ heterogeneity, with slightly different distributional effects and strong dependency of change in general price on the coarse tolling model applied. In the agent- and activity-based simulation, variation in $\eta$ ratios have similar effects as of $\mu$ ratios, with results of both scenarios being almost indistinguishable in many cases.

6.2.2 Multi-modal scenario with varying levels of bus service

Studies of pricing and optimisation policies based on simplified, unimodal models can help to gain valuable fundamental insights into the system dynamics and understand the interdependencies between various, underlying forces. However, considering transport modes in isolations neglects a number of important factors acting on transport demand and behaviour in a context of dense, modern urban environments. Shifting the focus of transport planning practice towards individual well-being, opportunities and inclusion, requires comprehensive integrated system modelling methods and approaches. The presence of alternative transport modes can have a significant impact on welfare and distributional effects of pricing policies. Hence, moving to a multi-modal approach with high degree of realism is crucial for the provision of valuable and applicable insights to planners and practitioners. Making one step in this direction, this work investigates the impact of public transport service, or in particular bus service with physical interaction between cars and buses and modelling of crowding effects on welfare effects of congestion charging.

Figure 6.3 provides an overview of heterogeneity effects of congestion pricing on social and consumer welfare in a multi-modal context and for varying levels of bus service, which are discussed in following.

No congestion pricing (NCP)

Providing a bus service as an alternative mode of transport significantly amplifies implications of heterogeneous user preferences on social and consumer welfare. Next to departure time
choice, the availability of a bus mode as a viable alternative to car travel represents an additional dimension for self-organization of heterogeneous travellers. Thus, as bus provides an imperfect substitute for car and has different travel cost attached to it, the increasing heterogeneity in traveller’s preferences allows for a socially more efficient demand distribution across the two available transport modes. Due to such self-organization social and consumer welfare increase with a higher heterogeneity of user preferences, with as well as without congestion pricing. Most striking however, is the convex shape of the welfare function and the degree to which it increases as the spread in values of time grows. For the highest level of heterogeneity with factor $n = 5$ social welfare in the NCP scenario reaches similar or even higher levels as in case of homogeneous users. More efficient distribution of travellers between the available modes offers a significantly higher potential for welfare gains as only the optimisation of departure time. With heterogeneous preferences, lower value of time users can switch to bus without significant losses in welfare. This also enables travellers with high values of time to disproportionally profit from the easing of congestion.

**Congestion pricing (CP)**

Social welfare and consumer changes in presence of congestion pricing follow a similar pattern as in the NCP case and exhibit convex increase with growing degree of heterogeneity. However, the increase is less steep and thereby the gains from congestion pricing policy gradually diminish as highly heterogeneous travellers show strong self-organization behaviour in presence of alternative mode, even without congestion pricing.

**Bus service frequency**

The extent to which social and consumer welfare increase with growing degree of heterogeneity depends on the service quality of the bus service along the corridor. Increasing the bus headway and therefore providing additional capacity allows more users to switch to the bus mode without experiencing significant utility losses associated with overcrowding. This enables a more efficient self-organization without tight crowdedness constraints, strengthening its effect and increasing the overall welfare.

Another, probably the most critical effect of increased bus headways is the change in consumer welfare with introduction of the congestion pricing policy. Dependent on the level of bus service, average consumer welfare change can be either positive or negative, which is crucial for a potential public support of or opposition to the implementation of the policy.

The constant distribution and spread of $\mu$ and $\eta$, which is independent from the degree of proportional heterogeneity, explains the more pronounced differences between proportional and ratio heterogeneities in $\mu$ and $\eta$ for small degrees of heterogeneity and diminishing significance.
Figure 6.3: Effect of congestion pricing on social and consumer welfare for different degrees of proportional heterogeneity, defined by factor $n$ and additional ratio heterogeneities of $\mu$ and $\eta$.

of $\mu$ and $\eta$ variations for the large $n$ - factor values. The impact of $\alpha$ ($\mu$) and $\gamma$ ($\eta$) heterogeneity is thereby even less noticeable in the multi-modal scenario than it was in the unimodal case.

Figure 6.4 visualises changes in social and consumer welfare after introduction of the congestion pricing policy dependent on the bus service frequency and the degree of proportional heterogeneity. Highlighting the findings depicted above based on Figure 6.3, it offers a new perspective on the sensitivity of the pricing welfare gains to the two parameters. For scenarios with homogeneous users, maximal social welfare gains from congestion pricing policy are achieved not given the presence of longest congestion delays occurring in absence of any public transport service, but in case of moderate congestion and low frequency bus service with 10 min headway. Given user homogeneity and thereby absence of mode choice self-selection,
low frequency bus service can not attract enough users without road pricing in place. With growing self-selection as heterogeneity increases and bus headways become shorter, gains from congestion pricing decrease. For a large spread of values of time and schedule delay, the largest welfare gains are observed in the scenario with the biggest congestion externality and total absence of bus service. From the twisted shape of the graph at the transition from low bus frequencies to no bus service at all, the interplay of two opposing effects becomes apparent: the increasing gains from pricing under proportional heterogeneity in a unimodal scenario, combined with stronger effect of self-organization from spread in values of time, and hence decreasing welfare gains from pricing heterogeneous users under multi-modal conditions. Lowest pricing gains can be observed, as to be expected for the combination of the high user heterogeneity and the high bus service frequency.

The changes in consumer welfare before the redistribution of toll revenues are shown in Figure 6.4 (b). In this graph the non-linear dependency on bus frequency and heterogeneity is striking. In absence of any alternative to the car mode, the spread in values of time has profound effects on consumer welfare. Given everyone that has the same preferences and no alternative mode is available, consumers suffer a substantial loss from the pricing policy. However, with growing gap between higher and lower travel time valuations consumer losses diminish or even turn into consumer gains. The same occurs for increasing levels of bus service, where already lower frequency service substantially reduces aggregate losses from the congestion pricing policy. These effects create a plateau, where given availability of an alternative, even with low level of service, results in fading impacts of heterogeneity on average consumers welfare. However, it is important to note, that though for scenarios with high heterogeneity and no alternative mode of transport as well as for scenarios with rather homogeneous user preferences and frequent bus operations, congestion pricing seems to have no impact on average consumer welfare, the underlying distributions of disaggregated welfare changes are fundamentally different.
This crucial aspect of distributional effects on congestion pricing is addressed in detail in the following section.

**Relation to relevant, multi-modal studies**

The problem of congestion pricing for heterogeneous travellers in the presence of an alternative mode of transport adds an additional degree of model complexity and makes an analytical approach to the analysis of such scenarios even more challenging. Hence, only few publications tried to shed some light on it. van den Berg and Verhoef (2013) looked at congestion pricing on road and rail network, modelling two dimensions of heterogeneity. For the road efficiency and distributional effects of congestion pricing were in line with the previous findings for only a single mode problem van den Berg and Verhoef (2011a,b). In contrast, no or even opposite distributional effect of pricing were found for the train.

The dominance of variations in $\alpha$ among agents and its impact on social welfare and consumer benefits is a consequence of the activity-based approach, where consumer benefits is a sum of utilities from activity performance and (dis)utilities of travelling (see section 3.1). Following the concept introduced in section 4.3, the individual values of time affect not only the travel cost but also the utility of activity performance, which represents the largest contribution to the total benefit of a consumer.

As van den Berg and Verhoef (2013) observe, the effects of congestion pricing on social welfare in presence of alternatives and under heterogeneous preferences depend on the assumption of alternatives being perfect or imperfect substitutes. This strongly relates to the investigation of the impact of the availability and the service level of an alternative transport mode on congestion pricing effects presented here. As public and private transport have different cost functions, a bus represents an imperfect substitute for a car. However, increasing bus service level reduces the cost difference between car and bus travel, enabling users to change their behaviour without suffering substantial losses.

In presence of an alternative mode, relative gains from congestion pricing decrease with the increasing proportional heterogeneity. However, the absolute values of average social welfare under congestion pricing increase, following the same pricing effect on travellers with different values of time as described in section 6.2.1. The welfare increasing effect of proportional heterogeneity under pricing is offset by self-sorting effects in presence of three imperfect mode substitutes: car, bus and walk. Comparable to the self-organization of departure times without congestion pricing in the presence of $\mu$ and $\eta$ heterogeneities in the unimodal NCP scenario, self-organization of travellers with proportional heterogeneity takes place across different choice dimensions such mode choice in a multi-modal case. It becomes more pronounced as the level and the capacity of bus service increases. These results match observations by
van den Berg and Verhoef (2013), who find relative efficiency of welfare maximising road toll decreasing in presence of a rail alternative. They are also do not contradict findings by van den Berg and Verhoef (2011b), who observe the relative efficiency gains from proportional heterogeneity for parallel tolled and non-tolled alternatives. Here, the welfare gains from pricing under heterogeneity are still observed, but are dominated by the welfare increase without congestion pricing.

6.3 Distributional Effects

Pricing policy as a demand management tool aims to provide incentives for travellers to change their travel behaviour in such a way, that the welfare or well-being of society as a whole improves. However, given the complexity of the transport system, with variety of individual preferences and vast number of interdependent factors affecting users travel experience and travel cost, a strictly pareto-optimal policy outcome is hardly ever reachable. Yet, understanding the distributional effects of transport policies and identifying winner and loser groups enables decision-makers to take countermeasures towards a more equal distribution of policy benefits.

Just as important as the understanding of the distributional policy effects, is their external communication and public discussion. Even the best policy can fall through before its actual implementation if the planned measures are not supported by the public. During the decision making process in favour or against a specific demand regulating policy, redistribution effects often play a central role. The public debate around winners and losers can dominate the arguments based on only aggregated economic indicators and heavily influence the public opinion and the planning process. Even in case of an optimal policy design without considerable losses for any of the populations subgroups, communication and demonstration of policy impacts before actual policy implementation is crucial for formation of public opinion in favour of the proposed measures. In particular, in the context of pricing policies, the direct policy impact on welfare of population subgroups characterised by a set of socio-economic attributes or spatial distribution of their residence and activity locations, can be much more evident at the first glance than the long-term and often indirect policy gains or redistribution benefits.

For the Corridor scenario, in which traveller’s values of time and schedule delay are directly linked to the household income, the distribution of changes in consumer welfare can be analysed dependent on the actual income. This makes visualisation of redistribution effects of the congestion pricing policy under question more tangible. As discussed in section 5.1.3, the impact of varying values of the heterogeneity factor \( n \) allows for two interpretations: changes in actual income distribution or the varying impact of income on the individual’s
valuation of activity and travel time. Plotting the distribution of welfare changes caused by the introduction of congestion pricing as the function of the initial household income, lets the second interpretation to appear more intuitive. Furthermore, additional advantage of using income bins for the identification of winners and losers is the constant bin size and the identical agent composition in each bin across all scenarios.

Figure 6.5 visualises gains and losses of different income groups for all scenarios with varying degrees of heterogeneity and levels of bus service. In all scenarios high income groups with high values of time are better off as compared to lower income groups. As it has already been observed above, the sign and the extent of changes in consumer welfare are strongly dependent on the availability of an alternative travel mode and the degree of heterogeneity. Another characteristic of the distribution of welfare changes is the narrowing gap between winners and losers with a better availability of an alternative mode. Figure 6.6 presents the box plots of consumer welfare changes dependent on varying degree of heterogeneity for scenarios without bus service and with bus operating at 2 min intervals. The size of the box (2nd and 3rd quartiles) stays almost constant for all degrees of heterogeneity, but with spread being noticeably smaller in the scenario with the frequent bus service provision. Furthermore, given provision of a bus service, number of outliers disproportionally affected by the congestion charging policy and the scale of extreme gains and losses remains limited, even in case of a highly heterogeneous population.

