DISS. ETH NO. 21496

DESTINATION CHOICE MODELING OF DISCRETIONARY ACTIVITIES IN TRANSPORT MICROSIMULATIONS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ANDREAS HORNI

Master of Science ETH in Computer Science

born on 22.11.1977

citizen of
Leutwil, AG

accepted on the recommendation of

Prof. Kay W. Axhausen, examiner
Prof. Darren M. Scott, co-examiner

2013

Contents

List of Figures

List of Tables

Abstract

Transport of persons and goods brings benefits and costs for individual actors and for the community. A main goal of transport planning is the maximization of net benefit or social welfare through influencing the transport system with due consideration of its environment. Naturally, the definition of net benefit is highly complex and subject to societal discussion. A first step for efficient control and understanding of transport system is transport modeling. While aggregate transport models, such as the 4-step procedure, still dominate the practice and, according to lex parsimoniae, probably will continue to play a significant role, disaggregate models (also called second generation models), with the individual as basic modeling unit, are becoming more and more important. They are able to address infrastructure management issues rather than being focused on infrastructure extension as are the first generation models.

A prominent instance of disaggregate modeling are microsimulations that explicitly model the interactions of micro-units, here individuals or vehicles. Due to their conceptual appeal, the large research body, the continuously increasing computational power, and a large availability of microsimulation software packages, microsimulations have the potential to become state-of-practice in efficiently complementing the aggregate approach. A prerequisite for the exploitation of this potential is progress in terms of several crucial issues. On a general level, this concerns the methodically correct and computationally efficient handling of random variability in large-scale scenarios. On a level more specific for spatial choices, such as destination choices, the consistent choice set specification is a crucial problem so far missing a consistent solution.

This thesis' goal is contributing to such progress by providing an operational shopping and leisure destination choice module, implemented for the multi-agent transport simulation MATSim as an example, efficiently applicable for large-scale scenarios, and easily adoptable for other similar simulation models. The main contribution is the formulation of an efficient procedure to generate quenched randomness, i.e., to enable consistent handling of randomness in iterative large-scale frameworks.

Zusammenfassung

Verkehr, d.h. die "*Ortsveränderung von Personen und Gütern (und Nachrichten) in einem definierten System*" (Ammoser and Hoppe, 2006, S.21), ist essentiell für die arbeitsteilige Gesellschaft, da Produktion und Konsum an verschiedenen verteilten Orten stattfinden. Dabei entstehen sowohl für das Individuum als auch für die Gesellschaft Kosten und Nutzen. Die Hauptaufgabe der Verkehrsplanung ist die Wohlstandsmehrung unter sorgfältiger Abwägung dieser Kosten und Nutzen und durch gezielte Einwirkung auf das Verkehrssystem (inkl. seiner Akteure) mit Berücksichtigung seiner Umwelt. Eine gezielte Einwirkung ist angewiesen auf das Verständnis des Systems, im Speziellen auf eine möglichst präzise Abschätzung der Systemantwort auf Veränderungen, z.B. auf Massnahmen. Dazu ist, wie bei jeder systematischen Suche nach Verständnis, Modell-Bildung unverzichtbar. Während aggregierte Verkehrsmodelle, wie z.B. der 4-Stufenansatz, immer noch die Planungspraxis dominieren und gemäss der Lex Parsimoniae wohl auch in Zukunft eine signifikante Rolle behalten werden, nimmt die Bedeutung disaggregierter Modelle laufend zu. Nur diese Modelle können die Fragen beantworten, welche in westlichen Zivilisationen zunehmend an Relevanz gewinnen, weil sich der Planungsschwerpunkt stetig vom Infrastrukturausbau in Richtung Infrastrukturmanagement mit ganz neuen Herausforderungen wegbewegt.

Ein prominenter Vertreter disaggregierter Modelle sind Computerprogramme, welche i.d.R. sowohl die Entscheidungen als auch die Interaktionen von Individuen explizit modellieren, sogenannte Mikrosimulationen. Durch ihre konzeptionelle Stärke, die kontinuierlich wachsende Rechenleistung und eine mittlerweile beachtliche Zahl von verfügbaren Mikrosimulations-Software-Paketen besitzen sie das Potential, das Spektrum der planerischen Standardwerkzeuge zu ergänzen. Entscheidend dafür ist aber weiterer Fortschritt bezüglich folgender Probleme: Die Interpretation und das Verständnis als auch der effiziente Einbezug von stochastischen Elementen in Mikrosimulationen sind zwar wichtig, stehen aber erst ganz am Anfang. Im Weiteren ist das für die Zielwahl und andere räumliche Entscheidungen bis dato ungelöste Problem der Spezifikation von Alternativensätzen dringlich.

Diese Dissertation berichtet über Fortschritt in diesen Bereichen, wobei die Erkenntnisse im Zielwahl-Modul für Einkaufs- und Freizeitverkehr in der Multi-Agenten-Verkehrssimulation MATSim, als Beispiel umgesetzt werden. Hauptbeitrag ist eine Methode zur effizienten und konsistenten Generierung von stochastischen Elementen in iterativen Prozeduren.

Chapter 1

This Dissertation: Research Goal, Steps and Structure

The ambition of this thesis is to contribute to microsimulation destination choice modeling. Destination choice modeling encompasses a large number of sub-problems ranging from computational to behavioral issues. The global ambition thus needs to be transformed into a concrete, revisable goal by limiting the problem in breadth or in depth. Given the clear need for an operational destination choice in the well-used MATSim framework and similar frameworks and the interdisciplinary orientation of the research group, the focus is laid on application and integration of basic destination choice findings into a microsimulation software. Many of the numerous existing simulators are based on utility-maximization and an agent-concept in some form or another. Choosing the MATSim framework as an example, hence, promises good generalizability of the thesis' findings.

Thus, this thesis' goal is implementation of a MATSim destination choice module for shopping and leisure activities efficiently applicable for large-scale scenarios and easily adoptable by other similar simulation models.

All simulators developed so far made progress along the lines we follow. However, the complexity and extent of the problem still asks for relevant improvements, in particular as many problems have been approached very preliminarily, given the different foci and possibly resources assigned. In our context, computability is a major issue, thus, main contribution is efficient and methodologically sound computation of quenched randomness so far not yet treated consistently.

Further progress is made along the following lines, representing the cornerstones of the thesis.

I. **Incorporating destination choice in microsimulations:**
   (a) **Adding destination choice to the MATSim choice process as an example**: Chapter 2 analyses the context of transport microsimulations and embeds MATSim therein. Focusing on the discrete choice modeling framework (McFadden, 1978), Chapter 3 identifies relevant choice determinants, required data, and previous methods applied in similar microsimulations. Chapter 4 provides an operational, efficient and easily generalizable destination choice module for MATSim. It is focused primarily on shopping trips, where, for sake of completeness, i.e., to make it operational, leisure trips are included as well.
   (b) **Further analysis of destination choice processes and specification of destination choice sets:** While for small alternative set problems choice set specification is natural, for problems with a large universal choice set, individual choice set specification becomes a challenging computational and behavioral problem. Chapter 5 presents destination choice model estimation with the example of MATSim. A probabilistic choice set model is tested and a survey introducing some methodological and technical innovations and laying a base for future approaches to the destination choice set problem, is presented.
   (c) **Modeling agents' destination interactions and spatial correlation of alternatives**: Similar to interactions on the road network persons' interactions at destinations influences destination choice; for leisure activities they are sometimes the only reason for a trip. For shopping, often interactions occur as competition, for example, during parking search. Chapter 7 presents models of destination interactions and spatial correlation of alternatives.

II. **Computability:** Computing power constantly increases due to technical progress, but at the same time problem range and size modeled ever increases. Thus, scalability is and always will be a major focus for microsimulation development. Chapter 4 introduces a set of combined techniques to handle the computational problem of huge choice sets usually present in destination choice. Iterative models, such as MATSim, intensify the computational issues. A procedure to consistently do random draws in very large-scale scenarios, usually termed *quenched randomness*, is the main contribution of this dissertation and applicable in a broad range of stochastic iterative problems.

III. **Interpretation of microsimulation results:** Microsimulation results span a weighted possibility space, where one simulation run represents a sample point in this space. With the incorporation of destination choice usually associated with large heterogeneity, variability analysis is necessary. Chapter 6 discusses microsimulation variability and analyzes MATSim results variability. Main contribution is the attempt to show, that variability is essential for microsimulations, where this view is not yet the common thinking in the microsimulation community.

Evidentially, there are more issues relevant for destination choice; they are proposed as future work in the Chapters 8 and 9.

1.1 Generalizability

Generalizability is a vital requirement of this thesis. Often, generalizability is achieved most efficiently by following a deductive approach, i.e., by developing a system of theories and then applying it to a large range of problems. Here, this would be a ubiquitous destination choice module applicable in all simulators. However, the high complexity of modern large-scale microsimulations and the many specific characteristics of each model, require their thorough inspection before any generalization step can be successfully performed. An example is quenched randomness, which is crucially required for iterative frameworks, but a non-issue in non-equilibrium models. Heuristics for achieving computability are another example. This means, that an inductive approach is more efficient or maybe even required for microsimulation development and this thesis.

The two different approaches, deductive and inductive, usually condense in different choices of projects, papers and thesis titles. Deductive approaches lead to titles starting with “Microsimulation ...” right from the beginning, whereas for inductive approaches, titles starting with the name of the specific simulator (“MATSim ...”) can only be legitimately replaced by general titles after a substantial generalization step. This rule is followed in this thesis. As its research—in particular the handling of quenched randomness, choice set specification, random variability, and agent interactions—clearly goes beyond the MATSim domain, the thesis title encompasses microsimulation destination choice in general.

1.2 Inclusion of Leisure Traffic

On the time scale of common transport micro-simulators, usually modeling an average day, home, work and education activity locations can be assumed to be stable. Furthermore, relevant data are often available at the person level with high spatial resolution. Consequently, they are often handled as exogenous input to the model.

In contrast, the time scale of *shopping* and *leisure* destination choices, and their relatively strong dependency on travel times suggests to treat them *both* endogenously. To ensure a certain methodological depth, this investigation focuses on shopping trips, where leisure trips are handled to have an operational model, or in other words, for sake of completeness. This is due to the important fact, that a large part of leisure activities are social, where sometimes social interaction is the only activity purpose (more than 20% of all leisure trips have purpose visiting friends (Swiss Federal Statistical Office (BFS), 2006)), meaning that, social structures crucially need to be captured. These models are currently researched for MATSim, but they are not quite ready for productive large-scale application. However, they will, most probably, explain a substantial part of the to date unobserved heterogeneity in our model.

1.3 Chapters

Following list informs about the published base of the chapters:

- **Chapter 1:** -
- **Chapter 2:** -
- **Chapter 3:** Horni and Axhausen (2012a)
- **Chapter 4:** Horni et al. (2009a,b, 2012c, 2011d)
- **Chapter 5:** Horni et al. (2011a)
- **Chapter 6:** Horni et al. (2011c,b)
- **Chapter 7:** Horni et al. (2009a, 2012a); Horni and Axhausen (2012a); Horni and Ciari (2011, 2009); Horni et al. (2012b)
- **Chapter 8:** combination of all references in this list
- **Chapter 9:** Horni and Axhausen (2012a)

Chapter 2

An Overview

The goal of this chapter is to present the MATSim basic principles and to embed MATSim in the transport modeling context. Section 2.1 introduces the central concept in transport planning methodology, the transport system equilibrium. Section 2.2 provides an overview of transport modeling, and in Section 2.3 transport microsimulations are introduced. Section 2.4 describes the MATSim basics and sketches its underlying principles.

2.1 Transport System Equilibrium

Transport of persons and goods brings benefits and costs for individual actors and for the community. A main goal of transport planning is the maximization of net benefit or social welfare through influencing the transport system spanning the infrastructure and its actors and in due consideration of the system's environment. Naturally, the definition of net benefit is highly complex and subject to societal discussion. Important questions are, among others, which factors require inclusion in the calculation or the weighting of individual and collective interests. Widely accepted criteria for guiding these political discussions are, e.g., *Kaldor-Hicks efficiency* or *Pareto efficiency*.

The first step for efficient control and understanding of transport system is transport modeling. Derived from a general economic perspective, a demand-supply equilibrium paradigm can be adopted in transport modeling (Boyce and Williams, 2003, p.38), (Bates, 2000; Patriksson, 1994), (Ortúzar and Willumsen, 2001, p.7). The equilibrium assumption can be formulated as follows (see also e.g., Ortúzar and Willumsen (2001, p.7f). Transport infrastructure, interpreted as the supply side, provides a service, whose usage costs increase with demand, e.g., travel time is higher for higher loads. Under the reasonable assumption that demand is dependent on these costs, naturally, formation of a demand-supply equilibrium can be expected.

Specific transport decisions are usually made by actors (e.g., travelers) optimizing *their individual* benefit, thus, the equilibrium described above is termed *user equilibrium*, reflecting Wardrop's first principle (Wardrop, 1952; Correa and Stier-Moses, 2010). The efficiency of this descriptive state can be compared to the normative *system optimum*, described by Wardrop's second principle.

Although this perspective, at first sight, looks straight-forward, transport planning is a very complex task. Above general perspective needs differentiation by different types of equilibria (Section 2.1.1). Furthermore, it suffers from severe problems such as e.g., from the Braess paradox (Braess, 1969) and the (Pigou-Knight-)Downs(-Thompson) paradox (e.g., Downs, 1962), where adding supply can, counter-intuitively, increase usage costs.

2.1.1 Types of Equilibria

Transport infrastructure is not static and, at least from modeler's perspective, also not deterministic. Extension of user equilibrium (UE) by randomness and dynamics leads to stochastic user equilibrium (SUE) and dynamic user equilibrium (DUE), respectively. Although, conceptually, this spans all choice dimensions, in early models, not all dimensions were included. Often, only route and time choice are subject to equilibration where mode, destination and activity chain choice are exogenous to models. Reasons are mostly of practical nature; models simply cannot be comprehensive from their beginnings.

Costs are usually given as *generalized costs*. Traditionally, they are composed of individual time and money expenditures (Bates, 2000, p.12), where, clearly, many further components exist. Externalities, i.e., costs for non-users, such as immissions, ever become more important, hence, modern models should also be able to compute cases, where users monetarily compensate these external costs, in other words, where these costs are internalized.

Deterministic User Equilibrium (UE): Deterministic user equilibrium is formulated by Wardrop's famous first principle Wardrop (1952) as: "*The journey times on all routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.*"

According to Boyce et al. (1988, p.162), the first-known statement of user equilibrium dates back to Pigou (1920) and was also discussed in Knight (1924). A rigorous mathematical formulation of user equilibrium is given in the seminal book of Beckmann et al. (1956). Their optimization problem formalization made possible efficient algorithms for computation of user equilibrium, but this was not recognized immediately by transport modelers (Boyce and Williams, 2003, p.26). Surprisingly, the close relation to the Nash equilibrium (Nash, 1951, 1950) [1] was not mentioned by Wardrop but only nine years later by Charnes and Cooper (1961).

Stochastic User Equilibrium (SUE): Clearly, travelers are *not* perfectly informed, and from modelers' perspective some behavior looks stochastic, generating *unobserved heterogeneity* in the surveyed data. These effects can be taken into account in the model by adding random error terms to the users perception, where, still each traveler is assumed to minimize his individual *perceived* travel costs (Ortúzar and Willumsen, 2001, p. 363) leading to stochastic user equilibrium (Daganzo and Sheffi, 1977).

Dynamic User Equilibrium (DUE): Traffic is highly dynamic. One approach to take this into account and thus to increase model resolution, is to build independent time slices, for example for peak and non-peak hours, and to assume a static equilibrium for each of these periods. A more elegant approach is extension of the equilibrium formulation as follows. Conceptually, the user equilibrium can be made dynamic relatively straight forward. Instead of only taking into account route choice, departure time choice can be included. This means that no user can improve his performance by unilaterally changing his route or departure time (see e.g., Friesz, 2010, p.411). Despite the conceptual straightforwardness, implementation of dynamic equilibrium models is complex (see Section 2.2).

2.1.2 Existence, Uniqueness, Stability and Behavioral Basis of Equilibria

For the design of algorithms to compute equilibria and the interpretation of results, knowledge about the qualitative characteristics of the targeted equilibrium—such as existence, uniqueness and stability—usually are productive. Hence, these characteristics of aforementioned equilibriums have been researched intensively, Dafermos (1971); Smith (1979, 1983) for the UE, Daganzo (1983) for the SUE and Smith (1993) for the DUE, to name only a few.

An essential component of an equilibrium formulation is its temporal and spatial domain. Equilibria, and its qualitative characteristics, can be

[1] As succinctly put by Correa and Stier-Moses (2010) "*a Wardrop equilibrium can be viewed as an instance of a Nash equilibrium in a game with a large number of players*".

local and global as well as short-term and long-term (Ortúzar and Willumsen, 2001, p.8). Furthermore, equilibrium can span multiple dimensions, including destination and activity type choice. This leads to day plan equilibria, which are *not* necessarily in equilibrium for every single choice dimension considered in isolation. A day plan equilibrium, for example, does not necessarily include a Wardrop equilibrium, if, in some situation, further driving generates a higher utility than waiting at the destination, while paying parking costs (Peter Vovsha, personal communication, July 2012).

Clearly, the concept of equilibrium, in particular the UE, is an abstraction from, and thus, an approximation to reality and it is discussed controversially (see e.g., Patriksson, 1994, p.58ff). The fundamental question concerns existence of equilibria in reality and their evolution from non-equilibrium states (Peeta and Ziliaskopoulos, 2001, p.254). Horowitz (1984), for example, investigated a simple 2-link network in terms of SUE and did not find dominating stability. He concludes that *"the validity of the standard assumption about the achievement of equilibrium appears to be highly questionable"*. Holden (1989, p.251), in a theoretical paper, criticizes UE assumption in context of a potentially chaotic system and in absence of a stringent behavioral basis. On the other hand, Friesz et al. (1994), later, investigated day-to-day adjustment processes toward a static Wardrop user equilibrium, and demonstrated that, eventually, an equilibrium state is reached. Mahmassani (1989) empirically found the same in laboratory experiments.

For microsimulations, very little is known about the targeted equilibria. These models are highly dynamic, stochastic and disaggregate with many user classes and behaviorally rich decision principles. In Section 2.4.2.3, the MATSim equilibrium is discussed. As, in general, for the interpretation of results but also for model development (see equilibration discussion in Section 9.2.2), knowledge about the characteristics of the equilibrium searched is helpful, further research is strongly suggested.

2.2 Transport Modeling

2.2.1 Modeling

Modeling is a universal and basically omnipresent and inevitable process while acting in this world. Already perception, creates an abstraction of reality, in philosophical terms, a simulacrum, and in more practical terms, a *model*. More focused on the explicit and conscious part of modeling, Epstein (2008, Section 1.9) lists 16 reasons for modeling, where in our opinion prediction, understanding, and experimentation are central. The modeling enterprise is sketched in Figure 2.1 loosely based on Petty (2010, Figure 10.2). Modeling starts with observation and measurement of reality (in Petty (2010) called "*simuland*") for acquisition of knowledge (in Petty (2010) called "*referent*"). Model creation—in a strict sense usually referred to as *modeling*— is based on the modeler's knowledge about the world. Based on a conceptual model, an executable model is implemented and calibrated. The executable model is evaluated in a verification step in terms of "*was the model made right?*" Petty (2010, p.332). Validation compares results with the referent in the sense of "*was the right model made?*" Petty (2010, p.332).

Final purpose of modeling is knowledge generation, shown on the right of Figure 2.1, and often forgotten in similar depictions. The crucial question is "*can the model surprise us in a reasonable way?*", as Eric Miller asked the examinee in a PhD defense. A model that cannot surprise the modeler, i.e., increase the knowledge, is of only little use. Due to the overwhelming complexity of simuland and the newness of the microsimulation approach, many microsimulation studies make progress in a verification rather than a validation perspective. In other words, modelers are often already satisfied if models can be told something rather than modelers being told something by the model (see also Chapter 9).

Finally, another important requirement for modeling is generalization. Robert Herman, as cited by Mahmassani (1988), asks for more generalization and abstraction in transport planning: "*We may not have had Ohm's law if Ohm was overly concerned with the detailed paths of the electrons and what these electrons were doing before crossing the resistor.*"

Important transport model approaches are presented in Section 2.2.2, calibration, verification and validation are the topic of Section 2.2.3.

2.2.2 Model Types

2.2.2.1 The 4-Step Procedure

Still the main method in transport planning practice and the basic structure of most modern executable planning models (see Figure 2.2) is the 4-step procedure also known as urban transportation planning procedure (UTP) (Bates, 2000, p.17ff), (TRB, 2007, p.2), (McNally, 2000). The 4-step procedure was developed in the 1950ies in the Detroit Area Transportation Study and Chicago Area Transportation Study (CATS). A very detailed history of transport models including its political dimension is presented by Weiner (2008). The UTP belongs to the aggregate, sometimes called

Figure 2.1: Modeling process

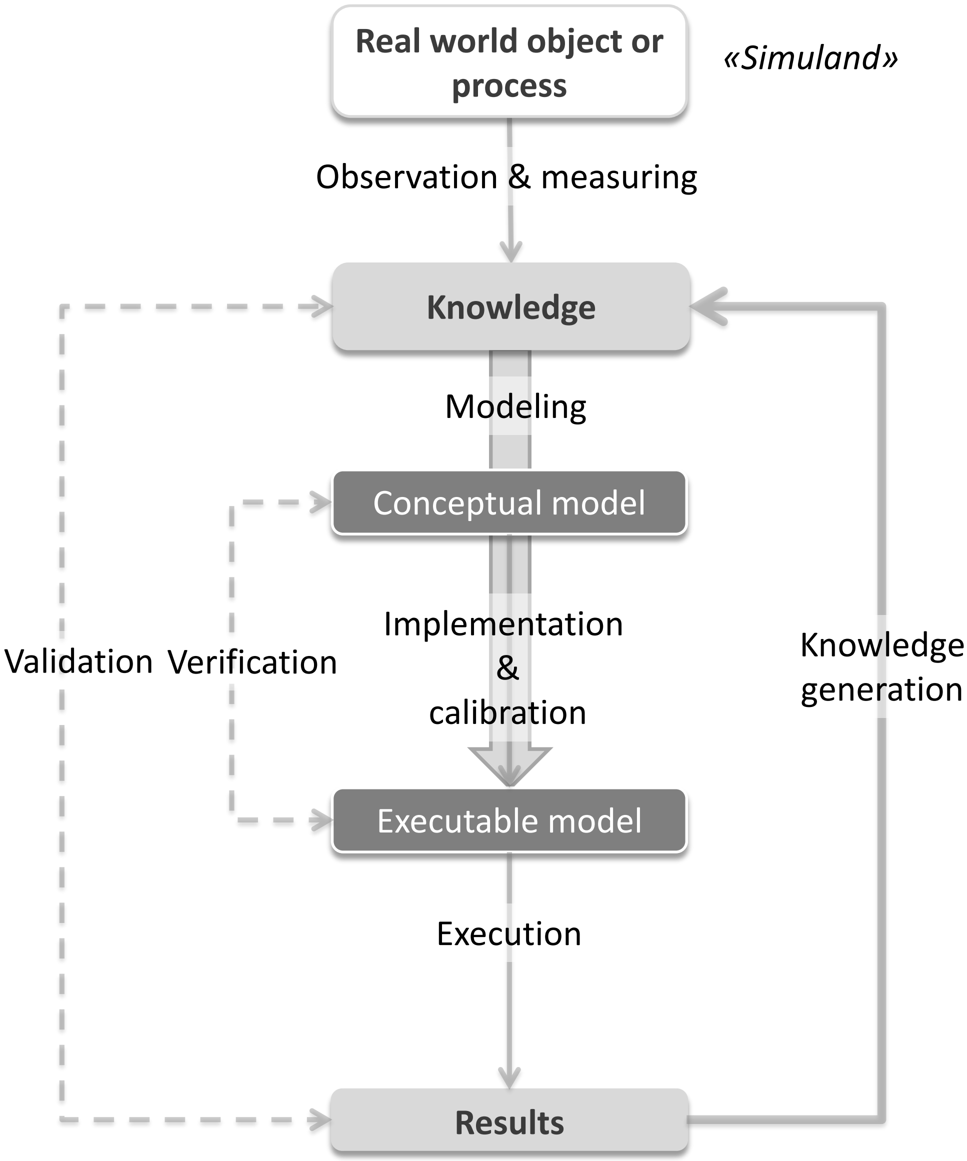
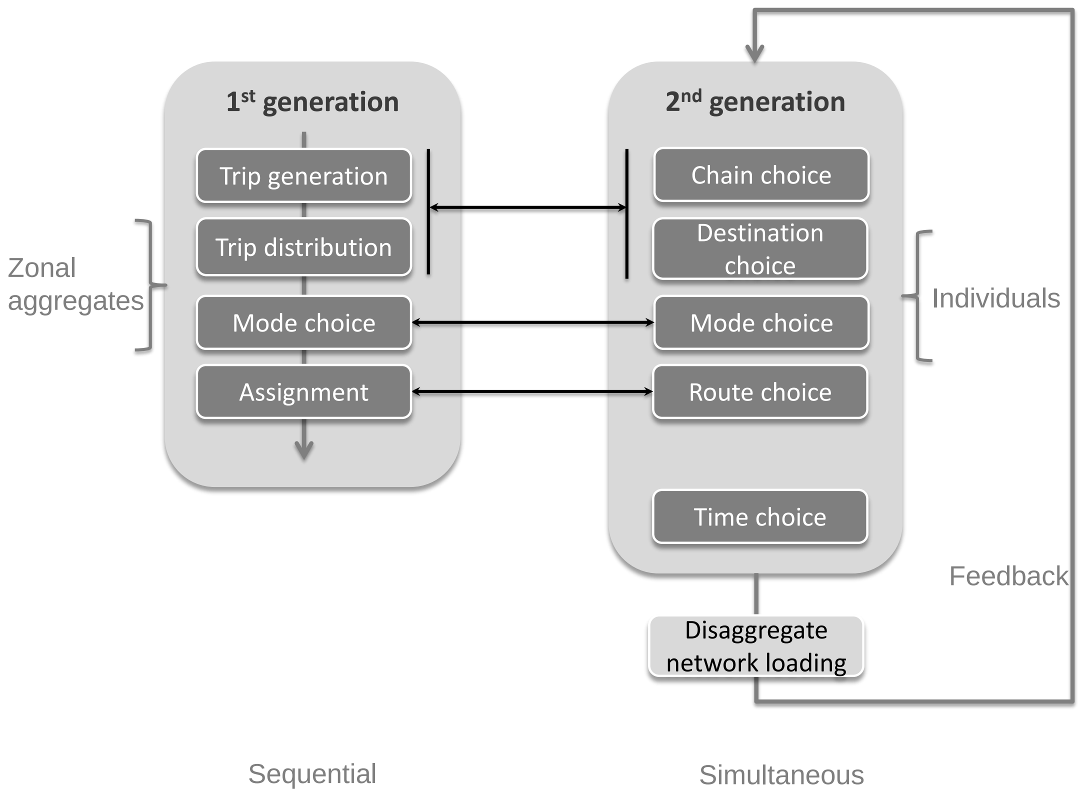


Figure 2.2: Comparison of first and second generation transport models



first-generation methods (Ortúzar and Willumsen, 2001, p.20ff). In the classic model, the demand is assigned to the infrastructure by trip generation and attraction, trip distribution, mode choice and assignment. In the early implementations, feedback was only present within the assignment step but not between steps. Boyce and Williams (2003, p.27) report on the problem of tenuous behavioral basis for the linking of the single steps. The UTP is an efficient and suitable approach for supporting the making of decisions relevant at the time of its invention, which was infrastructure extension (Kitamura, 1996, Section 2) ("car was king" (Daly, 2013)). With the shift from extending to managing the infrastructure, disaggregate, also called second generation, models were proposed.

2.2.2.2 Disaggregate Models and the Activity-Based Approach

Bhat (1998, Slide 4) depicts the evolution of travel demand forecasting techniques, from aggregate trip-based to disaggregate trip-based approaches and then to the activity-based paradigm. "Disaggregate" means that the basic modeling unit is the individual (see e.g., Ben-Akiva (1974); Domencich and McFadden (1975); Ben-Akiva and Lerman (1985), Ortúzar and Willumsen (2001, p.20)). The dominant paradigm of disaggregate modeling since the late 1970s is discrete choice modeling (Boyce and Williams, 2003, p.30).

The improvements of the second generation models over first generation models, listed at many places (see e.g., Axhausen (2006), (Sbayti and Roden, 2010, p.3), McNally and Rindt (2008, Section 2.2), (Bhat and Koppelman, 2003; TRB, 2007)), are summarized as follows.

- *Individual persons instead of zonal aggregates:* This allows to directly and naturally apply the choice models for all dimensions on individual level, not requiring aggregation, potentially making the behavioral basis stronger.
- *Activity chains instead of single independent trips, considering trip chaining*
- *Applying Feedback:* In the basic first generation models, feedback only appears in the fourth stage (e.g., in the method of successive averages). Although some improved models of the first generation already contain feedback of the network conditions to earlier stages, essentially the feedback, spanning the complete process is a feature of the second generation models. Clearly, taking into account feedback strongly improves the results quality.
- *Simultaneous choice instead of sequential procedure*
- *Higher temporal resolution:* Basically the first generation models are static. Some dynamics have been added (e.g., hourly origin-destination matrices) in later models. Nevertheless, this is always associated to usually quite large and fixed time steps. Explicitly including the time choice dimension, as in second generation models, promises a much higher temporal resolution, where, even though the *results* are usually again given as hourly aggregates, the results quality is expected to be higher if aggregation is done at the very end of the modeling.
- *Planning equilibrium instead of Wardrop equilibrium:* If further dimensions than only route choice are taken into account while searching for a user equilibrium, this equilibrium is not necessarily associated to a Wardrop equilibrium (i.e., network equilibrium). Also, the existence and the uniqueness of a user equilibrium is not assured anymore. However, targeting at a planning equilibrium is behaviorally more sound.

The most prominent instance and extension of disaggregate modeling are activity-based models. They are based on the fundamental principle that "*travel demand is derived from activity demand*" (Jones, 1979; Bowman, 2009a,b; Bhat and Koppelman, 2003; Ettema and Timmermans, 1997; Bowman and Ben-Akiva, 1996, 2001).

The roots of this approach are wide-spread. The literature names multiple seminal papers as roots of the activity-based analysis, such as Hägerstrand (1970); Chapin (1974); Fried et al. (1977) or Kreibich (1979) (published in Germany in 1972), cited in Axhausen and Herz (1989) and by Miller (1996, p.165). McNally and Rindt (2008, Section 3) see the linkage of activity and travel participation established already by Mitchell and Rapkin (1954). Despite these numerous early papers, it took very long until operational models were available (Boyce and Williams, 2003, p.31). Bowman (2009a, Figure 2) provides a timeline of US activity-based implementations. Today, many microsimulations are implemented within the activity-based framework.

2.2.2.3 Assignment Methods: Successors of Beckmann et al.

For traffic assignment, non-equilibrium (Matsoukis, 1986) and equilibrium methods (Matsoukis and Michalopoulos, 1986; Patriksson, 1994) exist. A major equilibrium algorithm to solve the static assignment problem is the "Method of Successive Averages" (MSA) rooted on Robbins and Monro (1951) and Blum (1954) [2] . An important issue for the MSA is the re-assignemnt share of link flows, where reaching a Nash equilibrium is guaranteed by decreasing the re-assignemnt share (see, e.g., Powell and Sheffi, 1982; Sheffi, 1985)). Efficiency of the algorithm can be improved by optimizing this share with the Frank-Wolfe algorithm (Frank and Wolfe, 1956).

Besides methods that search for the deterministic user equilibrium, procedures for performing stochastic traffic assignment Dial (see e.g., 1971); Sheffi and Powell (see e.g., 1981); Willumsen (see e.g., 2000); Correa and Stier-Moses (see e.g., 2010), and dynamic traffic assignment (Peeta and Ziliaskopoulos, 2001; Lin et al., 2008; Chiu et al., 2010; Friesz and Bernstein, 2000) were developed.

Besides a plethora of practice models based on the 4-step procedure, there is a strand of mathematical models relatively slowly entering practical models. They can be called the successors of Beckmann et al. (1956) as they apply relatively complex mathematical techniques for computation and qualitative analysis of equilibria, where mathematical programming, optimal control, and variational inequality formulations are dominating (Kinderlehrer and Stampacchia, 1980; Dafermos, 1980; Smith, 1979; Dafermos, 1983; Smith, 1993; Friesz et al., 1993; Smith, 1993; Nagurney,

[2] A clarification of the intricate adoption of the MSA in transport modeling is currently undertaken by Boyce and Williams (forthcoming), where relevant further information can also be found in Smock (1962); LeBlanc (1973); Nguyen (1974); LeBlanc et al. (1975); van Vliet (1977); Sheffi and Powell (1981)).

1993; Friesz, 1996; Nagurney, 2001a; Bierlaire and Crittin, 2006; Lin et al., 2008; Friesz, 2010; Harker and Pang, 1990; Noor et al., 1993). An interesting framework, combining the advantages of both dynamical systems (e.g., Jin, 2005) and variational inequalities, are projected dynamical systems, which are suitable for studying dynamic traffic assignment, and its non-equilibrium states (Nagurney and Zhang, 1996; Nagurney, 2001b; Dupuis and Nagurney, 1993).

A substantial gap exists between these "analytical" approaches (Peeta and Ziliaskopoulos, 2001, p.234), and simulation-based approaches in terms of equilibrium analysis. Although it may be very difficult if not impossible to specify large-scale microsimulation equilibria with specific mathematical terms such "convex, finite, compact, coercive, continuous, montone" etc., we nevertheless stick to Peeta and Ziliaskopoulos (2001, p.243) saying that "*[...] an ability to analyze the system properties even under simplified assumptions can be insightful in generating future directions to address problems*".

2.2.3 Calibration, Verification and Validation

In modeling, *calibration* is the process of adjusting model parameters to increase consistency of model outputs and observed target values (Hollander and Liu, 2007, p.348) (see also Trucano et al., 2006). Hollander and Liu (2007, Table 1) list numerous studies that each calibrate a specific transport microsimulation. Further examples are Smith et al. (2008); Kim et al. (2005); Rutter et al. (2009), microsimulation calibration guidelines are provided by Milam and Chao (2001); Wegmann and Everett (2008); Dowling et al. (2002). Hollander and Liu (2007, Table 2) describe measures of goodness-of-fit, that are productive for calibration. Due to the usually large number of model parameters, an automated process is favorable as far as possible. Essentially this is an optimization process (Hollander and Liu, 2007, p.353), for which various established procedures exist (e.g., Zhang and Ma, 2008, p.41ff). For MATSim, an automatic procedure adapting the plans to road counts was developed by Flötteröd et al. (2008). It is unclear however, if a certain loss of behavioral soundness is caused by adapting plans according to statistical matching. On the other hand, it is unclear anyway, to date, if the MATSim relaxation transitions should be given a behavioral meaning.

Verification is the procedure to test if a "*product is consistent with its specifications [...]*" Petty (2010, p.330). In verification, a perfect match can be achieved comparing the conceptual and the executable model (see Figure 2.1) in contrast to validation, where the model is always an approximation to reality (Kleijnen, 1995, p.145). According to Petty (2010, p.331), "*validation is the process of determining the degree to which the model is an accurate representation of the simuland.*" Validation is difficult to standardize due to the variety of models and model purposes. Some measures, tests, and applications relevant to transport modeling are given by Milam and Chao (2001, Table 2), Lima & Associates (2006), Kurth et al. (2006, p.155), Pendyala and Bhat (2006, p.157), Wegmann and Everett (2008, p.8), Milam and Chao (2001); Roorda et al. (2008); Hawas and Hameed (2009); Sadek et al. (2003); Goulias and Kitamura (1992), Cambridge Systematics, Inc. (2008, p.25), Kleijnen (1995, p.145) (see also David (2009), Sbayti and Roden (2010, p.56), Schiffer and Rossi (2009)). While for the 4-step procedure some validation standards have emerged (e.g., Barton-Aschman Associates, Inc.and Cambridge Systematics, Inc., 1997), a lack of standardization exists for activity-based models. Pendyala and Bhat (2006) say that "*despite the appeal of these models,*" [activity- and tour-based travel demand modeling systems] "*their widespread implementation appears to be hindered by the absence of a detailed validation and assessment of this new wave of model systems. Many MPOs will not adopt such models until they are tested.*" Kurth et al. (2006) cites a statement made by Chandra Bhat and Frank Koppelman in a DRCOG e-mail discussion: "*Researchers and practitioners have not thought carefully enough about the criteria for validation of models. Researchers have the habit of asking practitioners to believe that activity- based methods will produce better impact assessment and forecasts because such models more appropriately represent the actual decision process (we plead guilty to this charge). There is a good basis for this line of thought, but researchers need to go beyond this argument. They need to develop clear validation criteria and demonstrate the value of activity-based methods in ways that are easily understood.*"

Often neglected, but important, is performing sensitivity analysis (sometimes dubbed "what-if analysis" (Kleijnen, 1995, p.155)) (Kurth et al., 2006; Cambridge Systematics, Inc., 2008; TRB, 2007). Sensitivity analysis is similar to assessing elasticity of a variable (Wegmann and Everett, 2008, p.3f) and it tests reaction of the model to changed parameters including model input. This includes both testing the range of parameters for a given point in time, and analysis of the system's fore- and backcasting abilities (e.g., TRB, 2007, p.56), (Cambridge Systematics, Inc., 2008). As forecasting is a vital objective of most transport models, this test is crucial. Pendyala and Bhat (2006, p.158) puts it succinctly: "*There is no doubt that any model can be adjusted, refined, tweaked, and—if all else fails—hammered to replicate base-year conditions.*" and concludes that "*the quality of a travel demand model system is better judged on its ability to respond to a range of scenarios and policies of interest.*" In MATSim, a natural and interesting sensitivity test would be to compare the MATSim forecasts with the current actual state of Zurich network after addition of the bypass "Westumfahrung" in 2009 (Balmer et al., 2009; Baudirektion Kanton Zurich, 2008).

As mentioned above, models are in general flexible enough to be calibrated to target data. Thus, validation *must* be performed using a different data set than for preceding modeling steps (Cambridge Systematics, Inc., 2008, p.1), (TRB, 2007, p.56), (Ortúzar and Willumsen, 2001, p.18). In statistics, this is called cross-validation. It is particularly important for forecasting models, which need to be general enough to capture temporal changes. Calibration and validation should thus be strictly separated, however, in microsimulation practice, according to the author's opinion, they are (too) often mixed, sometimes due to the vast amount of data required for model implementation and calibration. In MATSim, for example, after model calibration only road count data is left for validation (Horni et al., 2009b). New data sources, such as road speed analyses based on GPS (Hackney et al., 2007), should be included.