The most surprising effect, however, appears to be the disappearing benefits and increasing losses of low-income groups in the multi-modal scenarios as the degree of heterogeneity increases. This effect, clearly visible in Figures 6.5(c) and 6.5(d) is solely due to the increasing bus ridership and associated crowding effects. With the increasing degree of heterogeneity, a growing share of population with low values of time leads to more bus users, as the monetary cost of bus trips is lower than the cost of driving. Hence, the bus mode share among low-income, low value of time population group for the highest degree of heterogeneity \( n = 5 \) reaches nearly 100%, even for the scenario without congestion pricing policy (Figure 6.7(a)). The introduction of congestion pricing results in a number of car drivers from medium income groups switching to bus (Figure 6.7(b)). With the further increase in the bus ridership crowding induced delays cause losses to the existing bus riders. The ability to adequately capture this effect originates from the detailed simulation of physical passenger - bus interaction, including dynamic dwell times and bus capacity constraints, as presented in section 3.2.

Figures 6.8(a) and 6.8(b) show boardings, alighting and occupancy of one bus vehicle during the morning peak hour for the whole length of the corridor and for the scenarios with (CP) and without (NCP) congestion pricing. The horizontal red line indicates the maximum vehicle capacity (90 passengers). While in the scenario without congestion pricing policy (NCP) the maximal bus occupancy for the particular service never reaches more than 80% of the total
Figure 6.5: Changes in consumer welfare dependent on income after introduction of congestion pricing for different levels of proportional heterogeneity.
Figure 6.6: Boxplot of consumer welfare changes as a result of congestion pricing for two different levels of bus service.

(a) proportional heterogeneity, no bus service  
(b) proportional heterogeneity, 2 min headway

Figure 6.7: Mode shares of car trips before and after congestion pricing (proportional heterogeneity, 2 min bus headway, last iteration).

(a) NCP  
(b) CP

vehicle capacity, in the scenario with congestion pricing (CP) the bus departs fully occupied from Stops 14, 15 and 17, therewith denying boarding to potential travellers waiting at these stops. This highlights the importance of the integrated multi-modal approach for the evaluation of distributional effects of pricing policies.

6.3.1 Second heterogeneity dimension - varying $\mu$ and $\eta$ ratios

As observed above in section 6.2, adding a second dimension of heterogeneity to already present proportional heterogeneity has only minor impact on average social and consumer welfare. Effects of heterogeneous ratios $\mu = \frac{\alpha}{\beta}$, also referred to as $\alpha$ or ratio heterogeneity, and $\eta = \frac{\gamma}{\beta}$, also referred to a $\gamma$ - heterogeneity, are mostly noticeable in the unimodal scenario
Figure 6.8: Boardings, alighting and occupancy for the east bound bus service starting at 07.56 a.m. (proportional heterogeneity, 2 min bus headway, n = 3).

![Graphs showing boardings, alighting and occupancy for NCP and CP services.](image)

and are negligible in a multi-modal context. However, on a disaggregated level varying $\mu$ and $\eta$ ratios can potentially impact the distribution of gains and losses from congestion pricing policies. Figure 6.9 shows a heat map of changes in consumer welfare, for $\alpha$ and $\gamma$ heterogeneity. It uses the scenario with the largest spread in absolute values of time and schedule delay ($n = 5$), as the most distinct welfare change differences are expected in that case. The shape of the heat map results from the existing combinations of value of time $\alpha$ and schedule delay $\beta$ and $\gamma$, as defined by the joint distribution presented in section 5.1.3. As seen in the case of the proportional heterogeneity, value of time plays a determining role for the level of individual gains and losses from the congestion pricing policy, while the impact of varying $\mu$ and $\eta$ ratios is rather minor. Figure 6.9 shows monetary changes for the highest degree of heterogeneity ($n = 5$), as here given a large spread in absolute values the most distinct differences in welfare changes are to be expected.

Previously, considering joint distribution of proportional and $\mu$ ratio heterogeneity, van den Berg and Verhoef (2011b) observe the greatest losses being incurred not by drivers with lowest values of time, but for drivers with intermediate values of schedule delay early $\beta$ and lowest possible $\alpha$ for the particular $\beta$. Though, the presented set-up does not exactly confirm this observation, negative effects of high values in schedule delay early $\beta$ for travellers with intermediate $\alpha$ are clearly visible (Figure 6.9(a)). Travellers with intermediate $\alpha$ and high values of $\beta$ suffer losses on the same scale as travellers with lowest VOT, while travellers with the same intermediate VOT but low $\beta$ experience welfare gains from congestion pricing. Interpretation of $\beta$ value as a measure for the rigidity of opening times constraints proved to be an intuitive explanation for this phenomena. In contrast, in case of $\gamma$ heterogeneity, no meaningful impact of $\eta$ ratio on consumer welfare gains and losses can be identified Figure 6.9(b). Though variations in $\gamma$ values can again be interpreted as rigidity of activity scheduling
Figure 6.9: Consumer welfare changes from congestion pricing in $, for the unimodal scenario.

Source: For this visualisation a Gridded Bivariate Interpolation for Irregular Data from *akima* R-Package is applied (Akima et al., 2015).

6.4 Fare and toll revenues

As already observed in the unimodal scenario, the revenues from congestion pricing decrease with the increasing degree of heterogeneity (Figure 6.10(a)). Lower toll revenues are the consequence of an increased disparity of values of time as higher heterogeneity implies more travellers with rather extreme valuations on both ends of the spectrum. Hence, already a comparatively low congestion charge is sufficient to trigger a behavioural shift in drivers with low time valuations and bring about free flow conditions on the road.

Figure 6.10(a) also illustrates the dependency of congestion charge revenues on the availability and the capacity of the bus service along the corridor. Here, the increasing bus service frequency has two major effects: decreasing congestion charge revenues and decreasing...
sensitivity of revenues to the degree of proportional heterogeneity. The first effect is expected as given a more frequent bus service, the bus becomes much more attractive and gains mode share without congestion pricing policy in place. Given less cars on the road, toll levels required to prevent congestion are lower. The second effect seems less obvious. As highlighted above, with a better availability of an alternative mode, self-organization effects in presence of heterogeneity policy become more significant, reducing potential gains of congestion pricing. At the same time however, available capacity of public transport enables a more efficient distribution of demand with only a minor market intervention. Incentivising travellers to switch to the bus mode, given an adequate service quality, speed and price, appears to be a lot easier than inducing schedule changes without availability of any alternative modes. Hence, the degree of proportional heterogeneity and its beneficial self-organising effects on departure times with pricing, plays a diminishing role in a multi-modal context.

The total revenue from bus fares is shown in Figure 6.10(b). Given a constant fare of 2$ per trip in the presented scenarios, total revenues dependent only on the bus ridership. Thus, increasing bus mode share with higher degree of heterogeneity, better bus service and introduction of congestion pricing, leads to an increase in the total revenue, as observed in the graph.

More interesting however, is the comparison of revenue levels in Figure 6.10 and bus operations cost in Table 5.1. It becomes apparent that given the fare level of 2$ per trip the revenues do not cover the operational cost of the bus service provision for any of the scenarios. This implies that public transport operations would require cross-subsidies. Such an outcome is realistic and is expected. Public transport subsidies are a widely spread policy measure in many cities all over the world. Supported by numerous research publication, public transport subsidies of 50% or more of operating cost are considered as welfare improving and are justified by economies of scale and reduction road transport externalities (Santos et al., 2010; Parry and Small, 2009). Optimal pricing of public transport will be subject of more detailed discussions in the following chapter. However, at this point, it is worth noting, that in case of congestion pricing scenarios, the toll revenues are sufficient to cover cost of public transport provision for 10 and 5 min headway scenarios, but are insufficient for financing 2 min bus operations in this particular set-up.

6.5 Scheduling preferences

The dynamic bottleneck model predicts an ordering effect of congestion pricing on departure times, according to traveller’s values of time. Under a first - best tolling, travellers with highest values of time exit the queue close to their desired arrival time and pay the highest tolls, while travellers with low value of time avoid tolls, but face schedule delay penalties for early or late arrivals (Small and Verhoef, 2007).
In an activity-based modelling framework with spatially distributed activity locations, departure rates from activities dependent on the time of day, individual’s value of time and spatial characteristics of a trip represent relevant study dimensions. For the Corridor scenario, Figure 6.11 visualises agent departure rates through the day, aggregated in 10 min time bins and compared to averaged values of time in each of the time bins. This is presented for two different degrees of proportional heterogeneity: \( n = 1 \) and \( n = 5 \). As only commuting trips are included into the Corridor scenario, distinct peaks in departure rates emerge during morning and evening rush hours. The degree of proportional heterogeneity appears to have no effect on the departure rates. For both heterogeneity factors (\( n = 1 \) and \( n = 5 \)), a noticeable feature of the departure rate distribution is the difference between the morning and the evening peak. Without congestion pricing, the maximal departure rate in the evening is nearly twice the maximal departure rate in the morning. The evening peak is observed around 7pm, which represents the closing time of work facilities, after which no utility from performing a work activity can be earned. In contrast to the classical bottleneck mode, introduction of congestion pricing does not lead to a constant departure rate. A major reason for this difference is the presence of the spatial dimension and the distribution of home and activity locations along the corridor (see Figure 5.1). For the morning commute, congestion pricing does not reduce the maximal departure rate, but shifts it towards a later time. For the evening commute, the effect of congestion pricing is significantly different, as it reduces the maximal departure rate by more than a half - exactly to the level of the morning peak (\( \approx 3900 cars/hr \)). Again due to the spatial distribution of activity locations this is significantly more than the maximal link flow capacity of 2400 vehicles per hour. Furthermore, the efficiency gains of congestion pricing appear to mostly result from the optimisation of departure order, as no significant spreading of departure times to earlier or later timings is observed.

The ordering effect of congestion pricing becomes evident, when looking at average values of time for travellers departing during each specific time bin. For the scenario without pricing, average value for all time bins is the same: \( \approx 15.5 \$/hr \). The outliers during very early or
late departure time bins are random and due to the fact that only a few individuals depart during these time bins. For congestion pricing scenario, an apparent departure order emerges: travellers with low values of time depart rather early or late, while travellers with high values of time depart closer to the optimal departure time window.
In this context, it is interesting to analyse congestion charges dependent on the traveller’s departure time. Figure 6.12 shows the average congestion charge paid by car drivers dependent on the trip start time. The charges are aggregated and averaged over 10 min time bins through the day. In general, the congestion charges exhibit a triangular shape as also predicted by the analytical models (e.g. see Small and Verhoef 2007). However, the level of heterogeneity as well as spatial distribution of home and work locations have a visible impact on congestion charging profiles of each commuting peak. Due to the self-organising effects described above, the level of charges in the scenario with highly heterogeneous users is significantly lower compared to the scenario with a lower degree of heterogeneity. This has also been observed based on total revenues from congestion charge in Figure 6.10(a). It is also interesting to note, that the average congestion charge profiles for the morning and evening commute differ in both scenarios. Given the spatial distribution of home and work locations and the set of behavioural parameters presented above, the morning and evening commutes are not symmetrical, as can be seen from the average departure rates as well as the average congestion charges (Figures 6.11 and 6.12). In practice, the average congestion charge profile, as emerged from the implementation of the first-best pricing approximation, can serve as a useful reference for design of second-best link-based congestion pricing schemes.

6.6 Travel times and travel speeds

Average travel times and travel speeds are often used as measures and indicators of transport network performance. Figure 6.13 displays the mean travel times and Figure 6.14 the mean travel speeds for car and bus mode, according to the varying degrees of heterogeneity and for scenarios with different bus headways. Solid lines indicate scenarios without congestion pricing, while scenarios with congestion pricing are represented by dashed lines. As expected travel times for cars decrease and speed increases for the NCP scenarios as the proportional heterogeneity increases. This is due to the increasing bus mode shares (see Figure 6.16) and associated reduction of congestion. This also leads to an increase in bus travel speeds, as the benefits of improved road speeds outweigh the longer dwelling times, which arise with the higher vehicle occupancy. However, in the scenario with congestion pricing, effects of the longer dwelling process can be observed based on the slight reduction of bus speeds as the number of bus passengers increases with higher degree of heterogeneity.

It is also important to note, that the dynamic bottleneck model predicts free flow congestion in the presence of congestion charging, which should result in the same car travel speed in all congestion pricing scenarios. However, for scenarios with increased bus operation frequency and thereby smaller car mode share, car speeds under congestion pricing are slightly higher. This indicates that even in presence of congestion charging, interaction between vehicles still
Figure 6.13: Mean travel times for different degrees of proportional heterogeneity $n$ and levels of bus service, before and after congestion pricing. Base case ($n = 0$) is the scenario with homogeneous travellers.

Figure 6.14: Mean travel speeds for different degrees of proportional heterogeneity $n$ and levels of bus service, before and after congestion pricing. Base case ($n = 0$) is the scenario with homogeneous travellers.

occurs. Such interactions represent a realistic condition as even with a second-best pricing scheme in place, traffic delays on the road still occur.

### 6.7 Effects of trip distance and travel mode

Distance, as an essential characteristic of each trip, has a strong impact on the mode and departure time choice as well as the total cost associated with a particular trip. The distance of a commuting trip is clearly determined by the home and work location and the route the commuter decides to take to get from one location to the other. Spatial distribution of activity locations and the individual route choice can be highly dependent on distinctive properties of
Figure 6.15: Departure time vs. avg. in vehicle distance, aggregated over 10 min bin for morning and evening rush hours (proportional heterogeneity, $n = 1$, no bus service).

the city or the area of interest, its land-use and network topology. These factors also determine the degree, to which travel distance correlates with the total trip cost and the trip duration.

With the Corridor scenario representing an arterial road of a mono-centric city with housing on the one side and work locations on the other and given the distribution of home and work locations along the corridor (see Figure 5.1), commuting distances between few hundred meters and up to 20km are represented in the population. Given this experimental set-up, almost all agents have to pass through the central part of the corridor as no alternative route is available to bypass this bottleneck.

Figure 6.16 compares the trip distance dependent on changes in consumer welfare after introduction of congestion charging for scenarios with homogeneous and highly heterogeneous user populations. The relevant mode shares before and after congestion pricing policy are indicated in the top part of each scenario graph.

As discussed above, in the unimodal scenario with homogeneous users, the majority of travellers experience welfare losses from the introduction of congestion pricing. However, these losses are minimal for the travellers with shortest trip distance. Consumer welfare losses increase gradually as the trip distance increases and remain constant for trips longer than 10km. As congestion pricing is operating along a larger section of the central part of the corridor and not only as a single-link cordon charge, consumer losses become larger with the increasing travel distance as travellers use larger sections of the priced road. This can be observed despite the fact, that commuters with longer trip distances tend to avoid the peak hours and travel
earlier during the morning and later during the evening commutes (Figure 6.15). However, including proportional heterogeneity into the model, eliminates the trip distance dependency of congestion pricing effects on consumer welfare of car drivers. A high degree of heterogeneity among travellers leads to individual values of time becoming a key factor for determining the individual’s departure time choice and therewith effect of congestion charges on individual’s welfare. Given the time choice as the only choice dimension available in the unimodal scenario, trip distance becomes irrelevant (Figure 6.16(b)).

In the multi-modal scenario, crowding on public transport emerges as a decisive factor influencing gains and losses of users with different travel distances. Also here, a detailed model of bus operations plays a significant role in the evaluation of effects from the pricing policy. In the heterogeneous case, self-organising effects lead to a significantly higher share of public transport users, with and without congestion pricing. Whereas bus users with longer trips are the main beneficiaries of congestion pricing for the homogeneous user case, an opposite situation can be observed for the scenario with heterogeneous users. This situation is particularly interesting, as with the elimination of severe congestion in presence of congestion charges, bus speed increases in all scenarios (see Figure 6.14). However, the losses of existing bus riders with heterogeneous values of time mainly result not from the longer dwelling times, as those are more than compensated by the reduced station to station travel times. Much rather the negative welfare changes of bus travellers are due to boarding denials at bus services operating at the maximum capacity. This can be considered as another externality, which given a constant bus fare is not accounted for by the markets. Hence, internalisation of the crowding externalities in form of dynamic bus fares represents a relevant policy instrument. Scenarios including dynamic public transport pricing based on marginal external costs are studied and discussed in the following chapter.

Interesting policy implications emerge from analysing welfare changes with regards to individual’s travel distance and value of time. In particular for scenarios with low degree of heterogeneity, scale of distance-dependent changes is similar to the scale of value of time dependent changes. This implies that depending on the public transport service quality and the degree of user heterogeneity, travel distance and value of time can be equally important for the individual gains or losses from a congestion charging policy. It is also important to note, that in case of travellers with low values of time and long travel distances by bus, the losses resulting from both factors are superimposed and can result in a double loss for a particular traveller. In contrast, high value of time users with a long car commute gain disproportionally.
Figure 6.16: Travel distance dependent consumer benefits of congestion pricing: homogeneous case vs. proportional heterogeneity ($n = 5$).
Chapter 7

Public Transport Pricing: Simulation Results and Discussion

"Nobody goes there anymore. It's too crowded."

Yogi Berra (Berra, 1998)

In an urban environment, the role public transport has two major elements: it extends network capacity and it increases accessibility. Collecting dozens or hundreds of travellers, going into a similar direction, in a single public transport vehicle yields to a much more efficient use of space dedicated to transport infrastructure. Thus, it provides an efficient way to significantly increase the overall capacity of a transport network. At the same time, it also allows promotion of equal access opportunities to places of economic and social activity by reducing individual cost of mobility and preventing exclusion of disadvantaged social groups. However, similar to road congestion, public transport operations can suffer from negative effects of temporally and spatially concentrated transport demand, which at peak times can push the existing transport network capacity to its limits. An overview of the external effects in public transport operations, such as travel delays and discomfort arising from crowding on public transport vehicles, has been given the section [2.1.3] of Chapter [2].

This chapter addresses effects of dynamic pricing of public transport usage based on the internalisation of external costs and given users heterogeneous values of time. It uses the same multi-modal Corridor scenario, as presented in Chapter [5] with socially optimal flat fare scenario as a benchmark. Addressing dependency of potential welfare gains on initial demand, behavioural parameters and interaction with congestion pricing policies, dynamic fare optimisation aims to maximise social welfare, ideally without making consumers worse off.

As highlighted in Chapter [2] public transport fares are traditionally considered from operational
financing point of view. However, profit oriented pricing of public transport infrastructure usually does not lead to a socially optimal outcome. Decreasing social cost in combination with economies of scale and external benefits of public transport, provide a strong case for subsidies. Focusing on eliminating inefficient distortions in match of supply and demand by considering external costs, public transport fares become less of a source for operation financing and more a tool for demand management.

This chapter is structured as follows: first, in order to gain a better understanding of the role of public transport pricing, public transport flat fares in the range between 0$ – 5$ are analysed based on their effects on social and consumer welfare. The approach to modelling of crowding disutility and its inclusion into the agent-based simulation framework are presented in sections 7.2 and 7.3. Experimental finding and a discussion on impact of dynamic pricing of public transport, based on internalisation of external cost, on social and consumer welfare as well as its distributional effect follow thereafter. The section 7.6 addresses the importance of initial status quo and the central role of balance and distribution of demand between available modes, as determined by populations behavioural characteristics and the total cost of each mode. The final section 7.7 discusses the transferability of results from the simplified Corridor scenario to more realistic urban networks. To this end effects dynamic congestion and public transport pricing policies are evaluated based on the enriched Sioux Falls scenario, as it was presented in section 5.2.

As a convenient visualisation tool, an interpolated heat-map is frequently used in this chapter. All heat-maps in the following sections are produced using Gridded Bivariate Interpolation for Irregular Data from \textit{akima} R-Package (Akima et al., 2015).

### 7.1 Optimal flat fare

Flat, trip-based public transport fare represents a simple scheme of charging for usage of public transport and as of 2015, is applied in many cities around the globe (e.g. Rome, San Francisco, Mexico City). Being easy to understand and offering cost transparency to the customer, flat fares offer only limited potential for demand management. With same monetary charge for any direction and time of travel, the level of fare can be only used to exercise control over overall modal split as well as associated level of revenues and share of self-financing of public transport operations.

In the previous Chapter 6 a flat public transport fare of 2$ was applied in all scenarios. As showed in section 6.4, this level of usage charges could not recover capital and operational cost of public transport (see Table 5.1 and Figure 6.10(b)). Before addressing the benefits of dynamic public transport pricing, it is worth to investigate a potential of public transport
flat fare. To this end, a set of ensemble simulation runs with varying public transport fares and degrees of user heterogeneity is conducted. A scenario with a bus headway of 5 min and proportional user heterogeneity is used as a case study.

Figure 7.1 shows the heat map of average social welfare per person, dependent of the fare level and the degree of proportional heterogeneity, for scenarios with and without congestion pricing (CP and NCP). For the NCP case, highest level of social welfare can be observed for the bus fare of 0$, for all degrees of heterogeneity. Furthermore, social welfare also increases with growing heterogeneity factor, which is consistent with observations discussed in Chapter 6. The benefits of free public transport mainly result from the optimisation of modal split as the total cost of bus travel decreases. Given a fixed operational bus frequency, the increasing fare leads to a mode shift from bus to car. As more commuters choose to drive, they contribute to congestion and thus to the externalities associated with it. The change in mode shares associated with the increasing bus fares is shown in Figure 7.2(a). This implies an optimal subsidy for the presented scenario of 100% of operational cost. In reality, however, making travel by public transport completely free may cause a number of side effects, which are not directly incorporated into the model. Such negative implications of free public transport may include significant number of induced trips, which otherwise would not have been conducted at all, low users valuation and perception of public transport as well as fairness concerns over large subsidies.

Adding the congestion pricing policy to the scenario results in a lower sensitivity of social welfare to the level of bus fares and the degree of user heterogeneity. This highlights the central role of the modal split. In the presence of a dynamic road congestion charge, road charges rise in parallel with the increase in bus fares. This counteracts the mode shift from public to private transport, ensuring a more optimal distribution of trips between the two modes.
However, the congestion charge alone cannot ensure an optimal modal split. Implemented as an approximation of first-best pricing, the congestion pricing policy does not fully eliminate congestion. As the bus fares increase, significant share travellers still shifts to from bus to alternative modes. Figure 7.2(b) visualises changes in modal split and the growing number car travellers as the bus fare increases. This results in longer car travel times and a less optimal overall state of the system.

Another reason for decrease in social welfare due to the modal shift from public to private to transport lies in the shortcoming of welfare calculation approach. To this end, it is important to recall the method used for the calculation of the social welfare. The welfare is based on individual choice sets. The choice-set for each individual is generated based on the user’s chosen departure times during the last iteration. Travellers who switch from bus to car due to higher bus fares, have lower values of time as compared to the existing car travellers. Hence, the new car users will rather avoid congestion charges and travel outside of the peak-hours, as discussed in section 6.5. Though from the individual consumer perspective the cost difference between bus travel and off-peak car travel might be marginal, the overall utility of individual’s choice-set, generated based on an initial off-peak departure time is expected to be lower than of the choice-set generated based on departure times, which are closer to the individual’s desired or optimal departure times. In certain cases, the rule-based generation of alternatives can lead to the creation of extremely unfavourable and maybe unrealistic daily schedules. Though the impact of this effect appears to be minor, this clearly represents one of the shortcomings of the rule-based choice-set generation.