Having said that, validation of a large-scale transport simulation is very difficult. Many central and comfortable characteristics of systems known from natural sciences are only seldom available for the social science, such as path-independence, decomposability, isolation, and on top of that repeatability of experiments. As a result, there is still a debate if social science actually can provide something similar as laws. Abel (1976, p.107ff) lists and discusses the 12 claims of the "*Verstehen Position*"; although, he finds contrary arguments to every claim, nevertheless, something definitely remains true, making social science model validation exceptionally difficult. For microsimulation results interpretation and model validation, it helped me to visualize the following example. A microsimulation forecast (or backcast) regarding the construction of the "Westumfahrung Zürich" provides a probability distribution of scenarios, and it is essentially an exercise in Monte Carlo sampling. 4 years later we have exactly one actual state, and there is no way to assess the forecasted (or backcasted) probability distribution beyond checking that this actual state is contained in the probability distribution, and hopefully with high probability. There is nothing like Monte Carlo sampling when it comes to aggregate real system states. In other words the existing state is *unique*. In essence, we thus compare an observed Dirac impulse with a computed probability distribution, which is a difficult undertaking.

2.3 Transport Microsimulations

Microsimulation is the modeling of the temporal development of a real-world system or process by explicitly considering the interactions of micro units such as individuals or vehicles. For concise definitions and further information see e.g., Miller (1996, Section 2) or Banks (2001, p.3), Bossel (2004) or Orcutt (1957), who is often referred to as the inventor of microsimulation.

According to this definition, only the program components that model transition processes should be termed microsimulation. Strictly speaking, in MATSim for example, only the network loading simulation is actually a simulation. However, the delineation is difficult and, thus, in the transport planning community, the term *microsimulation* is ambiguously used. Sometimes it actually denotes only the simulation of traveling persons and vehicles in the assignment step—as a replacement of volume-delay functions in aggregate models (see e.g., Nagel and Barrett, 1997, p.508). More often, it additionally includes the preceding choice processes (Kitamura, 1996; Liu et al., 2006). In this sense, microsimulation includes the activity-based demand modeling part and the dynamic traffic assignment (for a detailed discussion of combination of these two parts see Balmer (2007, p.10ff)).

Microsimulations are a consistent implementation of the disaggregate paradigm and offer a variety of benefits as for example listed by Miller (1996); Vovsha et al. (2002); Nagel and Barrett (1997); Bonabeau (2002a); Charypar et al. (2007). In our opinion, the most important ones are, first, the high precision in computing the network loading (Nagel and Barrett, 1997, p.508/524), second, the reproduction of complex interactions (as occurring for example for parking traffic) and, thus, the appropriateness for capturing emergent phenomena (Bonabeau, 2002a), and, third, the conceptual consistency by using the individuals as simulation units throughout the complete modeling process. Most prominent alternative to microsimulations are probably structural equations (e.g., Kitamura, 1996; McNally and Rindt, 2008).

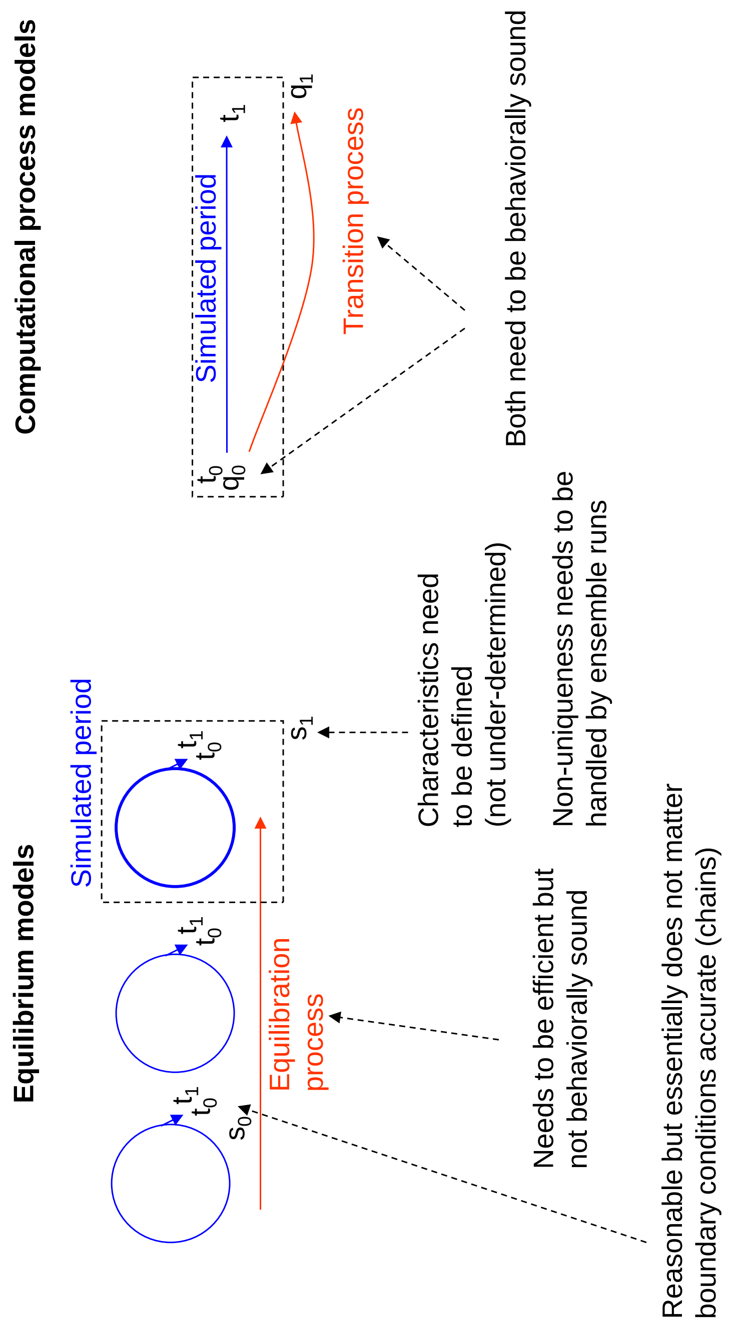
Clearly, the high model resolution and sensitivity comes at a price. Microsimulations are highly demanding in terms of data, as, naturally, they need to have the equally high resolution. Large-scale microsimulation scenarios are thus associated to financial, privacy, but also methodological issues. Falling back to disaggregation procedures applied on aggregate data, such as Balmer and Rieser (2004), naturally, strongly reduces the resolution. However, many correlations are still consistently captured in the day plans. The complexity of large-scale microsimulators requires a broad knowledge of the specific microsimulation interna for calibration, verification, validation, and finally application (e.g., Smith et al., 2008, p.273). Furthermore, with the currently given data base for large-scale scenarios, the conceptual advantages of microsimulations over traditional aggregate models are difficult to show (Lemp et al., 2007).

A common distinction of microsimulations is between utility-maximizing equilibrium-based models and computational process models (Hunt, 2006; Arentze et al., 2001). As shown in Figure 2.3, the paradigms differ as follows. Computational process models concentrate on the transition process leading from a reasonable starting situation to an essentially unknown outcome or end situation. This transition process is thereby made as behaviorally sound as possible. Equilibrium-based models claim to know the structure of the outcome, namely a demand-supply equilibrium. Start point and transition process, or in this case equilibration process, are only relevant in terms of efficiency but not with regard to content. Consequently, equilibrium models are inherently iterative, where computational process models are based on sequential procedures. For adequate handling of randomness, both paradigms need to be based on ensemble runs.

Microsimulations come at different levels of detail, lying between mesoscopic and submicroscopic simulations and ranging from complex car following models to cellular automaton to efficient but relatively rough queue-based models (as used in Gawron (1998)) (for detailed reviews on traffic flow models see also Hoogendoorn and Bovy (2001), Darbha et al. (2008)). In MATSim, a queue-based, event-based mobility simulation is employed per default (Charypar et al., 2007). Amongst other operational models, microsimulation examples are shown in Section 3.3.

Various microsimulations, among them MATSim, adopt the multi-agent approach (p.172ff Gilbert and Troitzsch, 2005), (Bonabeau, 2002b; Sanford Bernhardt, 2007; Sun, 2006), being a subfield in artificial intelligence research. In the words of Bonabeau (2002a, p.7280) agent-based modeling is the "*canonical approach to modeling emergent phenomena [...]*". An agent according to Wooldridge (2009, p.21) "*is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its delegated objectives.*".

Figure 2.3: Comparison of equilibrium and computational process models (bounded rationality)



2.3.1 Basic Procedure of Equilibrium-Based Microsimulations

The basic procedure of utility-based (econometric) microsimulations, such as MATSim or TRANSIMS is depicted in Figure 2.4. A comprehensive (discrete) choice model is applied to a population for a specific choice situation. The choices are forwarded to a network load simulation (sometimes also called the physical simulation or mobility simulation). This network simulation takes into account constraints, such as network capacities. Generalized travel costs, calculated in the simulation, are fed back to the choice model. The choice model is also subject to constraints such as opening hours. The microsimulation is instantiated by census data for the population, travel surveys to estimate the models and infrastructure information to define the constraints. This instantiation or application is described for the MATSim Zurich scenario by Horni et al. (2011e). In a very general sense, utility-based transport microsimulations do utility-maximization subject to constraints following the discrete choice methodology (McFadden, 1978).

The cycle in the middle of Figure 2.4 represents a systematic relaxation process (e.g., Balmer, 2007, Figure 1.3). In MATSim, the interpretation of the relaxation procedure is unclear. Sometimes the relaxation process is ascribed a behavioral interpretation, for example, day-to-day learning, where also the transition process and not only the final equilibrium has a meaning (Liu et al., 2006, p.128), (Nagel and Barrett, 1997, p.523). An opposite perspective exists, where the relaxation procedure is just a numerical method to compute the Nash equilibrium without behavioral basis of the transitions.

To reveal similarities with known mathematical problems and their solution approaches a more abstract formulation can be established.

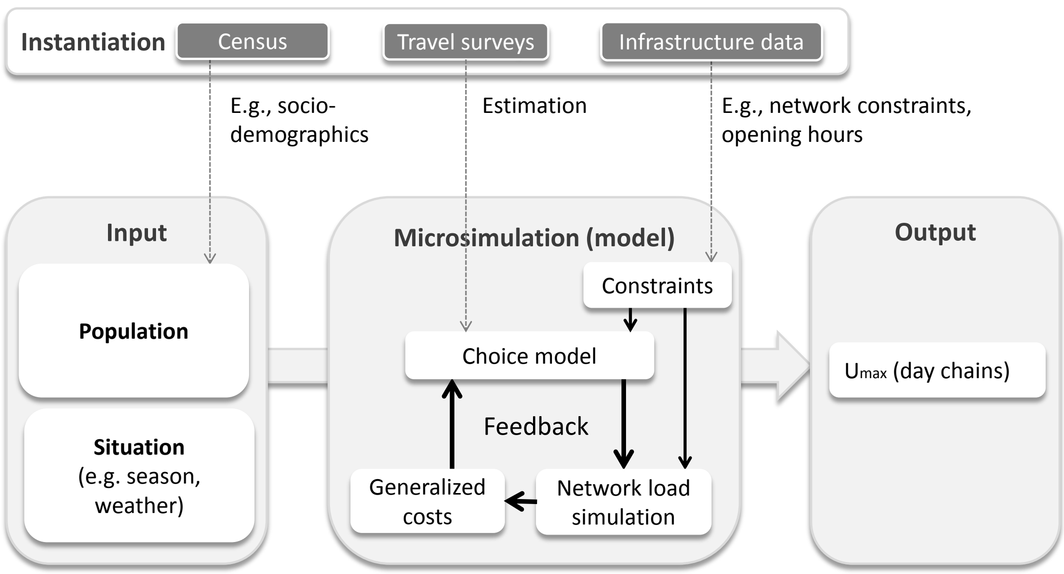
The choice model draw is given as follows.

$$\theta = h_0(\beta, y, \epsilon, \psi), \tag{2.1}$$

where θ are the choices, h_0 is the choice model function, β are the choice model coefficients, ψ are the generalized costs given by the infrastructure conditions (e.g., network conditions), y are further parameters such as the age of the decision maker and ϵ denotes the random error terms.

Utility-based transport microsimulations go beyond a single draw from a choice model. The parameters, such as the travel times, are an *endogenous* component of these microsimulations. The circular relation between choices and the generalized costs can be written as a (usually

Figure 2.4: Utility maximization subject to constraints as the basic principle of transport microsimulations



non-linear) system of equations:

$$\begin{cases} \theta = h_0(\beta, y, \epsilon, \psi) \\ \psi = h_1(\theta) \end{cases} \tag{2.2}$$

where θ are the choices, ψ are the generalized costs and h_0 and h_1 are mathematical maps. In microsimulations h_0 is implemented by the choice models and h_1 is the network loading simulation plus the succeeding conversion of infrastructure conditions into generalized costs.

This can be rewritten as

$$\theta = h_0(\beta, y, \epsilon, h_1(\theta)) \tag{2.3}$$

or in a more general form:

$$\theta = \varphi(\theta, \beta, y, \epsilon) \tag{2.4}$$

This is a fixed point problem (see e.g., Ramadurai and Ukkusuri (2008, p.6), Bierlaire and Crittin (2006); Kaufman et al. (1998); Ramadurai and Ukkusuri (2010)). The fixed points are found by iteratively applying Equation 2.4. In the microsimulation context, *iteratively*, the outcomes of the comprehensive choice model are directed to a network load simulation, whose outcomes (the infrastructure conditions) are in turn fed back to the choice model. But naively doing this most probably leads to very bad convergence behavior.

From numerics it is known that fixed point problems can be transformed such that convergence behavior is improved. This is explained with an example. In numerics root finding problems given as $f(x) = 0$ are often transformed into fixed point problems as

$$x = g(x) \tag{2.5}$$

where the fixed points are the solutions of the root finding problem. As fixed point problems can be transformed, infinitely many possibilities exist for the choice of $g(.)$. For example the root finding problem $f(x) = x^2 - 2x - 3 = 0$ can be transformed by using $g_0(x) = \sqrt{2x + 3}$ or $g_1(x) = \frac{3}{x-2}$ or $g_2(x) = \frac{x^2-3}{2}$ to name a few. The choice of $g(.)$ thereby has a crucial impact on the convergence behavior, i.e., dependent on the form of $g(.)$ and the initial point θ_0 the iterations may converge to a fixed point with different speed, are attracted by an orbit or may also move through the state space in a completely chaotic manner. In the example above, starting from $x_0 = 4$, $g_0(x)$ converges, $g_1(x)$ also converges but slower and $g_2(x)$

diverges. In numerics, the Picard method (Vogt, 2001, p.2ff) is improved by e.g., the Newton-Raphson method (Vogt, 2001, p.28ff), which employs an efficient mapping for $g(.)$.

A very similar approach is maybe productive for microsimulations. In other words, $\varphi(.)$ (which corresponds to $g(.)$) should be chosen reasonable, such that fixed points are found efficiently. For first generation models, much is known about existence, stability, uniqueness and computation of fixed points. Thus, by continuing the lines of the first generation models, one can hope to produce something reasonable also for second generation models, but this has to be investigated further. In the next section, a possible implementation of $\varphi(.)$ is presented for MATSim, which, in practice, has shown to efficiently lead to behaviorally sound fixed points.

2.4 The Multi-Agent Transport Simulation MATSim

2.4.1 The Basics

The development of the multi-agent transport simulation MATSim (MATSim, 2013; Balmer et al., 2006) has started approximately a decade ago as a collaborative effort of Prof. Nagel (now: TU Berlin) and Prof. Axhausen (ETH Zurich). The roots of MATSim lie in the transport simulation TRANSIMS (Raney et al., 2002), which was developed by Prof. Nagel as research team leader at the Los Alamos National Laboratory, and they lie in Axhausen (1988) as well. MATSim has been applied by local research groups world-wide for different regions (MATSim-T, 2013), such as Berlin (Balmer, 2007, p.67ff), Switzerland (Meister et al., 2010) (with Zurich as a more detailed sub-model (Balmer et al., 2009)), Singapore (Erath et al., 2012), Toronto (Gao et al., 2010), Gauteng (Joubert et al., 2010), Tel Aviv (Bekhor et al., 2011), Shanghai (Wang et al., 2013), and Padang (Lämmel, 2011).

MATSim is an activity-based, extendable, multi-agent simulation toolkit implemented in JAVA. It is open-source and can be downloaded freely (MATSim, 2013; SourceForge, 2013). The framework is especially designed for large-scale scenarios, meaning that, the features of all models are generally stripped down to efficiently handle the base functionality, where emphasis has been also been laid on parallelization. For the network loading simulation, for example, a queue-based model is implemented, leaving out the very complex car-following behavior. Due to the modular approach, modules such as the network loading can easily be replaced.

MATSim is based on a co-evolutionary principle. While being in a competition for space-time slots on the transportation infrastructure with all the other agents, every agent iteratively optimizes its daily activity chain. This is done by running through the MATSim loop as depicted on the left in Figure 2.6.

Every agent possesses a memory of a fixed number of day plans, where each plan is composed of a daily activity chain and an associated utility value (in MATSim called *plan score*). For now, MATSim is conceptually designed to model a *single day*, a common unit of analysis for activity-based models (see, for example, the review in Bowman (2009a)). In other words, basically, MATSim is a cross-sectional model. Nevertheless, in principle a longitudinal model could be implemented (e.g., Horni and Axhausen, 2012b).

In every iteration, prior to the simulation of the network loading (e.g., Cetin, 2005), every agent selects a plan from its memory. This selection is dependent on the plan utility. A certain share of the agents (often 10%) is allowed to clone the selected plan and modify this clone. For the method of successive averages (MSA) usually a decreasing share of travelers is reallocated to a new route to avoid oscillations. For MATSim, it has been shown that a variable replanning share can be productive as well and *"increase overall performance of the system by a factor of three or more"* (Charypar et al., 2006, p.7f). For the network load microsimulation step multiple simulations are available and configurable (Horni et al., 2011e, p.10f).

Plan modification is implemented in the *replanning* modules. Four choice dimensions are considered for now: time choice (Balmer et al., 2005), route choice (Lefebvre and Balmer, 2007), mode choice, and destination choice. If an agent ends up with too many plans (configurable), the plan with the lowest score (configurable) is removed from the memory of this agent. The agents which have not undergone replanning select between existing plans. The selection model is configurable; in many MATSim investigations, a model that generates a logit distribution for plan selection is used.

An iteration is completed by evaluating the agent's day described by the selected day plans (termed *scoring*). The basic MATSim utility function was formulated by Charypar and Nagel (2005) from the *Vickrey* model for road congestion as described in Vickrey (1969) and Arnott et al. (1993). Originally, this formulation was established for departure time choice. However, several studies (e.g., Balmer et al., 2009) indicate that the MATSim function is also appropriate for modeling time choice and route choice; it has thus also been adopted as the starting point for destination choice.

The utility of a plan U_{plan} is computed as the sum of all activity utilities $U_{act,q}$ plus the sum of all travel (dis)utilities $U_{trav,q}$:

$$U_{plan} = \sum_{q=1}^{n} U_{act,q} + \sum_{q=2}^{n} U_{trav,q}$$

The utility of an activity q is defined by:

$$U_{act,q} = U_{dur,q} + U_{late.ar,q} + \epsilon \ ,$$

where:

- $U_{dur,q} = \beta_{dur} \cdot t_{typ,q} \cdot \ln(t_{dur,q}/t_{0,q})$ is the utility of performing activity q, where opening times of activity locations are taken into account. $t_{dur,q}$ is performed activity duration, β_{dur} is marginal utility of activity duration for its typical duration $t_{typ,q}$ and $t_{0,q}$ is minimal duration, or in other words, the duration for which utility starts to be positive.
- $U_{late.ar,q} = \beta_{late.ar} \cdot t_{late.ar,q}$ gives the disutility of late arrival, where $\beta_{late.ar}$ is marginal utility of latency and $t_{late.ar,q}$ is latency compared to planned times given in the agent's day plan.
- ϵ is the random error term added in this thesis.

There may also be additional penalties for staying not long enough, departing too early, or (beyond the implicit opportunity cost of time) for waiting. These are not used in this thesis.

Travel disutility is given as

$$U_{trav,q} = \beta_{trav,m,q} \cdot t_{trav,m,q} \ , \tag{2.6}$$

where $\beta_{trav,m,q}$ is marginal utility of travel by mode m and $t_{trav,m}$ gives the mode-dependent travel time between location of activity $q - 1$ and q.

Note that travel receives an additional implicit penalty from the opportunity cost of time: If a travel time could be reduced by Δt_{trav}, the person would not only gain from avoiding $\beta_{trav} \cdot \Delta t_{trav}$, but also from making activities longer.

The iterative process is repeated until the average population score stabilizes, where the definition of the stopping criterion is subject of ongoing research initialized by Meister (2011); Nagel and Flötteröd (2009). Due to numerical problems with the log-form utility function for activity chain choice a new S-shaped function was researched for MATSim (Feil et al., 2009b).

2.4.2 The Underlying Principles of MATSim

MATSim is a utility-maximizing model and, thus, located within the discrete choice framework. Utility in discrete choice models is composed of a deterministic part and a random error term. The random error term represents the unobserved heterogeneity, i.e., it subsumes, both, truly, i.e., inherently random decisions and the modeler's missing knowledge about the choice and its context.

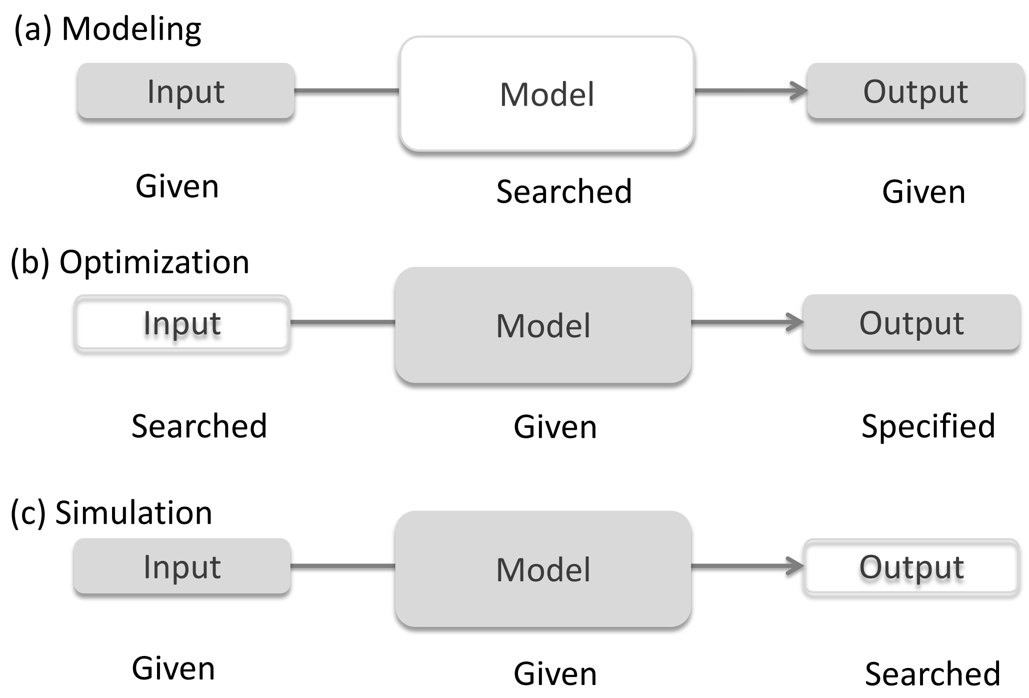
In MATSim, the utility function for route and time choice does not contain a random error term (yet). This can be regarded as a shortcoming of the model. However, this is at least partially compensated through the stochasticity of the replanning. First, route and time choices are usually subject to significant competition. The co-evolutionary algorithm of MATSim, detailed below, essentially assigns the resources in a random manner to the persons. For example, two identical persons may end up with different routes according to the order in which they undergo the replanning. Essentially, this means that a random term is present in the choice modeling. However, this randomness is introduced implicitly and not in a systematic manner.

In other words, choice outcomes do not only depend on implemented choice model $h_0(.)$, but are also implicitly influenced by the implementation of the algorithm to find the solutions of the utility-maximization (denoted as $\varphi(.)$ above). This is difficult to interpret, and, furthermore, replanning did up to now not add enough unobserved heterogeneity to destination choice. Thus, an explicit random error term, held stable over the iterations, is added as shown later.

2.4.2.1 System Analysis

System analysis helps embedding MATSim in modeling theory, providing a communication basis for further model development and application. System analysis generally distinguishes three types of problems, namely, *modeling* (or *system identification*), *optimization* and *simulation* (e.g., Eiben and Smith, 2003, p.8ff). As shown in Figure 2.5, for modeling problems (a) a model for known inputs and outputs is created, whereas for optimization problems (b) the model and some information about the output is known, and solving the problem means finding the input values. Simulations (c) produce outputs for known inputs and a given model. In MATSim, all of the three problem types are present. *Modeling* is done during MATSim model creation and calibration. The choice model parameters are estimated based on surveyed input and output data. *Simulation* is applied in MATSim as infrastructure load simulation.

Figure 2.5: Problem categories of system analysis



Optimization takes place, when individuals increase their individual day plan over the course of the iterations.

Clearly, for most models not the complete range of choice dimensions can be handled endogenously from the beginning, but needs to be fed into the model as input. Successively, during model development, more choice dimensions are included, in Nagel and Axhausen (2001) called "endogenising". On the scale presented in Nagel and Axhausen (2001, Figure 2), MATSim ranges from queue-based traffic flow to demand generation. Time, route, mode and destination choice are endogenously modeled. Activity chains [3] and locations for home and work activities are given as model input. Joint activities and rides, sometimes seen as an additional choice dimension, are researched intensively in the context of social networks and households but not yet part of the MATSim model. Projects combining MATSim with land use models, here with UrbanSim (Nicolai et al., 2011; Nicolai and Nagel, 2011; Schirmer et al., 2011; Waddell, 2010) are underway.

Potential measures of interest (for example population measures versus

[3] Activity chain choice is available in an experimental instance only.

population segment measures), their scale, and necessary number of ensemble runs to reach given confidence are a crucial topic discussed in Chapter 6. In MATSim context, population segmentation for analysis has been done by (Balmer et al., 2009), looking at potential losers and winners of the new Zurich bypass road.

2.4.2.2 Implementation of $\varphi(.)$: A Co-Evolutionary Algorithm

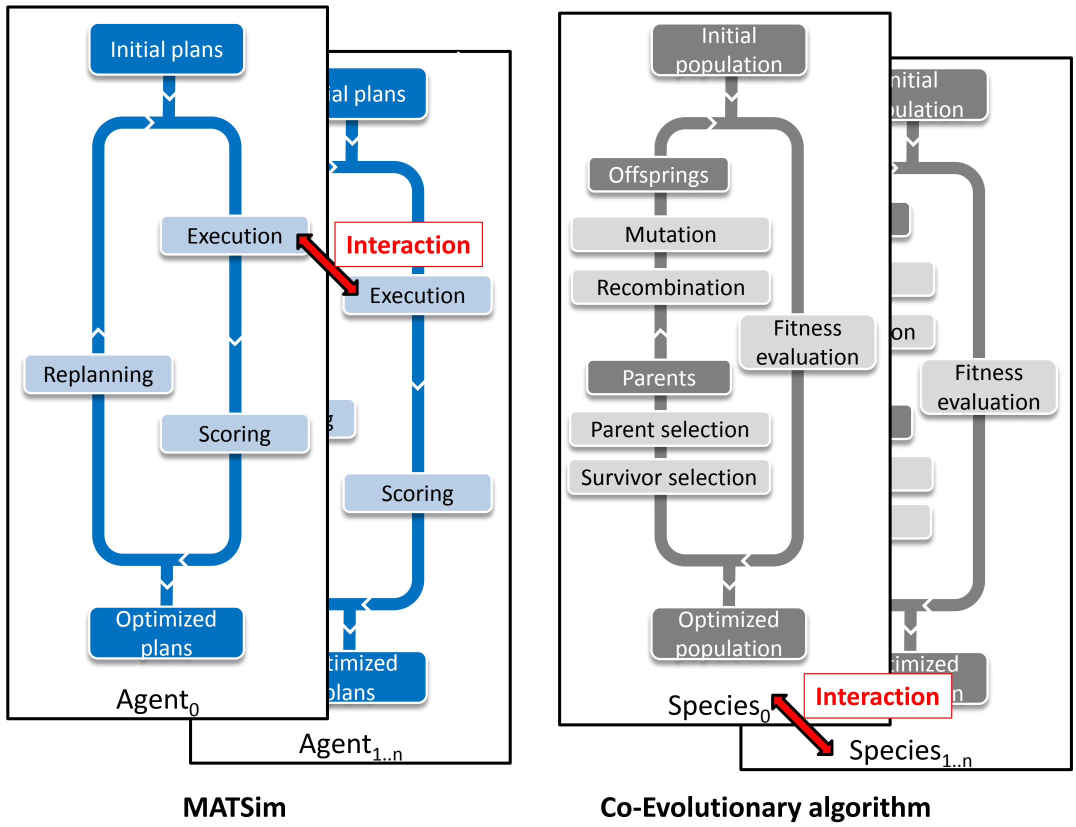
In first generation models, the fixed points are searched by implementing $\varphi(.)$ with iterative numerical algorithms such as the *method of successive averages* (MSA) (Ortúzar and Willumsen, 2001, p.342f) or the Frank-Wolfe algorithm (Frank and Wolfe, 1956), (see also Correa and Stier-Moses, 2010, p.4). Every person searches its optimum subject to competition leading to an equilibrium state. This is achieved in the MSA by moving a certain share of the flow to the cheapest route, while taking it away from the remaining routes per OD-relation. In second generation models, such as microsimulations, the fixed point problem not only encompasses route choices but *all* choice dimensions mentioned above. However, it is still possible to apply similar procedures for φ to find the fixed points.

As illustrated in Figure 2.6, a *co-evolutionary algorithm* is applied in MATSim. In co-evolutionary algorithms, different species co-evolve subject to interaction (e.g., competition). In MATSim, the individuals are represented by the plans of a person, where a person represents a species. By applying the co-evolutionary algorithm, optimization is performed in terms of agents' plans. Eventually, an equilibrium is reached subject to constraints, where the agents cannot further improve their plans unilaterally. When speaking in strict terms, there is a difference between application of an evolutionary algorithm and a *co*-evolutionary algorithm. An evolutionary algorithm would lead to a system optimum as optimization is applied with a global (or population) fitness function. The co-evolutionary algorithm instead leads to a user equilibrium as optimization is performed in terms of *individual* utility functions and within an agent's set of plans. At the moment, the MATSim co-evolutionary algorithm only includes mutation; recombination may come into play when joint day plans of family members, for example, are included in the future.

2.4.2.3 MATSim Equilibrium

Not much is known about the characteristics of the equilibrium found by the MATSim co-evolutionary algorithm. The Nash equilibrium and the

Figure 2.6: Adopting a co-evolutionary algorithm in MATSim



Wardrop equilibrium as its instance in transport planning are formulated for unilateral changes. But, in the co-evolutionary algorithm of MATSim, per iteration multiple persons (a specified share of the population) are allowed to modify their plans.

Due to, first, potential individual variations in utility function parameters, second, the implicit unobserved heterogeneity introduced by time and route choice, and third, the explicit error terms applied for destination choices the travelers in MATSim do perceive the travel costs differently. Thus, the equilibrium searched for in MATSim is rather a stochastic (SUE) than a deterministic user equilibrium (UE). Additionally, it is dynamic (DUE). The classification of the MATSim equilibrium as a *Mixed Strategy Nash Equilibrium* (MSNE) is often discussed in personal communications. This type of equilibrium is based on decision-makers which do not chose one specific alternative but a stable probability distribution over the alternatives.

An argument for the MSNE is as follows. The final state of MATSim with respect to the selected plans in one iteration is never perfectly stable if replanning is turned on and if this replanning is based on random mutation. This is the case for the very common replanning dimension *time choice*. However, this argument is weak as the share of replanners is usually relatively small leading to small intra-run fluctuations. Furthermore, one could specify the result to be based not on the modified selected plan of the current iteration, but on the best plan of each agent. Although the stability of the best plans is not researched extensively, experience tells that for the vast majority of the agents the best plan is stable after having reached the global relaxed state.

Another argument for the MSNE can be, at first sight, the selection mechanism of MATSim. For some configurations, the plans are selected for replanning and execution according to a logit-type probability distribution. But having a closer look, this argument is also weak. The plans in an agents memory converge all toward the best plan over the course of iterations, as with every replanning stage the worst plan is removed.

In conclusion, this means that no probability distribution is involved from which the agents draw in the equilibrium state, meaning that the arguments for an MSNE are weak. Clearly, the results over *multiple runs* with different random numbers represent a distribution, but this is not related to the discussion here.

The development of static flow equilibria to dynamic particle equilibria and finally agent-based equilibria is presented in Nagel (2012); Nagel and Flötteröd (2009), representing an interesting connection point for a methodological discussion of MATSim system characteristics. Another starting point is provided by Meister (2011), who researched stopping criteria in relation to the MATSim user equilibrium.

2.4.2.4 MATSim Ergodicity and Emergence

As mentioned earlier, the driving forces of the transition from non-equilibrium states to an equilibrium and its empirical and behavioral basis are not well understood (Section 2.1.2). If it happened to be the case, that there is no strong driving force, then the maintaining of the assumption of a dominating equilibrium state, would require ergodicity of the system (see also Holden, 1989, p.252), meaning that the system is actually able to reach all states of the state space. MATSim ergodicity is discussed in Flötteröd (2012).

Emergence, the formation of complex patterns generated by interaction of comparatively simple individual units, appear in many systems including transport system. Often, emergent effects are highly significant, such as phantom traffic jams. Consequently, the ability to reproduce them is crucial to modeling these systems. Multi-agent-based simulations are due to their structural similarity to the modeled multi-part systems, expected to be particularly suitable to capture these effects. A glimpse at emergence in MATSim is taken in Horni and Montini (2013a,b).

Chapter 3

Destination Choice Analysis

This chapter gives a short destination choice overview focused on analysis methods, choice determinants (Section 3.1), data availability (Section 3.2) and large-scale operational demand forecasting tools (Section 3.3). Overall goal is defining a parsimonious set of key determinants of shopping (and partly leisure) destination choice and establishing a theoretical basis for MATSim model development described in Section 3.4.

3.1 Destination Choice Research Fields, Methods and Choice Determinants

Shopping and leisure destination choice spans multiple research fields, amongst others, transport and urban planning, marketing and retailing science, economics, geography and psychology. This has led to a huge and heterogeneous body of literature, which makes comprehensive quantitative analyses in terms of single choice determinants difficult. As the large-sale data availability is limited anyway, it was decided to only make a qualitative literature overview, and, hence, to leave out a quantitative meta-analysis, which nevertheless, might be a fruitful future undertaking in its own right.

3.1.1 Methods

Various methods for modeling destination choice behavior have been designed in the numerous research fields. Most common distinction is between aggregate and disaggregate models (see e.g., Innes et al. (1990, p.127)). A popular example for an early aggregate model is the gravity model, derived from Newton's law of gravity, and applied for retailing (Reilly, 1931; Converse, 1949) and later for trip distribution in the 4-step model (Casey, 1955). The model in general is composed of an attracting term (analogously to mass in Newton's model) and an impedance term (the distance in Newton's model). For the base model and its derivatives (Huff, 1960, 1963), the attracting term is often operationalized with some function of population size, or infrastructure supply, where for the deterring term some function of distance or travel time is usually applied.

While aggregate models in average efficiently capture the main influences, clearly, at the same time, the zonal averaging or smoothing is a drawback as relevant peaks potentially are averaged out (see also Innes et al. (1990, p.127) and the discussion on aggregates in Section 6.1.1). The dominating method in transport planning is probably discrete choice theory (McFadden, 1978; Horowitz, 1985). In this framework, decisions are modeled as a utility maximizing choice from a finite set of alternatives, the choice set. The method has shown to be productive and has thus also been broadly applied in operational planning models. MATSim is based on utility maximization and thus strongly related to discrete choice theory.

Furthermore, a plethora of higher-resolution methods have been developed for, or applied on, the destination choice problem; examples are factor analysis (Koppelman and Hauser, 1978; Recker and Kostyniuk, 1978), principal component analysis (Ibrahim, 2002), discriminant analysis (Innes et al., 1990, p.128), neural networks (Davies et al., 2001), decision trees (CART) (Arentze and Timmermans, 2005), Bayesian probability theory (Burnett, 1977), computational process models (Gärling et al., 1994, p.356), the repertory grid method (Timmermans et al., 1982), information integration approaches (Timmermans, 2008), household production approaches (Odland, 1981), hazard models (Popkowski Leszczyc et al., 2000), structural equations (Prayag, 2009), the polythetic-division method (Uncles, 1996), and control theory (Venter and Hansen, 1998). Comparisons of different methods and methodological reviews are given by Barnard (1987); Recker and Schuler (1981); Berry et al. (1962).

3.1.2 Choice Determinants

As mentioned in Section 1.2, this thesis is primarily focused on shopping trips as it can be reasonably assumed that they are associated with less unobserved or even unobservable heterogeneity. The idea is to blow a breach for detailed activity type handling, i.e., utility function adaptation. Leisure trips are in the operational model handled for sake of completeness as another type of discretionary activity.

Studies, investigating numerous choice determinants, exist in great quantities. Examples investigating large attributes sets are Ibrahim (2002, list on p.281/282), Koppelman and Hauser (1978, list on p.160), Timmermans et al. (1982, table 1 on p.194), McCarthy (1980, table 1 on p. 1270), Oppewal et al. (1997, table 1 on p.1074 and figure 2 on p.1081), Recker and Kostyniuk (1978, table 1 on p.22), Simma et al. (2004, table 4 on p.18), Batt (2009, table 3), Colome and Serra (2000, table 1 on p.25), Erath (2005, table 6 on p.20). Important, frequently analyzed factors are the following.

- **Trip attributes:**
  - **Travel time, distance and cost**: Access, usually operationalized by travel time, distance and cost, is a cornerstone of transport planning and, thus, a key choice determinant in all models. Travel time is the base rationale for any kind of traffic assignment.
    Quantitatively, however, the situation is less clear. While many studies report strong influence of these factors on destination choice Brunner and Mason (e.g., 1968); Recker and Kostyniuk (e.g., 1978), others only find relatively small effect of access (Lademann, 2007, p.154/155) and (Innes et al., 1990, p.135f). Timmermans (1983, 1980, p.449f) argues that travel distance should not enter the utility function but should rather be included as a constraint for the choice set. Furthermore, values for travel time savings vary substantially, even when adjusted for purchasing power. For discrete choice models, these variations, representing an inconsistency, might stem, at least partly, from unsolved choice set definition issues as detailed in Section 5.2.
  - **Trip chaining and multi-stop, multi-purpose shopping trips:** Bernardin et al. (2009); Delleart et al. (1998, 1997); Kitamura (1984); O'Kelly (1983); Cirillo et al. (2003).
  - **Mode:** Ibrahim (2002); McCarthy (1980); Timmermans (1996); Yang et al. (2009); Handy and Clifton (2001); Ibrahim (2002).
- **Store attributes:**
  - **Price level:** Bell et al. (1998); Chamhuri and Batt (2009).
  - **Store size:** Lademann (2007); O'Kelly (1983); Hubbard (1978).
  - **Opening hours:** Innes et al. (1990).
  - **Quality of goods and store:** Innes et al. (1990); Batt (2009); Chamhuri and Batt (2009).
  - **Product range or selection:** (Oppewal et al., 1997).

  Furthermore, interactions at the activity locations, for example, due to parking, influence destination choices (Axhausen, 2006, p.3), (Innes et al., 1990; Ibrahim, 2002; van der Waerden et al., 1998; Timmermans and van der Waerden, 1992). Agglomeration effects are another driving force of destination choices, in particular, for multi-stop, multi-purpose shopping trips (Bernardin et al., 2009; Teller and Reutterer, 2008; Timmermans et al., 1992).

- **Person attributes:** Traditionally, socio-demographics and socio-economics are used in transport models as they are broadly available in nation-wide censi Uncles (see also 1996); Shim and Eastlick (see also 1998); Krumme et al. (see also 2010). Usually, age, sex, residential location and sometimes household structure are surveyed.
  Clearly, these traditional person characteristics are of limited explicative power. Other classifications are thus researched, such as lifestyles (Salomon and Ben-Akiva, 1983), attitudinal characteristics (Recker and Kostyniuk, 1978), psychometric scales (Rieser-Schüssler and Axhausen, 2012)[1] , personal values, attitudes or ethnicity (Shim and Eastlick, 1998, p.142).

Besides these usual suspects, many further choice determinants and mechanisms have been researched. Temporal aspects, such as rhythms or routines, but also timing within a day are treated (amongst other topics) by Kahn and Schmittlein (1989); Kim and Park (1997); Ehreke (2008); Kitamura et al. (1998); Landau et al. (1982a); Krumme et al. (2010). Influence of similarities of alternatives are handled by Bekhor and Prashker (2008); Schüssler (2006). Store and area image is a topic in Hong et al. (2006); Bell (1999). Customer loyalty as a kind of hysteresis is examined by Rhee and Bell (2002); Landsverk et al. (2003); Innes et al. (1990); East et al. (1998); Knox and Denison (2000). Sands et al. (2009) look at effects of in-store events. The influence of very large retail facilities on destination choice and travel behavior in general is the subject of Buliung et al. (2007); Buliung and Hernández (2009). Influence of specific choice context is researched in van Kenhove et al. (1999). Relations between brand and store choice are investigated by Baltas and Papastathopoulou (2003). Household inventory costs contrasted with travel costs are analyzed by Bawa and Ghosh (1999). Pedestrian behavior especially route choice as a response to transport and retailing measures are simulated in Borgers and Timmermans (1986). Following a bounded rationality approach Cadwallader (1975) replace objective travel distance with perceived (or cognitive) distance. This model can be tested for Switzerland, as the Swiss Microcensus also offers estimated or remembered reported

[1] The cited study does not focus on destination choice.

travel times. Cadwallader (1995) present a meta-analysis of attribute interactions. A view on a broader range of consumer decision components is taken by Solomon (2009); Underhill (1999, 2004). A specific focus on supermarkets is taken in Orgel (1997). Destination choice studies focused on Switzerland are Carrasco (2008); Kawasaki and Axhausen (2009); Horni et al. (2012c); Simma et al. (2004); Eggenberger (2001); Erath et al. (2007); Erath (2005).

Leisure choice determinants are investigated by Simma et al. (2002); Schlich et al. (2002, 2003a,b); Schlich and Axhausen (2003b); Schlich et al. (2004); Zängler (2000); Institut für Mobilitätsforschung (2000); Hautzinger (2003); Pozsgay and Bhat (2001); Kemperman et al. (2002); Aldskogius (1977); Stauffacher et al. (2005); Pozsgay and Bhat (2001); van Middelkoop et al. (2004). Clearly, the boundary between shopping and leisure is fuzzy, Rajagopal (2006) for example, analyze recreational shopping trips. Social interaction is an integral part of leisure activities, thus, modeling should include social networks as researched, for example, by Hackney (2009); Marchal and Nagel (2005b,a); Arentze et al. (2011); Frei and Axhausen (2007); Kowald and Axhausen (2012); Illenberger et al. (2010); Axhausen (2012, 2007); Frei et al. (2009); Hackney and Axhausen (2006); Carrasco et al. (2008).

3.2 Data Availability

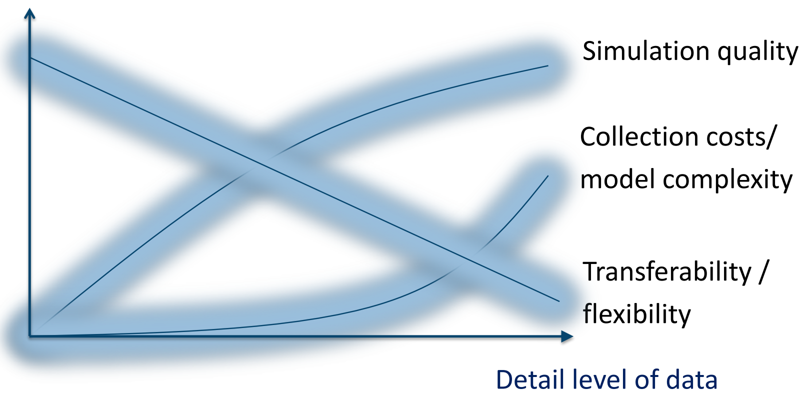
The main driving force for selecting model variables for application, naturally, is data availability.

For a flexible and general model (Patriksson, 1994, p.5) it is important that the main components, are either transferable, or that similar data is broadly available. Practically, this means that models should be applicable without extensive data collection and model estimation efforts [2]. In more detail, as illustrated in Figure 3.1, simulation quality, model complexity and data collection costs increase with increasing detailedness of employed data. In contrast, model transferability decreases with increasing level of detail of used data. This is irrelevant for a study performed for a specific region. But it is highly important, for designing a generalizable microsimulation module. It must be flexible enough to cope also with low levels of data detailedness.

On the other hand, higher-resolution methods, principally, offer finer analysis levels. One can assume, that they also offer higher sensitivity to parameters and input data, at least this is often pointed out as an

[2] For validation data, in contrast, more is always better (see e.g., McNally and Rindt (2008, p.7)).

Figure 3.1: Level of detail in model data



advance of these models (see e.g., Sbayti and Roden (2010, p.4) or Lemp et al. (2007). Due to this higher sensitivity, a sparse data base might annihilate the conceptual advantages of microsimulation models (Miller, 1996, Section 3).

These two antagonistic arguments imply a trade-off for data-collection, that urgently needs to be taken into account when developing a microsimulation or one of its modules.

Discussions of data requirements and collection methods are given by Axhausen (1997); Kitamura (1996), or Nagel and Axhausen (2001, p.6).

According to the authors' experience and personal communication, the Swiss data base mainly provided by BfS (2011) is comfortable; it can thus be interpreted as an upper bound for nation-wide data availability.

The main Swiss data sets used in MATSim are:

- Census of Population (Swiss Federal Statistical Office (BFS), 2000), a full survey, applied to create the MATSim population, including their home and work locations on hectare and municipality level respectively,
- National Travel Survey (Swiss Federal Statistical Office (BFS), 2006), a 30'000 person sample, used for MATSim demand creation (activity chains and times). Recently, the year 2010 was added (Swiss Federal Statistical Office (BFS), 2012).
- Business Census (Swiss Federal Statistical Office (BFS), 2001,

2008a), identifying enterprises at hectare level, utilized for creation of activity locations,

- Network Data, (TomTom MultiNet, 2011; NAVTEQ, 2011; Vrtic et al., 2005), spanning navigation and planning networks,
- Road Counts, (e.g., ASTRA, 2006), specifying hourly traffic volumes per lane, mainly applied for MATSim validation.

Data covering only Zurich region concern parking supply, signal programs (Balmer et al., 2009), store service hours (Meister, 2008) and public transport lines and schedules (e.g., Rieser, 2010, p.70ff).

3.3 Operational and Large-Scale Models

A large number of large-scale operational forecasting models exist; for reviews and summaries see Algers et al. (1998); Henson and Goulias (2006); Gärling et al. (1994); Jovicic (2001); Bowman (2009a,b); Burmeister et al. (1997); Axhausen and Herz (1989); Timmermans (2001). A lot of the state-of-practice models are created by US Planning Organizations (MPOs) (see e.g., Bradley and Bowman (2006)), where a large part of these tools is still based on the 4-step procedure (TRB, 2007, p.2). On search for inspiration for destination choice, one must look at the trip distribution step and to a certain extent to the trip attraction sub-step of the first step.

A huge number of models exist. As an extensive list is not available in the literature, it is started here, possibly providing an initial point for future literature reviews and consolidating work. Adler and Ben-Akiva model (Adler and Ben-Akiva, 1979, 1976), Aiumsun (AIMSUN, 2013), Alam-PSEM (Alam and Goulias, 1999), ALBATROSS (Arentze and Timmermans, 2007, 2000; Arentze et al., 2000), AMADEUS (Timmermans et al., 2002, 2000), AMOS (Pendyala et al., 1995, 1997; Kitamura and Fujii, 1998), AURORA (Joh et al., 2004; Timmermans et al., 2001), Axhausen model Axhausen (1988), Berg et al. model (Berg et al., 1976), Bowman and Ben-Akiva (Bowman and Ben-Akiva, 2001), CARLA (Jones et al., 1983), CATGW (Bhat and Singh, 2000), CEMDAP (Bhat et al., 2004; Pinjari et al., 2006), CentreSim (Kuhnau and Goulias, 2002), Cobra (Wang and Timmermans, 2000), Comrade (Ettema et al., 1995), Daily Activity Schedule (Ben-Akiva et al., 1996), Doherty and Axhausen Model (Doherty and Axhausen, 1998), Dynasmart (DYNASMART, 2013; Mahmassani et al., 1995), Ettema et al. models (Ettema et al., 1997b,a), FAMOS (Pendyala et al., 2004, 2005), FEATHERS (Arentze et al., 2006; Janssens et al., 2007), Flötteröd and Nagel model (Flötteröd and Nagel, 2006), Fotheringham et al. model (Fotheringham et al., 2001), GISICAS (Kwan, 1997), Han (Han et al., 2009), HAPP (Recker, 1995), Huisman and Forer model (Huisman and Forer, 2005), ILUTE (Salvini and Miller, 2005), Ma model (Ma, 1997), MADAM (Rossetti et al., 2002b), MASTIC (Dijst and Vidakovic, 1997), MERLIN (van Middelkoop et al., 2004), MIDAS (Goulias and Kitamura, 1992), mobiTopp (Schnittger and Zumkeller, 2004; mobiTopp, 2013), Mid-Ohio Regional Planning Commission (MORPC) model (Vovsha et al., 2003), NAVIGATOR (Gopal et al., 1989), New Yorks Best Practice Model (Vovsha et al., 2002), ORIENT (Sparmann and Leutzbach, 1980), PCATS (PCATS, 2011; Kitamura et al., 2005), PESASP (Lenntorp, 1976), PETRA (Fosgerau, 2001), Portland Daily Activity Schedule Model (Bowman et al., 1999; Bradley, 2005), Quadstone Paramics (Cameron and Duncan, 1996), RAMBLAS (Veldhuisen et al., 2000b), Rauh et al. multi-agent simulation (Rauh et al., 2007), SAMS (Kitamura et al., 1996), San Francisco Model (Jonnalagadda et al., 2001), SCAG (Bradley and Bowman, 2009), SCHEDULER (Gärling et al., 1989), SIMAP (McNally and Kulkarni, 2001), SMART (Stopher et al., 1996), SMASH (Ettema et al., 1996), STARCHILD (Recker et al., 1986a,b), Sivakumar et. al model (Sivakumar and Bhat, 2007), TASHA (Roorda et al., 2008; Miller and Roorda, 2003; Eberhard, 2002), TOUR (Kuipers, 1978), TransModeler (Caliper, 2013), TRASS (Lotzmann, 2009), TRANSIMS (Hobeika, 2005; TRANSIMS, 2009; Lawe et al., 2009; Nagel and Rickert, 2001; TRB, 2007), TSIS-CORSIM (McTrans, 2013), Vause's model (Vause, 1997), VISEM (Fellendorf et al., 2000), Wen and Koppelman model (Wen and Koppelman, 2000) .

Following frameworks are focused here for MATSim destination choice model development.

Transportation Analysis and Simulation System - TRANSIMS: MATSim is very similar to TRANSIMS (SimTRAVEL, 2013; FHWA, 2013; Nagel and Barrett, 1997). It was initiated at Los Alamos Laboratory and is now available open-source. The activity-based, agent-based framework searches for a utility-based equilibrium for a day period. The model contains a population synthesizer, an activity generator, a route planner and a cellular automaton traffic microsimulator. Feedback is commonly used between router and microsimulator (Hope et al., 2009; TRANSIMS Open Source, 2013) on the search for a Nash equilibrium in terms of agents' travel plans. However, further choice dimensions, such as time (Lee et al., 2010), mode choice (Lu, 2002), or destination choice (Hobeika, 2005, Chapter 4.8.5, p.69) are sometimes included in the iterations. Destination choice in the activity generator is done as described by Hobeika (2005, Chapter 4.8.5) and Ley (2008, p.29). The household home location is given. Destinations are chosen in a two-step procedure, based on discrete choice models, where first a zone and afterward an intra-zonal destination choice is done. Probabilistic intra-zonal destination choice is dependent on travel time and an attraction variable, specified by the user in external activity location tables. First, work locations are assigned and then locations for all other activities are chosen taking trip-chaining into account.

A Learning-Based Transportation Oriented Simulation System - ALBATROSS:
An example for an activity-based computational process model is ALBATROSS. Decision-making is based on decision-trees derived from activity diary data. Destination choice is based on heuristic rules and guided by a set of constraints defining the set of feasible alternatives. In Arentze and Timmermans (2007), the new destination choice model is presented, where, as a main innovation, detour time is introduced for a more accurate inclusion of travel time. A similar concept is applied for MATSim utility function estimation described in Chapter 5.

Prism-Constrained Activity-Travel Simulator - PCATS:
PCATS is an activity-based one day simulation framework that uses DEBNetS as a microsimulator (Kitamura et al., 2005; Kitamura and Fujii, 1998). Central component of the model are time-geographic prims constraining the potential activity space. Destination choice is done with a nested logit model combining mode and destination choice based on the attributes given in PCATS (2013, table 2), which are zone attributes (zone size, population density, number of commercial establishments), household/person attributes (age, sex, auto ownership) trip attributes (travel cost, travel time, number of transfers) and activity attributes (location type of current activity, location type of next fixed activity, day time of the activity).

3.4 MATSim Destination Choice

3.4.1 Choice Determinants

As shown above, numerous choice determinants are researched. In this section, we define a parsimonious set of key influence factors for application in the MATSim destination choice module, where generalizability and data availability play a major role.

3.4.1.1 A Note on Utility Scale in MATSim

Following a universal metaphor in economics, activity destination choice maximizes the difference of subjective benefit and subjective expenses, subject to constraints. Coming from the gravity model paradigm, the destination choice models, basically, contain attracting and deterring factors, which can be interpreted as generating revenues and expenses, respectively. Naturally, attracting factors are the one, which have a positive sign in the utility function and deterring factors are the negative ones. However, things, unfortunately, are more complex. For some factors, it depends on the scale if a positive or negative utility is contributed. Store comfort, for example, might be an expense, in form of stress in a uncomfortable store, or a revenue, if shopping is done in a nice ambiance. Discrete choice models are based on utility differences, in other words, on relative utilities, in general not requiring any further deliberation about utility scale. However, in MATSim, complete day plans spanning multiple choices are re-planned and evaluated. Thus, utility is transferred over multiple choices [3] , and its scale is thus relevant and requires attention during model estimation.

3.4.1.2 Chosen Determinants

In this thesis, following choice determinants are included in the destination choice model.

- **Trip attributes:** In the standard scenario used in this thesis and described in Section 8.1, *travel times* (and activity performing time) are a core choice determinant. Travel distances, after being converted from speeds and travel times, are only used for search space generation (Section 4.2.2). Monetary travel costs are not applied. *Trip chaining* is taken into account by simulating complete day plans. *Mode-dependency* has been added for the thesis' final runs (Section 8.6).
- **Store attributes:** *Price level* and *store size*, as two important and widely available attributes, are added in Chapter 5, further improving the model as shown in Section 8.3. *Opening hours*, which represent an important temporal constraint widely and publicly available, is a standard component of the MATSim Switzerland scenarios. *Spatial correlation* and associated agglomeration effects are covered in the model, implicitly, by inclusion of trip chaining, and

[3] If, for example, all destinations generate negative utility, then not the destination with still the highest utility is chosen, but instead the complete activity might be dropped, while this is not the case if at least one destination generates positive utility.

explicitly, by an agglomeration term τ_{agglo} as described in Chapter 7. Agent interactions in form of competition is included based on daily capacity estimates created by Meister (2008). Improved estimates are available for future investigations (Stahel, 2012).

- **Person attributes:** Consideration of person attributes, i.e., person heterogeneity, is probably one of the most powerful, but to date, only scarcely applied features in MATSim. As detailed by Horni et al. (2011e), person attributes are more or less only used for activity chain assignment (age: no children as workers) and mode choice (mobility tools and age). As argued later income, as a central component of any economic model, needs to be included urgently.

3.4.1.3 Some Remarks on Choice Determinants

In Switzerland, many large chains exist, which at least for grocery shopping offer a relatively broad range of product quality and price levels. Migros and Coop, the two largest chains in Switzerland, offer both low-budget (*M-Budget* and *Coop Prix Garantie*, respectively) as well as relatively high-quality lines (*Migros Selection* and *Coop Fine-Food*). Thus, it is assumed, that influence of price level and product quality is difficult to quantify.

For some attributes strong correlations must be assumed, for example for the price level and quality variables. Although factors like store comfort or image have an influence on prices, the free market usually ensures a strong correlation between price and quality. Due to this correlation, and also as it is difficult to collect the quality variable without using price as a proxy, it is omitted in our model for the moment.

A similar correlation might exist between store size and product range. Again it is difficult to collect for large-scale scenarios. Thus, it is also dropped for the moment.

However, the incorporation of attraction variables (as handled in Cadwallader (1975); Hubbard (1978); Robinson and Vickerman (1976); Koppelman and Hauser (1978)) is important as due to the price-quality correlation a search for a high-quality special product might be masked as an a priori search for high prices, leading to strong biases in estimates.

Many different sub-categorizations of shopping activities can be created, such as grocery and non-grocery trips, or long-term and short-term shopping (see also Section 9.2.1.3). Swiss Microcensus provides 4 categories of different shopping purchases (groceries, consumer goods, investment goods and leisure shopping). In Horni et al. (2009b) (see also Chapter 4) shopping is divided into grocery and non-grocery trips. Sub-classification is an important topic for future model improvement (see also Section 9.2.1.3).

3.4.2 MATSim's Relationship to Other Operational Large-Scale Models

As will be shown in Section 4.1.1, the first version of the MATSim destination choice module was based on time geography. However, different to PCATS, the actual choice within the activity space was not (yet) based on a discrete choice module.

A very close relationship between MATSim and TRANSIMS exist (see Section 3.3 and 2.4.1). From the documentation available, it is not clear to the thesis author, if destination choice actually is a standard choice in the iterations or if it is only performed in case day plan anomalies occur. Furthermore, and more important, it is unclear how consistent random draws over the course of the iterations are achieved. This is guaranteed for MATSim now as shown in Section 4.2. TRANSIMS thus might be improved as well by our work. On the other hand, the MATSim equilibrium discussion might be enriched by the TRANSIMS stopping criterion discussion reported by Kelly and Nagel (1998).

For problems related to within-day replanning (rather than end-of-day replanning), it is required to bring MATSim from utility maximizing closer to heuristic and situational choice making; detailed parking search modeling might be an example (see Section 7.1.2 and Horni et al. (2012b)). For this undertaking, rule-based models, such as ALBATROSS, can be consulted.

3.4.3 MATSim Model Improvement

Analyzing destination choice modeling methods (Section 3.1.1) and choice determinants (Section 3.1.2) and figuring out how to integrate them into large-scale microsimulation models (Section 3.3)—in other words, the design of an operational microsimulation model—is basically an *up-scaling* undertaking. Naturally, due to data availability and computational issues, theory and small-scale destination choice models can feature a much higher level of detail than large-scale microsimulation scenarios. Thus, for large-scale models, the choice process must be strongly reduced, while still keeping explanatory power as far as possible. The significance of the upscaling problem is often overlooked, when taking a strictly theoretical focus as its sources are rooted in practice. Thus, it is important to apply investigations on a practical example; here, MATSim is chosen.

Essentially, MATSim validation bases on road count data, prepared to represent an average working day (see e.g., Horni (2007); Horni and Grether (2007), Meister et al. (2010, p.8)). Substantial parts of demand, such as commercial traffic, are not yet completely covered by the MATSim base scenario. Furthermore, count data show a large temporal variance (see Figure 8.16). This variance is *not* taken into account, neither in MATSim initial demand creation nor in the productive ensemble runs. Consequently, model validation can only be relatively rough. As further discussed in Section 9.1.2, count data furthermore suffer from the following issue. There are two versions of the Zurich scenario (Horni et al., 2011e). In the older scenario, initial demand generates substantially too low, and in the newer scenario too high simulated volumes. This means that, any change substantially increasing or decreasing the volumes, respectively, appears as a model improvement. This practical constraint delimits the space for potential model improvements and urgently asks for further validation data.

Chapter 4

Basic Model

4.1 Earlier Approaches

4.1.1 Local Search Based on Time Geography

A first MATSim destination choice approach based on Hägerstrand's time-geographic framework (Hägerstrand, 1970; Landau et al., 1982a) is incorporated in Horni et al. (2009a). Inclusion of destination choice in the system results in an enormous search space that would be impossible to explore exhaustively within a reasonable time. Horni et al. (2009a) address these computational issues and show on the MATSim Zurich scenario that the system rapidly converges toward a system's fixed point if the agents' choices are per iteration confined to local steps. In Figure 8.1 it can be seen, that configuration 4 (the time-geographic approach) shows good convergence behavior.

However, if the score is only composed of travel time, then the closest destination of the correct type would always be chosen. Further attributes, thus need to be considered with this approach; in Horni et al. (2009a) activity location competition and store size was taken into account. But still, in destination choice a relatively large unobserved heterogeneity exists, which is not taken into account in this model. Thus, too short travel times and distances results over the course of iterations. A second approach as described in the next section was adopted to integrate unobserved heterogeneity.

4.1.2 Hollow Space-Time Prisms

An improvement to the time-geographic approach is incorporated in Horni et al. (2009b). Instead of optimizing in a circular potential path area, a ring form is used, which is why we called the method "hollow space-time prisms approach". The ring radii are drawn from empirical distance distributions given in the Swiss Microcensus. Assignment is also dependent on the activity duration as a positive (but weak) relation is found in the literature (Iragael, 2007; Kitamura et al., 1998). The approach actually improves simulation results as shown in Figure 9.1. Essentially this (calibration) procedure is similar to applying random error terms, but, the strong restraining of a person's choice set lacks flexibility, and it is not based on an established framework. Thus, this approach was eventually replaced by the solution described in the next section.

4.2 Random Error Terms

The current destination choice module is based on random error terms (Horni et al., 2012c, 2011d), not only improving the methodology of the destination choice module but of the complete MATSim framework. In MATSim, up to now, the randomness measured in empirical data was included implicitly and in an uncontrolled way through the stochasticity of the simulation process. For destination choice, this has led to a dramatic underestimation of total travel demand as mentioned above.

Unobserved heterogeneity is now added directly to the utility function through a random error term, making MATSim fully compatible with econometric discrete choice methodology (McFadden, 1978).

Importantly, these random error terms are *quenched*, i.e., they will be the same for repeated executions of the choice model (Section 4.2.1), which is not straight-forward to achieve. This holds true for all iterative equilibration procedures repeatedly performing random draws.

High-resolution destination choice for large-scale microsimulations raises several further technical issues; pragmatic engineering solutions have been developed or applied to cope with them. These solutions are described in technical detail below to assist in the further development of similar microsimulations.

Real-world simulation experiments for Zurich in Section 8.2 show that the current modifications substantially improve results.

4.2.1 Repeated Draws: Quenched vs. Annealed Randomness

4.2.1.1 Including Randomness in the Microsimulation

One assumed advantage of microsimulation is the conceptually straightforward inclusion of heterogeneity. In the first instance, one can, whenever it is needed, either

- randomly draw from a choice model given as probability distribution or
- randomly generate an ϵ_{piq} for every person p, alternative i and activity q and select an alternative i as $\underset{i \in choice\ set}{\text{argmax}}\ U_{piq}$.

However, problems with repeated draws must be solved. Repeated draws mean that the same individual p is repeatedly faced with an identical choice, a frequent situation in iterative models. Obviously, the ϵ_{piq} should remain fixed once they have been drawn for the first time. In physics, this would be called "quenched" randomness; all randomness is computed initially and then attached to particles or locations, rather than instantaneously generating it, which would be called "annealed" randomness.

4.2.1.2 Implementing Quenched Randomness

Quenched randomness can be achieved by applying one of the following two strategies:

- (a) Freezing the applied *global* sequence of random numbers, meaning that a Monte Carlo method with the same random seed is used before and after the introduction of a policy measure and over the course of iterations. Thus, the ϵ_{piq} should come out the same way *before* and *after* the introduction of the policy measure. Differences in the outcome can thus be directly attributed to the policy measure.
- (b) Computing and storing a separate ϵ_{piq} for every combination of person p, alternative i and activity q.

We reviewed relevant literature, but could not determine strategies applied in each case in other large-scale transport microsimulations. Through personal e-mail communication with the simulator authors, some answers emerged: in AMOS and OpenAMOS (OpenAMOS, 2011; Pendyala et al., 1997) (a) is applied. In Albatross (Arentze and Timmermans, 2000), both (a) and (b) have been applied. For the NYC activity-based microsimulation (Vovsha et al., 2008) in most cases (a) is used, although they recently switched to (b). The Tel Aviv model (Cambridge Systematics Inc., 2008) is based on (a). The Sacramento and Portland models (e.g., Bradley et al., 2010; Bowman et al., 1999) apply (a).

Both strategies have flaws. Approach (a) is only an option if one is completely certain about literally all aspects of the computational code. Importantly, one additional random number, drawn in one iteration but not in the other, completely destroys the "quench" for all decisions computed later in the program.

Thus, approach (b) is expected to be more robust in practice. However, for large numbers of decision makers and/or alternatives, storing error terms is difficult. For destination choice, one quickly has 10^6 decision makers and 10^6 alternatives, resulting in more than 4×10^{12}Byte = 4TByte of storage space.

One may argue that this should not be a problem, since a normal person will rarely consider more than the order of a hundred alternatives in their choice set, reducing the computational problem. Aside from the necessity of storing every decision maker's choice set, this converts the computational problem into a conceptual one, since a good method to generate choice sets then needs to be found. With more conceptual progress, this may eventually be an option, but at this point, a conceptually simpler approach is preferred.

As far as we know, this set of problems has not been discussed in existing literature. The computational problem associated with approach (b) is solved by Horni et al. (2012c) as follows.

Instead of storing these error terms directly, requiring an infeasible storage effort, the same *stable* error term can be *re*-calculated on the fly by using stable random seeds $s_{piq} = g(k_p, k_i, k_q)$, where, k_p, k_i and k_q are uniformly distributed random numbers associated with p, i, and q. That is, for each person p a random number k_p is generated and stored, and the same is done with each destination i. The value for the activity q can be stored in the person or else derived from its index in the plan possibly combined with the person's value k_p. This reduces the storage space dramatically from $\bar{n}_q \cdot n_p \cdot n_i$ to $\bar{n}_q(n_p + n_i)$, where n_p is the number of persons or agents and n_i is the number of destinations and $\bar{n}_q$ is the average number of discretionary activities in an agent's plan. This means that storage space for a typical large-scale scenario is reduced to approx. $2 \cdot 4 \cdot 10^6$Byte = 8MByte, which can be easily stored on any modern machine. The distribution of these seeds is essentially irrelevant; any error term distribution can be generated from any basic seed distribution. In this work $g(k_p, k_i, k_q) = (k_p + k_i + k_q) \ mod \ 1.0 \cdot v_{max}$ is used. v_{max} is the maximum (long) number that can be handled by the specific machine.

To evaluate utility for a person p visiting the destination i for activity q, a sequence of Gumbel-distributed random numbers seq_{piq} is generated on the fly for every person-alternative-activity combination using the seed s_{piq}. Some random number generators have problems in the initial phase of drawing, e.g., the first couple of random numbers are correlated or do never cover the complete probability space. As in our procedure the random number generator is constantly re-initialized, for these technical reasons, the error term ϵ_{piq} is not derived from the first element but from the m^{th} element of the sequence $seq_{piq}[m]$. Here, m is set to 10. This procedure is valid as the set of all m^{th} elements of all different sequences is also a pseudo-random sequence following the same distribution as the sequences seq_{piq}; clearly, *true* random number generators relying on physical phenomena, such as hardware temperature, are not applicable. In this thesis, a standard Gumbel distribution is applied.

Having now a method at hand, that ensures stable error terms over the course of iterations, one can generate and assign the necessary seeds in a pre-processing step. The optimization is then performed as a deterministic search based on the resulting utility function. In fact, this can be seen as a return to the roots of random utility modeling; rather than absorbing the ϵ_{piq} into the choice model, they are now explicitly generated.

However, a way to efficiently perform this optimization for large search spaces is required.

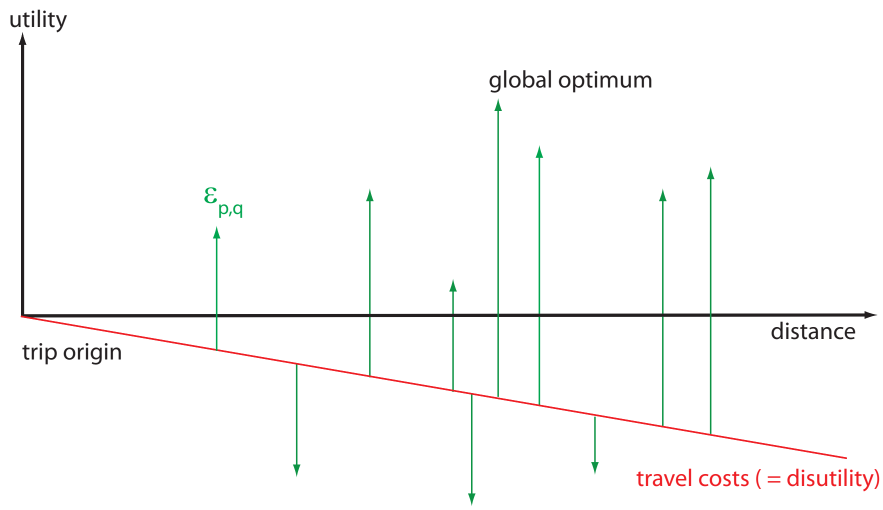
4.2.2 Search Space and Search Method

Conceptually, MATSim is based on random mutation embedded in a local search strategy. Exactly one alternative per choice dimension is evaluated in each iteration. However, the huge number of available alternatives for some choice dimensions makes the introduction of optimizing mechanisms indispensable, i.e., mechanisms that do not return a random alternative from the set of available alternatives but a “good” one, often called *best-response* in MATSim. (Within-iteration) optimization, rather than random mutation, is performed for route choice and sometimes for time choice (Meister et al., 2005). For destination choice, random mutation is also impossible not only due to the huge alternatives set, but also due to the search space characteristics. The discrete search landscape is characterized by random noise because error terms are not (or only locally) spatially correlated (see Figure 4.1(a)). For such problems—as opposed to continuous landscapes (see Figure 4.1(b))—efficient search methods, such as local search methods, generally do not work. Fortunately, as long as the change of travel costs between succeeding iterations is not too large, multiple search space destinations can be evaluated per person and per iteration. Normally, the relatively small share of agents who re-plan, keep the inter-iteration changes small. Thus, increasing the number of evaluated alternatives per iteration might be feasible. This reduces the number of iterations and substantial costs associated with simulation of network loading (see also the discussion on equilibration in Section 9.2.2).

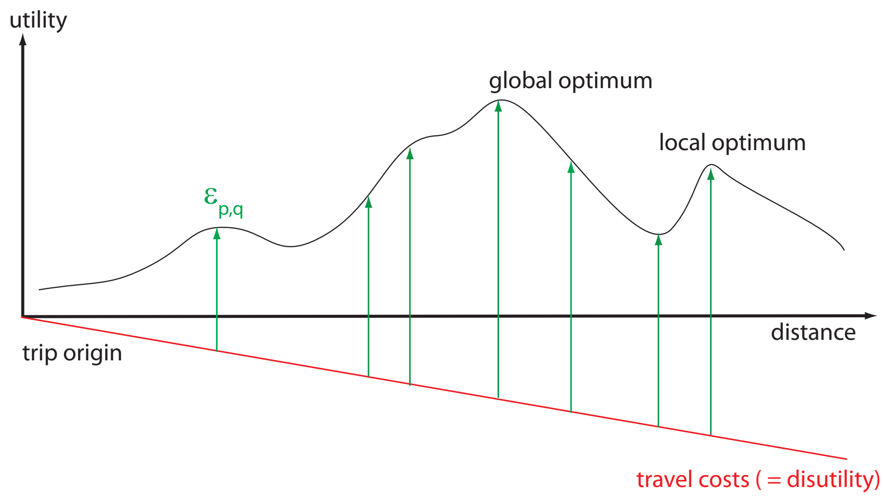
The next problem is the huge number of alternatives for destination

Figure 4.1: Search space: The search algorithm is required to be able to handle correlated but also uncorrelated error terms as given by the MNL model. Local search methods, such as hill-climbing algorithms are only able to handle continuous search spaces, thus, for situation (a) a best-response global search algorithm is required.

(a) Uncorrelated error terms



(b) Spatially correlated error terms



choice. In the model and in reality, the utility contributed by the error term is unlimited; the search space for potential destinations is hence unlimited. In consequence, exhaustive search for finding the optimal destination per iteration is not an option. It usually produces prohibitively large computation efforts for large-scale scenarios. Thus, the application of problem-tailored heuristics and approximations in the destination choice module is unavoidable.

A first attempt to narrow down the individual search space is as follows. In discrete choice theory, individual p chooses alternative i if it produces the maximum utility for activity q:

$$U_{piq} \geq U_{pjq}, \forall j \in choice\ set\,,$$

that is

$$V_{piq} + \epsilon_{piq} \geq V_{pjq} + \epsilon_{pjq}, \forall j \in choice\ set\,,$$

where V denotes the deterministic part of the utility function. V is usually composed of travel effort V_{trav} and utility for performing an activity V_{act}. Hence,

$$V_{act,piq} + V_{trav,piq} + \epsilon_{piq} \geq V_{act,pjq} + V_{trav,pjq} + \epsilon_{pjq}, \forall j \in choice\ set\,.$$

In MATSim, $V_{act,pjq}$ is equal for all destinations j if the performed activity time is equal. It can be omitted when searching for an upper bound for the accepted travel costs, as explained in the following: Clearly, V_{act} is larger for closer destinations – the longer the trip takes, the less time there is to perform the activity. In other words, utility decreases by traveling due to travel costs V_{trav} and the opportunity cost of time (loss of V_{act}). An *upper bound* for the maximum search space can thus be found by considering only the travel costs and the error term, i.e., by ignoring the opportunity cost of time (lost activity performing time):

$$V_{trav,piq} + \epsilon_{piq} \geq V_{trav,pjq} + \epsilon_{pjq}, \forall j \in choice\ set\,. \tag{4.1}$$

It can be seen that a person only travels farther from destination j to destination i if the additional travel effort is at least compensated by the error term difference (remember that destination choice is only done for flexible activities), i.e., if that effort produces a net benefit.

If we now assume that a flexible activity is dropped if it does not generate positive utility at least for one destination, i.e.,

$$V_{trav,pjq} + \epsilon_{pjq} \geq 0\,, \tag{4.2}$$

and if we assume that travel costs V_{trav} are always negative, then the maximum potential travel effort a person is willing to invest is constrained by the maximum error term per person and activity, i.e.,

$$-V_{trav,pjq} \leq \epsilon_{pq,max} \tag{4.3}$$

with $\epsilon_{pq,max} := \max_j \epsilon_{pjq}$. Note that $V_{trav,pjq}$ typically is negative, and thus the equation means that $V_{trav,pjq}$ cannot be more negative than ϵ_{pq} is positive. We thus continue with

$$\|V_{trav,pjq}\| \leq \epsilon_{pq,max} \ . \tag{4.4}$$

In this thesis, linear travel disutilities in terms of time are used as shown in Equation (2.6). To constrain the search space—but not for the final evaluation of the alternatives in the scoring phase—travel times are transformed into travel distances by the external parameter "crow fly speed" v. With that, the above equation translates to:

$$distance_{pq} \leq \frac{\epsilon_{pq,max}[utils] \cdot v[km/h]}{\|\beta_{pq}\|[utils/h]} \ . \tag{4.5}$$

This approach is promising, as very large values for Gumbel-distributed $\epsilon_{pq,max}$ are rare, meaning that a huge space must be searched for only a few persons.

A search space Γ_{pq} can now be constructed as follows: Let us assume that for the activity q of person p, a new destination l_{pq} has to be found. Γ_{pq} can then be defined as a circle whose center is the mid-point between the preceding activity l_{pq-1} and the succeeding activity l_{pq+1}. The radius of the circle is set to:

$$r_{\Gamma_{pq}} = (distance(l_{pq-1}, l_{pq+1}) + max(distance_{pq}))/\psi \ .$$

The most productive value for ψ is not yet apparent. For every discretionary trip, there is a $max(distance_{pq})$ that person p is willing to travel at most. Looking at an individual discretionary tour with fixed and identical destinations $l_{pq-1} = l_{pq+1}$ clearly $max(distance_{pq})$ includes the return trip and $\psi = 2$ is thus a natural choice. But, for consecutive multiple discretionary activities the search space is probably larger, and ψ is thus smaller. However, essentially, the value of ψ is subject to calibration and needs further research. In this paper, $\psi = 2$ is used.