As pointed out earlier, optimal distribution of transport demand between available modes is essential for maximising social welfare. Following the assumption of homo oeconomicus,
traveller’s mode choice is determined by the anticipated total cost of each available mode and is based on the individual’s trip characteristics. Among others, these costs consist of trip duration, corresponding mode and situation specific valuation of travel time as well as associated monetary expenditures. In simulation and scenario set-up presented in this study, the capability of MATSim for detailed modelling of public transport operations and interaction between passengers and vehicles on individual level is fully exploited. This leads to a relative high cost of public transport relative to the cost of car use, as compared to other similar studies (e.g. Kaddoura et al., 2014). Following aspects of scenario set-up and detailed public transport model contribute to the high cost of access and usage of public transport:

- full interaction between buses and cars, with buses being stuck in congestion,
- passenger friction due to occupancy dependent dwell time model,
- spatial distribution of activity location with distance dependent cost of access and egress,
- spatial distribution of bus stops (every 600m).

On the other hand, some drawbacks of the presented scenario such as lack of car ownership model with car ownership being equal to 100%, no direct consideration of car access and egress cost as well as assumptions about sufficient availability of parking space, also contribute to the bias towards the use of private transport.

The previously mentioned experiments on the agent-based optimization of flat bus fares and operational frequencies presented by Kaddoura et al. (2014), find optimal public transport fares slightly above zero and optimal operational frequency of around 5 min, based on the specific
scenario set-up and chosen parameters set. As highlighted above, detailed simulation of public transport operations as well as heterogeneous user population represent a major difference of this work to the study by Kaddoura et al. (2014) and explains the slight difference in results.

Another aspect, which contributes to the rise in social welfare with decreasing public transport fares, is the welfare evaluation methodology, as presented in section 3.4. Taking the cost of alternatives into the consumers logsum welfare calculations leads to positive effects from availability of non-chosen, low cost alternatives and therefore results in a higher welfare benefits as compared to the realised utility analysis.

As discussed in Chapter 2, maximisation of fare revenues does not necessary result in a maximisation of social welfare. This is also the case in the Corridor scenario discussed here. Figure 7.4 shows the total revenues from public transport pricing, with contour lines highlighting the same revenue level. In absence of congestion pricing on the roads (Figure 7.4(a)), fare leading to the revenue maximisation amounts to about 2.5 – 3.5$, dependent on the level of heterogeneity in travellers values of time. The revenue maximising bus fare increases as the heterogeneity and spread in values of time among user become larger. This is a consequence of the self-organizing effect, which as described in Chapter 6, originates from travellers with very low value of times shifting to the bus mode and therewith reducing road congestion. In a real world situation, where operators are allowed to pursue profit-maximisation given a certain level of service, higher public transport fares are to be expected in societies with higher degrees of value of time heterogeneity.

In scenarios with a high degree of user heterogeneity, commuters with low income and low value of time can not afford the relatively high cost of private transport and are left with choice between public transport or walking. Given the fact, that no trips can be dropped, a rather inelastic demand enables operators to charge higher fares. At the same time, this also leads the low-income part of the population, who are the main users of public transport, becoming the main losers of revenue maximising public transport fare schemes. Another interesting aspect to note, is that for scenarios without congestion pricing, there exists no fare which would enable public transport operator to recover its operational cost and therewith eliminate the need for subsidies. Even if the revenue maximising fares are enforced, only around 25 – 50% of operational cost can be recovered, dependent on the degree of heterogeneity. In order to be able to recover its operational cost, the operator would have to reduce operating bus frequencies and narrow time windows during which bus service is provided.

A different picture of revenue maximisation emerges in presence of congestion charges, as shown in Figure 7.4(b). As the rise in public transport fares is followed by an increase in congestion charges (Figure 7.5), non-motorized modes of transport become the only competitor to the bus mode. This leads to a continuous increase in operator revenue. The revenue grows up to a point where non-motorized modes of transport become cheaper for a large part of
Figure 7.4: Operator revenues dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

(a) Fare revenue without congestion pricing

(b) Fare revenue with congestion pricing

Figure 7.5: Congestion pricing revenue dependent on bus fare level and degree of proportional heterogeneity (5 min bus headway).

(commuters, who do not have an option of cancelling their trips. In case of the presented Corridor scenario with walking mode being the only alternative to private and public transport, the optimal revenue maximising bus fare from operators point of view lies significantly above the maximal simulated fare of 5$. However, already given the fare of 5$, with congestion pricing in place, operators can almost or even fully recover its operational cost for all degrees of heterogeneity.)
7.2 In-vehicle valuation of time under crowding

An important factor in personal valuation of travel time in public transport, are the conditions and the level of comfort, which an individual user experiences during her trip. Following an explosive growth of dense urban areas, crowding became one of the central factors influencing perception of comfort while travelling by bus or train. The necessity of accounting for crowding (dis)utility or, as it also referred to, value of crowding (VOC), when modelling and optimising public transport operations was discussed in detail in Chapter 2, Section 2.1.3. This section takes up on the previous discussion, presenting implementation of crowding disutility in an agent-based context and discussing its implications.

7.2.1 Crowding in an agent-based context

Agent-based modelling paradigm offers a highly suitable framework for an accurate inclusion of crowding disutility into a transport demand model. Together with a detailed simulation of public transport operations, it enables accurate calculation of crowding cost for each individual during the journey as well as precise allocation and redistribution of these cost among causing agents. In this work, a model investigated and estimated by Tirachini et al. (2013) based on rail and bus service data from Sydney, is implemented and applied within the agent-based MATSim framework. In their study of crowding externalities on public transport systems, Tirachini et al. (2013) estimate Multinomial Logit and Error Component models for 5 different utility specification, based on vehicle occupancy levels or vehicle load factors. The load factor is defined as the ratio between the actual number of passengers inside a vehicles and the total number of seats (Whelan and Crockett, 2009; Tirachini et al., 2013). As all of the models presented perform reasonably well, this work applies a load factors based specification, also mainly used by Tirachini et al. (2013) for the discussion. Alternatively, an occupancy based approach, which accounts separately for the proportion of passengers seated and the density of standees inside a vehicle, can provide additional level of behavioural resolution, but requires explicit tracking of standing and seating passenger, assumptions about seat choice behaviour and appears to be more dependent on the vehicle layout and configuration.

The utility specification with a crowding disutility term as applied by Tirachini et al. (2013) and implemented in this work, is presented in equation 7.2. The crowding disutility term is added to the utility function and contributes to the overall utility only for the load factors equal or above 60%. The choice of this threshold is based on the analysis of crowding studies on trains by Wardman and Whelan (2011). Tirachini et al. (2013) also note, that the model appears to be not very sensitive to the minor variation of this factor, as e.g. a lower threshold of 50% does not have a significant impact on the goodness-of-fit of the model. However, the authors also evaluate a threshold of 90% for their study.
Equation 7.1 presents the definition of the crowding utility term $V_{\text{crowd}}$, given a load factor $lf$ and a 60% threshold. In-vehicle travel time is represented by $t_v$ and $\beta_{lf60}$ indicates the sensitivity of the utility to the in-vehicle load factor.

$$V_{\text{crowd}} = \beta_{lf60} \max(lf - 0.6, 0) t_v$$

Including the crowding (dis)utility into the travel utility function, as defined by Tirachini et al. (2013), results in equation 7.2.

$$V_{\text{bus}} = \alpha_{\text{bus}} + \beta_a t_a + \beta_h t_{h\text{bus}} + \beta_v t_v + \beta_e t_e + \beta_c + \beta_{lf60} \max(lf - 0.6, 0) t_v$$

Table 7.1 presents parameters as used in equation 7.2 and indicated corresponding values as estimated by Tirachini et al. (2013) and next to it, corresponding parameters as defined for the Corridor scenario, applied in this work. In order to keep the scenarios with and without crowding parameters comparable, the original values of the Corridor scenario are retained and the crowding disutility term is added to the overall utility function. The value estimates by Tirachini et al. (2013) show, that for the load factor of above 60%, the disutility of crowding increases linearly with the load factor by $\beta_{lf} \approx 0.55 \beta_v$, a parameter which links load factor to the (dis)utility of travel time.

Transferring these findings to a set of different parameters, requires an assumption about the estimation methodology of the model without crowding disutility. More specifically, the question is, if the average disutility of crowding is already implicitly included in the mode specific $\beta$ for travel time, or whether the SP or RP data does not sufficiently capture the travel experience in a crowded conditions in the first place.

This work adopts the second assumption, including the disutility of crowding in addition to the existing disutility of travel time. The ratio between value of time and marginal utility $\beta_{lf}$, which scales linearly with the load factor over 60%, is assumed to be the same as estimated by Tirachini et al. (2013). For the activity-based utility model this transfers to $0.55 \cdot \text{opportunityCostOfTime} = 0.55 \cdot (\beta_{pt} - \beta_{act}) = -0.62$ [utils/h].

For the bus type Mercedes-Benz OC500LE with 34 seats and total capacity of 90 passengers, as used in the Corridor scenario (see section 5.1), the maximum value for the load factor is 2.65. This leads to the range of marginal disutility of crowding from 0 to 1.643. Expressed in values of time, for homogeneous users, this translates to the overall VOT range for public transport travel between 18.39 $/h$ and 44.86 $/h$, dependent on the level of crowding.
Table 7.1: Parameters for the scenario with crowding disutility.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value as estimated by Tirachini et al. (2013)</th>
<th>Value as applied in Corridor scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access time $\beta_a$</td>
<td>-0.9 [utils/h]</td>
<td>-1.401 [utils/h]</td>
</tr>
<tr>
<td>Headway $\beta_h$</td>
<td>-0.6 [utils/h]</td>
<td>-</td>
</tr>
<tr>
<td>Travel time PT $\beta_v$</td>
<td>-0.66 [utils/h]</td>
<td>-0.66 [utils/h]</td>
</tr>
<tr>
<td>Egress time $\beta_e$</td>
<td>-2.7 [utils/h]</td>
<td>-1.401 [utils/h]</td>
</tr>
<tr>
<td>Fare $\beta_c$</td>
<td>-0.148 [utils/$]</td>
<td>-0.062 [utils/$]</td>
</tr>
<tr>
<td>MSC bus $\alpha_b$</td>
<td>-2.681 [utils]</td>
<td>-0.124 [utils]</td>
</tr>
<tr>
<td>Travel time load factor bus $\beta_{lfb}$</td>
<td>-0.36 [utils/h]</td>
<td>-0.62 [utils/h]</td>
</tr>
<tr>
<td>Activity performance $\beta_{act}$</td>
<td>-</td>
<td>+0.48 [utils/h]</td>
</tr>
<tr>
<td>Waiting for PT $\beta_{wait,pp}$</td>
<td>-</td>
<td>-1.458 [utils/h]</td>
</tr>
</tbody>
</table>

7.3 Dynamic pricing of public transport externalities

Dynamic and differentiated pricing of public transport is very rarely applied in practice. As already pointed out in the section 2.1.3, the major focus in setting public transport fares lies in the financing of operations and degree of operational subsidies required for the desired level of service provision.

A large number of public transport pricing schemes are directly or indirectly based on the distance travelled between origin and destination. Hence, it is common to subdivide an urban area into zones, which often are shaped as concentric circles around the city centre. More recent fare schemes use smart card fare collection systems and charge ‘true’ distance-based fares derived from actual distance measures and independent of zones. However, even for smart card based fare collection schemes, most cities do not apply any further price differentiation according to temporal and spatial changes in demand in the course of the day. Such rigid pricing schemes exhibit only very limited potential for demand regulation and are often not very efficient or fair. Yet, alternative dynamic, time- or direction based fare differentiation and fare-based demand regulation are often considered as politically unacceptable.