It is crucial that Γ_{pq} can be computed fast and that all destinations actually accessible are contained. On the other hand, only computation times, but not the quality of the results, are influenced if destinations that are actually inaccessible are included in the evaluation. For that reason, it is possible to approximate the travel distance $distance_{pq}$ by the straight-line distance. This distance can then be computed once in a preprocessing step.

Following improvements are left for future work:

- First, while handling a plan in this work every discretionary activity is handled separately, i.e., destination choice is done sequentially. Clearly this is an approximation, as consecutive multiple discretionary activities may shall be handled in parallel.
- Second, if the preceding and the succeeding activity destinations are not identical, the space which can be accessed within a certain travel time or travel distance budget is elliptic and not circular. Thus, in the future research the specification of the search space Γ_{pq} as an ellipse should be analyzed for computational reasons. However, data structures for efficient spatial searches (such as *Quad Trees* (Finkel and Bentley, 1974)) do not yet exist for elliptical spaces.

4.2.2.1 Further Speed-Ups

Although, the reduction of the search space saves a lot of computation time, it is still infeasible and further speed-ups are necessary. Most computation time is due to calculation of travel times, i.e., due to routing, for evaluation of the alternatives in the search space. To reduce these huge routing costs, the following procedure is applied.

Let us assume that location l_q of activity q is changed, where all other plan activities are fixed. Travel times for routes between activity location l_{q-1} and all potential locations l_q can be exactly and efficiently computed by Dijkstra's algorithm because it efficiently computes the best routes from one location to *all* other locations in the network. Travel times of the best routes between activity locations l_q and l_{q+1} are computed by running Dijkstra's algorithm *backwards*, using an average estimated arrival time as initial time. This is an approximation, as the arrival time at l_{q+1} is different for different locations l_q.

To reduce possible approximation errors, a *probabilistic* best response is applied. Search space destinations are evaluated as described above; then a random choice weighted by these approximated scores is performed. The plan containing the new choices is finally simulated and eventually scored, based on *exact* travel times by the MATSim iteration scoring. This approach is justified by the assumption that, during the course of the iterations, the probabilistic choice probably reduces, or even compensates, the errors incurred by approximating travel times as described above.

However, the probabilistic choice brings back the problem of slow convergence. If every alternative in the search space is chosen with probability greater than zero, this still very large set again necessitates a very large number of iterations. For reasonable convergence, the probabilistic choice must be performed on a reduced choice set. Thereby, restraining the choice set to the ϕ destinations producing the highest approximate plan scores is natural. ϕ is essentially dependent on the approximation error done by estimating travel times. For the present thesis, after some trial runs, ϕ was set to 30.

With this procedure, the required computational effort is dramatically reduced, allowing application of destination choice to large-scale scenarios. One iteration of the 10% Zurich scenario takes roughly 5 minutes (instead of 10 days, when not applying any of the aforementioned speed-ups). The simulation is run with 10 parallel JAVA threads and approximately 15GB of RAM. The Linux server is equipped with an Intel Xeon(R) processor, 3.33GHz, with 24 cores and 96GB of RAM.

Parameters described above are documented for the software module in the javadoc section on MATSim (2013).

Chapter 5

Destination Choice Utility Function Extension

This chapter's topic is destination choice model estimation and application in microsimulations. Estimating a model with a standard tool and transferring the parameters into the MATSim utility function is not straight forward as activity duration utility is non-linear. The procedure applied in Section 5.1, contributes to the few estimations done up to now (Kickhöfer, 2009; Balmer et al., 2010; Feil et al., 2009a).

For destination choice model estimation, specification of choice sets is an unsolved issue (Section 5.2). Estimated parameters are sensitive to choice sets (Schüssler, 2010; Pellegrini et al., 2005) and at the same time no established choice set definition procedure exists and thus choice sets are highly dependent on the modeler, which is basically an exogeneity problem. A structurally similar problem exists for aggregate models, the "Modifiable Areal Unit Problem" (MAUP) (Gehlke and Biehl, 1934; Openshaw, 1984), where results are dependent on modeler's zoning specification.

In Section 5.3, an example of the strand of probabilistic models, combining endogenous parameter *and* choice set estimation in one procedure, is tested. The model of Zheng and Guo (2008) is based on the relatively weak assumption that choice set are continuous. This resolves the combinatorial problem associated with prohibitively high computation times, which are principally attached to probabilistic choice set models.

The test, however, has shown, that this approach also has its problems. Thus, to further approach choice set specification a choice set survey was performed, focusing on store visiting frequencies and store awareness (Section 5.4). First results are presented in Section 8.3.3. The data will be archived according to international standards in an on-line travel data archive (IVT and ETH Zurich, 2013) to enable future, and possibly collaborative, model estimation and further analyses. The used computer code will be made available in a consistent software central (SourceForge, 2013), and code was already shared with several research groups.

Model application is discussed in Section 5.5.

5.1 Destination Choice Model Estimation

Main goal of this section is computation of parameters applicable in MATSim standard scenarios. Before, testing a more elaborate approach in the next section, an MNL model is estimated based on the universal choice set, here defined as the set of all stores in Zurich. As mentioned above, destination choice set specification is an unsolved problem. For our setting, estimation based on the universal choice set is not a very elaborate but consistent solution as the later application is also essentially based on the universal choice set, where, additionally, the models are estimated for Zurich city, which is naturally very similar to the region for which it is applied, namely the Zurich metropolitan region. Furthermore, runtime for the universal choice set model is a non-issue, here, being for all models below 2 minutes. However, maybe only limited generalizability and transferability results, and, thus, in the long run, robust choice specification approaches are required for all of destination choice modeling.

For model estimation, Swiss Microcensus 2005 and 2010, providing complete trip data, is used. For this work, all grocery shopping trips starting and ending in Zurich city of the city residents are used. 600 persons with 634 grocery shopping trips are available for estimation.

The data generated in the shopping survey is focused on person's store frequencies and awareness over a longer time period. To keep the survey burden acceptable, it does not provide information about single trips, and hence it cannot be directly used for the estimation performed here. However, it can in the future be used to estimate probabilistic choice set models, that require additional person details such as Ben-Akiva and Boccara (e.g., 1995).

5.1.1 Model Specification

Goal of estimation using Biogeme (Bierlaire, 2013) are first indications about quantitative relation of MATSim time parameters and further choice attributes. Attributes used for estimation are *store size* (in categories, see Table 5.1), *price level* (in categories, see Table 5.1) and *additional linear distance* to the store as illustrated in Figure 5.1, similar to the detour distance defined in Arentze and Timmermans (2007). Distance is computed in kilometers. These attributes were chosen from the key destination choice determinants in Chapter 3.

Clearly, for direct application in MATSim, travel times rather than distances would be better, but this information is not available consistently. A minimal set of variables is chosen due to data availability and as the main goal is laying an instructive base for future MATSim utility function estimations and their application in the MATSim Zurich scenario. Alternative-specific constants are not assigned to destinations to prevent over-fitting [1] .

For extraction of the store sizes, the shopping destinations reported in the Microcensus are mapped to the Swiss Business Census 2008, which reports size in classes by the variables (52.11A-E and 52.12A). For price level assignment, the destinations are mapped to the stores set available in the survey, mentioning the affiliated chain, each chain having essentially homogeneous prices nationwide. The actual assignment of a price level to a chain is not easy as all chains cover a large range of levels (Section 3.4.1.3). The classification applied here, is based on Saldo (2002); Comparis (2011); for future analyses filialsuche.ch (2013) could be incorporated. The minority of stores, not belonging to a chain, are—in the first instance—assigned to the price level of *Spar* according to our every-day experience. Further price info about the large chains not yet used here is available (Schweizerischer Verband der Lebensmittel-Detaillisten, 2013). Code for most data preparation steps are available in the playground package *anhorni* downloadable at MATSim (2013).

In the interaction models, income (in categories, see Table 5.1) and age (in years) are used.

Estimation results are shown in Section 8.3.1.

5.2 Choice Set Definition and Destination Choice Process

For decisions with only few alternatives the standard discrete choice procedure—a draw from a finite set of alternatives—is natural and perfectly adequate. However, for spatial choice problems (destination choice, but also route choice (Schüssler, 2010; Frejinger et al., 2009), etc.), the number of available alternatives is huge, such that choice set formation becomes a crucial computational and methodological problem. While the decision rules for choosing an alternative of the choice set have reached

[1] For a discussion of alternative-specific constants see e.g. Bierlaire et al. (1997).

Figure 5.1: Additional distance: In case I additional distance is $d_1+d_2-d_0$, where in case II for a shopping round trip it is defined as $2d_1$

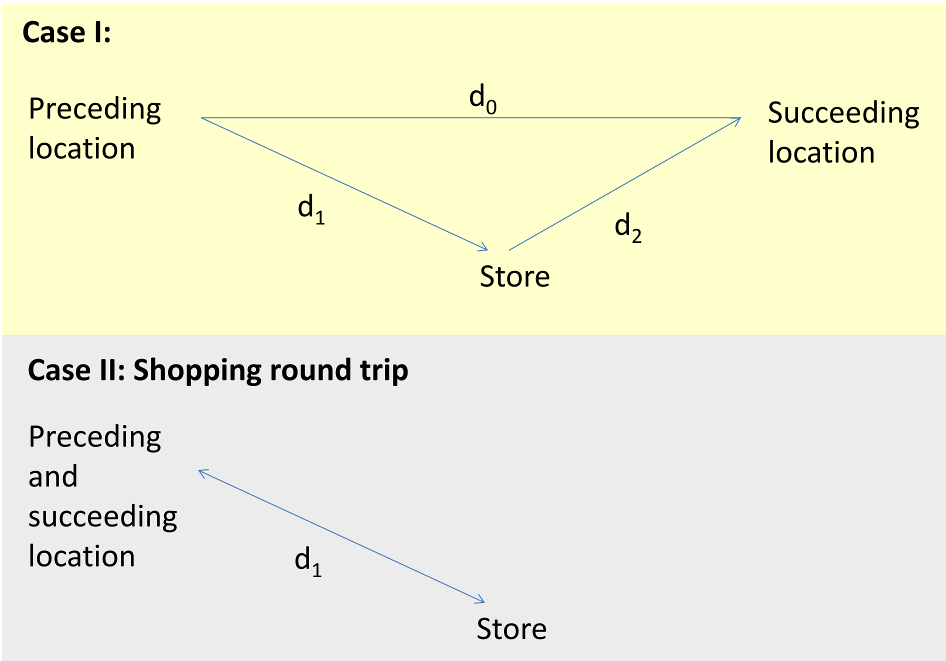


Table 5.1: Size, price and income categories

Original size category	Size category for estimation
"Kleine Geschäfte" ($< 100m^2$)	1
"Grosse Geschäfte" ($100 - 399m^2$)	2
"Kleine Supermärkte" ($400 - 999m^2$)	3
"Grosse Supermärkte" ($1000 - 2499m^2$)	4
"Verbrauchermärkte" ($> 2500m^2$)	5
"Warenhäuser"	6
Store chain	**Price category for estimation**
Lidl, Aldi	1
Denner	2
Migros, Coop	3
Spar and Other	4
Marinello, Globus	5
Monthly household income	**Income category for estimation**
$< 2000SFr.$	1
$2000 - 4000SFr.$	2
$4001 - 6000SFr.$	3
$6001 - 8000SFr.$	4
$8001 - 10000SFr.$	5
$10001 - 12000SFr.$	6
$12001 - 14000SFr.$	7
$14001 - 16000SFr.$	8
$> 16000SFr.$	9

a high level of sophistication, choice set specification is still a major unsolved issue. Even worse, as estimated model parameters in general are very sensitive to the specification of choice sets, a consistent choice set specification procedure is crucial.

In the literature, two main strands addressing choice set formation for problems with a large number of alternatives can be identified (for overview articles see Thill (1992); Pagliara and Timmermans (2009)).

The models of the first strand are based on a *deterministic* specification of choice sets, where the choice sets are an exogenous input to the estimation step. Examples range from early ad-hoc models (Gautschi, 1981; Weisbrod et al., 1984; Adler and Ben-Akiva, 1976; Miller and O'Kelly, 1983; Southworth, 1981) to nested and hierarchical choice-making approaches (Fotheringham, 1988) to the rather complex cognitive models (Chorus and Timmermans, 2009; Hannes et al., 2008; Mondschein et al., 2008; Arentze and Timmermans, 2004; Golledge and Timmermans, 1990) and include models of the time-geographic approach (Hägerstrand, 1970; Landau et al., 1982b; Thill and Horowitz, 1997; Scott, 2006; Scott and He, 2012).

The second strand, which is often called the *probabilistic* approach, was founded by Manski (1977); Burnett and Hanson (1979); Burnett (1980) and integrates the choice set formation step into the estimation procedure by jointly estimating selection of a choice set and the choice of a particular alternative of this choice set (Kaplan et al., 2009; Pagliara and Timmermans, 2009; Manski, 1977; Swait, 2001; Horowitz and Louviere, 1995; Ben-Akiva and Boccara, 1995; Swait and Ben-Akiva, 1987; Swait, 2001; Martínez et al., 2009; Cascetta and Papola, 2009; Bierlaire et al., 2009; Scrogin et al., 2004; Manrai and Andrews, 1998; Ansah, 1977).

Probabilistic choice set formation is conceptually appealing as choice sets are, in principle, not restrained a priori by exogenous criteria as in deterministic choice set specification. However, the procedure is in general associated with combinatorial complexity, making it computationally intractable. As a consequence, the practical approaches also require mechanisms to reduce the complexity of the choice set specification problem (see e.g., Ben-Akiva and Boccara (1995, p.11)). Zheng and Guo (2008), for example, make the moderate assumption of continuous choice sets (i.e., sets without "holes") around the trip origin, while the random-constraints model of Ben-Akiva and Boccara (1995) exploits additional information on alternatives' availability for individuals. The approach by Zheng and Guo (2008) is tested below.

Besides consistent choice set specification, two further methodological problems exist. It is not clear if the (possibly rule-based and temporally extended) learning process, that precedes the final destination choice, can be adequately modeled with a discrete choice model, which essentially represents simultaneous choice-making. A related problem, influencing choice set specification, is interpretation of discrete destination choice modeling either as a statistical or behavioral approach. These two issues are discussed in the following two Sections 5.2.1 and 5.2.2.

5.2.1 Modeled Decision Horizon

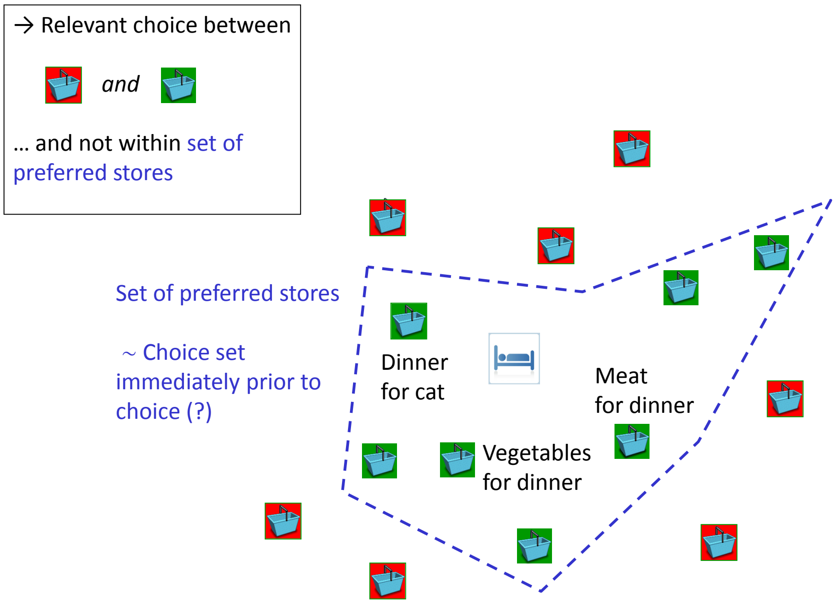
In discrete destination choice models, usually, the decision made *immediately prior* to the purchase is focused. The (implicit) behavioral premise or story underlying the model is an instantaneous utility maximizing choice from a set of alternatives, which are simultaneously evaluated. However, destination choices vary in terms of cognitive consumer involvement, leading to different types of decision behavior, where for some of them the above premise—and hence also the modeling procedure—fits badly and becomes fictive.

Marketing research differentiates several types of consumer decision behavior, (e.g., Solomon, 2009; Kroeber-Riel and Weinberg, 2003; Foscht and Swoboda, 2007). The classification proposed by (Solomon, 2009), for example, distinguishes between *extensive problem solving, limited problem solving* and *routine response behavior*. As the names imply, the categories are characterized by a decrease in cognitive consumer involvement. This decrease is caused, for example, by low purchase costs not worth much deliberation, a strong *emotional* stimulus that dominates cognition, but also by familiarity with the specific decision-making situation. During repeated extensive decisions on the same subject, approved purchasing criteria (i.e., previous knowledge) can be established that lead to limited decisions or even routine behavior.

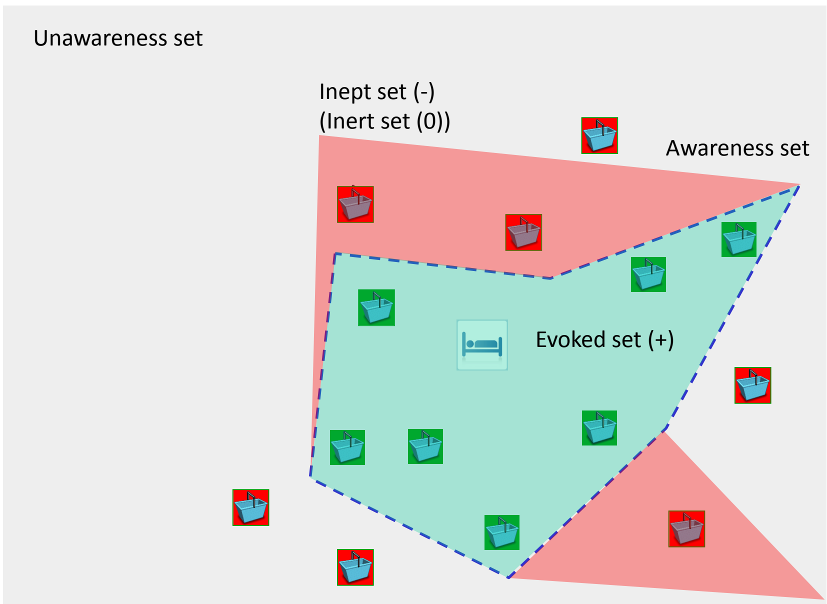
Consequently, this means that if the investigation is limited to actual decisions made *immediately prior* to the purchase, there is a high probability of missing the relevant part of the decision. The accumulation of previous knowledge, i.e., the learning process as a prerequisite for the final decision, is not included in such a model. To make it even clearer: Assuming (as a key-hypothesis of this chapter) that for the regular grocery shopping trips people have a persistent set of frequently visited stores, the choice immediately prior to the trip is probably done *within* this set, and thus completely driven by current needs and *not* by general preferences (Figure 5.2(a)). This means, that the actually interesting subject for modeling and even more for planning tools is *not* the final choice from this set, but the distinction between this set and the set of never or only seldom

Figure 5.2: Store sets involved in shopping destination choice

(a) Regular grocery shopping destination choice example: Preferred and frequently visited set of stores



(b) Store sets as proposed by Narayana and Markin (1975)



visited stores. In the survey presented later, the set of frequently visited stores is surveyed.

A further refinement of this distinction is proposed in marketing models such as Narayana and Markin (1975); Howard and Sheth (1969);

Crompton (1992); Shocker et al. (1991); Spiggle and Sewall (1987). In these models, in addition to the set of options considered immediately prior to the choice (often termed the evoked set, consideration set or choice set), additional higher-level sets that are relevant for the decision-making process are defined. Narayana and Markin (1975) for example, introduce the following sets: *unawareness set, awareness set, inept set, inert set* and *evoked set* (Figure 5.2(b)). As the name implies, the awareness set consists of all options that the consumer is aware of. The awareness set is further divided into the inept set (for which the consumer has a negative evaluation), the inert set (for which the consumer has neither a positive nor a negative evaluation) and the evoked set (for which the consumer has a positive evaluation). The process of how the evoked set is derived from the awareness set remains unspecified.

As it can be reasonably assumed that the positive evaluation of the stores in the evoked set leads to frequent visits, a strong similarity between the evoked set and the set of frequently visited stores can be implied. The survey described later tries to get first empirical insights also on these further sets. A model is not yet created, though.

5.2.2 Are Discrete Destination Choice Models a Statistical or a Behavioral Tool?

The interpretation of discrete destination choices is relevant for correct method usage both for estimation and application, where it is particularly important for choice set specification. When applying discrete choice models as a purely statistical tool without any behavioral meaning, then a strictly numerical approach can be chosen for choice set specification. Applying behavioral rules to form the choice sets is not expected to be very productive in this case. Assuming, parameters to be estimated stabilize to a certain extent when making the choice set larger and larger, one could simply define a stability threshold similar to the stopping criteria in numerical methods. However, research on this stabilization behavior of the parameters and the associated constraints is rare and computability might be an issue, especially for the application in operational models.

Following a behavioral perspective, the overall goal is to define the choice set in such a way that it is *easy to survey* (and easy to generate in forecasting models) *and* so that it *actually plays a well-defined role* in the process of coming to and making a destination decision. Space-time prisms of time geography, derived from travel time budgets, among others, provide an appealing approach to specify choice sets that play a role in the final decision before undertaking the respective trip. However, the specification of individual travel time budgets is subject to the same empirical and methodological problems pointed out earlier, so that searching for travel time budgets is essentially a *proxy problem* to specifying choice sets.

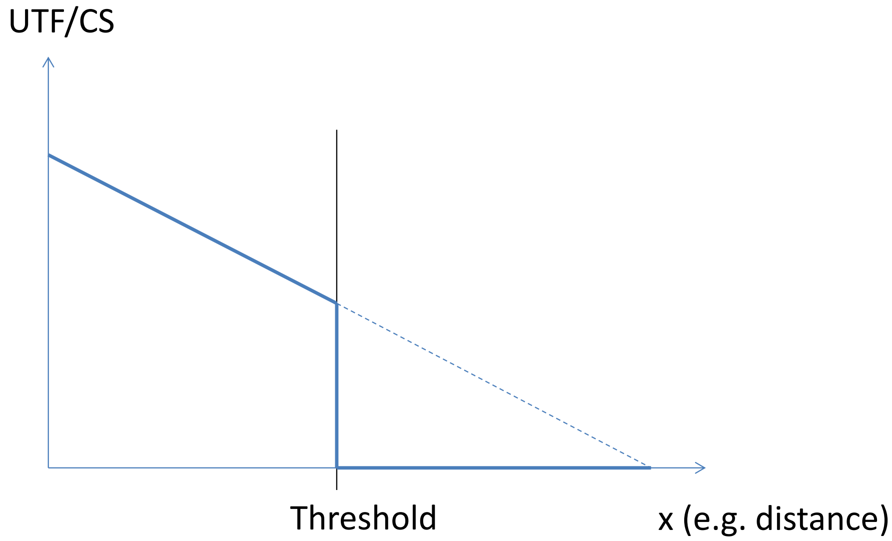
Practically, therefore, no interpretation can be favored. Theoretically, the discussion is equally undecided. As shown by Tversky (1972), decision problems with a large number of alternatives are associated with *non-compensatory* decision behavior. This means that, when facing a complex decision situation, many alternatives are eliminated by the decision-maker on the basis of a limited information search and evaluation. It follows that routine or habitual behavior is not necessarily derived from *extensive* decisions, but can also be the result of preceding decisions that were guided by heuristics. In either case, the routine behavior is the result of preceding decisions. In other words, a sequential learning process is present. Modeling a *sequential* (potentially non-compensatory) process using a *simultaneous* utility-maximizing (i.e. compensatory) model is a strong argument for the purely statistical interpretation of discrete destination choice models. On the other hand, for choices with smaller choice sets, the underlying premise, i.e., the behavioral interpretation is powerful. Where exactly, when ranging from few-alternatives to many alternatives problems, the common premise breaks down is unclear, but a very interesting subject for future research, potentially guiding to new improved stories and models.

5.3 Probabilistic Choice Set Model Estimation

As argued above, for behavioral soundness as well as practical computability, individual choice sets usually need to be restrained. Cutting choice sets essentially means setting the eliminated alternatives to a probability of zero or the utility to minus infinity as shown in Cantillo and Ortúzar (2006, Figure 3, p.683) and in Figure 5.3. Goal of choice set estimation thus is to identify these alternatives with a very low choice probability anyway and then removing them (and only them) from the choice set. The promising approach of Zheng and Guo (2008), doing this, is tested. From this model, choice set distance thresholds can be extracted, which theoretically could also be applied in MATSim for defining choice sets.

The tested approach is structurally related to the well-known nested logit model commonly applied in destination choice modeling (Domencich and McFadden, 1975; Sobel, 1980; Fotheringham, 1986), where the two approaches are convertible by using geographical zones as nests.

Figure 5.3: Constrained choice sets: Cutting choice sets corresponds to setting the removed alternatives to zero utility



5.3.1 Model Specification

The empirical setting in Zheng and Guo (2008) is based on TAZ, while, here, individual stores are used. Estimation is done for individual trips derived from the Swiss Microcensus 2005 and 2010.

Similar to Section 5.1.1, coefficients β for store size, price level and additional linear distance to the store are estimated. Additionally, the density of stores for a constant additional travel distance d_{add} of 1km area is incorporated. Attributes are given as x_n. Person attributes y_n are the person's age and household income. Corresponding coefficients, including a constant, are labeled α.

The probability function for alternative i and person n (adapted from equation (16) in Zheng and Guo (2008)) is given as follows.

$$P_n(i) = \sum_{l=1}^{L} \left\{ \frac{e^{\beta' x_{n,i}}}{\sum_j e^{\beta' x_{n,j}}} \cdot \left[\Phi\left(\frac{\theta_{n,l+1} - \alpha' y_n}{\delta}\right) - \Phi\left(\frac{\theta_{n,l} - \alpha' y_n}{\delta}\right)\right]\right\} \quad (5.1)$$

Zheng and Guo (2008) is based on zones, where $\theta_{n,l}$ gives the distance from the origin zone to zone l. Zones are ordered according to distance, with L being the index of the zone farthest away from the trip origin. This approach is adopted here for different additional travel distance bands as follows

$$\theta_{n,l} = \Delta d_{add} \cdot l,$$

with $\Delta d_{add} = 500\ m$.

The expression in brackets [.] gives the probability that the choice set distance threshold actually is between $\theta_{n,l+1}$ and $\theta_{n,l}$. Φ is the cumulative standard normal function.

The distance threshold for the choice sets is given as

$$t_n^* = \alpha \cdot y_n + \epsilon_n$$

with $\epsilon_n = N(0, \delta)$. The spread parameter δ is estimated as well.

Parameters are estimated by maximizing the following function with the maximum likelihood method.

$$f = log\left(\prod_{n=1}^{N} P_n(choice)\right) \quad (5.2)$$

where N is the population size.

The estimation code is set up as an object-oriented MATLAB script.

Careful parameter scaling and choice of initial values is important such that the values of the exponential functions in above equation and the natural logarithm of the maximum likelihood method remain in feasible regions. The following scaling has shown to work well:

ζ_{const}	:	1.0
$\zeta_{density}$	:	0.1
ζ_{age}	:	0.01
ζ_{income}	:	0.1
$\zeta_{addDistance}$	:	0.001 (conversion to km)
ζ_{size}	:	0.1
$\zeta_{price\ level}$	:	0.1
ζ_{δ}	:	1.0

Constraining of the optimization is only done for δ (being a positive value). Equation 5.1 is adapted from Zheng and Guo (2008) by removing the denominator as the common usage of cumulative standard normal is expected to make this normalization obsolete.

Estimation results are reported in Section 8.3.2.

5.4 Choice Sets Survey

To get further insights on the choice set problem a Web-based survey tool was created by Horni et al. (2011a). For the time being, it is implemented and parameterized for Zurich as an example, but it was designed to be easily adaptable for other regions. As reporting on destination choices is known to be very challenging for the respondents (Thill, 1992) the tool was designed to provide consistent support of a graphical map-based survey method.

Revealing first empirical insights on customers' store visiting frequencies and store awareness was the main goal of the survey.

5.4.1 Survey Design Summary

Details of the survey design are provided by Horni et al. (2011a). Here, a summary is given. The survey tool's scope was the workday shopping trip in the city of Zurich with a purchase amount greater than 20 Swiss francs. Only persons living in Zurich were surveyed. To limit the effort, only grocery shopping was investigated.

As depicted in Figure 5.4, the survey consisted of five sections and an entry and exit page. It was map-based and made extensive use of *Google Maps* and its application programming interface (API) (Google, 2010a,b) (for an example store see Figure 5.5). The survey language was German.

Figure 5.4: Survey overview

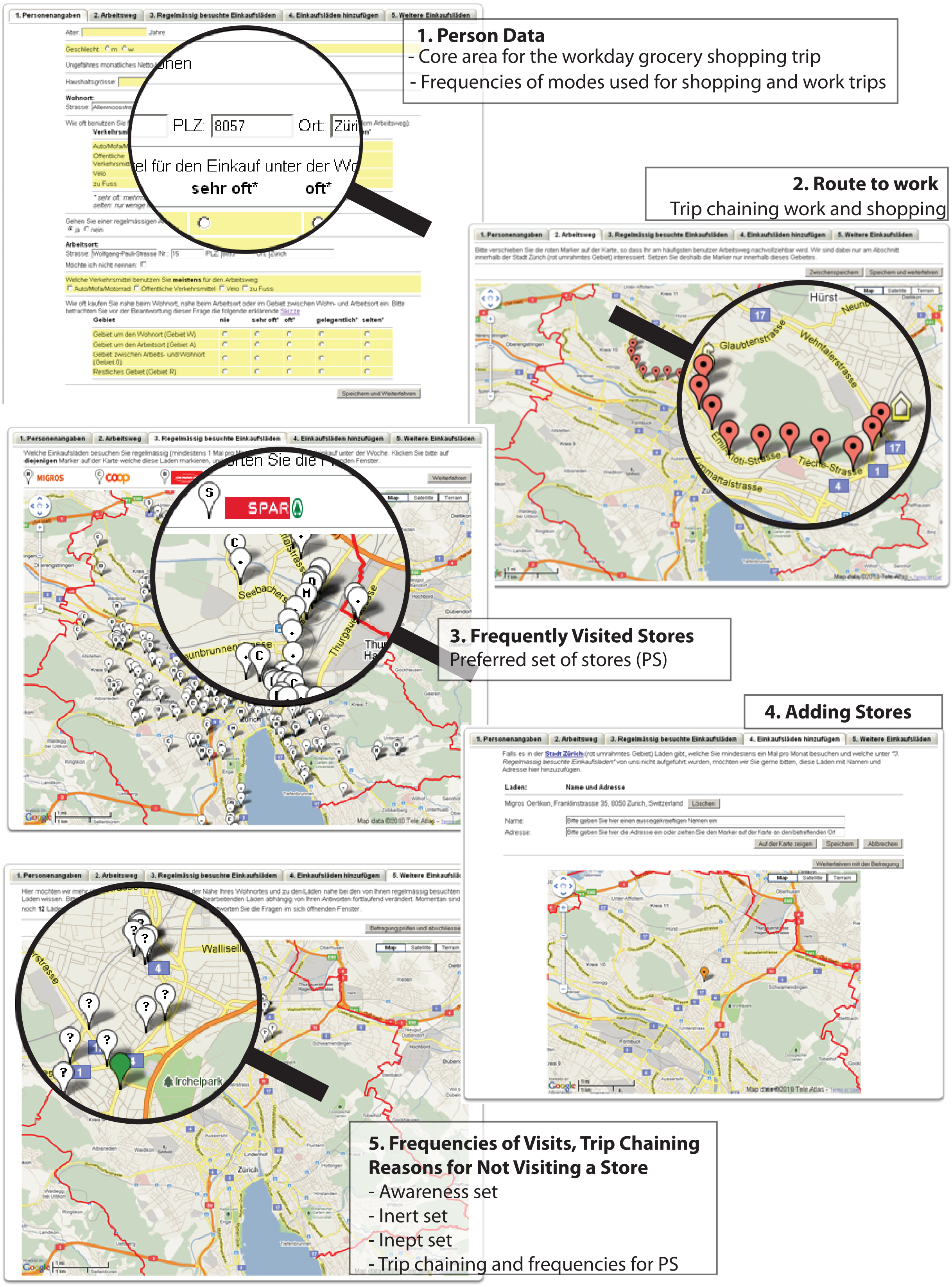


Figure 5.5: *Google street view* perspective of a store site and survey questions (German)

Dhillons Quartiermarkt

Kennen Sie diesen Laden?

- (•) ja ( ) nein

Wie regelmässig haben Sie diesen Laden in diesem Jahr besucht?

- (•) nie
- ( ) sehr oft (mehrmals pro Woche)
- ( ) oft (ca. 1 Mal pro Woche)
- ( ) gelegentlich (wenigstens 1 Mal pro Monat)
- ( ) selten (wenigstens 1 Mal pro Jahr)

Welches sind die Gründe, warum Sie diesen Laden selten oder gar nie besuchen (Mehrfachantworten möglich)

- [ ] Weiss nicht, habe ich mir noch nicht überlegt
- [ ] Mir wichtige Produkte fehlen
- [ ] Unpraktisch, weil für mich keine Umsteige- oder Endhaltstelle (Reise mit dem ÖV)
- [ ] Ist mir zu weit weg / liegt nicht an meinem Weg
- [ ] Schlechtes Preis-Leistungsverhältnis / zu teuer
- [ ] Schlechte Parkiermöglichkeiten
- [ ] Einkaufsatmossphäre gefällt mir nicht (Produktpräsentation, Platzverhältnisse, Personal etc.)
- [ ] weitere Gründe:

Speichern und Fenster schliessen

The two main sections of the survey asked about the frequently visited stores (a.k.a. the evoked set), the actual visiting frequencies and reasons against visiting stores in the vicinity of either home location or the stores actually visited, which gives indications about the inert and inept set. A base for future estimation of multi-set models, as described in Section 5.2.1, was thus provided.

5.4.2 Descriptive Analysis and Comments

Due to organizational reasons, the survey ran in two waves, with two pre-tests. The first wave was conducted around beginning of 2011 and the second at the end of 2012. In the first wave, 300 letters were sent out with 34 respondents, 22 undeliverable ones and 4 drop-outs. In the second wave, 599 letters were sent out, with 66 persons completing the survey, 92 undeliverable ones and 6 drop-outs. An incentive of 20 Swiss Francs for completed surveys was offered. Both times, no reminder was sent.

In total, 100 respondents are achieved, which means a total response rate of 11.3%.

The computed response burden according to Axhausen and Weis (2010) is approximately 189 points, where this value substantially varies as the survey is adaptive according to previous answers. The response rate compared to the response burden is comparatively low.

For specific analyses, up to 89 persons can be used. However, after restrictively filtering persons with missing data in terms of age, income or gender, and missing information about the awareness of stores, only 42 persons remained. Descriptive analyses and a comparison with the Microcensus 2005 (derived from Fröhlich et al. (2012, p.62)) are given in Figures 5.6, 5.7 and 5.8, showing an oversampling of low-household-income persons and an under-sampling of middle-household-income persons, maybe partly caused by the monetary incentive. Designing the survey as a web-survey was expected to create an age-bias, due to the expectedly higher share of non-computer owners amongst elderly persons. To reduce this bias it was planned to offer home support. However, this support was not needed and the age shares do not show a substantial underestimation of the oldest age class compared to Microcensus.

First survey analyses are given in Section 8.3.3.

5.5 Model Application

Goal of this section is making the simulation ready to consistently handle more attributes in the utility function than used until now. As shown in Sec-

Figure 5.6: Descriptive analysis: Gender

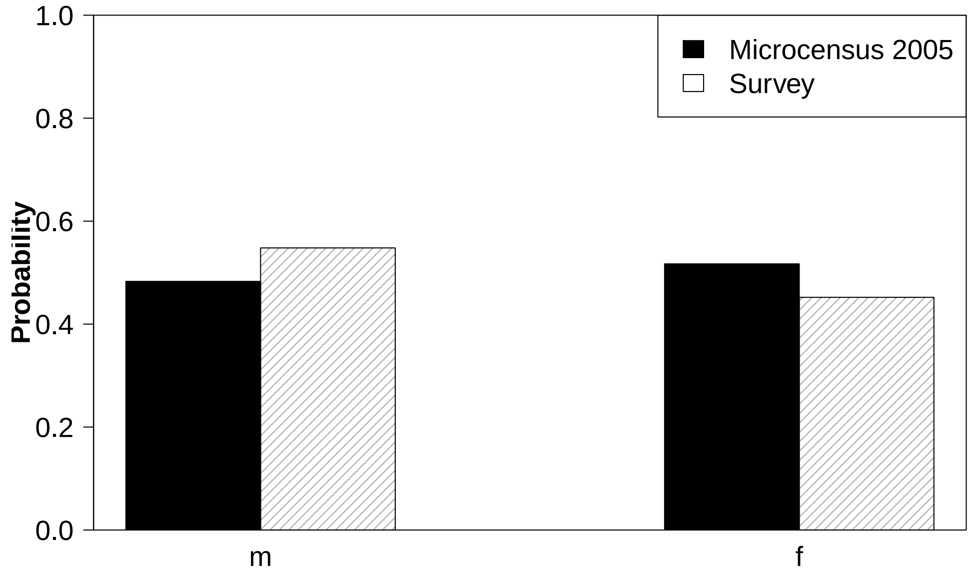


Figure 5.7: Descriptive analysis: Age

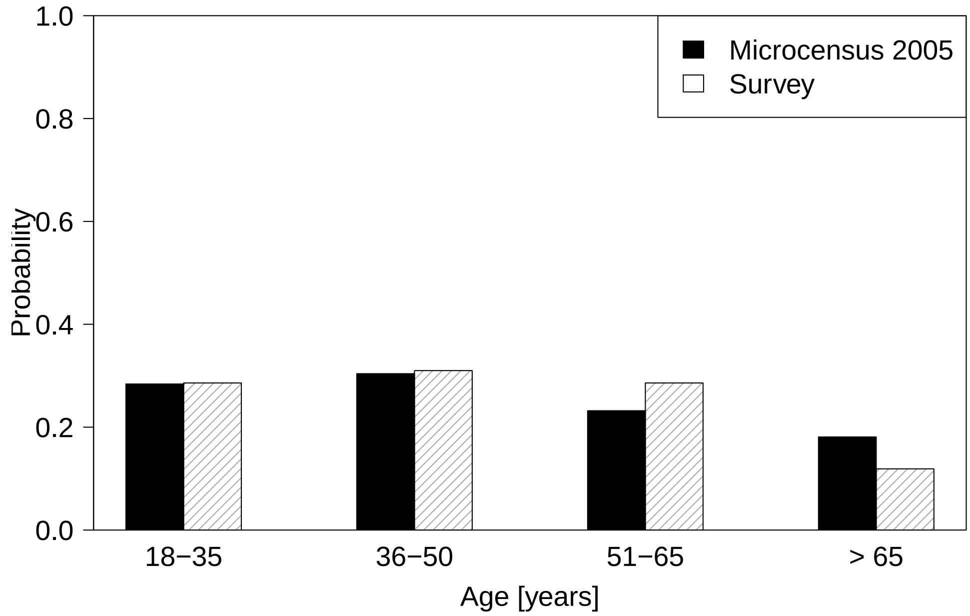
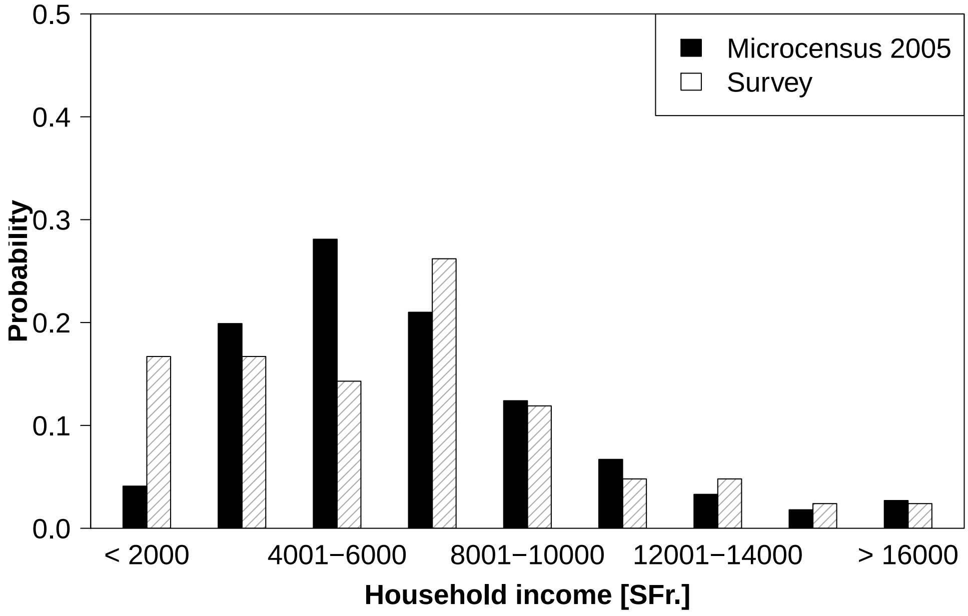


Figure 5.8: Descriptive analysis: Household income



tion 8.3.2, the conceptual problems of the tested probabilistic choice set model do not allow to use them for application in MATSim. This means that no method to endogenously generate destination choice sets is available yet. Thus, application is based on the BIOGEME models described at the beginning of this chapter in Section 5.1.1. These estimations can be included in the base model (Section 4.2) as follows. The term describing activity utility is extended by $\beta_{size} \cdot size$ and $\beta_{price\ level} \cdot price\ level$ yielding

$$U_{act,q} = U_{dur,q} + U_{late.ar,q} + \beta_{size} \cdot size + \beta_{price\ level} \cdot price\ level + \epsilon\ ,$$

For computation of the search space, the maximum error term $\epsilon_{pq,max}$ is used up to now. For the extended model, instead, the maximum destination score s_{max}, including size and price level, has to be used. Attributes, that do not change over the course of iterations can be handled in a straightforward manner in the pre-processing step, where s_{max} is computed. For attributes, that actually do change in the iterations, the value that creates the maximal search space needs to be considered in the sense of an upper boundary for the search space. An example for such an attribute is facility crowdedness influencing the activity score as shown in Chapter 7.

The technical application of the estimated parameters is described below.

5.5.1 Configuration and Parameters

The application of estimated parameters in the MATSim utility function is not straight-forward. Estimated models are often linear and dependent on travel distance rather than on travel time as this information is usually not available from surveys. On the other hand, the standard MATSim utility function is logarithmically dependent on activity duration and linearly dependent on travel time. To include the estimated parameters a conversion between estimated utilities $utils_{estimation}$ and their counterparts in MATSim $utils_{MATSim}$ is required, meaning that, a suitable conversion factor between linear travel distance utility and non-linear activity duration utility needs to be found. Obviously, this can only be approximate.

MATSim activity performing utility is defined as shown in Section 2.4.1.

$$U_{dur} = \beta_{dur} \cdot t_{typ} \cdot \ln(t_{dur}/t_0) \ [utils_{MATSim}]$$

where β_{dur} is marginal utility of activity duration for its typical duration $t_{typ,q}$. $t_{0,q}$ is the duration for which utility starts to be positive. Clearly, due to constraints not necessarily all activities can be performed up to their typical duration. However, for this calculation, it is a plausible assumption, where β_{dur} can then be set to the (individual) value-of-time (Charypar and Nagel, 2003). The logarithm can then be interpreted as a weighting factor $v(t)$ of the value-of-time leading to

$$U_{dur} = \beta_{dur} \cdot t_{typ} \cdot v(t) \ [utils_{MATSim}].$$

The conversion of estimated parameters for application in MATSim $q_{estimation \rightarrow MATSim}$ is as follows. Estimation utilities $utils_{estimation}$ can be converted to their counterparts in MATSim $utils_{MATSim}$ as a function of $v(t)$

$$utils_{MATSim} = q_{estimation \rightarrow MATSim} \cdot utils_{estimation}.$$

The investigation is based on the standard MATSim parameter set (Balmer et al., 2009, 2010; Charypar and Nagel, 2005, 2003) with $\beta_{dur} = +6.0 \left[\frac{utils_{MATSim}}{h}\right]$ and $\beta_{trav,car} = -6.0 \left[\frac{utils_{MATSim}}{h}\right]$. For converting the estimated parameters, the estimated distance parameter $d_{add} = -1.59 \ \left[\frac{utils_{estimation}}{km}\right]$ is

Table 5.2: Converted parameter values

$\nu(t)$	$q_{estimation \rightarrow MATSim}$	$\beta_{price\ level}$	β_{size}
1.0	0.50	0.18	-0.17
2.0	0.75	0.27	-0.25
3.0	1.01	0.36	-0.34
4.0	1.26	0.44	-0.42
5.0	1.51	0.53	-0.50

used as follows.

$$q_{estimation \rightarrow MATSim} = \frac{(-6.0 \cdot \nu(t) - 6.0) \left[\frac{utils_{MATSim}}{h}\right]}{-1.59 \left[\frac{utils_{estimation}}{km}\right] \cdot \bar{v} \left[\frac{km}{h}\right]} \left[\frac{utils_{MATSim}}{utils_{estimation}}\right]$$

The nominator is composed of -6 $utils_{MATSim}$ per hour for traveling and another $-6 \cdot \nu(t)$ $utils_{MATSim}$ for opportunity costs, which is the lost time due to traveling (loss of $t_{perform}$). For converting time to distance, $t\bar{v} = d$ is used with an assumed average speed $\bar{v}$. Again, this is approximate. The estimation combines different travel modes, thus defining $\bar{v}$ is not trivial. A range of coefficients is thus tested here (see below), where mode-specific estimations have to be performed in future. Assuming $\bar{v} = 15 \left[\frac{km}{h}\right]$ following conversion factors are computed.

$$q_{estimation \rightarrow MATSim} =$$

$$\frac{(-6.0 \cdot (\nu(t) + 1.0)) \left[\frac{utils_{MATSim}}{h}\right]}{-1.59 \left[\frac{utils_{estimation}}{km}\right] \cdot 15 \left[\frac{km}{h}\right]} \left[\frac{utils_{MATSim}}{utils_{estimation}}\right]$$

$$= 0.252 \cdot (\nu + 1) \left[\frac{utils_{MATSim}}{utils_{estimation}}\right].$$

Using the model 0 (Table 8.2) and applying $q_{estimation \rightarrow MATSim}$ to $\beta_{price\ level}$ and β_{size} for different values of $\nu(t)$ yields the MATSim size and price level coefficients parameters shown in Table 5.2.

As mentioned above, the calculation is approximate; different reference values are thus provided for different values of $\nu(t)$ and in the experiments reported in Section 8.3.4, several of these values are tested. Logarithmic and time-based estimations are required in the future.

5.5.2 Imputation of Missing Store Attribute Values

Missing values are a common problem for large-scale scenarios. As a first approach, one could, per iteration, only consider attributes in the utility function, which are available for the complete choice set. If an attribute is missing for at least one alternative, recursively, a simpler utility function is applied. However, this approach is complex as choice sets may change during the iterations and thus the utility function changes, which makes convergence an issue. As a second approach, one could draw the unknown values from observed distributions, where uncertainty is handled by doing ensemble runs. This very expensive approach is approximated here, as a third approach. If the actual value is unknown the average attribute value is used, computed for the universal choice set.

Chapter 6

Variability and System Specification

Microsimulation system design as well as concrete microsimulation studies require specification of the measures of interest [1] . For microsimulation design, they need to be general and broadly available, count data are an example. For specific experiments and purposes other measures might be added; Kitamura (1996), for example, lists measures relevant for emissions modeling. Considering their scale is important for model development, but a certain lack of research exists in this regard. Nagel and Axhausen (2001, Section 2.2), for example, say: "Another question regarding scales is which scale is necessary to answer which question. There is wide-spread intuition but currently little hard knowledge. Rules-of-thumb, such as to include one level of resolution below the level of interest, are just rule-of-thumb."

Scale of the measures of interest is also relevant for results production, as different scales or resolution levels usually lead to different levels of *variability* and, thus, to different study costs in terms of required computation effort. Transport microsimulations are usually stochastic. Randomness is, for example, introduced by the error terms of discrete choice models, a common component in utility-based microsimulations. This leads to random variability in results. Parameters or population statistics, such as averages, thus, need to be estimated by random sampling. Microsimulations are thus essentially a sampling tool (Wolf, 2001), where one run represents one sample unit (in statistical terms one *realization of a random variable*). This makes clear, that the whole toolbox used for other statistical methods, must be applied also here. As a first step, required sample size—here, the number of runs—to ensure a given confidence in the

[1] Formal system specification (and verification) is discussed e.g., in Fisher and Wooldridge (1997); Bourahla and Benmohamed (2005).

calculated averages needs to be calculated. In this context, variability is often seen as something tedious, because higher variability leads to larger minimal sample sizes, with usually relatively high costs per sample point. But, this view falls short. Clearly, unobserved variability—modeled as random variability—should be replaced to the extent possible (by explaining it). However, when looking at the very large proportion of temporal variability actually present in simuland (Figure 8.16), a substantial part of this variability is—even if it was actually explicable—much more efficiently handled by including randomness, as model complexity would be prohibitively large otherwise. In this sense, microsimulations are a tool to capture these real-world fluctuations (Newman and Barkema, 1999, p.11), (Esser and Nagel, 2001, p.704). This means also that the focus should not only be on averages, but also on variance incorporated in the calculations of statistical confidence. The following sections broaden the microsimulation variability analysis. A closer look at temporal variability is presented in Section 6.2. Up- and downsampling for computational efficiency is discussed in Section 6.3.

6.1 Microsimulation Variability Analysis

Multiple possibilities to categorize microsimulations variability exist; some categories are described by Horni et al. (2011c). Often a distinction between endogenous (model) variability and exogenous (input) variability is made. Equally suitable one can distinguish systematic and random variability. The experiments reported below mainly focus on *random, endogenous* model variability. Random variability stems from inherently random choices and from actually systematic choices not recognized as systematic by the modeler.

Similar analyses were done in a few other microsimulation studies: Veldhuisen et al. (2000a) for RAMBLAS, Lawe et al. (2009); Ziems et al. (2011) for TRANSIMS, Castiglione et al. (2003) for the San Francisco Microsimulation Model, Cools et al. (2011) for FEATHERS, and Hackney (2009); Horni et al. (2011c,b) for MATSim. The investigations focus on the required number of microsimulation runs to reach "stable results". Random seeds are mutated, while inputs are held constant. Consensus is, that—for the measures, and their resolutions and the choice dimensions analyzed—random variability is relatively small, and consequently, only small numbers of simulation runs are required for reliable results. Further interesting papers on microsimulation variability are (Hale, 1997; Gibb and Bowman, 2007; Vovsha et al., 2002; Milam and Chao, 2001).

6.1.1 Aggregation and Random Variability

For variability analysis, aggregation is very important, as it defines the level of resolution. Although, aggregation is very prominently used in transportation science, it is mathematically not strictly defined and means essentially *composing parts to a whole*, which can be done by *averaging* or *summing* up the parts. We discuss both operations.

It is intuitively clear, that aggregation helps reducing variability and, thus, the number of required random runs or sample points. However, as far as we know, no statistical law perfectly explains this mechanism. A weak relation can be established to the *Law of the Large Numbers* but this law focuses on the average and not on the variance. In the field of filtering and smoothing, variance reduction techniques by moving averages or kernel averages are well-known (Vucetic et al., 2000), (Perrone, 1993, p.27); however, a concise mathematical explanation of the underlying processes is also missing there.

Here, we try to expand the intuition a little further by means of sampling theory and a small example.

As stated by sampling theory, a sampling process is associated with a sampling error. We assume an arbitrary probability distribution given by the density function $f(x)$ with finite mean μ and *finite variance* σ^2. Let us, as an example, focus on the mean μ; then the standard error or sampling error σ_s is:

$$\sigma_s = \frac{\hat{\sigma}}{\sqrt{n}}$$

where $\hat{\sigma}$ is the estimated sample standard deviation of $f(x)$ and n is the sample size. For a derivation of this formula see Hutchinson (1993) or Horni et al. (2011c, p.4).

This error appears in the more common formula for confidence intervals: The confidence interval CI for the parameter θ of $f(x)$ is usually given as:

$$CI = [\hat{\theta} \pm \psi]$$

where $\hat{\theta}$ is an estimate of θ (here $\theta := \mu$) and ψ is the margin of error given as:

$$\psi = q(\alpha)\frac{\hat{\sigma}}{\sqrt{n}}$$

or, when inserting the sampling error:

$$\psi = q(\alpha)\sigma_s$$

$1 - \alpha$ is the confidence level, and $q(\alpha)$ is the α-quantile of $f(x)$. Clearly, the confidence interval is broader for small sample sizes n and higher *population variability* estimated by $\hat{\sigma}$. The sample size (here the number of runs) is specified by the modeler.

If resources only allow to perform few simulation runs, the population variability must be lowered somehow to reach the same level of confidence. One way of doing this, is by aggregation as shown by the following generic example.

Let us assume that decision makers face two alternatives. The choice of person i for one of these alternatives can be described with a Bernoulli variable X_i which takes the values 1 for one alternative and 0 for the other alternative. The choice probability for the first alternative shall be p, for the other alternative $1 - p$. The mean is $\mu_i = p$ and the standard deviation is $\sigma_i = \sqrt{p(1-p)}$.

For an aggregate of $\tilde{n}$ decision-makers, each described by X_i the following holds.

6.1.1.1 Mean of an Aggregate

The *mean* of this aggregate is a random variable X_{avg} with $\mu_{avg} = \frac{1}{\tilde{n}}\tilde{n}p = p$ and standard deviation

$$\sigma_{avg} = \sqrt{Var\left(\frac{1}{\tilde{n}}\sum_{i=0}^{\tilde{n}} X_i\right)}$$

Assuming independent choices with $Cov(X,Y) = 0$ this gives:

$$\sigma_{avg} = \sqrt{\frac{1}{\tilde{n}^2}\sum_{i=0}^{\tilde{n}} \mathrm{Var}(X_i)}$$

$$\sigma_{avg} = \sqrt{\frac{1}{\tilde{n}^2}\tilde{n}\,\mathrm{Var}(X_i)}$$

$$\sigma_{avg} = \sqrt{\frac{1}{\tilde{n}^2}\tilde{n}p(1-p)}$$

$$\sigma_{avg} = \frac{\sqrt{p(1-p)}}{\sqrt{\tilde{n}}} = \frac{\sigma_i}{\sqrt{\tilde{n}}}$$

The standard deviation of a single person's decision is σ_i. The standard deviation of an aggregate of decisions is smaller by $\tilde{n}$, i.e., variability decreases with aggregates' size, meaning that fewer random runs n are required to reach a given error level for the aggregate than for an individual person.

6.1.1.2 Sum of an Aggregate

The *sum* of an aggregate is a random variable X_{sum} with $\mu_{sum} = \tilde{n}p$ and standard deviation

$$\sigma_{sum} = \sqrt{Var\Big(\sum_{i=0}^{\tilde{n}} X_i\Big)} = \sqrt{\tilde{n}p(1-p)}$$

A sum of Bernoulli trials is described by the Binomial distribution. Showing that the required number of runs is reduced with larger aggregates for sums is more complicated than for aggregates' averages. The variance of the sum grows linearly with $\tilde{n}$. The standard deviation of this sum grows with $\sqrt{\tilde{n}}$. However, standard deviation can be normalized with the estimated parameter using the following argument. When defining a confidence interval for a population statistic, the margin of error ψ is reasonably chosen *relative* to the this statistic. In other words, the margin of error is given as a relative percentage of the estimate.

Here, normalizing the standard deviation by the mean (to be estimated) gives:

$$\sigma_{sum,normalized} = \frac{\sigma_{sum}}{\mu_{sum}} = \frac{\sqrt{\tilde{n}p(1-p)}}{\tilde{n}p} = \sqrt{\frac{1-p}{\tilde{n}p}} \tag{6.1}$$

The normalization for an *individual* decision described by X_i gives:

$$\sigma_{i,normalized} = \frac{\sigma_i}{\mu_i} = \sqrt{\frac{1-p}{p}}$$

Joining the last two equations gives:

$$\sigma_{sum,normalized} = \frac{\sigma_{i,normalized}}{\sqrt{\tilde{n}}}$$

It can be seen that, with respect to the mean, relative normalized standard deviation $\sigma_{sum,normalized}$ decreases for the aggregates compared to individual decisions described by $\sigma_{i,normalized}$.

Concluding, the required number of runs n, due to reduced population variability, is decreased by increasing aggregate size for both the average and the sum. Clearly, the applicability of these statements is reduced by correlation between the observations.

Extensive aggregation to reduce variability might, nevertheless, be inefficient in situations where model aggregates are much larger than naturally observable aggregates (when aggregating over a city and its surrounding agglomerations for example). A stratified approach strictly guided by empirics might be more productive.

6.1.2 MATSim Variability

Until recently, the utility function of MATSim was deterministic, i.e., it did not contain random error terms. Now, as part of the recent destination choice integration for discretionary activities, the random error terms have been added (see Section 4.2), potentially adding large variability to MATSim.

Endogenous MATSim choice dimensions, contributing to inter-run variability, currently consist of time (Balmer et al., 2005), route (Lefebvre and Balmer, 2007), mode, and destination choice for discretionary activities. Besides, the random error terms also the co-evolutionary algorithm adds some randomness as it essentially assigns limited resources to persons in a *random* manner. This means, for example, that two identical persons with the same origin and destination may end up with different routes or start times, according to the random order in which they undergo the replanning. The meaning of this implicitly added variability is not yet fully understood in MATSim.

The experimental results for MATSim random variability are shown in Section 8.4.

6.2 Temporal Variability

Transport system temporal variability is substantial. Its driving forces span the whole range from the decision maker's preferences and needs to the alternatives' attributes as well as the choice situation, for example season or weather conditions. Analyses of temporal variation are Hanson and Huff (1988b); Buliung et al. (2008); Pas and Koppelman (1987); Jones and Clarke (1988); Huff and Hanson (1986); Kitamura et al. (2006); Susilo and Kitamura (2005); Schlich (2001); Schlich and Axhausen (2003a); Pas (1988); Hanson and Huff (1982, 1988a); Golledge (1970); Chikaraishi et al. (2010); Burnett (1977); Axhausen et al. (2002).

Different approaches to model temporal variability can be imagined. As illustrated in Figure 6.1, approach (a) and (b) employ a cross-sectional model, while approach (c) uses a longitudinal model for computation of temporal averages.

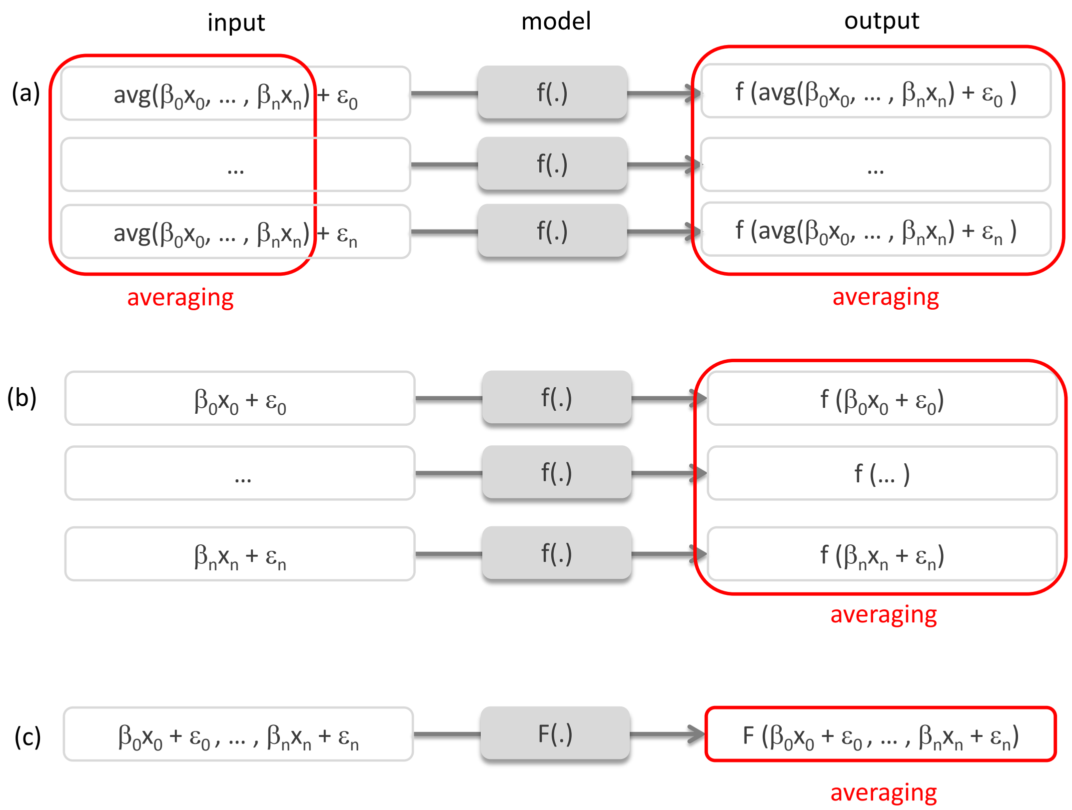
Method (a) feeds average systematic input ($avg(\beta_i x_i)$ for day i) to the model $f(.)$ and runs it multiple times with different random error terms ϵ_i, in other words with different random seeds. In principle, this is a valid procedure, however, as for nonlinear models $f(.)$ usually $f(y_0 \otimes y_1) = f(y_0) \otimes f(y_1)$ does *not* hold true, it cannot be expected that procedure (a) leads to the correct average of multiple days i measured in simuland. As the variability generating mechanisms of the model and simuland are very different, it is also difficult to interpret the generated variability. The adoption of this approach, probably also has led to a short-coming in past variability analysis. As mentioned before, for most studies, random variability turned out to be a non-issue. But as shown in Figure 8.16 (temporal) real-world variability is actually large. Future studies, thus, should extent the scope by also considering temporal variability actually present in simuland. Concluding, microsimulation variability should be an issue.

Configuration (b) represents a next step toward a longitudinal model. Temporal variability is not only generated by the error term variation, but also by temporal variation in systematic input. This approach is conceptually fully adequate. However, this means that all temporal variations and correlations need to be neatly modeled in pre-processing and feed into the model. This can be a problem in daily modeling practice, which may favors approach (c), a longitudinal model $F(.)$.

Extension of MATSim simulation horizon to one week with varying agents' preferences is done by Horni and Axhausen (2012b). Destination choice was not yet applied in this scenario, thus it is not reported in detail here.

An approximate approach to integrate a longitudinal scenario into a single simulation run by oversampling is discussed in the next section.

Figure 6.1: Temporal variability: β_i are the choice model coefficients and ϵ_i are the random error terms. $f(.)$ is a cross-sectional model and $F(.)$ a longitudinal one. For the computation of average results, the averaging is done at different stages during model application.



6.3 Sampling and Oversampling

In microsimulation practice, very often not the complete population is simulated but a sample of it. In MATSim often 10% samples are drawn. This saves computation time with the hope that approximation error is not too large; systematic analyses are outstanding for MATSim. In this context but going in the opposite direction, an idea often being a topic in informal scientific discussions is oversampling. Idea is to "*change the sampling rate instead of the number of runs. [A] sampling rate greater than 1 is equivalent to averaging across multiple runs.*" says Bradley (2005). Walker (2005) who investigates a household microsimulation, writes that *"one can synthesize more than the full population [in one run] to decrease the sampling error even further [...]"*.

Using an oversampled population in one run means that temporal (i.e., intra-personal) variability is handled by creating clones of a person, where each clone has a different daily activity chain of that original person. The number of clones is thereby chosen according to the frequency with which the activity chain is observed in longitudinal empirical data. In conclusion, different activity chains of a person are superimposed and simultaneously simulated in one single simulation run. This represents efficient sampling, potentially bringing great computational savings. *Efficient sampling strategies or variance reduction techniques*, such as stratified sampling or importance sampling, are well-known in other research fields and have shown to be highly productive (e. g., Marnay and Strauss, 1991; Kahn and Marshall, 1953; Wang, 2008; Oliveira et al., 1989).

Efficiency gain in sampling is assumed as follows. In the sequential approach, in one run, *one* combination of individual day plans is simulated. Due to the large number of agents and the extensive choice space, there is a huge number of potential combinations. In other words, to cover the possibility space of outcomes, potentially, a huge number of runs is required; efficient sampling helps here. The investigated simultaneous method may do this, as it efficiently samples every individual's choice space and combines all these individual samples in one big run. It can be assumed, that the results space is thus also covered efficiently.

The computation costs associated to this run might are substantially smaller than for sequentially simulating (a sample of) the huge number of combinations. Computation costs for an example of n persons choosing between k alternatives are as follows. Sample enumeration of all possible outcomes generates costs $O(k^n)$. Random sampling even increases this costs. Costs for a sample of sequential runs are dependent on variance of the outcomes distribution, the desired accuracy and confidence as discussed above. Oversampling is associated to $O(k \times n)$. If the variance of the outcomes is high, then oversampling might be an efficient solution.

However, although the idea is conceptually appealing, according to our analysis, it has following severe shortcomings.

- *Linearization Error:* Networks are usually characterized by non-linear effects, for which it holds: $g(x_0 \otimes x_1) \neq g(x_0) \otimes g(x_1)$, where $g(.)$ is an arbitrary non-linear function and $\otimes$ is an arbitrary operator. In other words, the oversampling method is accompanied by a substantial linearization error.
- *Emergent Behavior:* Due to non-linear relations microsimulations are often characterized by emergent behavior (Bonabeau, 2002a). It must be expected that the emergent behavior in the single runs is substantially different from emergent behavior for the oversampled run and there is no estimate for the deviation.
- *Self-Interaction:* To a certain extent persons interact with their own clones, which is a process that does not occur in sequential microsimulation runs. However, due to the relatively small number of clones per person, this problem is expected to be negligible.
- *Missing Path-Dependency:* Definitely path-dependency is not captured as a person lives multiple days simultaneously.

In conclusion, the approach, although promising at first sight, has severe limitations speaking against further investigations. Maybe it can help in getting a very first impression about the extent of possible outcomes space. However, due to completely different emergent effects, the range of outcomes might look very different than for sequential runs. Having said that, in the opinion of the author, it is still worth to research *variance reduction techniques* for microsimulations.

Chapter 7

Agent Interactions in Activities Infrastructure and Spatial Correlations in Choices

Customer interaction effects at the activity locations (Section 7.1), such as competition for parking lots (Section 7.1.2), as well as spatial arrangement of alternatives (Section 7.2), play a significant role in customer destination choices. As argued in Section 2.3, microsimulations provide high-resolution and thus potentially high-precision network loading information. This line of argumentation can be extended to infrastructure loading in general, i.e., including both network and activity locations.

As an aside, in Section 7.3 supply side interactions and interdependency of customers' destination choices and retailers' location choices are discussed.

7.1 Person Interaction Effects at Activity Locations

The influence of interaction in *transport* infrastructure for people's route and departure time choice has been recognized early (e. g., Pigou, 1920; Knight, 1924; Wardrop, 1952).

Similarly, it can be reasonably assumed that agent interaction in *activities* infrastructure affects destination choice (Axhausen, 2006). Marketing science provides ample evidence that agent interactions influence utility of performing an activity, where it can have both, positive or negative influence. (Baker et al., 1994, p.331), (Eroglu and Harrell, 1986; Eroglu and Machleit, 1990; Eroglu et al., 2005; Harrell et al., 1980; Hui and Bateson, 1991; Pons et al., 2006). Presence of other people at recreational places such as bars, discos or party locations usually contribute positive utility, whereas competition for parking lots (see also Section 7.1.2) or crowdedness in shops, clearly, reduce utility (Albrecht, 2009, p.119ff).

In transport microsimulations, demand-supply equilibration is usually limited to the transport infrastructure. Rare counter-examples are Vovsha et al. (2002); Horni et al. (2009a). Consideration of *positive* interaction effects in microsimulation is not known to us; for MATSim this was investigated by Stahel (2012). For estimation, demand-supply equilibration is usually completely neglected, but de Palma et al. (2007) does it for residential location choice and Vrtic (2005) for route and mode choice.

7.1.1 A First Approach: A Singly-Constrained Dynamic Model

In Horni et al. (2009a) a singly-constrained model is presented that introduces competition for space-time slots on the activity infrastructure. The actual load is coupled with time-dependent capacity restraints for every activity location and incorporated explicitly into the agent's destination choice process as detailed below.

Activity location load, computed for time bins of 15 minutes, is derived from events that are delivered by the Mobsim. The load of one particular iteration combined with time-dependent activity location capacity restraints is considered in the agents' choice process of the succeeding iteration. In detail, this means that the utility function term $U_{act,q}$, described above, is multiplied by $max(0; 1 - f_{load\ penalty})$ penalizing the agents dependent on the load of the location they frequented. $f_{load\ penalty}$ is a power function, as this has shown to be a good choice for modeling capacity restraints (remember that the well-known cost-flow function by U.S. Bureau of Public Roads (1964) is a power function). To introduce additional heterogeneity regarding the activity locations, an attractiveness factor $f_{attractiveness}$ is introduced that is defined to be logarithmically dependent on the store size given by the official census of workplaces.

Likewise for demonstration purposes, capacity restraints are exclusively applied to shopping locations, where in principle leisure activity locations could be handled similarly. However, deriving capacity restraints for leisure activity locations is expected to be much more difficult than for shopping locations because data availability is much smaller for leisure locations and capacity restraints vary much more between different leisure locations than between different shopping activities (hiking versus going to the movies might be an illustrative example).

The model allows the assignment of individual time-dependent capacities to the activity locations. For the sake of demonstration, the capacities of all shopping facilities are set equal, where the values are derived from the shopping trip information given in the National Travel Survey of 2005. The total daily capacity is set so that the activity locations located in the region of Zurich satisfy the total daily demand with a reserve of 50%. In detail, the capacity restraint function for a location i is as follows:

$$f_{load\ penalty,i} = \alpha_i \cdot \left(\frac{load_i}{capacity_i}\right)^{\beta_i}$$

with $\alpha_i = 1/1.5^{\beta_i}$, $\beta_i = 5$. $f_{load\ penalty,i}$ is the penalty factor for location i as described above.

The simultaneous computation of the score reduction for all agents avoids the last-record problem discussed in Vovsha et al. (2002). Therein, a sequential choice process is proposed where alternatives are removed from the choice set of the later travelers if the locations are already occupied by the earlier travelers. Thereby, the order of the travelers is specified arbitrarily and thus the last-record problem (the last travelers have to travel far to find an available location) is not negligible when modeling heterogeneous travelers.

As expected, our constrained model improves results' quality by reducing the number of implausibly overcrowded activity locations as detailed in Section 8.5.

7.1.2 Parking

An important example of agent interactions is parking search. It generates a significant share of traffic (Shoup, 2005; Young et al., 1991), and, thus, it is identified as a relevant destination choice determinant and essential for any microsimulation. Parking is intensely researched by van der Waerden et al. (2009, 2006); Marsden (2006); Widmer and Vrtic (2004); Anderson and de Palma (2004); Golias et al. (2002); Hensher and King (2001); Gerrard et al. (2001); Baier et al. (2000); Albrecht et al. (1998); van der Waerden et al. (1998); Axhausen et al. (1994); Axhausen (1988); van der Waerden et al. (1993); Glazer and Niskanen (1992); Topp (1991); Axhausen and Polak (1991); Arnott et al. (1991); Polak and Axhausen (1990); Feeney (1989); Miller and Everett (1982); Gillen (1978, 1977);

Maley and Weinberger (2011); Bonsall and Palmer (2004).

Supposedly, the additional travel time, i.e., the search time, is most accurately modeled with a simulation model such as Benenson et al. (2008); Gallo et al. (2011); Thompson and Richardson (1998); Dieussaert et al. (2009); Young (1986); Young and Thompson (1987). In Horni et al. (2012b), we propose a cellular automaton agent-based microsimulation to model parking search with a high temporal and spatial resolution. This stand-alone model is designed for easy integration into MATSim, which is a future task. It is programmed in MATLAB and open-source (LaHowara & Commander Spock, 2013).

A similar approach in the literature is PAMELA (van der Waerden et al., 2002), who links a parking search model with ALBATROSS (Arentze and Timmermans, 2000) and uses a cellular automaton for the parking search. Another related approach is proposed by Kaplan and Bekhor (2011), who base their model calibration on GPS data, as we intend to do for our model in future research. One-week GPS data, suitable for microsimulation calibration, is currently being surveyed and analyzed for Zurich region (Montini et al., 2013, 2012).

The cornerstones of our model are traffic and parking assignment with a cellular automaton-based microsimulation and parking choice modeling adopting a weighted random walk. Agents remember to some extent free parking spaces during driving and adjust the random choice accordingly. Thus, the approach further exploits the agent-based approach by using an agent short-term memory.

The cellular automaton that we implemented is based on Nagel and Schreckenberg (1992), which has shown to be able to predict urban flow patterns (Wu and Brilon, 1997, p.1). The main technical innovation is probably that the CA update process is essentially reduced to iteration *over agents* instead of simply iterating over all nodes, links and cells in every time step. This is achieved by using auxiliary data structures, which dynamically manage agents' positions by means of waiting queues. Obviously, this generates a substantial speed-up.

The model was tested on a small-scale chessboard scenario, and a first verification step was performed for a real-world scenario for the town center of Zurich (see Section 8.5.2). Results show, that the model basically is able to replicate empirical observations, but that future work is still required to have it ready for the Zurich scenario and similar large-scale scenarios.

7.2 Spatial Distribution of Destinations

Besides customer interactions at activity locations, also the spatial arrangement of alternatives influences destination choices. Clustered destinations help minimizing travel effort for multi-purpose, multi-stop shopping trips (see e.g., Bernardin et al. (2009, /p.144), Arentze et al. (1994, /p.89), Arentze et al. (2005); Popkowski Leszczyc et al. (2004); Messinger and Narasimhan (1997); Oppewal and Hoyoake (2004). Also for planned single-purpose shopping, agglomerations might be beneficial due to the reduction of risk of not finding specific products at the chosen location. Agglomerations, thus, usually generate utility beyond the sum of single opportunities (see also Teller and Reutterer (2008); Teller (2008)). Including these effects in a model, and, thus, capturing frequencies at large shopping malls and extensive nightlife areas better, is particularly important for weekend scenarios.

Agglomerations, or in more general, spatial correlations of alternatives, can be modeled by correlated error term models, which can be efficiently estimated with copulas (Bhat and Sener, 2009). Agglomerations can also be treated by inclusion of an explicit agglomeration term τ_{agglo} as discussed in Horni et al. (2012a) and applied in Section 8.6. The later approach is similar to the aggregate model of Fotheringham (1985), extending the gravity model. Further examples considering spatial distribution of destinations are Fotheringham (1985, 1983a,b); Fotheringham et al. (2001); Timmermans et al. (1992); Berry et al. (1962).

7.3 Including Supply-Side Interactions

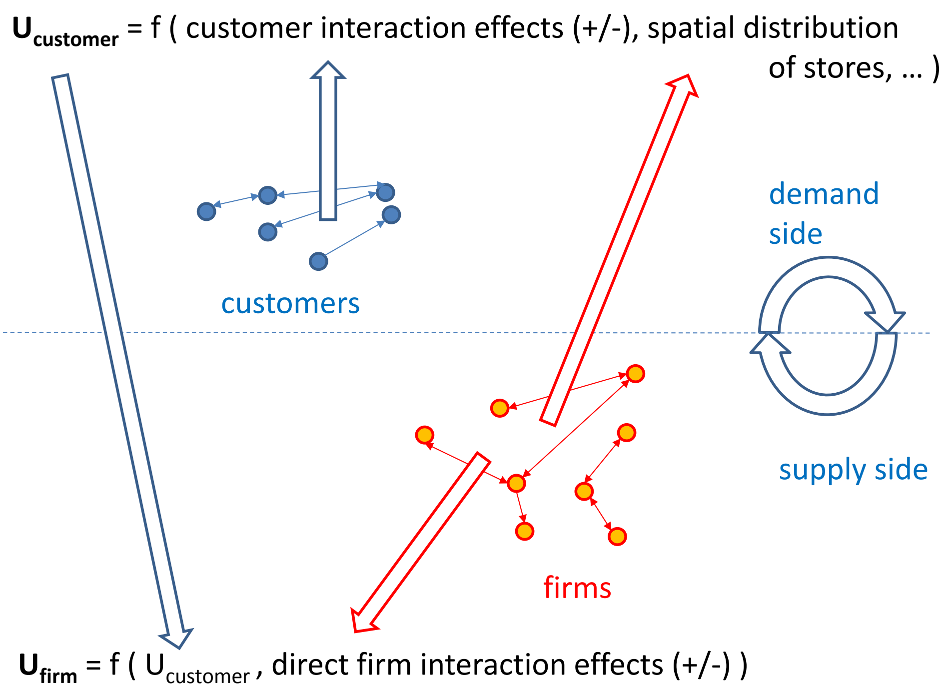
An approach to simulate different types of agents in MATSim, namely customers and retailers, is discussed in Horni et al. (2012a); Horni and Ciari (2011, 2009), where we analyze the combination of the MATSim customer destination choice module and the MATSim retailer location choice module (Ciari and Axhausen, 2011). Methodological progress in coupling demand and supply side is increasingly important as the coupling of MATSim with land-use models (e.g., Urbansim (UrbanSim, 2011)) will be intensified in the near future (Nicolai et al., 2011). Furthermore, and more relevant for this thesis, the MATSim destination choice module is thereby embedded in a larger context.