The emergence of digital technology and smart card based fare collection systems enable exploration of new approaches to the design and optimisation of public transport fares. Singapore is one of the cities, which adopted a distance-based public transport pricing scheme. Simultaneously, Singapore is also experimenting with a basic time and direction dependent public transport pricing, by offering free train rides into the city centre during the early morning.
hours (Land Transport Authority 2015a). However, no attempts have been made to extend the use of smart card based fare collection systems to the internalisation of public transport externalities and apply it as a demand management tool on a larger scale.

As discussed in section 2.1.3, this work investigate optimal marginal public transport pricing based on two major externalities: (i) travel time delays arising from dwelling process and capacity constraints and (ii) higher values of time associated with discomfort and disadvantages of crowded conditions from an individual point of view. Next to effects of optimal pricing on social and customer welfare, this work focuses on the interaction between optimal pricing of private and public transport. Often considered separately, a joint consideration of the two policies aims to provide a better understanding of dynamics relevant for design of integrated urban mobility pricing schemes. In the following sections, effects from internalisation of public and private transport externalities are analysed based on the Corridor scenario, as it was presented in Chapter 5 and studied in Chapter 6.

7.3.1 Agent-based cost internalisation approach

An agent-based simulation approach with its detailed representation of public transport interactions enables the implementation of marginal pricing of public transport based on the internalisation of externalities on a level of an individual traveller. As presented in section 3.2, the MATSim framework includes a modular implementation of public transport simulation at a high level of resolution. Direct integration of individual public transport vehicles and their physical characteristics into the queue-based mobility simulation, combined with an extended and highly detailed model of passenger-vehicle interaction (Sun et al. 2014) at each stop or station allow detailed representation of single vehicle occupancies and travel time delays resulting from boarding and alighting times as well as vehicle capacity constraints.

Crowding disutility

The total number of seats of any public transport vehicle in MATSim is given by the definition of a vehicle type and its configuration as part of the simulation input. Furthermore, total vehicle occupancy as well as individual agents inside a vehicle can be tracked at any given time step of the simulation. This enables accurate computation of the load factor and thereby calculation of the (dis)utility of crowding for each individual, according to the equation 7.1. For the implementation of the load factor based crowding disutility model as specified above, the following assumptions on calculation of individual’s travel disutility are made:

- each passenger in a vehicle experiences the same level and disutility of crowding
• each passenger contributes equally to the total crowding externality
• total crowding externality is divided among all passengers
• passengers boarding or alighting do not experience crowding disutility during the dwell time process
• crowding cost of passengers who remain inside the vehicle during the dwell time process is divided among these passengers

Per definition externalities of crowding in a single vehicle are equal to the total disutility experienced by all the passengers inside the vehicle. Given the assumptions above, in the event of crowded condition with load factor over 60%, the fare each passengers has to pay is based on the external cost she is responsible for and equals the crowding disutility she experienced during her trip.

Given the assumptions above, MATSim’s modular architecture allows addition of data structures for tracing of each agents experienced crowding disutility during the simulation and its addition to the individuals utility (scoring) function.

**Dwell time delay and capacity restriction**

As discussed in section 2.1.3 next to changes in individual’s valuation of time due to crowded travel condition, longer travel times resulting from extended boarding and alighting process as well as longer waiting times as a consequence of vehicle capacity constraints represent another significant externality. Internalisation of external travel and waiting time delays was previously addressed and analysed by Kaddoura et al. (2015). In their study, Kaddoura et al. (2015) consider three types of externalities:

• delay caused by boarding and alighting process
• extended waiting time if a passenger cannot board a service due to public transport vehicle being full
• extended waiting time for passengers due to dwell times at previous stops

As Kaddoura et al. (2015) demonstrate, internalizing all these effects has positive effects on the average realised utility, used as a welfare indicator. In this work, only the first two types of **

---

1During the implementation process parts of Java code previously implemented by Guillaume Rérat and Sergio Arturo Ordóñoz Medina were reused.
2Implementation applied in this work is partially based on the Java-Classes developed by Ihab Kaddoura and published on GitHub and SourceForge within the MATSim project.
public transport externalities are taken into account: travel time delays from the dwell process and the waiting time delays due to vehicle capacity constraints. The third type of externality, extended waiting time for passengers due to dwell times at previous stops, strongly depends on operation type and the arriving pattern of passengers at the stop or station. As Turvey and Mohring (1975) note, if no schedule is in place and passengers arrive randomly at the stop or station, these delays will on average cancel out between passengers who had to wait longer and shorter for their service to arrive. In this work buses leave the first stop of the line in the predefined intervals (2 min, 5 min or 10 min) and thereafter serve the route without specifically following any schedule. The public transport route planning and therefore arrival timing of passengers is based on the stop or station arrival times, as it was observed in the previous iteration (Ordóñez Medina and Erath, 2013). Given the absence of fixed stop departure time, individual’s waiting time depends on a variety of factors, which either delay the expected service or lead to its early departure, causing the individual to wait for the following service. Quantifying it as an externality is ambiguous, as these factors include traffic conditions on the one side, but also number of passengers boarding and alighting during previous stops as well as the random component of the dwell time calculations.

As mentioned above, MATSim allows accurate tracking of vehicle arrival and departure time at stops and stations as well as individual passengers, who board, alight or are travelling inside the vehicle. Hence, externalities of each stop can be accurately calculated and distributed among causing agents. Here, the following assumptions for in-vehicle passenger delay externality are made:

- each boarding and alighting passenger contributes equally to the delay externality
- each in-vehicle passenger is affected equally by the delay
- externality is calculated based on an average value of time
- for buses, delay from vehicle deceleration before and acceleration after the stop is assumed to be 8 seconds and time to open and close the doors 1 second each. The resulting additional delay of 10 seconds per stop is evenly distributed among all boarding and alighting passengers.

In case of the capacity constraint, the time cost of a passenger who has to wait for the next service to arrive, is distributed equally between all passengers inside the vehicle, which departed from a stop or station with full occupancy, leaving the affected traveller behind. Similar to the in-vehicle time delay externality, the monetary cost are calculated based on an average value of time.
7.3.2 Challenges in the welfare evaluation

One of the major challenges in generating a choice set for each individual is the time needed to evaluate the cost of available route alternatives. The solution to this problem chosen in this work, is based on the use of a pseudo-simulation approach and was presented and discussed in detail in section 3.4. One major difficulty, however, arises from the underlying idea of the pseudo-simulation and the internalisation of public transport externalities. As developed by Fourie et al. (2013), pseudo-simulation approach relies on the use of travel times from the last iteration and saves computation time by avoiding a detailed calculation of queue-based traffic dynamics and simulation of public transport operations. Without a detailed model of public transport dynamics, however, external cost of an alternative choice can not be evaluated. This requires implementation of a solution for externalities approximation, similar to the one used by pseudo-simulation for travel time estimation.

This lead to an implementation of a public transport externality estimator, which is based on externalities observed during the last iteration of the full simulation run. The observed external public transport costs are aggregated and averaged over 5 min time bins, based on passengers boarding times. Similar to the travel time lookup, external congestion cost and therewith the fare of a public transport alternative evaluated with a pseudo-simulation are determined by using data structures created from the last full simulation run.

7.4 Social and consumer welfare

The impact of adding crowding disutility and dynamic public transport pricing to the Corridor scenario are evaluated based on social and consumer welfare. The monetary revenues as well as the distribution of actually paid public transport fares also provide important insights into effects of internalisation of crowding externalities.

7.4.1 Impact of crowding disutility

Adding a crowding disutility term to the utility specification, inevitably raises the cost of travel, making public transport less attractive and lowering social and customers welfare. This effect, as shown in Figure 7.6 for the scenario with free public transport usage, is to be expected. It is interesting to note, that in the scenario with congestion pricing (CP), the relative as well as absolute loss from consideration of crowding disutility is substantially lower as in the scenario without pricing. Initially, in view of larger public transport mode shares in the CP scenario as compared to the NCP case, this result appears to be all the more surprising. Only part of
the welfare losses, however, results from the increased cost of bus travel in the presence of the crowding disutility. Another significant contribution to lower welfare comes from the increase in car trips and therefore rise of congestion and externalities associated with it. As congestion charge internalises the external congestion cost and significantly reduces congestion, only crowding related losses of bus travellers contribute to the lower welfare in the CP scenario.

For the scenario with congestion pricing and crowding disutility, the consumer welfare remains unaffected by the introduction of congestion charging policy. Without crowding disutility, however, and given a high degree of heterogeneity, congestion pricing policy leads to a minor losses from traveller’s point of view. This difference can be explained with same argument, as used for the social welfare in the preceding paragraph above. The degree of congestion in the presence of crowding disutility is higher compared to the scenario, where the crowding disutility is not considered. This results in a CP equilibrium, where gains from internalisation of congestion cost dominate the increased cost of bus travel.

### 7.4.2 Benefits of dynamic, marginal cost pricing

Figures 7.6 and 7.7 present social and consumer welfare for 4 varying public transport cost and pricing scenarios, each with and without congestion pricing. As already discussed in the previous section, Figure 7.6 shows effects of accounting for the increased valuation of travel time in crowded conditions. In comparison to it, Figure 7.7 visualises social and consumer welfare for scenarios with dynamic public transport pricing based on the internalisation of external costs. The first (VOC intern.) scenario only includes the internalisation of crowding externality. The second (VOC intern. + TTD intern.), combines the internalisation of externalities from crowding disutility as well as delays from dwelling process and vehicle capacity constraint.

For the case without congestion pricing (NCP), the scenario without consideration of crowding disutility shows the highest welfare. More surprisingly, however, is the fact that the internalisation of public transport externalities has almost no effect on social welfare. The scenarios with dynamic fares based on crowding and travel delay externalities, perform only marginally better than the scenario with free public transport (Figure 7.6(a) vs. Figure 7.7(a)). This result can be explained by the modal split in favour of private transport found in the base Corridor scenario. As mentioned above, this is partially intended, partially due to the lack of car ownership model, a detailed model of public transport passenger interaction as well as the behavioural parameter specification. As demonstrated in section 7.1, without demand regulating pricing policies, the bus ridership remains under the socially optimal level, which results in an average social welfare loss from road congestion. Hence, operation of public transport without any usage charges becomes a socially optimal public transport pricing.
strategy. Internalisation of externalities raises the average public transport fare above zero. Though the internalisation of crowding and travel time delays by marginal cost pricing helps to spread the peak demand, its benefits can hardly compensate for the losses from an increase in road congestion due to commuters switching to private transport. As a result, increase in congestion externalities in absence of congestion regulating pricing policy eliminates almost all benefits of marginal cost pricing of public transport.

A very different picture emerges, if next to external costs of public transport, also road congestion externalities are internalized using dynamic, link-based congestion charges, as presented in Chapter 6. As observed above, congestion pricing increases social welfare in all scenarios. As one would expect, the scenario with free public transport usage and inclusion of crowding disutility exhibits the lowest level of social welfare (Figure 7.6). Once the crowding externality is internalised and congestion pricing policy is in place, social welfare rises above the base scenario, which does not include crowding disutility (Figure 7.7). This,
initially paradox situation, can be explained by the complimentary nature of crowding and travel delay externalities. As discussed above, both of these externalities are correlated. First, dynamic, crowding based fares reduce average vehicle occupancy and facilitate the spread of the demand peak during morning and evening commutes. Second, due to the implementation of a detailed, non-linear dwell time model (see section 3.2), travel time delay caused by boarding and alighting processes is directly related to the vehicle occupancy. Hence, internalising crowding externalities, simultaneously contributes to the reduction of travel delay externalities. Furthermore, as discussed in section 7.1 in the presence of dynamic congestion charging, social welfare is rather insensitive to the cost of public transport usage. Congestion pricing ensures, that the increased cost of public transport does not cause a shift from public to private transport and the distribution of demand between modes remains close to optimal. This, however, remains the case only as long as the cost of both motorized modes do not get too high and trigger a significant mode shift to non-motorized modes of transport.