As illustrated in Figure 7.1, on demand as well as supply side, interactions between actors, i.e., customers and firms exist. The interdependency is observable in spatially correlated decisions and materializes in store agglomerations, a very important topic in economics.

On the supply side, firm location choice is a very important business decision (Fox et al., 2007; Rogers, 2007). It greatly influences business success. At the same time, the decision is overwhelmingly complex. Summaries of the numerous retailer location choice determinants and location search strategies can, for example, be found in Ciari et al. (2008); Bodenmann (2005); Brown (1994); Hale and Moberg (2003); Hernández and Bennison (2000); Hernández et al. (1998); Löchl (2008). Complexity of the choice process is generated, inter alia, by latency of choice determinants and their temporal variability. Profit, for example, being a main optimization objective is—although some indicators might be known—essentially latent and varies with time (see also Maier and Toedtling, 2006, /p.23). Due to their high complexity, location choice strategies are driven by *heuristics*, where the whole range from hard (or rational) to soft (or emotional) criteria is considered (see e.g., Scherer and Derungs (2008, /p.12/23/26/41), (Schmidt, 1980, /p.60)). This usually leads to approximate, suboptimal, but efficiently computable solutions. Furthermore, relocation of stores is associated with very high costs. Thus, search is basically non-iterative, meaning that ex-post optimization (through relocation) is relatively weak. A certain global location choice optimization exists due to business failure (Maier and Toedtling, 2006, /p.33).

All these factors make integration of a *behavioral* location choice in MATSim very difficult, even more as MATSim's time horizon is one (average) day, where location choices are done on a much larger timescale. A *normative* model, however, as proposed in Ciari and Axhausen (2011) fits well into MATSim and can be coupled with destination choice (for a similar approach see Huang and Levinson (2011)).

Figure 7.1: Demand and supply side interaction effects: Customer utility $U_{customer}$ is influenced positively and negatively by interactions with other customers and by the spatial distribution of the stores. Firms' utility (e.g., profit) U_{firm} is dependent, on the one hand, on customers utility. If a store does not generate enough utility for a customer, he or she chooses a different one. On the other hand, firm utility is dependent on direct competition and possibly agglomeration effects generated between firms. In a somewhat simplified perspective, this demand-supply side interdependency guides the spatial arrangement of stores, possibly generating spatial agglomerations.



Chapter 8

Results, Conclusions and Future Work

This chapter summarizes main results of the destination choice improvements described in the previous chapters and reported in more detail in the author's publications. For the specific topics, conclusions are discussed and future work is sketched here, where a more general discussion including future research avenues are presented in Chapter 9.

The results were generated over a time period of six years. Global refactorings concerning many MATSim components have been necessary. In the last 4 years, since the MATSim core-package was created, more than 2200 revisions were made to this package, where a systematic evaluation of the modifications to the scoring module, which is now based on MATSim events rather than the agents' plans, is open. Some reproduced results are thus not perfectly (but structurally) identical with the ones presented in the original papers. In Table 8.1 the software version for results generation is detailed.

8.1 Zurich Scenario

The Zurich scenario, described in technical detail by Horni et al. (2011e), is frequently used in MATSim development but also for projects in Swiss planning practice (e.g., Balmer et al., 2009). Scenario demand is derived from the Swiss Census of Population 2000 (Swiss Federal Statistical Office (BFS), 2000) and the National Travel Survey for the years 2000 and 2005 (Swiss Federal Statistical Office (BFS), 2006). As illustrated in Figure 8.3, a 10% sample of car traffic (excluding cross-border traffic), crossing the area delineated by a 30 km circle around the center of Zurich (Bellevue), is drawn, which results in almost 62'000 agents simulated.

The activity location data set, comprising more than 10^6 home, work, education, shopping and leisure locations, is computed from the Swiss Census of Population 2000 and the Federal Enterprise Census 2001 (Swiss Federal Statistical Office (BFS), 2001). The planning network from the Swiss National Transport Model (Vrtic et al., 2003) consists of 60'492 directed links and 24'180 nodes (see Figure 8.4). A single day is simulated, with 3.35 average number of trips per agent. In total, 25'896 shopping activities and 40'971 leisure activities are performed. Comparable data is available in most countries from official sources, such as censi, national travel diary studies and commercial sources, such as navigation network providers, yellow pages publishers or business directories.

8.1.1 Road Count Data

MATSim focuses on the regular workday. The count data are prepared as follows. A couple of filtering steps are applied: only Tuesdays, Wednesdays and Thursdays are included, while any public holidays are excluded. The days between Christmas and New Year are also filtered out and finally, only count values greater or equal than zero are included. Traffic count data for 2004-2005 from automatic national, cantonal and municipal count data stations (e. g. ASTRA, 2006) are taken into account, resulting in 600 unidirectional links measured for Switzerland and 123 for the center of Zurich (defined as a 12 km radius area around the Bellevue).

8.2 Error Term Runs

Simulation results are shown for the Zurich scenario as an example. The utility function, in particular, the error terms are calibrated with respect to trip distance distributions, taken from the National Travel Survey for the year 2005. Two calibration parameters $f_{shopping}$ (here set to 0.5) and $f_{leisure}$ (here set to 1.15) are applied to the error terms, for which a standard Gumbel distribution, scaled to produce a 1.0 standard deviation, is chosen.

The following two configurations are simulated:

- **Configuration 0:** $U = f(t_{activities}, t_{travel})$, i.e., excluding unobserved heterogeneity, where $f(.,.)$ refers to the standard MATSim utility function described earlier.
- **Configuration 1:** $U = f(t_{activities}, t_{travel}) + f_{shopping\ |\ leisure} \cdot \epsilon$, i.e., including unobserved heterogeneity.

Figure 8.5 shows that calibration of travel distances can be productively performed with error terms; the dramatic underestimation of travel demand (configuration 0) is corrected (configuration 1). Figure 8.6 shows for city center link volumes, in terms of validation, that the median relative error of daily volumes (averaged over the 123 links) is reduced from approximately -45% to approximately -30%. Results are shown for iteration 1000, representing a sufficiently relaxed state for this scenario according to common MATSim practice.

By applying the travel time approximation described earlier, the scenario is computable in reasonable time. It takes roughly 5 minutes per iteration, with 20% replanners doing, either time choice, or combined route and destination choice.

Future work concerns more comprehensive and systematic analysis and calibration of the various parameters.

8.3 Utility Function Extension

8.3.1 MNL Estimation Based on the Universal Choice Set

Four MNL models are estimated with Biogeme; three models include interaction terms. All models use 634 observations, converge, and derive variance-covariance from the finite difference hessian.

Model 0 reported in Table 8.2 shows significant estimates for store size, price and additional distance (as specified earlier) to reach the respective store. Following utility function is used.

$$U_0 = \beta_{addDist} \cdot d_{add} + \beta_{size} \cdot size + \beta_{price\ level} \cdot price\ level + \epsilon$$

Coefficients have the expected sign and plausible values.

In model 1 (Table 8.3), the price level is interacted with household income (scaled with household size) as follows.

$$\beta_{price\ level,income} \cdot \left(1 + \beta_{income} \cdot \frac{income}{household\ size}\right) \cdot price\ level + \epsilon$$

The interaction coefficients are not significant, which is unexpected, given the usually strong influence of income on many areas of life. Additionally, the β_{income} is positive, which is unexpected, as income seems to increase the price sensitivity here.

In model 2 and 3 (Table 8.4, 8.5), interactions of household income and additional distance, respectively age and price level are done, both yielding insignificant estimates.

Although model fit is satisfying—ρ^2 is greater than 0.32 for all models—similarity issues need to be analyzed (Schüssler, 2010; Schüssler and Axhausen, 2007) in future. Possibly, aggregation (Ben-Akiva and Lerman, 1985, p.253) or sampling (Ben-Akiva and Lerman, 1985, p.261) of alternatives may improve results further.

8.3.2 Probabilistic Choice Set Model Estimation

The probabilistic choice set model is estimated for the same data as the linear MNL Biogeme model above. It takes approximately 3 days of computation time. Equation 5.2 needs to be maximized, which is done by minimizing its negative using the MATLAB function *fmincon*. Following parameter estimates result.

α_{const}	:	3.32
$\alpha_{density}$	:	-0.23
α_{age}	:	0.26
α_{income}	:	-0.85
$\beta_{addDistance}$	:	0.26
β_{size}	:	4.12
$\beta_{price\ level}$	:	-3.43
δ	:	2.24

The null model leads to an undefined initial log-likelihood, i.e., setting all coefficients to 0 in Equation 5.2 generates "not-a-number" (NaN). A work-around to estimate the initial log-likelihood is applying random sets of coefficients (see e.g., Steenbergen (2003, p.8)). Here, 10 runs varying the coefficients uniformly in the following ranges is used: $\alpha_{const} = [-4..4]$, $\alpha_{density} = [-1..1]$, $\alpha_{age} = [-1..1]$, $\alpha_{income} = [-1..1]$, $\beta_{addDistance} = [-1..1]$, $\beta_{size} = [-5..5]$, $\beta_{price\ level} = [-4..4]$, $\delta =]0..3]$. This leads to an average initial log-likelihood of -3634. The final log likelihood is -2844 leading to $\rho^2 = 0.22$. The computation of the t-statistics require estimation of every coefficient's sample standard deviation. This can be done by running the model multiple times with different sub-populations. This task is left for future work, due to the huge computation time (3 days for the full population estimation) and the relatively low predictive power of the model.

Besides these restrictions in results' confidence, the data additionally show the following issue. In general a positive distance parameter $\beta_{addDistance}$ is estimated. One reason might be that the choice probability increases with decreasing choice set as alternatives appear in the denominator of the choice probability calculation. Thus, estimation of choice sets underlies minimization. This potentially leads to underestimation of $\beta_{addDistance}$ as an absolute value. In other words, also this approach probably does *not* produce the true preference values.

Nevertheless, a consistent estimation *without exogenous information* may be provided for the non-spatial variables. After scaling the values (with factor $\zeta_{price\ level} = \zeta_{size} = 0.1$ (see Section 5.3.1)), the price level and store size variable (here -0.34 and 0.41) are consistent with the Biogeme estimation reported earlier (-0.33 and 0.35).

8.3.2.1 Spatial Indifference

An interesting line of argumentation, completely resolving the above problem of positive distance parameter, is given in Timmermans (1983, p.450). Based on the postulate of spatial indifference, it is argued that within the “indifference zone, indicated by the consumer’s idea of a reasonable travel time, shopping centers are evaluated only in terms of their non-locational attributes.” Distance is with this conceptualization “no longer considered as a factor which contributes to the overall utility of an alternative, but rather as one of the constraints which define the choice sets of individuals.”

As this postulate is not well-known and controversial, improvement of our model is stopped here. Further steps for approaching choice set specification are given in Section 5.4.

8.3.3 Survey

This section summarizes the survey presented by Horni et al. (2011a). It looks at three things: the persons’ store awareness, their set of frequently visited stores and their common area of shopping. Survey data will be archived according to international standards in our online travel data archive (IVT and ETH Zurich, 2013) to enable future model estimation and further analyses.

The small sample size requires caution in results’ interpretation.

8.3.3.1 Store Awareness

By looking at awareness of close-to-home stores, the continuity of the choice sets can be assessed, being the cornerstone of the probabilistic choice set model tested in the previous section. In the survey, a home-set containing the 10 closest grocery stores around the person’s home location is constructed. The respondents are asked to answer a couple of questions about the stores of this set (see e.g., Figure 5.5).

Figure 8.7(a) shows the number of aware or known stores in the home set. On average, 6 stores out of 10 are known, rendering the continuity assumption, made for the probabilistic choice set model, questionable.

Figure 8.7(b) shows the distance to the farthest aware store in the home set, where mean distance of the farthest aware store is 750m (median = 600m) with a maximum of persons knowing the farthest store between 500m and 750m.

The estimated average choice set distance threshold (in terms of additional distance d_{add}) as given by the probabilistic choice set model estimation is:

$$\bar{t^*} = \alpha \cdot \bar{y} + \bar{\epsilon}$$

$$\bar{t^*} = \alpha_{const} + \alpha_{density} \cdot \zeta_{density} \cdot \bar{y}_{density} + \alpha_{age} \cdot \zeta_{age} \cdot \bar{y}_{age} + \alpha_{income} \cdot \zeta_{density} \cdot \bar{y}_{income} + \bar{\epsilon}$$

yielding

$$3.32-0.23\cdot 18.95\cdot 0.1+0.26\cdot 50.51\cdot 0.01-0.85\cdot 4.12\cdot 0.1+0\,km = 2.67km.$$

This is significantly higher than the above calculated store awareness distances. However, d_{add} cannot be directly compared to the round trip distances for the home set of the survey. Thus, this comparison cannot be more than a very first indication.

8.3.3.2 Frequently Visited Stores

Figure 8.8(a) shows each person's number of frequently visited stores (at least one visit per month) [1]. Figure 8.8(b) shows the respective distribution of the distance to home. The urban setting might explain the short distances, but further analysis is required also here. Figure 8.8(c) shows the number of frequently visited stores for the home set.

The collected data allow for two interesting future investigations. First, the spatial indifference assumption (Section 8.3.2.1) can be further tested. Second, destination choice models not only based on the set of frequently visited stores (a.k.a. preferred set) but also on the awareness set could be estimated, where a natural extension of the standard discrete choice mechanism could be to assign the role of the *observed choice* to the preferred set, where the role of the choice set could be assigned to the awareness set. Practically, this model could be estimated by an exploded logit model or by simulating the *n* decisions of one person with *n* persons, each making one decision. However, this line of modeling is generally subject to the fundamental problem, depicted in Figure 8.9. Decision-making (including preceding learning processes) can be seen as a process during which the decision-maker, starting with the awareness set and ending with the final decision, successively reduces the number of alternatives. For any non-trivial decision-making problem, the more one retrogrades in this process, the more dominant become the random influences. For example, the membership of a store in a person's awareness set can be caused by one single random visit of this person in a bar close to the store in question, whereas the reasons that a store belongs to the person's preferred set are much less random. However, the awareness set is in fact an important component of the decision-making process and, thus, means have to be developed to adequately model these random influences.

[1] One person has no frequently visited store, but manual checking did not detect an error in his or her survey completion.

8.3.3.3 Area of Shopping

The usual area for the grocery shopping act is surveyed also (Figure 8.10). This will help to speed-up the destination choice module as the search-space can potentially be more precisely tailored for shopping activities. At the moment, the search space of the destination choice module equally covers the work and home area, whereas the comparison of the Figures 8.11(a), 8.11(b), and 8.11(c) clearly shows, that most shopping trips are done in the home area.

In conclusion, the survey data allow a couple of analyses, not yet done here, and not yet present in the literature. For publishing survey data and thus enabling their further exploitation, a very important topic awaiting more elaborate approaches is anonymization (Golle and Partridge, 2009; Hoh et al., 2010) of high-resolution spatial data and their presentation.

The data make it possible to investigate the distance and travel time influence on choice while taking trip chaining into account. As these are prominent factors in destination choice models, both in the utility function and for choice set specification, the results can be used to further develop, calibrate and validate existing models, in particular, the very promising time-geographic models (Scott and He, 2012) and models moving in the direction of mental map models (Chorus and Timmermans, 2009; Hannes et al., 2008; Mondschein et al., 2008; Arentze and Timmermans, 2004; Golledge and Timmermans, 1990).

8.3.4 Application of Estimated Utility Function in MATSim

To investigate the effect of the estimated attributes, four MATSim runs with $\beta_{size} = 0.0, 0.2, 0.5, 1.0$ and $\beta_{price\ level} = 0.0, -0.2, -0.5, -1.0$ were performed. Results are shown for iteration 1000. The run with $beta_{size,price\ level} = 0.0$ is roughly calibrated to match the Microcensus values for mean and median distances, where, clearly, a perfect match for mean and median at the same time is impossible, as the simulated and observed distributions do not have an identical functional form.

Variability due to different settings can be neglected as the distance analysis is done at the population level, featuring only low random variability as shown for the average plan score in Section 8.4.1.1.

Assessing the effect of the estimated attributes is not straight-forward due to a lack of high-resolution validation data (see also discussion in Section 9.1.2).

The results show that only $\beta_{size} = 1.0$ and $\beta_{price\ level} = 1.0$, being much higher than the estimated parameters in Section 8.3.1, have a substantial effect. They increase the mean and median values from $4400m$ to $4626m$ and from 2173 to $2327m$ [2], which means, a priori, that the error terms can be reduced and that, a larger part of heterogeneity is now explained. In other words, from modeler's perspective, the randomness of the destination choice is reduced. Attributes are only applied to shopping destinations, however, due to trip chaining leisure trips are also weakly influenced (the mean distance is increased from $5347m$ to $5389m$ in the extended model).

To match the distance distributions of the base model the shopping error term has to be reduced, here from $f_{shopping} = 0.4$ to $f_{shopping} = 0.5$. Experiments with this extended and adapted model (configuration b) show that adding the size and price attributes to the base model (configuration a) generate a substantial decrease of the relative error in count data (Figure 8.12). Given the small effect on travel distances, as reported above, this clear success is surprising and requires further inspection. A first explanation might be as follows. The relative error is essentially reduced by generating more traffic on roads with traffic counters. The city center has a higher coverage with counters, both in our scenario and in reality. At the same time, many large stores are located in the center, toward which the size attribute in the improved model attracts more traffic. Thus, a better matching with count data is plausible. Another more sober explanation is the following. As will be argued later, simulated and counted volumes only have one degree of freedom (they can only increase or decrease) and, thus, the chance for spurious correlations is high.

8.3.5 Discussion

Technically, the software is now ready to handle any number of attributes by reading the coefficients and the attribute values in two xml files. Person-specific coefficients are now possible allowing application of mixed logit models and similar.

Clearly, the incorporated variable set represents a rough approximation to the shopping decision factors. As usually, it must be expected that many aspects are hidden not only in the random error terms but also in the coefficients. The size coefficient, for example, can be assumed to catch motivations such as searching for a specific product not available in small stores, which naturally only offer a limited set of products. Even so, price level coefficient might mix search for a luxury product (increasing

[2] The mean and median travel distances are substantially higher for the Microcensus; very long trips cannot be modeled here due to the limitation to a 30km radius area, cutting the distance distribution tail.

the coefficient) and the search for best price (decreasing the coefficient). Thus, mixed logit and latent class models should be tested in a next step. An explicit exclusivity index or contexts/sub-purposes as investigated by Arentze et al. (2013) should be considered.

The price level definition is also rough. The intra-store price variance for the large chain stores is large and thus average price level is relatively homogeneous between the chains (see Comparis (2011)). This contradicts to some extent the author's daily experience. When buying without much attention to price, one faces substantial price differences between different chains.

One reason might be, that although similar products are available in all chains, accessibility of the cheap or luxury line products as well as store atmosphere and product presentation differs between chains leading to different motivations to actually buy specific products. Thus, systematic and comprehensive shopping basket studies, taking into account the subjective choice processes and context, might substantially increase model quality.

Estimation should be extended in the future to differentiate non-grocery and grocery shopping trips (according to Swiss Microcensus covering 25 % and 75 % respectively, of all shopping trips). A distinction between round trips and intermediate shopping stops should be made. Furthermore, the estimated model should be improved to be able to capture the income effects.

Here, linear models are estimated. However, linearity is seldom in reality (dos Santos and Porta Nova, 1999, p.502). When looking at the travel distance or times distributions, the idea to fit travel decisions with linear models seems inappropriate. Furthermore, especially when estimating leisure destination choice models, sub-purpose must be considered for defining functional form and thresholds of the utility function. E.g., hiking in the mountains usually does not provide large utility below 2 hours, where marginal utility is close to zero above 3 hours when going to the movies.

The shopping survey provides frequency information. In the future, this information can be incorporated into model estimation. Importantly, choices are then not described by binary variables as in single-observation-models but by continuous variables calculated from the reported visiting frequencies. Another exploitation of the survey's information might be the models similar to Ben-Akiva and Boccara (1995) (see also Cantillo and Ortúzar (2006, p.683)), making use of additional information about the alternative or the person for choice set specification.

Besides probabilistic choice set generation approaches mentioned above, the *multiple-discrete-continuous extreme value model*, brought to transportation by Bhat (2005), is particularly suitable when multiple alternatives—not being perfect substitutes—are chosen simultaneously. Bhat (2005) uses such a model for modeling choice of discretionary activities and duration choice. Similarly, choice of stores and respective expenses could be modeled based on our survey data. Furthermore, frequency data allow to estimate exploded logit (Chapman and Staelin, 1982) or ordered logit and probit models for rank and ratings data respectively.

8.4 Variability Runs

Section 8.4.1 summarizes the variability analyses for the 10% Zurich scenario. The experiments were done by Horni et al. (2011c) and repeated by Datye (2012). Temporal variability of road count data and results for an illustrative small-scale scenario are shown in Section 8.4.2.

8.4.1 Zurich Scenario

As done in most previous studies, endogenous random variability is investigated, meaning that inputs are held constant while the random seeds are varied, corresponding to the *method of replication* as described in Benekohal and Abu-Lebdeh (1994). Random seeds, in this work, influence time and route choice (both implicit) and destination choice (explicit). All simulation random seeds are varied simultaneously.

30 simulation runs of the Zurich scenario are performed with 200 iterations each, where the replanning share was 20%.

For results reporting, the coefficient of variation (CV) is used, which is the relative standard deviation, σ/μ, where σ is the standard deviation and μ the mean of the output distribution. According to statistics, in this thesis, the standard deviation is corrected with the *sample* standard deviation.

8.4.1.1 Utilities: Person and Population Level

At person level, the mean CV of the agents' executed plan utilities is approximately 3% and median is approximately 0.9% (Figure 8.13). At population level, as expected, there is little variability between simulation results; mean utility (averaged over agents) of all executed plans of the final iteration 200 has a very low CV of 0.087 %, showing that, as discussed earlier, population aggregation actually reduces variability.

8.4.1.2 Link Volumes

The 123 links in the center of Zurich, which have count stations are analyzed. Network link volumes are a very common and important aggregate measure in transport planning.

Every link's CV over the 30 runs is plotted in boxplots and in scatter plots, meaning that every link is compared with *itself* over the 30 runs. In the scatter plots, the abscissa represents the *average* value over multiple runs (inter-run variability analyses) or multiple iterations (intra-run variability analyses), where on the ordinate the individual observed values are plotted.

Variability of *daily* volumes is shown in Figures 8.14(a) and 8.15(b). Consistent with previous work, relatively little variability exists at this resolution level.

Variability for *hourly* volumes is shown in Figures 8.14(b) and 8.15(c). Selectively, different hours are included; values for other hours are very similar. For this resolution, the variability measured is an issue. This is not in-line with previous studies. A direct comparison with previous studies, however, is difficult. In this study, focus was to generate first indications about variability of MATSim results. To compare these results with previous studies model resolution needs to be adapted such that comparisons are possible. Random variability, clearly, depends on the spatial and temporal resolution and the choice dimensions included in the model. For example, taking only route choice and daily volumes into account strongly reduces the degrees of freedom in the model leading to smaller variability. Veldhuisen et al. (2000a); Cools et al. (2011); Castiglione et al. (2003) limited their analyses to *daily* measures. Cools et al. (2011) investigated the population level. Ziems et al. (2011); Lawe et al. (2009) evaluated hourly measures, but only included route choice. Hackney (2009, p.128ff) applied only time and route choice and results are given for daily measures. Likewise MATSim analyses with reduced choice dimension settings should be applied as a future task for comparisons.

Figures 8.14(c), 8.14(d), and 8.15(d) show relatively large *intra*-run variability over the last 10 iterations of a run. A large intra-run variability could indicate that the system has reached a utility plateau with many user equilibria close to each other, or that it has not yet reached equilibrium although the score is relatively stable. Intra-run variability might also be created by the replanning modules based on random mutation. This raises the issue of reducing the MATSim replanning share, when approaching equilibrium. In any case, intra-run variability is a component of inter-run variability and, thus, it requires future investigation.

8.4.2 Temporal Variability

8.4.2.1 Road Counts

Figure 8.16 shows measured real-world link volumes given for both, the complete year and single months, meaning that a single point in the box plot represents temporal variability of a single network link, either for the whole year, or for a specific month. The hours 11-12 and 17-18 are shown as examples; similar patterns can be observed for all hours. Daily volumes are also reported.

The plots show that temporal variability in reality is substantial. It can also be seen that variability over the whole year is larger, than for single months. This is due to general rhythms of life induced, for example, by season. These rhythms can be interpreted as *temporal correlations* in the population, substantially influencing *aggregate* results' variability. Illustrated with an example and in more mathematical terms this is as follows. Given two random variables X_0 and X_1 representing an arbitrary time-dependent decision of individual 0 and individual 1, i.e., $X_0 = f_0(t)$ and $X_1 = f_1(t)$, the variance of two random variables is $Var(X_0 + X_1) = Var(X_0) + Var(X_1) + 2 \cdot Cov(X_0, X_1)$. The covariance is non-zero for correlated variables; the covariance is greater than zero if variables are equidirectional as for example given by general life rhythms. There are also decisions where correlation is negative i.e., $Cov(X_0, X_1) < 0$. An example might be the avoidance of demand peaks, such as not visiting certain skiing resorts during school holidays. But, by analyzing the count data, it can be seen that the positive correlation predominates, increasing temporal aggregates' variability. An illustrative example is given in Figure 8.17. These correlations can be best captured with a longitudinal model (Figure 6.1 (c)), where it is, assumedly, very difficult to model them with pure random variability (Figure 6.1 (a)).

8.4.2.2 Small-Scale MATSim Simulation Scenario

A small-scale toy scenario is used for illustrating the temporal variability and correlations discussion in Section 6.2. Its configuration is depicted in Figure 8.18.

1'000 persons, living in the two residence zones h_0 and h_1 perform exactly one shopping activity with a desired duration of 90 minutes per day. The shopping activity can be performed in the residence zones or in the city zone (locations 4-9). 5 consecutive working days are simulated. Over the week, a variable share of persons does a working activity in the city zone (at location 3). The daily shares are 0.9, 0.9, 0.9, 0.8, 0.5 (from Monday to Friday). All workers have a desired working activity duration of 9 hours, except on Friday. Then they only work 7 hours. Time, route and destination choices are performed.

Two different configurations are simulated.

- Configuration 1: An average working day is simulated, i.e., as in the actual MATSim run, averaging is done on the input side. An average working share of $(0.9 + 0.9 + 0.9 + 0.8 + 0.5)/5 = 0.8$ is applied. $0.5 \times 1000 = 500$ of $(0.9+0.9+0.9+0.8+0.5) \times 1000 = 4000$ or a share of 12.5% are short working days. I.e., 12.5% of all workers are short workers on the average working day. This corresponds to model (a) in Figure 6.1.

- Configuration 2: Temporal correlations are explicitly taken into account. 5 single days are simulated without doing any averaging. This corresponds to model (c) in Figure 6.1.

For both configurations 250 runs using different random seeds are performed ($250\ runs = 5\ days/week \times 50\ weeks$). For configuration 1 the week contains 5 identical days (the average working day) but the runs are based on different random seeds.

The hourly departures from the origins of all agents are plotted in Figure 8.19; the situation is very similar for the arrivals at the destinations. Configuration 2, which explicitly models the temporal rhythms shows substantially larger variability than the average configuration. (Configuration 1). Interestingly, the averages over the runs are very similar. However, in general this cannot be expected as usually microsimulations contain many non-linearities.

8.4.3 Discussion

Results of this investigation are partly in line with previous work and partly contrary. Similar to previous studies, *daily* link volumes and agents' utilities show little variability such that, actually, few runs are necessary to achieve stable results. *Hourly* volumes, in contrast, show substantial variability. This is initially surprising but not implausible. The resolution is higher and/or there are more degrees of freedom in this experiment than in previous studies, suggesting that a higher variability must be expected.

Future investigations should encompass MATSim intra-run variability and population fractions γ. As mentioned earlier, due to computational reasons, often only population fractions are simulated. In the calculation of the CV, a scaling factor γ cancels out because γ is applied to both numerator ($\hat{\sigma}$) and denominator ($\hat{\mu}$). However, the fact, that in fractional scenarios, each agent represents a group of agents leads to a discretization error, with a potential influence on variability. This might be particularly significant for low-volume links, which tend to have larger relative inter-run variability (Figure 8.14(b)).

A comparison of sampling error and other types of errors, such as measurement errors, should also be included in future work.

Ideally, microsimulation results should be accompanied by a confidence interval. For a given error level, the required number of runs can be derived. This is straight-forward at high aggregation levels. At *low levels*, however, this is non-trivial. For example, the investigation in this paper encompasses 123 links, each with 24 hourly volumes. Every hour on every link has its own variance, meaning that, in the extreme, for every link and every hour, a confidence interval has to be given. Furthermore, it is not clear, which of these intervals defines the required number of runs. Methods to analyze (and also present) numerous confidence intervals need to be developed in the future for the microsimulation context.

Finally, the body of previous studies lacks continuity, i.e., most studies—and also ours—have a relatively strong focus on the specific analyzed simulator. Seamless continuation and repetition of studies would increase knowledge about this really complex and important topic.

8.5 Interaction Runs

8.5.1 Singly-Constrained Dynamic Model

The effect of explicitly incorporating the activity location load, coupled with capacity restraints, into the agents' destination choice process is visible in Figure 8.20 and 8.21. Figure 8.20(a)), shows a strong shift toward the activity locations with a high attractiveness factor for configuration 2 without capacity restraints. Figure 8.20(b) shows that this generates many implausibly overloaded facilities and is hence unnatural. Applying a constrained model (configuration 3) avoids this unnatural situation.

In addition to affecting the daily activity facility load, the constrained model can be used in a dynamic context. If one has dynamic constraints data, for example, operation schedules, the activity location load actually can be adjusted to the time-dependent capacity restraints (Figure 8.21).

Concluding, a singly-constrained dynamic model of interaction effects might be a suitable calibration means to improve behavioral realism of microsimulations. However, validation of improvement is outstanding and awaiting suitable high-resolution data, for example, dynamic customer count data.

While the inclusion of destination interaction effects in microsimulations is conceptually expedient, before intensively continuing this line of research, the magnitude of interaction effects, i.e., their significance, needs to be quantitatively researched. Furthermore, capacity data is required but difficult to collect. For Switzerland, disaggregate employment information (given by the number of full time equivalents) (Swiss Federal Statistical Office (BFS), 2008a) is available. The work started by Meister (2008), deriving rough capacity estimates for employment data should be continued and refined for the dynamic context. Future analyses must also answer the question if activity infrastructure load actually should be microsimulated (analog to the network loading simulation) or if the typical aggregate cost-load-curves can be applied approximately.

As started in Stahel (2012), different utility-load relationships should be investigated. In relatively static contexts with sharp capacity limits (e.g., in a small restaurant), also the utility-load function should show a sharp decrease at capacity limit. In more variable contexts, for example, inside stores or on very large parking sites, a softer form of utility-load function can be expected.

8.5.2 Parking

Our parking simulation was tested on a small-scale chessboard scenario (Section 8.5.2.1) and on the Zurich scenario reduced to the town center (Section 8.5.2.2). The integration into MATSim is discussed in Section 8.5.2.3.

8.5.2.1 Small-Scale Toy Scenario: Chessboard

For efficient development, testing and basic illustration purposes, a toy scenario was created, named *chessboard* (Figure 8.22). It was simulated with 100 agents with different trip starting times and a desired activity duration of 30 minutes. Private parkers and transit agents were not included in this scenario. The agents start from bottom-left and top-right of the chessboard and try to park as close as possible to the chessboard center square. After finishing the activity they drive to the corner opposite to their origin.

Figure 8.23 shows the median search time [3] dependent on the number of parking spaces in the study area. A non-linear relationship between the median search time and parking supply was observed. As one does not yet know the functional relationship between search time, parking supply and demand, a direct comparison with Axhausen et al. (1994, p.308) investigating varying demand which is operationalized by parking lot occupancy, is impossible, but is a good source for a future investigation.

The non-linear trend, simulated here, should in a future analysis be contrasted with the work of Benenson et al. (2008, p.438), whose simulation confirmed the empirical finding by Shoup (2005) saying that average search times *"[...] hardly react to changes in parking supply as long as the demand/supply ratio is around one."*.

8.5.2.2 Real-World Scenario: The Town Center of Zurich

A very first verification step was undertaken for a real-world scenario in the town center of Zurich, defined here as the area within a 1.5km radius around Bellevue.

As the parking supply was expected to be local in nature, a detailed navigation network (see Figure 8.24) comprising 1,218 nodes and 4,750 links composed of 43,881 cellular automaton cells was used (derived from TomTom MultiNet (2011)). Parking supply data were gathered from

[3] Here the median was used instead of the average in order to account for outliers such as persons who had not yet found a parking space by the end of the simulation.

various sources, and 1,355 parking lots with a variable number of spaces were created.

Demand was derived from a MATSim Zurich scenario (Horni et al., 2011e) in which a total of 190,000 agents were generated for the whole day. For performance reasons, not a complete day was simulated but only the morning hours from 8-10 o'clock, whereby only the second hour was evaluated due to boundary or warm-up effects. Approximately 20 hours of runtime were required for a 100% run. Approximately half of the population was in transit and the other half was looking for a parking space in the study region. A share of 25% private parkers, who did not have to search for a public parking space was assumed.

As a first step, two 10% runs were performed. To reduce the complexity of the implementation, parking capacity was scaled, but not road capacity. The first run was performed with the actual parking supply available in Zurich (but scaled), and the second with doubled scaled supply. Note, that only the capacity but not the location, or in other words the spatial distribution of parking lots is changed. The search-time histogram looks similar to the one observed for road travel times (see Figure 8.25). Average values decreased from 3.9 minutes to 3.6 minutes when the supply was doubled, which is a smaller decrease than expected. A natural explanation might be, that both scenarios are located in the low search time elasticity area with respect to parking supply, in other words on the right side in Figure 8.23. Thus explanation is supported by the the recent analysis of Montini et al. (2012) reporting a relaxed parking situation in Zurich. Search times are then basically given by searching a parking lot, not a free parking space within such a lot. However, according to the authors' every day experience an undersupply of parking lots exists in the city center of Zurich.

In addition to the 10% runs, similarly, 100% runs were conducted, which clearly showed two major issues: First and foremost, serious deadlocks appear, resulting in unrealistic results. Link capacities require further investigation. In MATSim queue simulations, not only the flow capacity, i.e., how many cars can pass the link in a certain time, but also the storage capacity, i.e., how many cars fit onto a link, has always been a crucial issue. The same probably holds true here.

Probably even more important and a major lesson learned is the relevance of node capacities, or in more general intersection dynamics. The deadlock situation was relaxed (although not resolved) when increasing the nodes' capacities by decreasing the simulation time step; intersection-crossing is possible for one car per time step. This implicit modeling of intersections should be improved in the future. Interestingly, this does not necessarily require an explicit model but can be achieved by calibration of the implicit model. The update rules in Nagel and Schreckenberg (1992, p.2222) do not consider intersections, in other words, nodes are handled implicitly. Nevertheless, as stated by Wu and Brilon (1997, section 2.2, p.4), "for describing queuing systems, e.g., intersections of two urban streets, the standard Nagel-Schreckenberg cellular automaton delivers very good results compared with the real-world traffic conditions". In MATSim, on the other hand, too high urban travel speeds are usually observed by implicitly handling intersections, where experimental modules for green-lights were developed but not yet part of the standard simulation.

Besides improvement of intersection modeling, our software also requires parking decision models calibration and their enhancement by further choice determinants and mechanisms. An example is the look-ahead procedure mentioned earlier and described by Benenson et al. (2008, p.434). Very interesting data to improve modeling the search starting point are currently collected in a study, where drivers have to push a button in the car in the moment they start their active parking search (Johanna Kopp, personal communication, May, 2012).

This work is focused on verification, where, in the future validation steps are necessary. A one-week GPS survey conducted in the Zurich region (Montini et al., 2013) is suitable for parking behavior analyses as done by Montini et al. (2012) and, hence, direct comparison of simulation and survey figures, such as search times (Montini et al., 2012, p.12) [4] . Further valuable validation sources are a Swiss parking stated-preference survey (Weis et al., 2011), several municipal parking surveys (Planungsbüro Jud, 2010, 1990; DemoSCOPE und Planungsbüro Jud, 2007), parking count data (Waraich and Axhausen, 2012a), and road count data (e.g., ASTRA, 2006). The majority of the public spaces are subject to fee. Associated turnover figures might be another valuable validation source, not yet available to planning.

8.5.2.3 Integration into MATSim

The larger aim of the cellular automaton implementation was to improve MATSim destination interactions modeling. Furthermore, the runtime of the MATLAB model is very high. A migration to Java, which is usually associated with good parallelization capabilities, would probably be beneficial. For these two reasons it is planned (and now already started) to migrate the stand-alone MATLAB model to Java and then to integrate

[4] One possible operationalization of search time, to be more precise. Remember that the actual search time is latent.

it into MATSim. This integration, however, poses two main problems.

First, MATSim is an equilibrium model, which means that agents maximize utility, given the constraints imposed by competition with other agents. By adding the detailed on-the-fly parking search process a rule-based and short-term component enters the long-term equilibration paradigm. It is unclear whether the inclusion of short-term fluctuations ever leads to a stable equilibrium. This important but complex issue is also relevant for the inclusion of within-day (rather than end-of-day) adaptation of the agents' travel decisions as recently and increasingly practiced for MATSim (e.g., Dobler, 2013). This issue's analysis is beyond the scope of this thesis and should be performed as a separate future task.

Second, as MATSim is intended for large-scale applications, a high-resolution parking search model may be prohibitively expensive for practical use. This problem can be solved with following hybrid approach. In areas with high competition for parking lots (e.g., in city centers), the parking search can be microsimulated based on the cellular automaton approach. In regions with low competition (e.g., residential areas), either average search times can be derived from aggregate functions, or an existing simplified MATSim parking model approach, such as Waraich and Axhausen (2012a,b); Waraich et al. (2012a,b, 2013); Dobler and Lämmel (2012), can be applied. Obviously, the hybrid approach increases model accuracy (compared to deriving search times from aggregate data) and at the same time maintains feasible computation times for large-scale scenarios. The final MATSim model will be used to investigate the effects of parking on shopping destination choice. This is particularly relevant because a simulation of the MATSim Saturday scenario, with a higher share of shopping activities, is under development.

8.6 Synthesis Run

For sake of completeness, the above developed and tested destination choice improvements are applied in a multi-modal scenario derived from Vitins et al. (2012). It includes border crossing and freight traffic, both generated by disaggregation of origin-destination matrices provided by Vrtic et al. (2007) and Gottardi and Bürgler (1999) respectively. Non-car modes are teleported. Future versions might consider heterogeneity in walk-speeds (Dobler, 2013; Weidmann, 1992) and interactions between the modes (Dobler and Lämmel, 2012), both available in an experimental manner.

Agent interactions and spatial correlations, as described in Chapter 7, are included as well. Agents' competition is considered by a penalty factor as described in Horni et al. (2009a) and in Section 7.1. The activity score is multiplied with a factor $f_{load\ penalty}$ in the range [0,1] decreasing with overcrowding. To make pre-processing of the search space limits possible, the minimum penalty needs to be applied, meaning that the store load factor $f_{load\ penalty}$ is set to 1.0 for pre-processing. Choice of the best location per iteration is based on the load of the previous iteration.

If attracting factors are applied, respective upper boundaries must be defined. One example is applied in this thesis. Spatial correlations as an attracting agglomeration factor τ_{agglo} are included as described in Section 7.2. It is computed as follows. If more than 10 stores are located in a 300m radius neighborhood, then it is set to 1.0 else it is 0.0. This is similar to the estimated stores' density in Section 8.3.2.

Further technical improvements, made here, concern the inclusion of agent-specific preferences (such as typical or minimal activity durations and latest start and earliest end times) for scoring.

An iteration for this scenario takes roughly 45 minutes which is relatively long. 500 iterations are required to reach a relaxed state. For speeding-up the module, a sampling method is applied on the search space. 10% sampling is done here leading to an iteration time of approximately 10 minutes. While, sampling of alternatives leads to consistent estimates for MNL models, further analyses such as Nerella and Bhat (2004); Lemp and Kockelman (2012) need to be included when turning to more complex models.

As for the base model, distance distributions can be nicely calibrated (Figure 8.26). A problem identified here and requiring future effort concern validation with count data for the multi-modal Zurich scenario. Already, for the relaxed demand with too short travel distances, the link volumes are overestimated (Figure 8.27), although, a substantial part of commercial traffic is presumably still missing in the scenario. Destination choice, by correcting the distance distributions, makes the overestimation even larger. As described in Section 9.1.2, this makes count data for validating destination choice in this scenario problematic.

The destination choice module is technically ready for application in a multi-modal scenario and destination choice improvements can be applied without methodological barriers. However, extended experiments are required as a next step with more validation data.

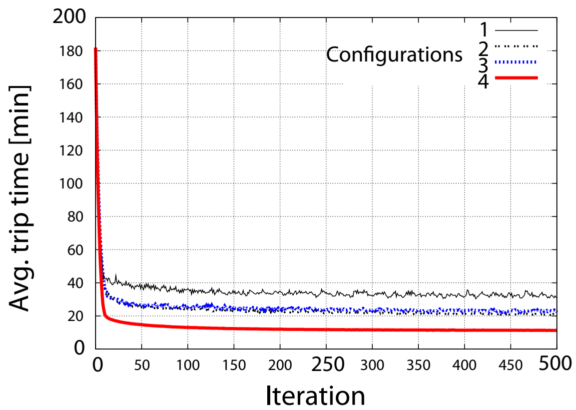
8.7 Figures and Tables

Table 8.1: Software versions of thesis results

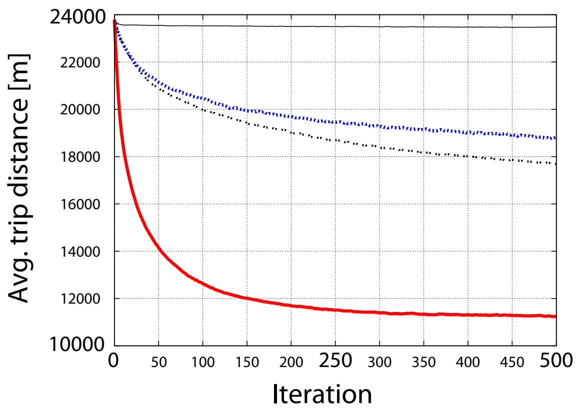
Section	Software version	Original paper
8.2 Error Term Runs	r23281 (2013-03-04)	Horni et al. (2011e)
8.3 Utility Function Extension	r23281 (2013-03-04)	unpublished
8.4 Variability Runs	r16230 (2011-07-25)	Horni et al. (2011c)
8.5 Interaction Runs (Singly-Constrained Dynamic Model)	r2468 (2008-07-20)	Horni et al. (2009a)
8.6 Synthesis Run	r24098 (2013-05-16)	unpublished

Figure 8.1: Testing the time-geographic destination choice approach: Configuration 1: No destination choice, no interactions; Configuration 2: random destination choice (RDC) in the universal choice set (UCS), no interactions; Configuration 3: RDC in UCS, with interactions; Configuration 4: time-geographic destination choice; with interactions

(a) Average trip travel times of agents' best plans



(b) Average trip travel distances of agents' best plans



(c) Average plan score of agents' best plans

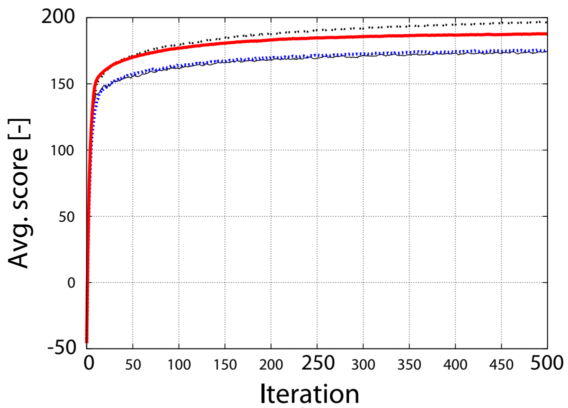
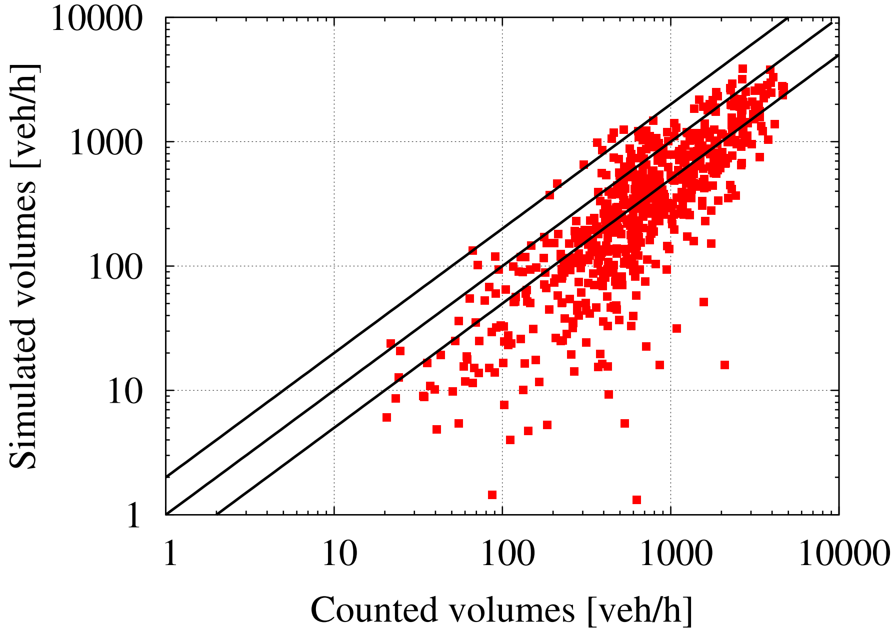


Figure 8.2: Counted volumes versus simulated volumes: evening peak (18:00 - 19:00)

(a) No hollow space-time prisms



(b) Hollow space-time prisms

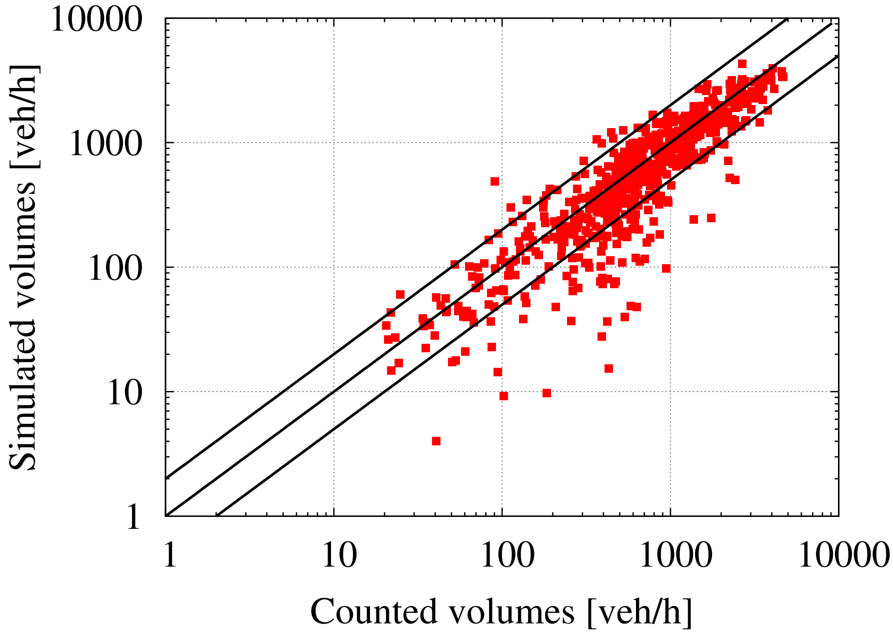
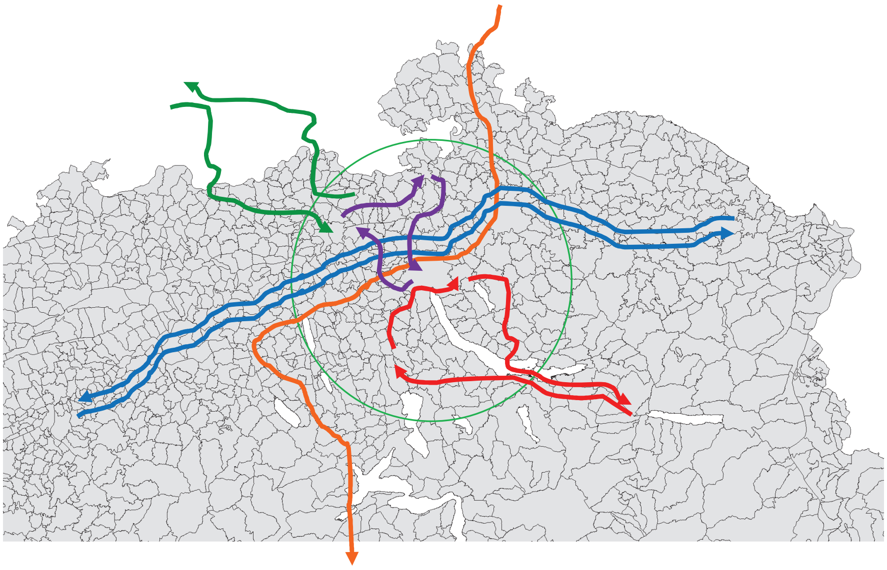


Figure 8.3: Zurich scenario demand

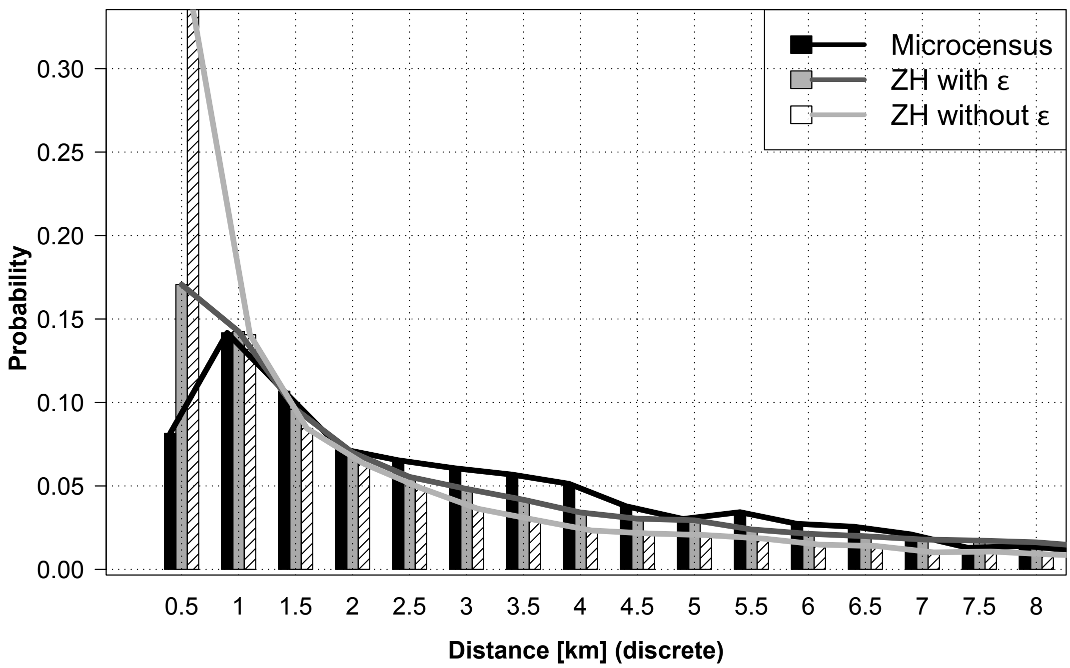


Source: adapted from Balmer et al. (2009)

Figure 8.4: IVTCH network, thinned out for the neighboring countries

Figure 8.5: Error Term Runs for the Zurich scenario

(a) Shopping trips



(b) Leisure trips

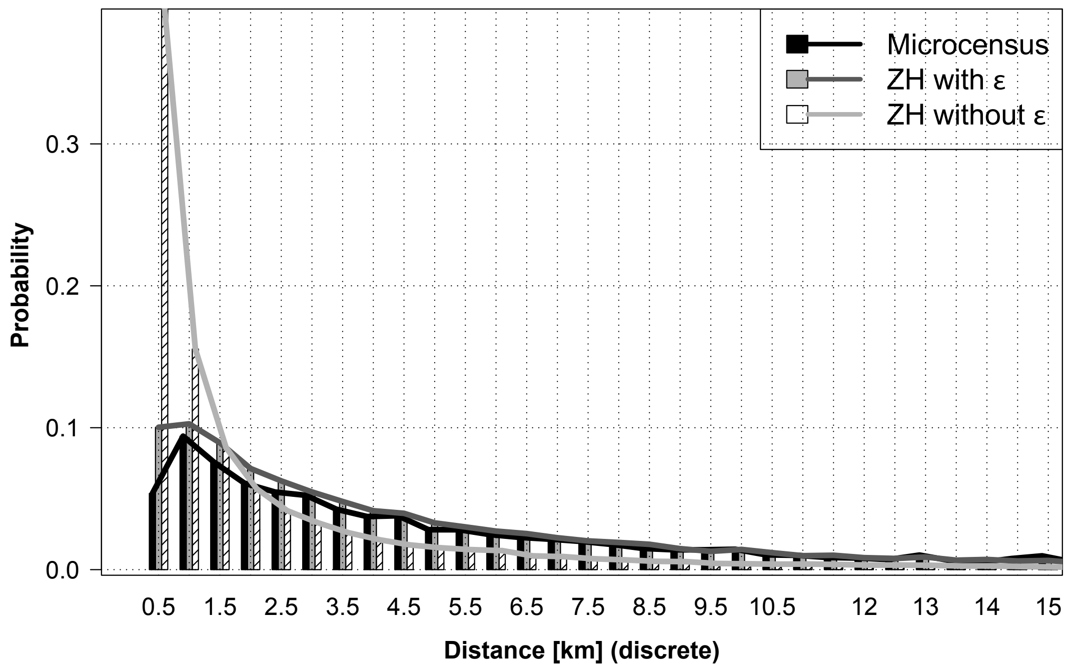


Figure 8.6: Daily traffic volumes for 123 links compared to traffic counts. Per link k the relative error is used, i.e, $(vol_{simulated,k} - vol_{counted,k})/vol_{counted,k}$.

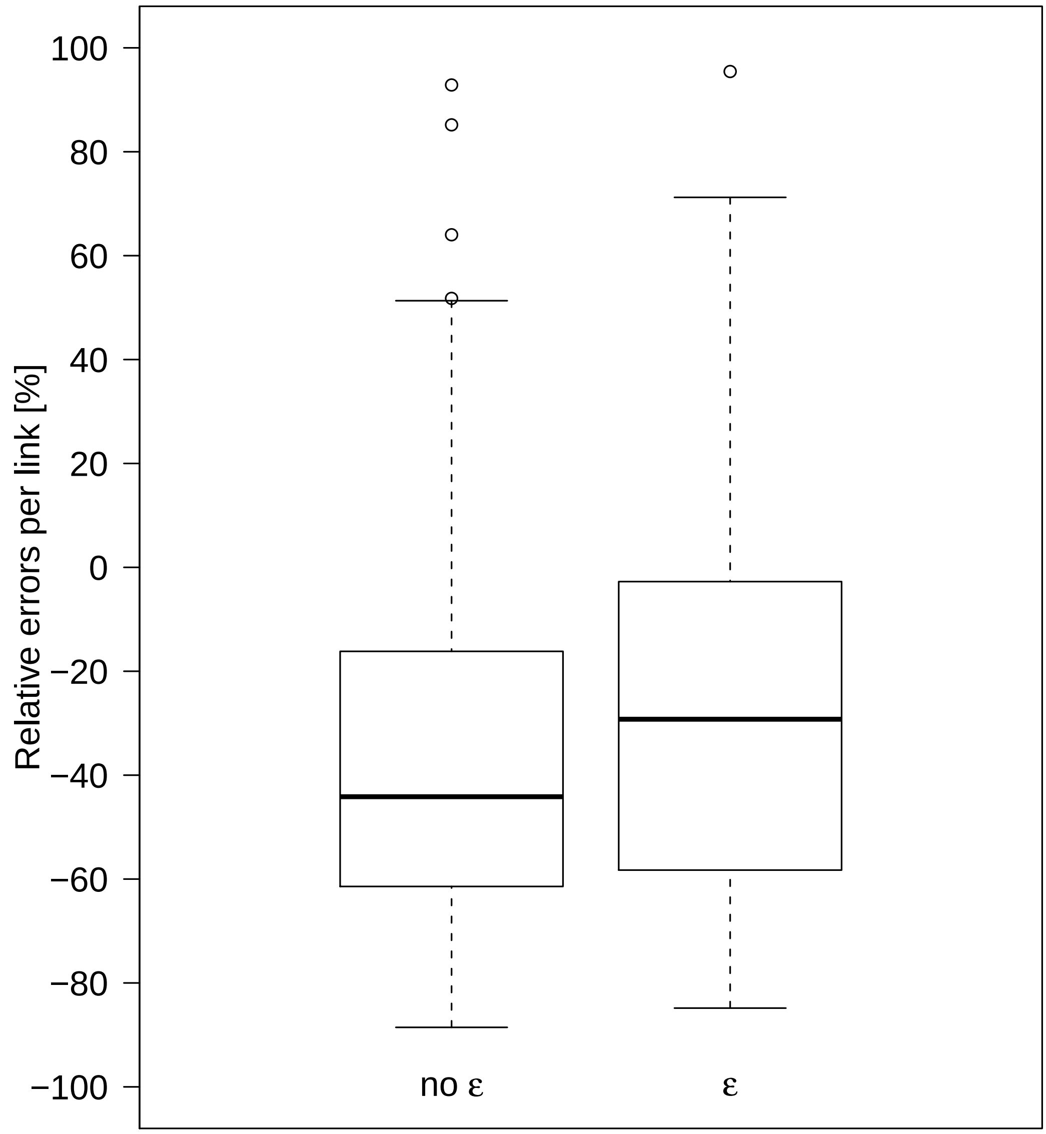


Table 8.2: Model 0

Number of estimated parameters	:	3
Null log-likelihood ($\mathcal{L}(0)$)	:	-3609.826
Cte log-likelihood ($\mathcal{L}(c)$)	:	-2949.469
Init log-likelihood	:	-3609.826
Final log-likelihood ($\mathcal{L}(\hat{\beta})$)	:	-2449.269
Likelihood ratio test ($-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})]$)	:	2321.114
ρ^2	:	0.321
$\bar{\rho}^2$	:	0.321
Final gradient norm	:	+1.330e-002
Iteration	:	53

Parameter number	Parameter name	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	$\beta_{price\ level}$	-0.333	0.0347	-9.60	0.00
2	β_{size}	0.353	0.0230	15.34	0.00
3	$\beta_{addDist}$	-1.59	0.0912	-17.44	0.00

Table 8.3: Model 1

Number of estimated parameters	:	4
Null log-likelihood ($\mathcal{L}(0)$)	:	-3609.826
Cte log-likelihood ($\mathcal{L}(c)$)	:	-2949.469
Init log-likelihood	:	-3609.826
Final log-likelihood ($\mathcal{L}(\hat{\beta})$)	:	-2446.162
Likelihood ratio test ($-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})]$)	:	2327.329
ρ^2	:	0.322
$\bar{\rho}^2$	:	0.321
Final gradient norm	:	+1.428e-002
Iteration	:	121

Parameter number	Parameter name	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	$\beta_{price\ level,\ income}$	-0.102*	0.0809	-1.26	0.21
2	β_{size}	0.354	0.0231	15.36	0.00
3	$\beta_{addDist}$	-1.59	0.0912	-17.47	0.00
4	β_{income}	0.940*	1.02	0.93	0.35

Table 8.4: Model 2

Number of estimated parameters	:	4
Null log-likelihood ($\mathcal{L}(0)$)	:	-3609.826
Cte log-likelihood ($\mathcal{L}(c)$)	:	-2949.469
Init log-likelihood	:	-3609.826
Final log-likelihood ($\mathcal{L}(\hat{\beta})$)	:	-2448.772
Likelihood ratio test ($-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})]$)	:	2322.108
ρ^2	:	0.322
$\bar{\rho}^2$	:	0.321
Final gradient norm	:	+1.438e-002
Iteration	:	52

Parameter number	Parameter name	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	$\beta_{price\ level}$	-0.333	0.0346	-9.63	0.00
2	β_{size}	0.353	0.0230	15.34	0.00
3	$\beta_{addDist,\ income}$	-1.48	0.217	-6.84	0.00
4	β_{income}	0.0315*	0.0597	0.53	0.60

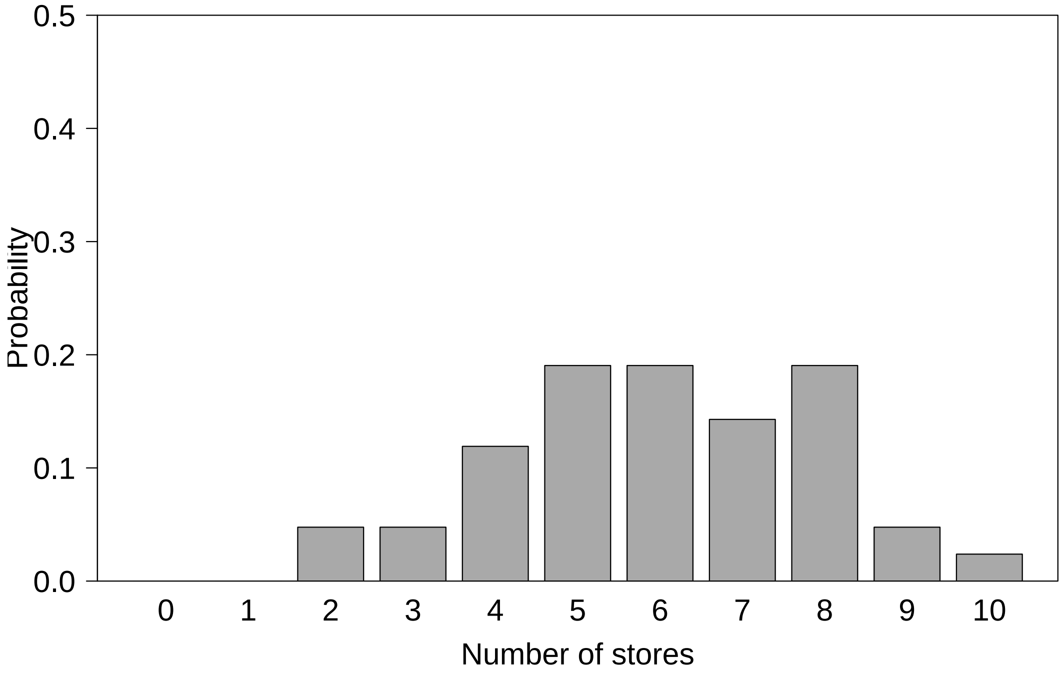
Table 8.5: Model 3

Number of estimated parameters	:	4
Null log-likelihood ($\mathcal{L}(0)$)	:	-3609.826
Cte log-likelihood ($\mathcal{L}(c)$)	:	-2949.469
Init log-likelihood	:	-3609.826
Final log-likelihood ($\mathcal{L}(\hat{\beta})$)	:	-2449.248
Likelihood ratio test ($-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})]$)	:	2321.156
ρ^2	:	0.322
$\bar{\rho}^2$	:	0.320
Final gradient norm	:	+8.307e-003
Iteration	:	157

Parameter number	Parameter name	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	β_{age}	-0.00132*	0.00441	-0.30	0.76
2	$\beta_{price\ level,\ age}$	-0.357	0.0938	-3.80	0.00
3	β_{size}	0.353	0.0230	15.35	0.00
4	$\beta_{addDist}$	-1.59	0.0912	-17.44	0.00

Figure 8.7: Home set: Aware stores

(a) Number of aware stores in the home set



(b) Distance to the farthest aware store in the home set

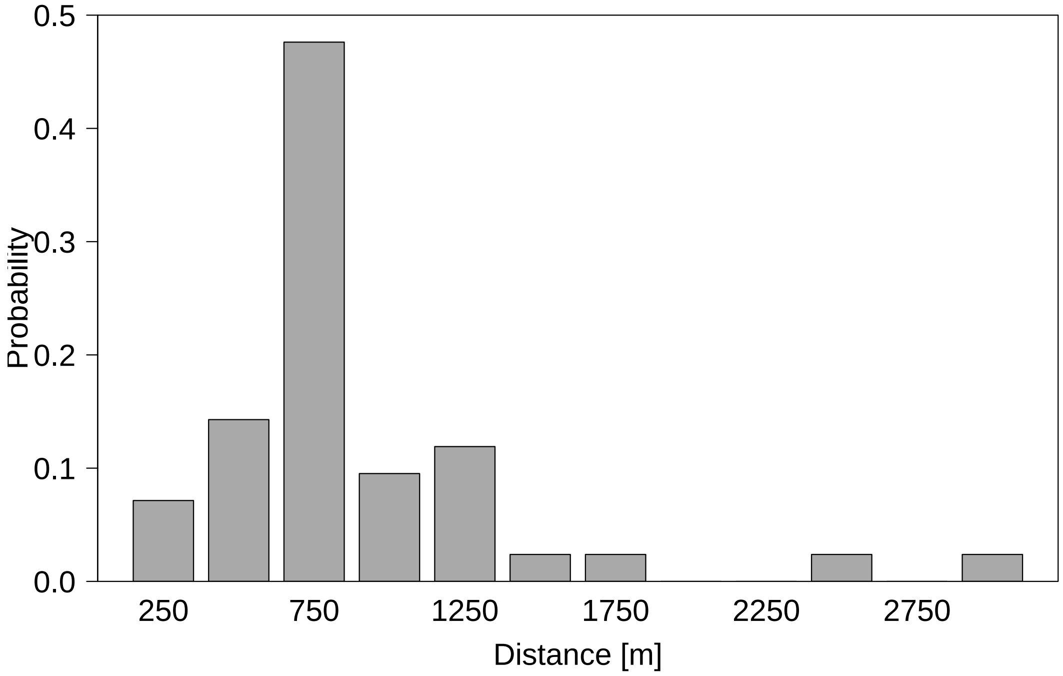
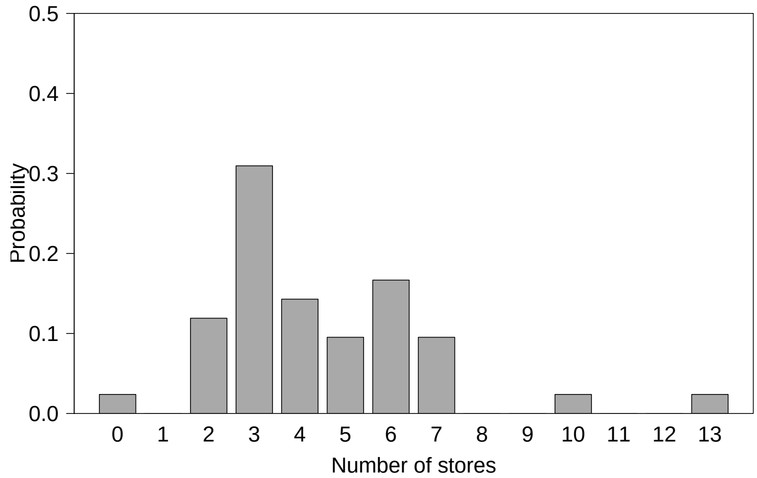
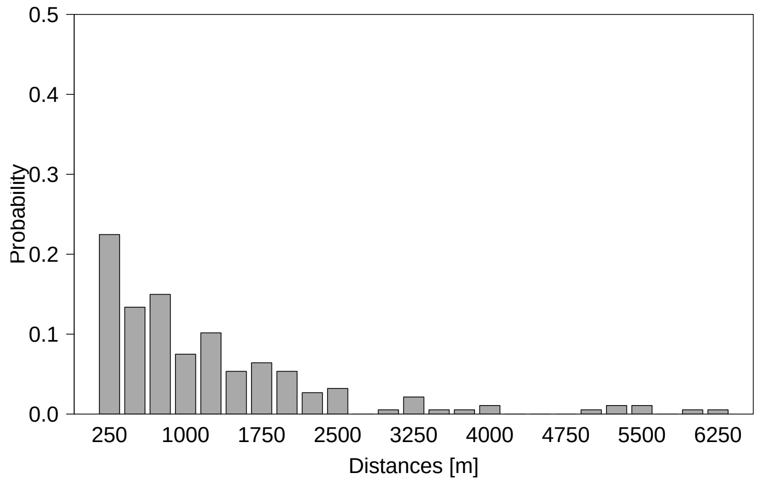


Figure 8.8: Set of frequently visited stores

(a) Number of frequently visited stores



(b) Distance distribution of frequently visited stores



(c) Number of frequently visited stores in the home set

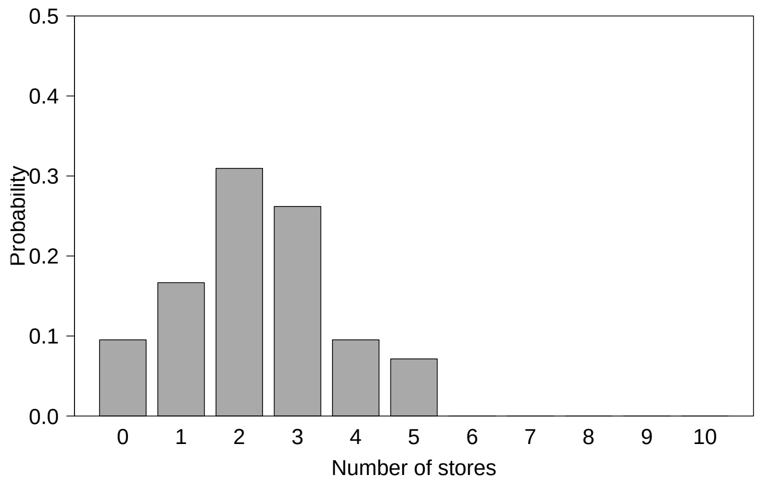


Figure 8.9: Different sets in the decision process and influence of chance. The processed set, shown as an example, is specified in Foscht and Swoboda (2007) as the set of stores for which the individual has gathered information.

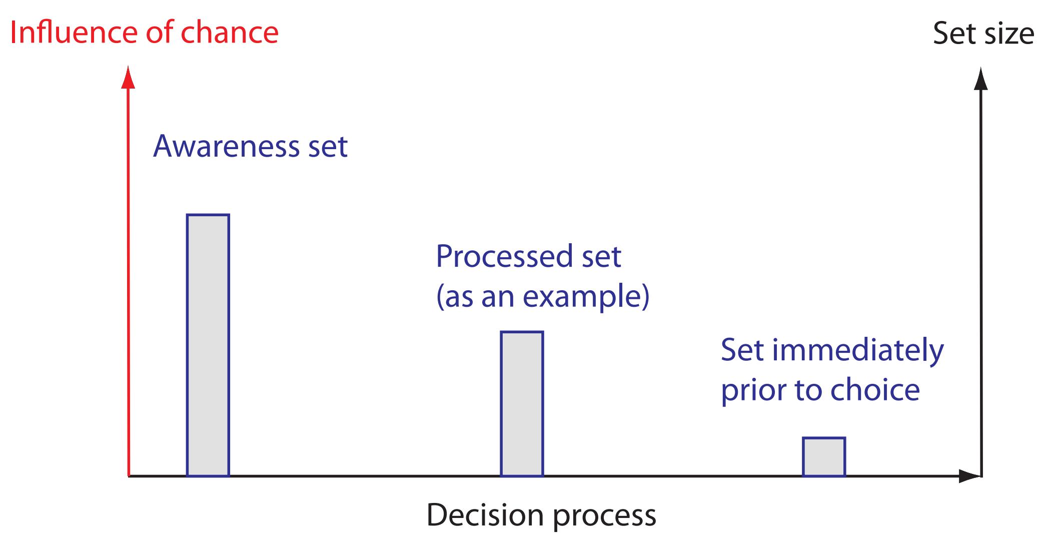


Figure 8.10: Main area of grocery shopping trips for commuters

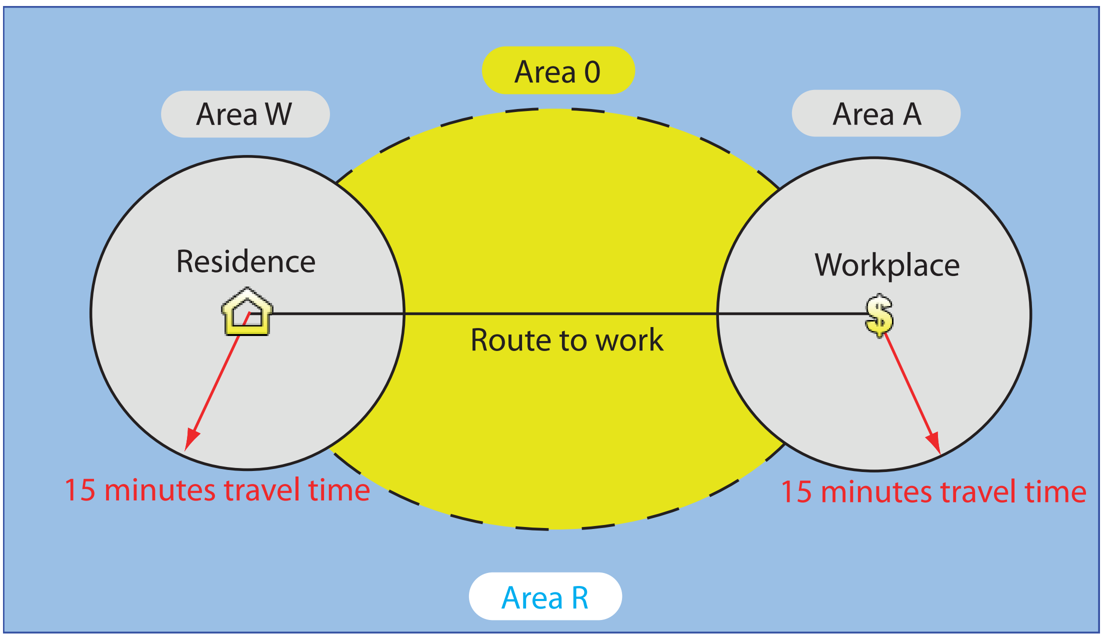
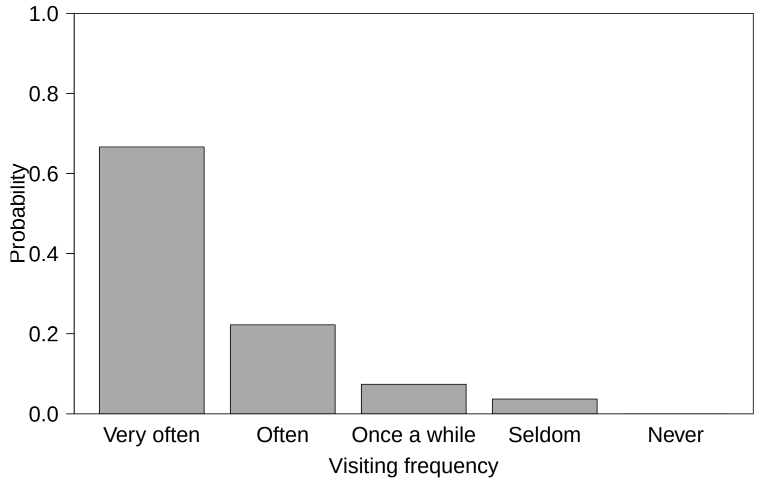
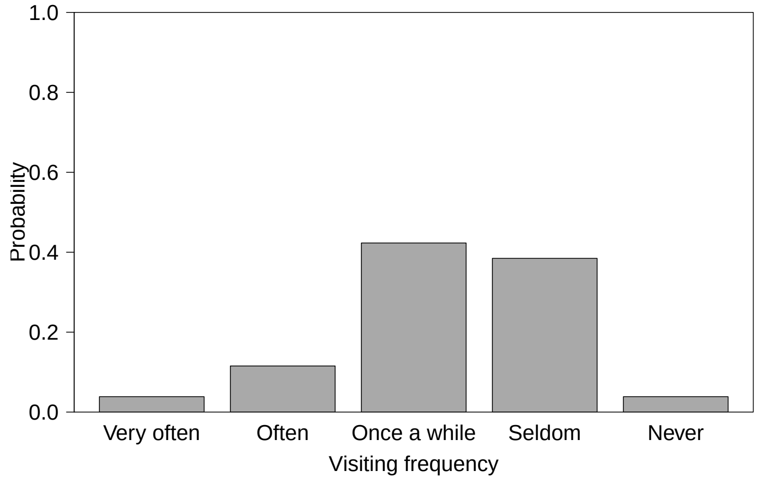


Figure 8.11: Area of shopping

(a) Visiting frequency around the home location



(b) Visiting frequency around the work location



(c) Visiting frequency between home and work location

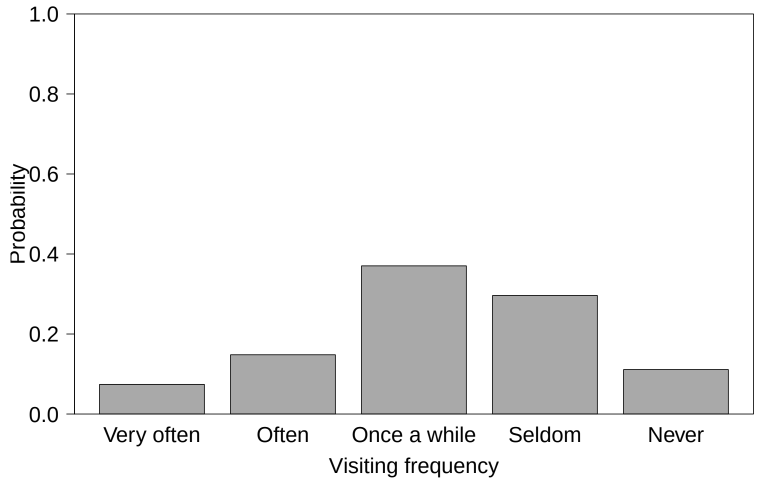


Figure 8.12: Daily traffic volumes for 123 links compared to traffic counts. Per link k the relative error is used, i.e, $(vol_{simulated,k} - vol_{counted,k})/vol_{counted,k}$.