Adding internalised external cost of longer in-vehicle travel times from extended dwell processes as well as additional waiting times from vehicle capacity constraints, on top of the already internalised crowding cost to the bus fares, increases social welfare even further. Again, dynamic congestion pricing controls for the optimal mode share, while internalisation of public transport externalities provides incentives to individual agents to adjust their departure times in a socially optimal way. Given the synergies with crowding externality, however, the welfare gains from internalisation of travel delays on top of crowding cost are comparably low.

Similar to the social welfare in the NCP scenario, internalisation of public transport externalities has only minor effects on the aggregated values of consumer welfare. As discussed in section 7.4.1 including crowding disutility into the model adds another term to the travel cost function and noticeably lowers social and consumer welfare. More relevant, however, is the slight decrease in consumer welfare after dynamic public transport fares based on crowding externality are introduced. Additional fare based on the internalisation of travel delay cost due to dwell time process and capacity constraints has a similar effect. Again, as discussed above, the lack of mode share regulating policy result in the situation, where gains from socially more optimal public transport demand distributions in presence of dynamic fares, are offset by losses from rising congestion delays.

In the presence of congestion pricing, the internalisation of crowding costs slightly increases consumer welfare. An additional fare increase associated with internalisation of both crowding and travel time delays, has a minor negative effect on consumer welfare. From the traveller’s point of view, the small welfare gains from the more efficient public transport operations do not compensate for the increase in public transport fares.

In all scenarios with internalisation of public transport externalities, the degree of heterogeneity has a similar impact on welfare as already observed in the Chapter 6 for scenarios with a
flat public transport fare. As user preferences become more and more heterogeneous, a more socially optimal modal split leads to an increase in social and consumer welfare. For social welfare in case of congestion pricing, this effect fades, resulting in diminishing benefits of congestion pricing policy as heterogeneity increases.

### 7.4.3 Monetary revenues

Figure 7.8(a) visualises monetary revenues, which internalisation of crowding cost (VOC) and travel delay externalities (TTD) generate. If both these costs are internalised and included into the dynamic public transport fares, their internalisation yields similar revenues. The revenue amount doubles after congestion pricing is introduced and increases with the degree of heterogeneity in users values of time. Still, even in that case the revenues from internalisation of external public transport costs are not sufficient to cover operational expenses.

Total monetary revenue from simultaneous internalisation of public transport and road congestion externalities, split in its components, is presented in Figure 7.8(b). As the scale of the y-axis significantly changes compared to the Figure 7.8(a), the scale difference between public and private transport cost becomes apparent. Revenue from congestion pricing is by factor 5-7 higher than the revenue generated from internalisation of public transport externalities. Only a fraction of congestion pricing revenue is required to fully subsidize public transport operations. In a more general economic context, a major part of it still could be used to e.g. lower income taxes and thereby minimize road pricing effects on the labour market. However, it is also important to note, that cost of car ownership, which are not explicitly modelled in the Corridor scenario, can have a significant impact on driver’s willingness to pay for the road usage and therewith on overall potential for revenues from congestion charging.

It is also interesting to note, that as revenues generated by congestion pricing decrease with the growing degree of user heterogeneity, the revenues from public transport increase. This is due to the self-organization effect discussed above. Increasing degree of heterogeneity pushes travellers with lower values of time to public transport, lowering the necessity to price roads in order to ensure optimal speeds. At the same time, higher public transport ridership generates higher crowding levels and travel time delays, driving up the revenues from marginal cost pricing of public transport.

Distributions of bus fares paid per person and per trip, shown in Figures 7.9 and 7.10 for lower and higher degrees of heterogeneity, demonstrate, that even in presence of dynamic public transport pricing a large share of commuters still uses public transport for free. The dotted, vertical red line in each graph represents the total mean fare paid per trip. In case of lower heterogeneity, mean fare per trip is 1.1$ without congestion pricing and about 2.2$ with it. For extreme heterogeneity ($n = 5$) average trip fares are a bit higher, with 1.3$ in the NCP case, and
Figure 7.8: Revenues from internalisation of public and private transport externalities for proportional heterogeneity (5 min bus headway).

(a) Revenues from marginal PT pricing

(b) Revenues from marginal PT and congestion pricing

Figure 7.9: Bus fare distribution for proportional heterogeneity with $n=1$ (5 min bus headway).

(a) No congestion pricing (NCP)

(b) With congestion pricing (CP)

2.0% with CP. More travellers paying high fares and increased average fare in case of congestion pricing policy seems logically, in particular based on the previous discussion. Comparing scenarios with lower user heterogeneity ($n=1$) and therefore less public transport travellers with scenarios with higher user heterogeneity ($n=5$) and larger number of bus passengers, the distribution and fare level hardly changes. Higher operator revenues, as observed in Figure 7.8, result from higher ridership and not from a significant fare increase.
7.5 Distributional effects of dynamic fares

As mentioned above, dynamic public transport pricing is an unusual and not widely accepted demand management measure. In order to gain public support for implementation of a fare scheme based on internalisation of public transport externalities, understanding the economic effects of such scheme on the individual user groups is key.

Figure 7.11 shows the changes in consumer welfare after introduction of dynamic public transport fares, with and without congestion charging policy in place. In both cases, for homogeneous as well as slightly heterogeneous users, consumer welfare remains almost unaffected by the internalisation of public transport externalities. From a user perspective, the average benefits from reduced crowding disutility and travel delay are offset by an increase in public transport fares. The situation changes for travellers with the highest degree of heterogeneity \( n = 5 \). Here, the travellers with a high value of time are the biggest losers of the dynamic bus fare policy. This is the case for scenarios with and without congestion pricing. This can be explained by the fact, that travellers with a high income and therefore high values of time are almost exclusively car drivers. A hike in public transport fares leads to a shift of middle income travellers from bus to car, which increases congestion delay or in case of congestion pricing policy, congestion charges. While car drivers experience decline in consumer welfare, additional fare payments of remaining public transport users are compensated by declining crowding disutility and faster travel speeds.
Figure 7.11: Changes in consumer welfare dependent on income after introduction of marginal, dynamic bus fares for different levels of proportional heterogeneity.

### 7.6 Status quo and the importance of modal split

Benefits of any demand managing pricing policy strongly depend on the level of externalities and the degree to which the system state differs from a socially optimal solution. The initial state of the transport system, the status quo, depends upon properties of transport supply and behavioural characteristics of transport demand.

The preceding analysis of dynamic public transport and congestion pricing is based on the certain assumptions about cost parameters for each mode. Based on these parameters, an emerging stable equilibrium state serves as a base line scenario for pricing policy evaluation. Naturally, some degree of congestion and crowding is needed to be present, in order to demonstrate effects from pricing these externalities. This is strongly related to the (mis)match between supply and demand one the one side, and parametrisation of user preferences on the other.

An important behavioural dimension, which is neglected in this work, is related to the implications of induced demand from altering social and economic activity choices and activity chains of individual urban dwellers. Adaptation of activity locations and travel behaviour
would be logically the next step for extended studies of pricing policy implications. Yet, in consideration of a multi-modal environment, the balance of existing demand between available modes also plays a central role for potential benefits of demand managing pricing policies. This section attempts to provide a better understanding of impact from different initial conditions on policy effects and outcomes from the multi-modal perspective.

To this end, an alternative scenario, with higher fixed car cost per trip is investigated. In the Chapter 5, the only fixed car cost considered were parking cost of 12 $ per day. However, as the model does not include a detailed car ownership and car access model, the car emerged as a preferable travel mode in the multi-modal Corridor scenario. Car trips in the equilibrium state account for between 53% and 90% of all trips, dependent on public transport service frequency, degree of heterogeneity and congestion pricing policy. In order to investigate the sensitivity of pricing models to the specific base case scenario, the fixed car cost per trip in the following scenario are increased three folds to 36 $ per day. In the absence of detailed model of car ownership, parking space availability and accessibility as well as urban design factors, altering fixed car cost allow for different interpretations. Next to the increased cost of parking as reflection of true land-value, it can reflect other factors or forms of taxation, mentioned above. As interaction with these markets is not taken into account in this model, the fixed car cost are not added to the social welfare. This leads to lower welfare of the base case scenario, but does not affect analysis of relative efficiency and benefits from demand regulating pricing policies.

In following sections, analysis based on the altered scenario with higher fixed car cost is presented in a similar manner as it was done in the preceding sections. The discussion focuses on the effects of increased car cost and differences to the case reviewed above.

### 7.6.1 Flat public transport fare

Analysis of social welfare for a scenario with flat bus fares in the range of 0 – 5$ and varying degree of income heterogeneity presented in Figure 7.12, reveals that with the increased car cost, higher public transport mode share and lower degree of congestion, the effects of both varying level of bus fare as well as degree of heterogeneity diminish. Ranging from 159 – 169$ per agent in the NCP scenario with low car cost (see Figure 7.1(a)), average social welfare varies only between 152 – 155$ per agent in the scenario with high car cost. The welfare level also remain stable for flat bus fares in the range of 0$ – 1$, slowly declining for higher fares. The socially optimal flat fare appears to lie slightly above zero, but with only small welfare variations and incorporation of fare revenue into the social welfare calculation, the optimal point does not appear to be very distinct.

The insensitivity of demand to the degree of heterogeneity and bus fare becomes even more
clear, when looking at the revenues from public transport fares in Figure 7.13. While in the previous case with cheaper car usage (Figure 7.4(a)), a revenue maximising fare was clearly identifiable, for high car cost scenario the revenue contentiously increases with a higher public transport fare, indicating a potential saturation point beyond a maximal studied fare of 5$ per trip. This applies to the scenario, with and without congestion pricing policy.

### 7.6.2 Dynamic public transport and congestion pricing

High car cost also show strong effects on benefits from congestion pricing. Given free public transport usage, social welfare increases only slightly after congestion pricing is introduced.
Figure 7.14: Social and Consumer welfare with crowding disutility and internalisation of public transport externalities for proportional heterogeneity and fixed car cost of 36 $ per day (5 min bus headway).

(Figure 7.14(a)). This is mainly due to the high public transport mode share and absence of severe road congestion. Effects of self-organization and more efficient distribution across modes, as a result of high degree of user heterogeneity are still visible, but are less pronounced.

Almost in all settings, consumers experience a welfare losses from internalisation of congestion externalities. Largest losses occur for the homogeneous user case with inclusion of value of crowding into the model. Given already high level of occupancy on public transport, an additional increase in ridership after congestion pricing is introduced leads to disproportionally high losses from in-vehicle crowding. Interestingly, for a high level of heterogeneity these losses disappear, as on average the gains of high value of time car commuters compensate for crowding losses of bus passengers.

Figure 7.15 highlights the effects of internalisation of public transport externalities and their interaction with congestion pricing policy for the scenario with a high cost of car usage and therefore high public transport mode share. In contrast to the scenario with lower car cost, internalisation of bus crowding and travel delay externalities leads to a significant social welfare benefits, even without congestion pricing policy. Due to the correlation of both public transport externalities discussed above, internalisation of crowding disutility alone leads to a higher welfare as compared to the scenario which does not account for crowding disutilities at all (Figure 7.14(a) vs. Figure 7.15(a)). Also, as distortions and inefficiencies in public transport operations are eliminated, the impact of user heterogeneity increases. For a highly heterogeneous user population, internalisation of public transport travel and waiting time delays on top of the crowding disutility internalisation, has similar welfare benefits as congestion pricing. Furthermore, consumers also benefit from internalisation of public transport externalities, despite of the increase in bus fares during the peak hours (Figure 7.15(b) vs. Figure 7.14(b)).
Figure 7.15: Social and consumer welfare with crowding disutility and internalisation of public transport externalities for proportional heterogeneity and fixed car cost of 36 $ per day (5 min bus headway).

Figure 7.16: Changes in consumer welfare dependent on income after introduction of marginal, dynamic bus fare for different levels of proportional heterogeneity and fixed car cost of 36 $ per day (5 min bus headway).

The distribution of gains and losses from dynamic public transport pricing across different income groups and for varying degrees of heterogeneity is presented in Figure 7.16. For both scenarios, with and without congestion pricing policy, the average consumer welfare for all
groups notably increases. The benefits in case of congestion pricing are slightly larger due to the high public transport mode share. For scenarios with high degree of heterogeneity ($n = 3$ and $n = 5$), most significant benefits are experienced by middle-income travellers. While commuters with highest values of time use private transport and commuters with lowest values of time travel on public transport outside peak hours to avoid higher fares, commuters with medium values of time represent the large proportion of peak hour trips on public transport and therefore benefit most from the more efficient demand distribution.

Comparing revenues from simultaneous internalisation of public transport and congestion externalities, joint bus fare revenue collected from crowding and travel delay externalities surpasses the revenue from congestion charging and dependent on the degree of heterogeneity represents $53 - 67\%$ of total monetary revenue (Figure 7.17). Given the high public transport mode share and the high cost of private transport, dynamic public transport pricing based on the internalisation of external costs generates more than sufficient revenue for financing of public transport operations.

In conclusion, the scenario with high cost of private transport and therefore large public transport ridership demonstrate a very different relations between benefits of dynamic congestion and public transport pricing. The gains resulting from minimizing market inefficiencies with internalisation of external costs are strongly depended on the scale of externalities present in the initial system state. Wider consequences of these findings on mobility pricing policies in urban context are taken up again in the Chapter 8.

### 7.7 Large scale scenarios and alternative activity chains

Up to now, a rather simple Corridor scenario, as presented in section 5.1, was used to investigate the impact of user heterogeneity in values of time and effects of dynamic congestion and public
transport pricing policies. However, while providing a convenient set-up for demonstration of welfare and redistribution effects of dynamic pricing policies, it also contains some drawbacks and oversimplifications. For example, it only considers mono directional commute traffic, neglecting route choice and alternative trip purposes. The more complex, enriched Sioux Falls scenario, as presented in section 5.2, addresses some of these shortcomings by adding additional degrees of realism to the simulation. In particular, it contains a more complex road and bus network, enabling the inclusion of a trip-based route choice dimension. Furthermore, next to the peak-hour commuter demand with a typical home - work - home activity chains, it adds an alternative, home - secondary - home activity chain with a different, more flexible set of activity constrains. These trips, which represent shopping or recreational activities, are mostly performed outside of the peak-hours and are therefore affected by user heterogeneity and pricing policies in a different way.

Figure 7.18 illustrates the effects of user heterogeneity and dynamic mobility pricing policies on realised social and consumer welfare for the Sioux Falls scenario. It is worth noting, that only welfare indicators based on the Realised Utility, as defined in section 3.4, are used for the welfare evaluation in the Sioux Falls scenario. This is primarily due to the arising challenges in the choice-set generation of daily schedules, as the route choice becomes an additional relevant choice dimension. It leads to a potential explosion of user’s choice set size and requires an alternative approach to the choice set generation and evaluation. Though addressing this challenge lies outside the scope of this work, the correlation between Expected Maximum Utility (EMU) and Realised Utility (RU), as it has been observed in section 6.2, allows the use of RU as a reasonable welfare indicator. However, one of the major drawbacks of the welfare evaluation based on the RU instead of the EMU approach, is a larger sensitivity to mode shifts. This is due to the absence of the individual’s non-chosen alternatives in the RU’s evaluation process.

As discussed in section 3.4, the EMU evaluation approach keeps the size of each individual’s choice-set and therewith the number of mode alternatives constant. Hence, changes of mode shares affect social and consumer welfare only based on their impact on the performance state of the overall transport system. The RU approach, on contrary, only considers the chosen alternatives and therewith is more sensitive to the actual modal split. This becomes in particularly relevant in the presence of walking as a third mode. Representing a non-congestible alternative, walking attracts number of commuters, who shift away from private and public transport as prices of car and bus travel increase. This mode shift causes a decline in average social welfare, as cost of walking do not contain any charges or fares, which can be recycled to benefit the society (this work follows an idealistic assumption of the full return and recycling of monetary revenues for the social benefit). At the same time, other benefits of non-motorized modes of transport such as e.g. zero emissions or overall health benefits, are not modelled, resulting in decrease in social welfare, as number of walkers increases. This issue also
Figure 7.18: Effect of congestion pricing on social welfare and consumer surplus for different degrees of heterogeneity in the Sioux Falls scenario.

The effect of pricing policies on the social welfare in the Sioux Falls scenario, as shown in Figure 7.18, resembles the effect observed in the Corridor scenario and discussed in section 7.4. Also in case of the Sioux Falls scenario, the social benefits from congestion pricing are clearly visible. Furthermore, introduction of dynamic bus fares based on internalisation of external crowding and travel time delay cost has a similar impact as in the Corridor scenario. Without congestion pricing policy, dynamic public transport fares raise the cost of public transport usage. The associated mode shift outweighs the social benefits of internalisation of externalities. Congestion pricing policy, however, prevents an extensive mode shift. Hence, dynamic public pricing combined with congestion pricing policy leads to an overall increase in the social welfare. From consumer point of view, however, congestion as well as dynamic public transport pricing have negative effects and result in a net losses in consumer welfare as the price of travel increases. In general, similar impact of pricing policies in the Corridor and Sioux Falls scenarios indicates the transferability of results and conclusions from simplified scenario to larger transport network.

A very different picture emerges when considering the impact of user heterogeneity on the social and consumer welfare. Here, the results in the Sioux Falls scenario appear to diverge from observations made in the Corridor case. Both welfare indicators in the Sioux Falls scenario resemble a U-shaped curve, with welfare initially sinking as the degree of user heterogeneity increases. However, after reaching a minimum for moderate degrees of heterogeneity ($n \approx 2 - 3$), it rises again for populations with more extreme spread in values of time. This impact of heterogeneity on social and consumer welfare can be observed with and without presence of pricing policies.
Such, initially surprising behaviour, can be understood when analysing the welfare impact of the two groups of travellers separately: commuters and travellers with secondary activity trip chains (Figure 7.19). Commuters with home - work - home activity chains travel during the morning and evening peak and are the only group of travellers previously considered in the Corridor scenario. In the case of the Sioux Falls scenario, social and consumer welfare of commuters significantly increases as heterogeneity in values of time increases (Figures 7.19(a) and 7.19(b)). Similar to the Corridor scenario, the welfare increase is caused by two related effects: first, as degree of user heterogeneity increases, travellers with lowest values of time shift from car to public transport or walking. For these users the total cost of car usage outweighs the time gains associated with it. Second, the welfare gains from reduced congestion caused by changes in modal split are amplified, as the main beneficiaries of ease in congestion are the travellers with higher values of time. These travellers remain car drivers and contribute disproportionately to the aggregated welfare indicator. This effect is so dominant in the Sioux Falls scenario, that the social welfare of homogeneous commuters under congestion pricing policy is noticeably lower than the social welfare of highly heterogeneous commuters without congestion pricing.

Figure 7.20 shows the changes in mode choice dependent on the scenario’s degree of population heterogeneity. Though for home - work - home activity chain the modal split changes for lower degrees of heterogeneity are barely noticeable, reduction of car mode share only by few percentage points during the peak hours already leads to a substantial increase in social and consumer welfare due to large benefits of commuters with high values of time.

The increasing heterogeneity, however, has an exactly opposite effect on the welfare of travellers with home - secondary - home activity chains, who mostly travel outside the peak hours. Figures 7.19(c) and 7.19(d) show a clear drop in social and consumer welfare of these travellers, as the degree of heterogeneity increases. As can be seen in Figure 7.20(b), the mode shift among travellers with shop or recreational activities is dominated by increase in walking trips. Here, the travellers with lower values of time can not afford travelling by car or bus and increasingly choose walking as their main travel mode. However, as outside of the peak-hours, this change in modal split does not lead to any benefits from congestion or crowding reduction and results in a net welfare loss. Travellers with secondary activities spend significant larger share of their time on travelling instead of performing activities. Furthermore, as there is hardly any congestion and crowding during the off-peak hours, congestion pricing and dynamic public transport pricing has no effects on welfare of these travellers.

The superposition of welfare changes from the two groups of travellers leads to an overall U-shaped welfare graph as previously observed in Figure 7.18. For lower degrees of heterogeneity the welfare changes are dominated by losses of travellers with secondary activities, while the gains of commuters with higher values of time lead to an increase in overall welfare for highly
Figure 7.19: Effect of congestion pricing on social welfare and consumer surplus for different activity chains and varying degrees of heterogeneity in the Sioux Falls scenario.

(a) Realised social welfare (home - work - home)  
(b) Realised consumer welfare (home - work - home)  
(c) Realised social welfare (home - second. - home)  
(d) Realised consumer welfare (home - second. - home)

Figure 7.20: Sioux Falls mode share depended on the degree of heterogeneity and individual’s activity chain (NCP).

(a) Home - Work - Home  
(b) Home - Secondary - Home

heterogeneous population.

It is also important to note that changes in welfare of different travel groups are much more substantial than changes in the total aggregated welfare. This can be partially attributed to
the evaluation of Realised Utility as the welfare indicator. Taking only the actually executed daily schedule into account, RU is more sensitive to mode shifts as the Expected Maximum Utility evaluation approach (chosen alternative vs. logsum over a choice-set); nevertheless, this behaviour once again strengthens the case for a disaggregated modelling approach and inclusion of heterogeneity into the demand modelling process. It also demonstrates how activity patterns of individuals may play a major role for identification of winners and losers of pricing policies and highlights the need for further research.
Chapter 8

Conclusion and Policy Implications

In this thesis, effects of various mobility pricing policies based on the internalisation of external costs were investigated using an agent- and activity-based simulation. The agent-based implementation of dynamic congestion and public transport pricing policies in a multi-modal context enabled the analysis and evaluation of the impact of such policies on social and consumer welfare from aggregate and disaggregate perspectives. To this end, detailed models of interactions between public transport users and public transport supply on the one hand, and private and public transport on the other hand were applied. Such models have proven to be key for understanding a variety of forces impacting travel behaviour and leading to the emergence of stable equilibrium between transport demand and supply under dynamic mobility pricing policies. Important model features include non-linear dwell time model accounting for passenger friction during boarding and alighting process, public transport vehicle capacity constraints, interaction of buses and cars on the road as well as crowding disutility models based on in-vehicle passenger occupancy.

Aiming to demonstrate the importance of integrated approach to urban mobility pricing design and evaluation, this work put a special emphasis on welfare impact of heterogeneous user preferences, availability of alternative modes and interaction of multiple pricing policies in a multi-modal context. Addressing these questions with an agent- and activity based simulation approach, required the development of methods for economic analysis and linkage of the simulation-based results with findings of analytical approaches, as presented by a wide body of literature.

On the demand side, the behavioural parameters of travel population were borrowed from various studies and adopted in an activity-based context, aiming to achieve a realistic and representative study case. Income-based valuations of travel time were assumed to represent the main source of heterogeneous user preferences. Exploiting the central advantage of an agent-based approach, individual values of time were modelled on a level of individual agents.
and integrated into each agents utility function by altering the marginal utility of activity performance. As a result of this methodologically innovative approach, various impacts of monetary toll or fare charges emerge intrinsically as agents maximise their utility.

Focusing on mode and departure time choice, structures and algorithms for internalisation of congestion as well as public transport externalities were implemented and studied within a multi-agent transport simulation framework (MATSim). Link-based congestion pricing was implemented based on internalisation of marginal external congestion cost. Public transport externalities considered in this work and used for implementation of dynamic public transport fares include travel time delays from boarding and alighting processes during bus operations, extended waiting time due to full public transport vehicles as well as increased valuations of time in crowded conditions. As economic theory predicts, internalisation of external cost resulted in an increase in social welfare in all scenarios. The size and scale of social welfare benefits, however, showed significant variations.

### 8.1 Multi-modal modelling approach and user heterogeneity

The conducted experiments demonstrate the relevance of considering heterogeneous values of time when analysing the impact of pricing policies with regards to mode choice. Affecting the level of efficiency of transport demand distribution across available modes, user heterogeneity plays a major role for potential benefits from dynamic pricing policies. Using only one average value of time for all travellers does not allow for emergence of self-organising effects in case of multiple mode choice dimensions and can significantly over- or underestimate the social and consumer welfare changes and influence policy design.

Furthermore, integrated multi-modal approach to policy evaluation proves to be crucial for capturing distributional effects of mobility pricing policies. From social equity perspective, understanding who and how changes their travel behaviour as a consequence of a new demand management policy, is a decisive factor for accurate cost-benefit analysis of such measures. If, for instance, implementation of a pricing policy induces changes in the modal split, it is crucial to understand the availability, cost and attractiveness of alternative modes for users who are most likely to be affected by the policy in question. This work shows how an integrated multi-modal and agent-based simulation approach can be successfully applied to address such challenges in modelling and simulation.

Initially, based on the employed multi-modal Corridor scenario, it is demonstrated, that the evaluation based on a homogeneous user population leads to underestimating the mode share of modes with longer travel times, but potentially lower monetary cost. In most cities and urban areas public transport represents such a mode. Depending on the relation between
transport demand and supply, a misrepresentation of mode shares can have significant effects on evaluation of social and consumer welfare and potential benefits of demand regulating pricing policies. For example, evaluating the effects of public transport capacity extension or operational service frequency increase based on the assumption of homogeneous traveller population, may result in underestimation of welfare benefits from these measures. Vice versa, the gains from introduction of congestion pricing policy in a multi-modal transport network can be overvalued.

Another example that highlights the importance of user heterogeneity is observed in the distributional effects from public transport fares based on the internalisation of externalities. While adding dynamic public transport pricing in a homogeneous scenario might not significantly change the consumer welfare, in heterogeneous case higher fares can disadvantage car travellers. This is a consequence of a significantly wider range of values of time leading to a higher public transport ridership. More public transport travellers lead to a larger number of users switching to car, once the introduction of dynamic bus fares increases the cost of public transport. In absolute terms, higher number of public transport travellers changing to the car mode, triggers a disproportional increase in congestion delays. These findings strengthen the case for use of multi-modal models and heterogeneous values of time for urban transport policy design and evaluation.

Furthermore, the balance between the scale of inefficiencies present in each of the available choice dimensions has proven to be the determining factor for benefits from a pricing policy under study. For example, in the initial Corridor scenario with low car usage cost, introduced in Chapter 5 and applied through Chapters 6 and 7.1–7.5, car mode attracts a large number of commuters due to its speed advantages, which for the most travellers outweigh the associated monetary expenses. This makes congestion pricing policy significantly more efficient as introduction of dynamic public transport fares. While congestion pricing policy pushes a number of commuters from private to public transport and therewith corrects for an inefficient demand distribution between modes, dynamic public transport fares can only correct for departure time distribution inefficiencies. At the same time dynamic fares make the demand distribution between modes even more inefficient by increasing the average bus fare. Increase in the cost of car usage, as investigated in Chapter 7.6 reverses the situation and makes the internalisation of public transport externalities significantly more efficient policy measure in terms of social welfare than congestion pricing. Pricing externalities of private and public transport simultaneously results in a more efficient mode distribution independent of the initial situation and allows for additional benefits of departure time distribution of travellers for all modes. Scenarios with internalisation of road congestion, public transport travel and waiting time as well as crowding externalities, result, as to be expected, in the highest level of social welfare.
Relation between cost, capability to attract users and overall physical capacity of different modes of transport also determines the optimal level and the overall benefits of public transport subsidies. For situations, where public transport struggles to attract enough ridership due to comparatively low cost of car usage, no flat or dynamic fare might exist which can enable self-financing operations, given a compulsory minimal level of service provision. Alternatively, if car use is expensive or car ownership low, public transport ridership is less sensitive to the fare level. This might represent one of the reasons why public transport networks in cities like Singapore, where car ownership cost are extremely high, are able to attract a critical number of users in order to generate enough revenues for operation financing without subsidies. However, as high passenger numbers inevitably lead to crowding and associated externalities, dynamic public transport pricing based on internalisation of external cost can significantly improve operations efficiency and increase social and consumer welfare. Yet, before relying on assumptions of an inelastic demand, interaction with other markets, as e.g. labour market and other equity and accessibility related interdependencies have to be carefully considered.

From a consumer perspective, changes in individual’s welfare after introduction of congestion pricing policy exhibit a strong dependency on the availability of alternative modes and their service quality. In the case of comprehensive and efficient public transport provision, losses of drivers who are tolled off the road and need to divert to public transport, are minimised. For homogeneous users the combination between travel distance and travel mode turns out to determine the impact of congestion charges on individuals welfare. However, for highly heterogeneous population impact of the travel distance fades, leaving the value of time as the determining factor for change in consumer welfare.

Modelling of detailed road interaction between car and buses enables to account for policy welfare effects on users, who are not directly affected by the policy. One the one hand, a positive feedback loop between congestion reduction and public transport travel times may lead to public transport passengers, who continue to use buses after congestion pricing policy is introduced, profiting from the increase in bus travel speeds on uncongested roads. On the other hand, increased demand on public transport due to the implementation of congestion pricing may cause long distance bus commuters to loose disproportionally from the policy, even though they appear to be unaffected by the direct congestion charges. This effect becomes more distinct, once crowding disutilities are taken into account. Which of these effects will dominate the welfare changes in the end, strongly depends on the balance of transport demand and supply as well as the total cost of each mode.
8.2 Application to a large-scale scenario

Applying the same methodology to a more complex and realistic Sioux Falls scenario, provides new valuable insights into welfare effects of user heterogeneity and mobility pricing. While confirming beneficial impact of joint dynamic congestion and public transport pricing policies based on the internalisation of externalities, the presence of alternative activity chains affects welfare changes resulting from an increase in user heterogeneity. For commuters, travelling during the peak-hours, the mode shift from public to private transport induced through self-organisation of heterogeneous travellers has overall positive effect. Easing car congestion, it improves the overall state of the transport system. This has been also observed in the Corridor scenario. However, the opposite appears to be the case for the off-peak travellers. Here, the mode shift towards modes with smaller monetary cost, such as bus and walking, does not improve the overall state of the system simply because of absence of congestion and crowding. Hence, with the increasing heterogeneity social and consumer welfare for travellers with secondary activity chains who mostly travel outside of the peak hours decreases.

This effect is amplified by the differences in the welfare evaluation approaches between Corridor and Sioux Falls scenarios. In the Sioux Falls scenario, welfare indicator based only on the chosen alternative is applied, thus neglecting the potential benefits of available but non-chosen alternatives. While such approach provides qualitatively valuable results, the scale of welfare changes between the two approaches may differ. For instance, given the availability of walk as a third, non-motorised mode, dynamic pricing of road and bus users makes travellers with shorter routes switch to walking. However, as health and pollution benefits of non-motorised modes are not considered, welfare gains of dynamic pricing are likely to be underestimated. On the contrary, losses from managing and reinvesting monetary toll and fare revenues are also not accounted for and may lead to some overestimation of pricing benefits, in particular in areas with higher levels of corruption and oversized governing structures.

In summary, the Sioux Falls scenario shows, that welfare impact of self-organisation of heterogeneous users between available modes, is highly dependent on users’s activity chains, associated departure times and network congestion levels. The welfare evaluation approach also plays an important role and requires refinement and further research for large-scale networks with route-choice as one of the behavioural choice dimensions.
8.3 Planning practice and outlook

While considering the cost of each mode, it is crucial to note, that the real cost of transport usage does not only contain travel time and monetary expenditures, but also depend on a variety of other factors. The imbalance of attractiveness between private and public transport modes can be also a consequence of transport infrastructure design and urban landscape. The presence of overhead bridges, lack of escalators and elevators, inconveniently placed entries and exits to underground stations and public transport stops or unsafe walking conditions can all significantly contribute to the high cost and therewith unattractiveness of public transport. In the same way, easy access to parking, subsidised parking space, wider roads or traffic light circuits optimised solely for cars, can lower the cost of private transport beyond benefits reflected in measures of travel time or money. Hence, in a diverse, urban environment multi-modal mobility pricing is only one of many policy measures available to the policy maker for improving efficiency of the transport system and should be applied in an integrated context and as a part of a wider policy mix.

Applied efficiently, however, mobility pricing provides a powerful demand management tool, enabling the use of existing infrastructure in a most efficient way. This work highlights the importance of joint evaluation of congestion pricing policy and public transport operations. For example, given a fixed public transport fare, optimal operation frequency with heterogeneity can be higher than with homogeneous users. Furthermore, even if an increase in level of public transport service might have a negative effect on the social welfare when evaluated as a stand alone policy, packaging it with congestion pricing policy can not only increase social but also consumer welfare and turn out to be Pareto improving.

Yet, despite of the advantages of an integrated assessment approach with heterogeneous user population, as presented in this work, caution needs to applied when assessing benefits and potential drawbacks of a certain policy or an infrastructure project. In a model with heterogeneous user preferences, individuals with high values of time contribute more to the social welfare as individuals with low values of time. This gives high value of time travellers more weight in the social welfare calculation and might result in a situation where only a minority, consisting of high-value of time travellers, benefits from gains the project or the policy in question. Hence, disaggregated analysis of groups according to their socio-economic characteristics as well as activity- and travel patterns is crucial to ensure not only an efficient but also a democratic policy design. In addition, functioning governing bodies, efficient revenue management and fair redistribution mechanisms are key for the social and economic success of any mobility pricing policy.

One of the major challenges in implementation of models with heterogeneous user populations is associated with excessive cost of data collection, required for estimation of disaggregated
values of time based on socio-economic characteristics, activity types, trip characteristics etc. Furthermore, additional degrees of complexity arise in the presence of new behavioural dimensions, such as activity location choice and complex activity chains. Having potentially wider economic implications, effects of pricing policies on economic and social activities of individuals require further study and careful evaluation.

This work mainly focuses on the short-term effects of mobility pricing policies. However, from the urban development perspective long term effects, such as residential and work location choice or decision of car ownership, are at least as important as the more rapid behavioural response to mobility pricing. Here, agent-based frameworks coupled with land-use and residential location choice models (see Section 2.1.2), open a highly interesting avenue for the future research. Design of optimally balanced mobility pricing scheme under consideration of both, short-term and long-term effects, is one of a big challenges for transport and city planners.

In general, a set-up of a large-scale agent-based model is a laborious task with extensive data requirements on transport infrastructure, building stock, population statistics, travel behaviour as well as residential-, business-, work- and education locations (Erath et al., 2012a). However, once such model is established, it provides a variety of benefits for scenario-based analysis and due to the large amount of disaggregated data incorporated in it, allows a wide range of applications. As shown in this thesis, an agent- and activity based simulation approach opens up new prospects for pricing policy design and evaluation. In particular, its ability to account for travel demand patterns of economic agents on individual scale including their socio-demographic attributes, makes it a highly suitable and attractive tool for aggregate and disaggregate evaluation of transport-related policies in urban regions of any size or scale.
Bibliography


163