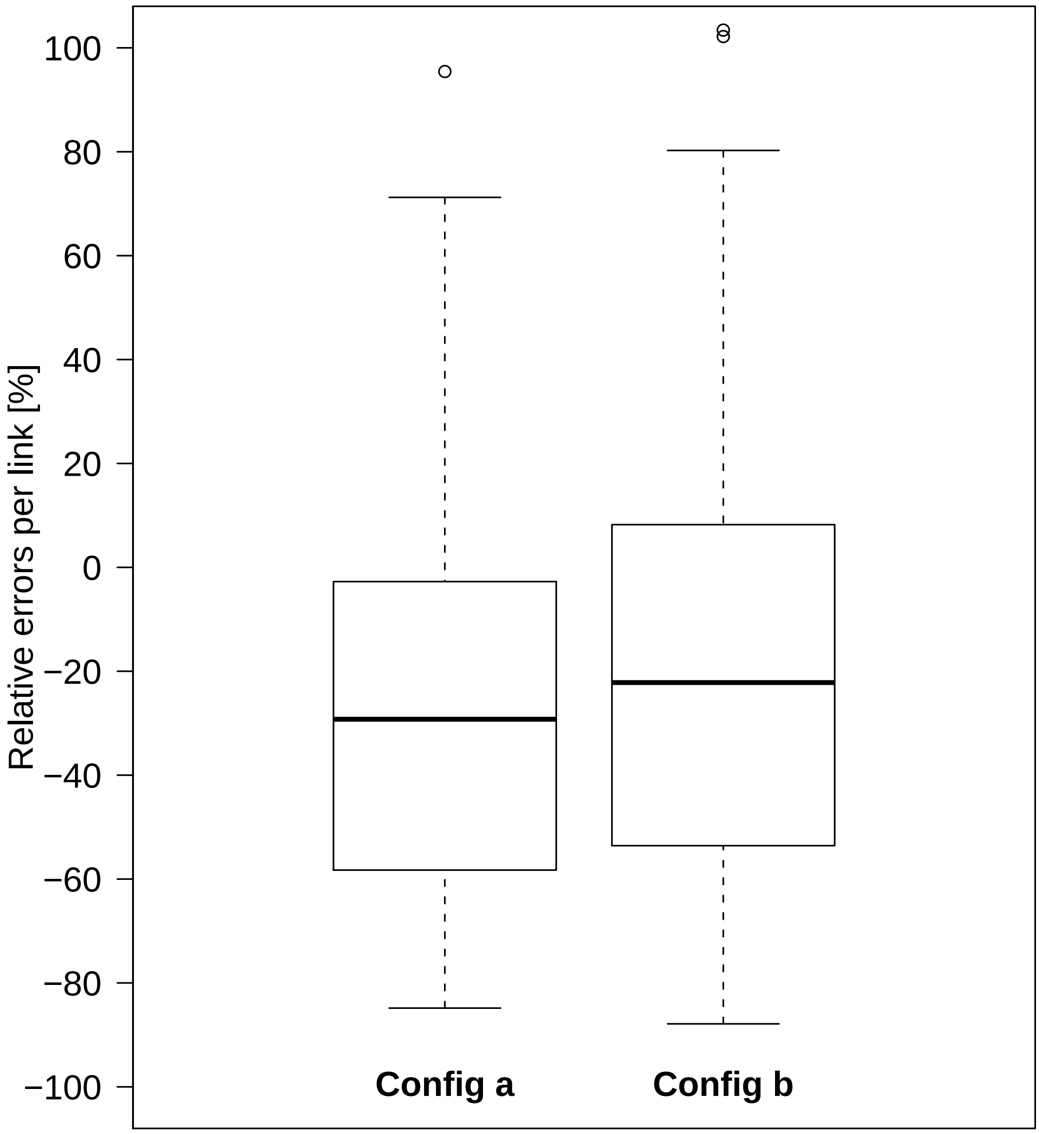


Figure 8.13: Executed plans scores

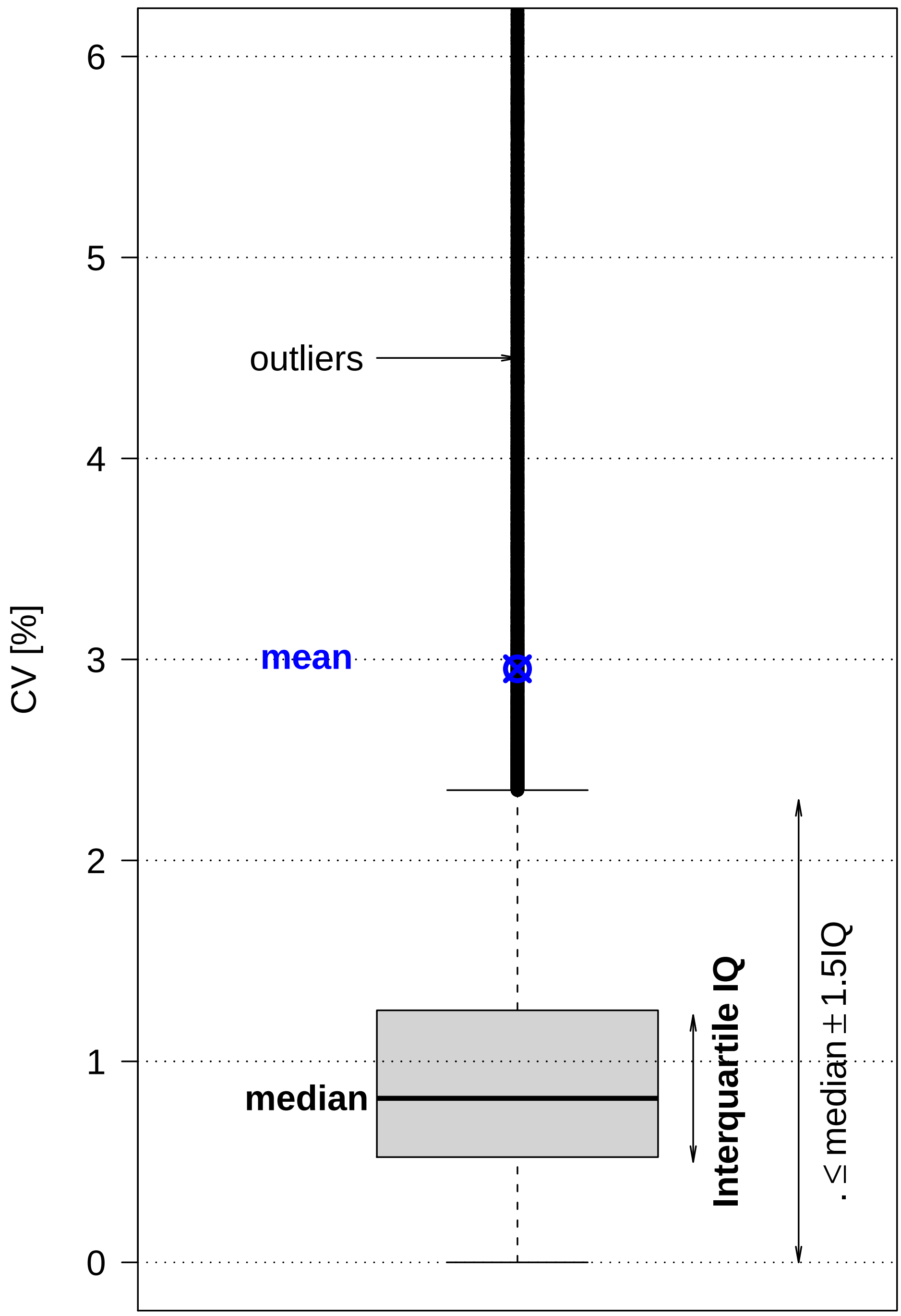
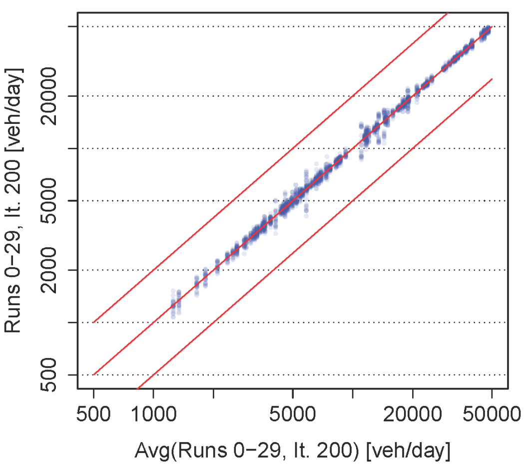
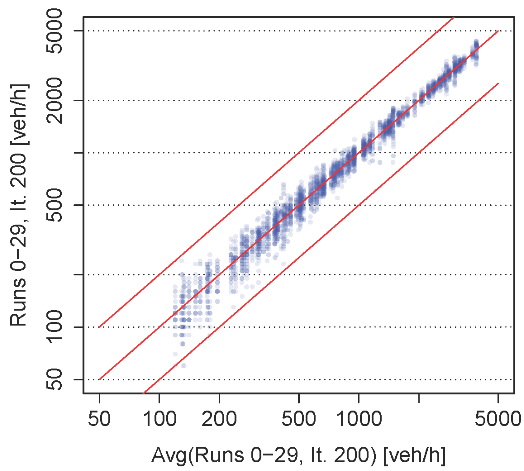


Figure 8.14: Simulated link volumes

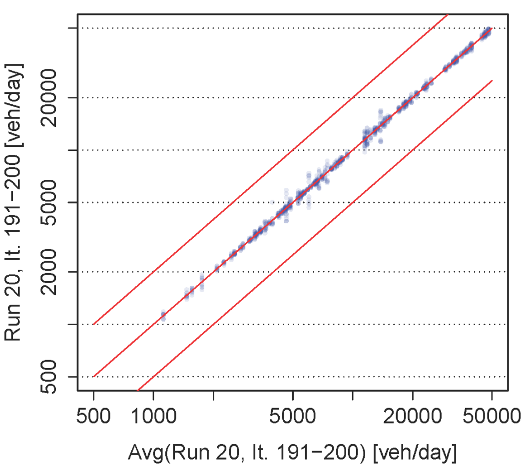
(a) Daily Volumes: : Inter-run Variability



(b) Hourly Volumes (Hour 17-18), Inter-run Variability



(c) Daily Volumes, Intra-run Variability



(d) Hourly Volumes (Hour 17-18), Intra-run Variability

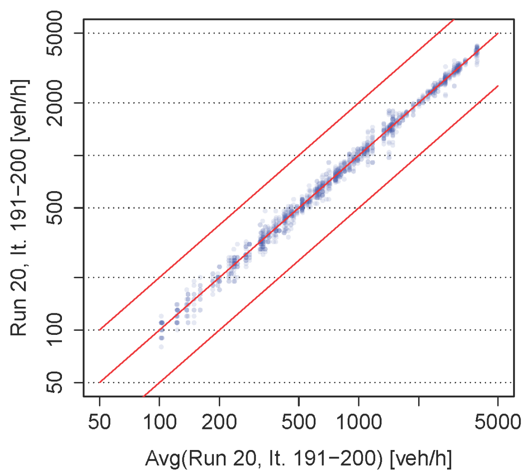
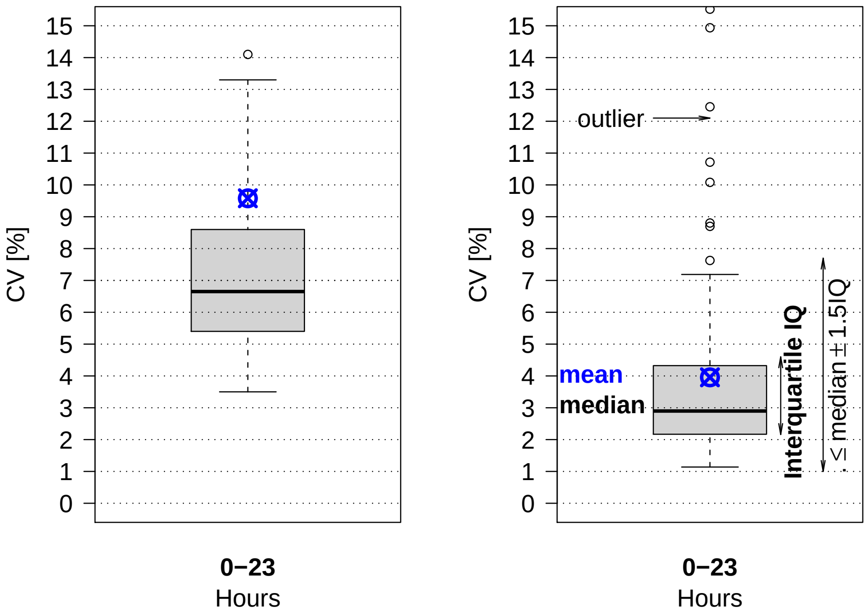


Figure 8.15: Simulated and measured link volumes

(a) Daily Volumes Measured Over One Year in the Region of Zurich.

(b) Simulated Daily Volumes: Inter-run Variability, Runs 0-29, Iteration 200



(c) Simulated Hourly Volumes: Inter-run Variability, Runs 0-29, Iteration 200

(d) Simulated Hourly Volumes: Intra-run Variability (Run 20, Iterations 191-200)

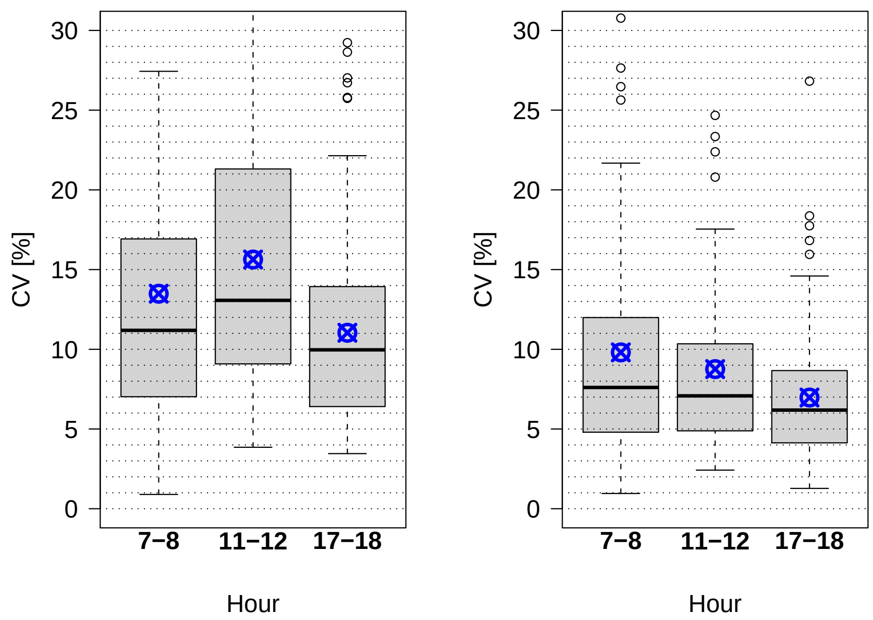
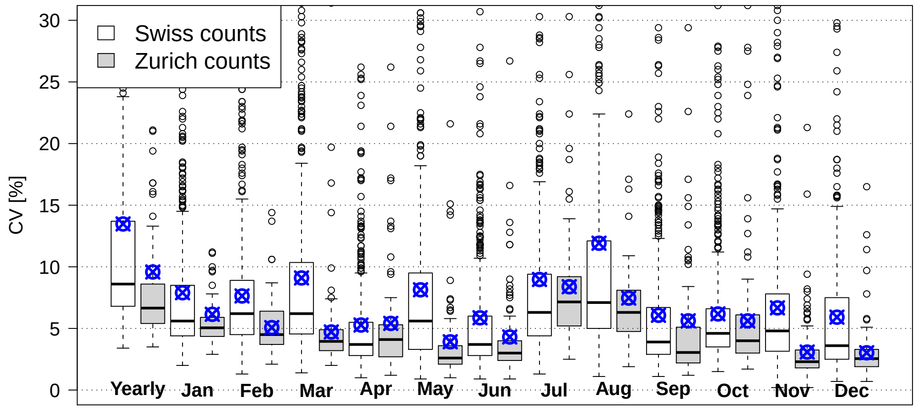
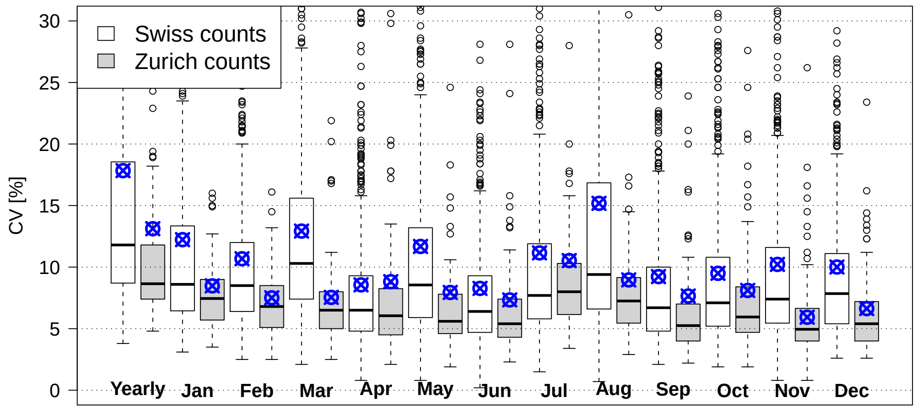


Figure 8.16: Measured volumes

(a) Daily Volumes



(b) 11:00-12:00



(c) 17:00-18:00

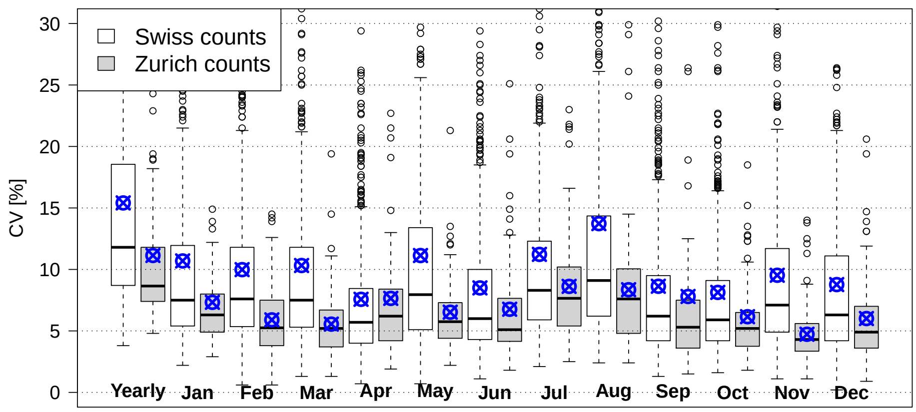


Figure 8.17: Illustrative example for temporal correlations: The fluctuations for small time slots (black) (for example months) are usually much smaller than the fluctuation over the complete time period (red) (for example a year) in presence of a global rhythm of life. y can be imagined as for example the load on roads, that are influenced by season, in the extreme, e.g., roads to skiing resorts.

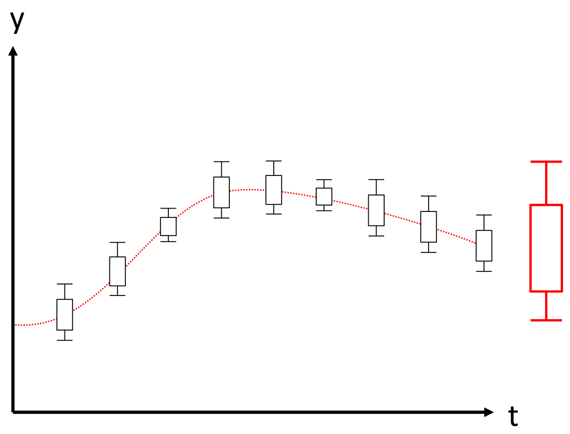


Figure 8.18: Small-scale scenario

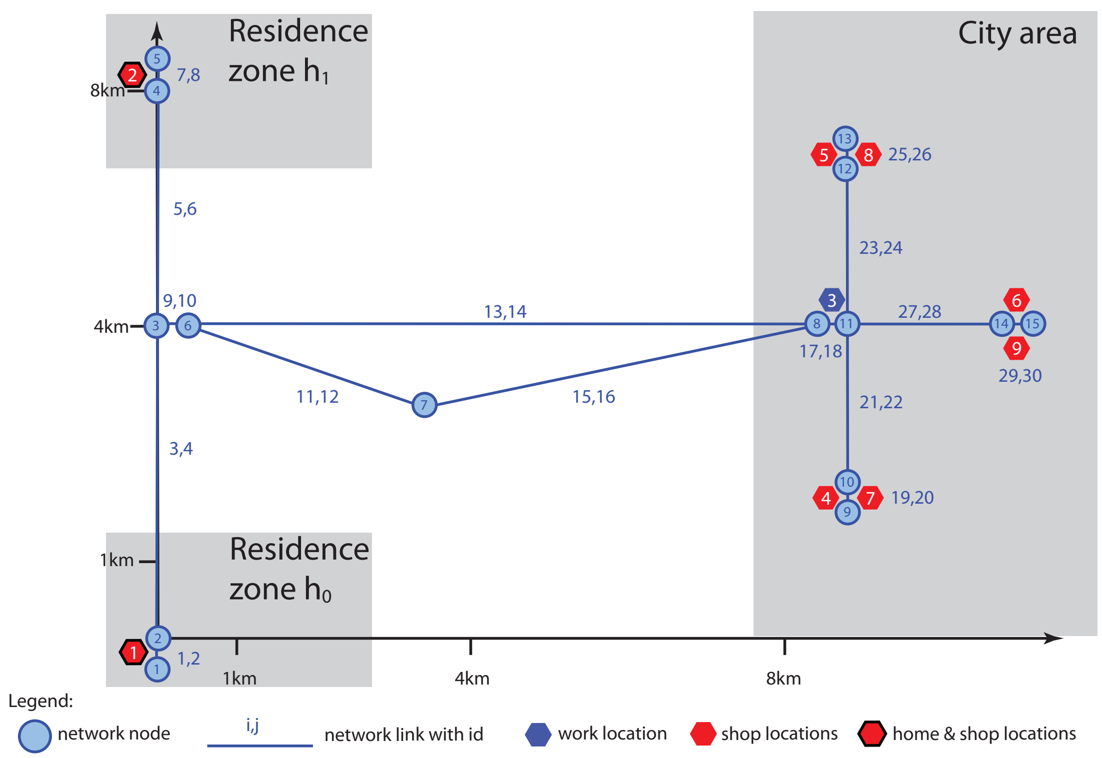
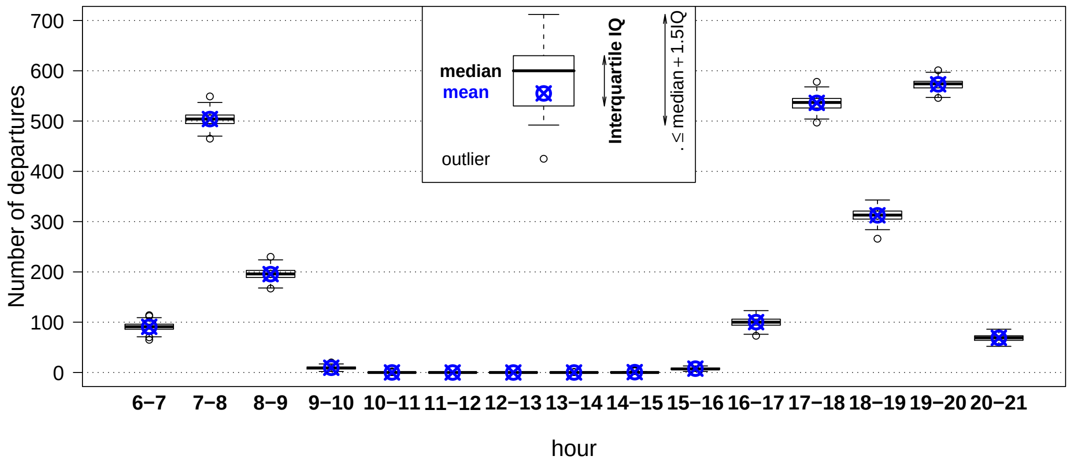


Figure 8.19: Departures for configuration 1 and 2 (relaxed state)

(a) Configuration 1



(b) Configuration 2

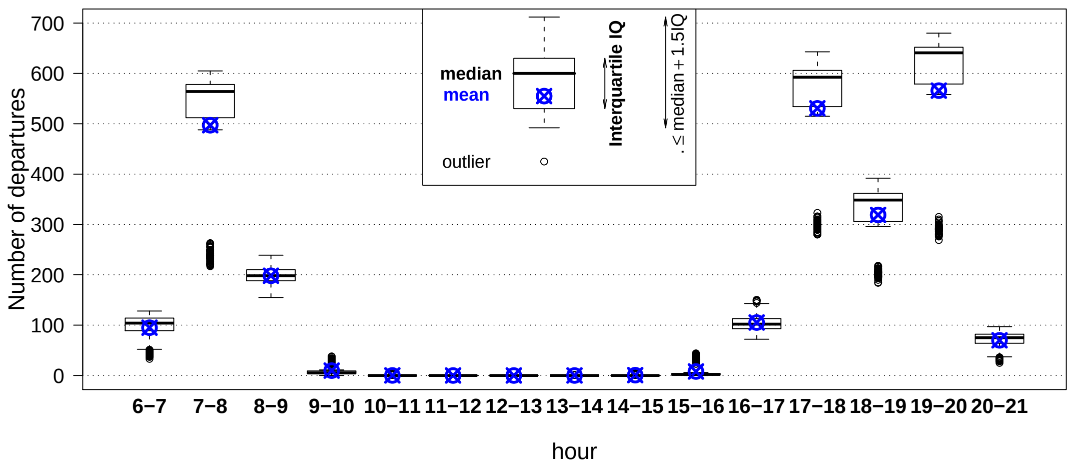
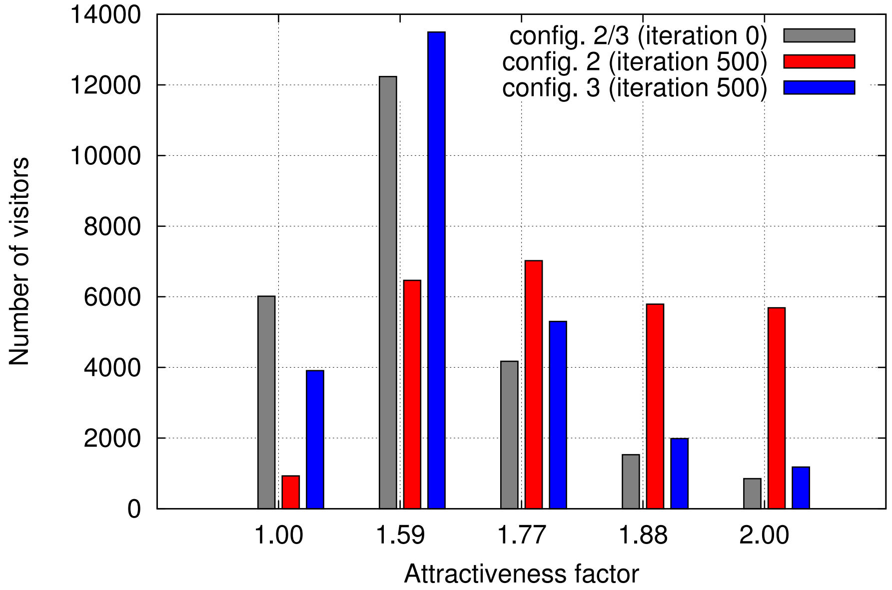


Figure 8.20: Activity location load

(a) Number of visitors at the activity locations with respect to $f_{attractiveness}$



(b) Number of visitors at the activity locations with respect to activity location load

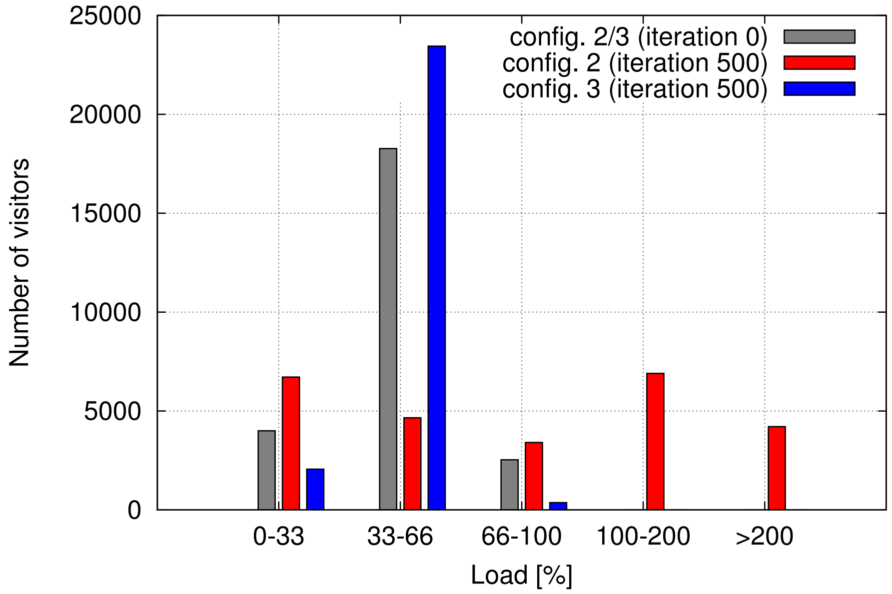


Figure 8.21: Aggregated hourly load of the shopping activity locations

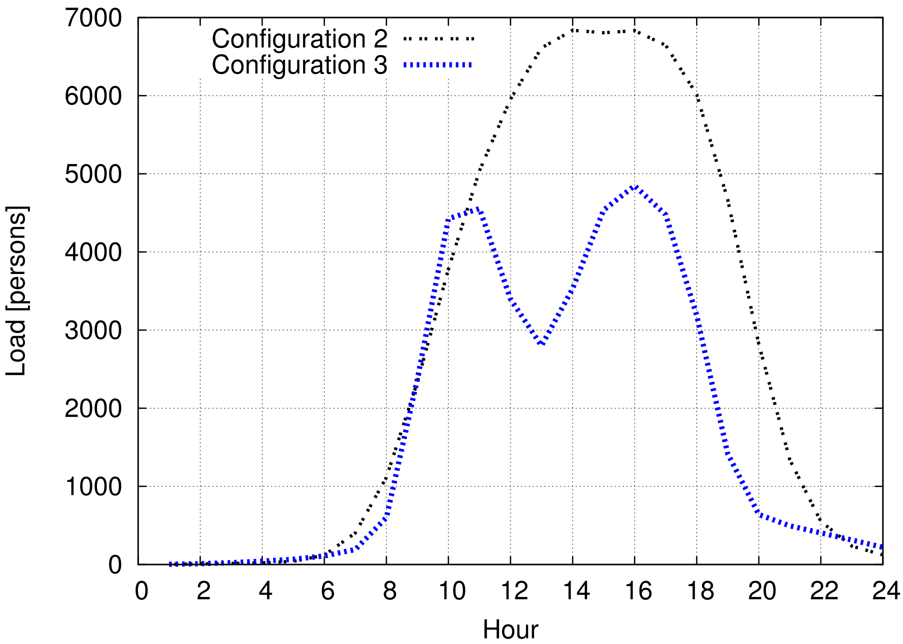


Figure 8.22: Chessboard scenario

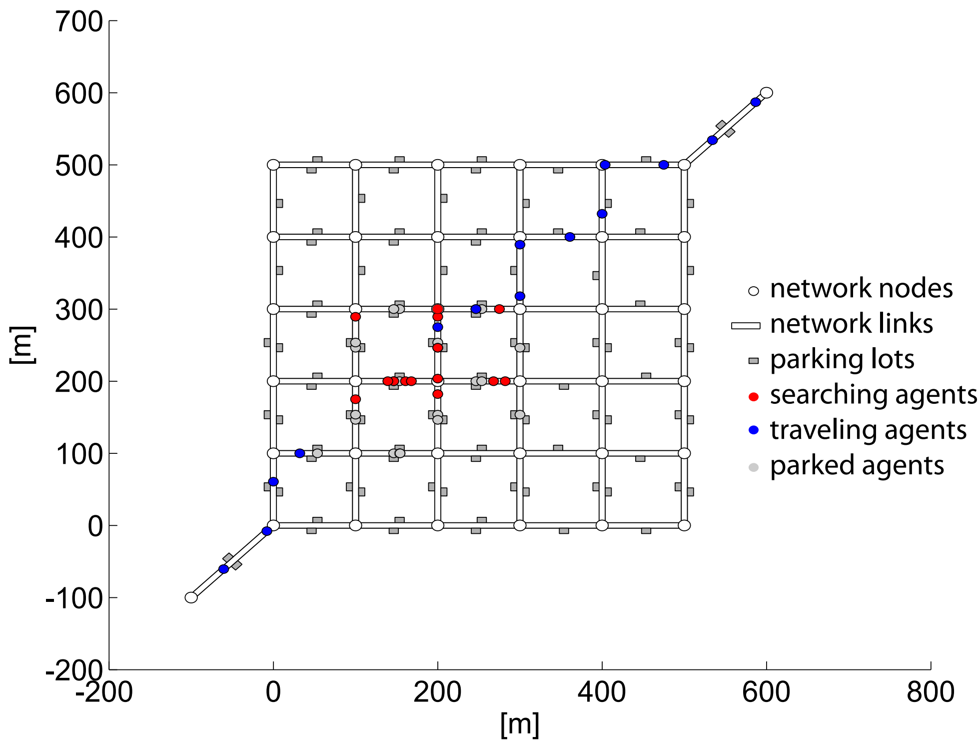


Figure 8.23: Aggregate search time model, chessboard scenario

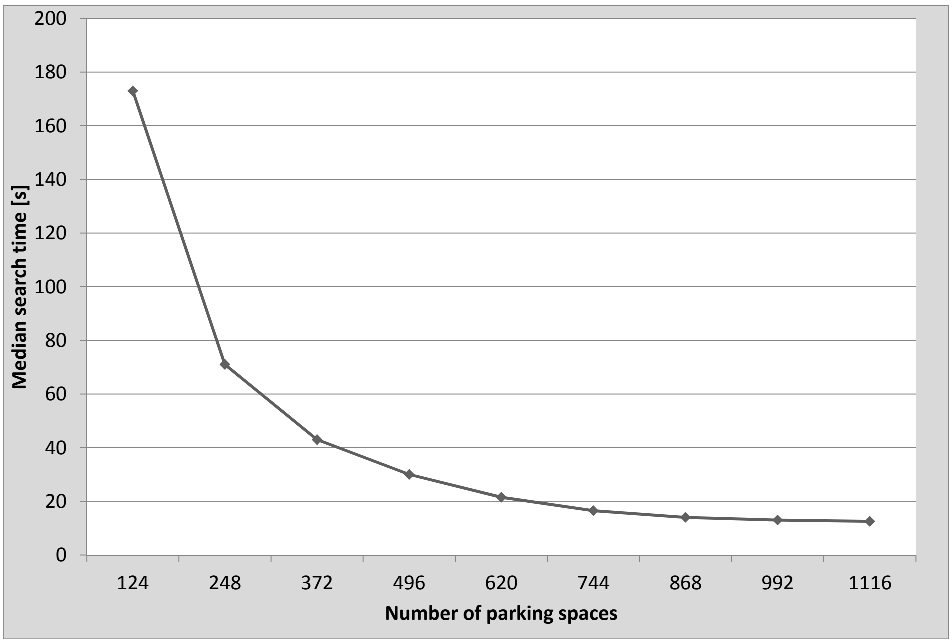
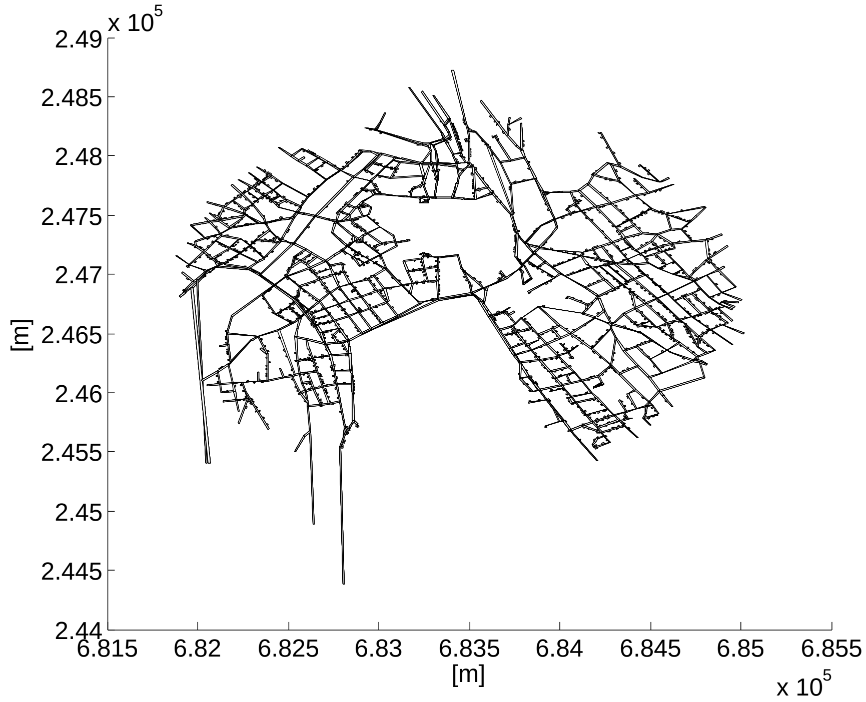


Figure 8.24: Zurich scenario

(a) Navigation network



(b) Planning network

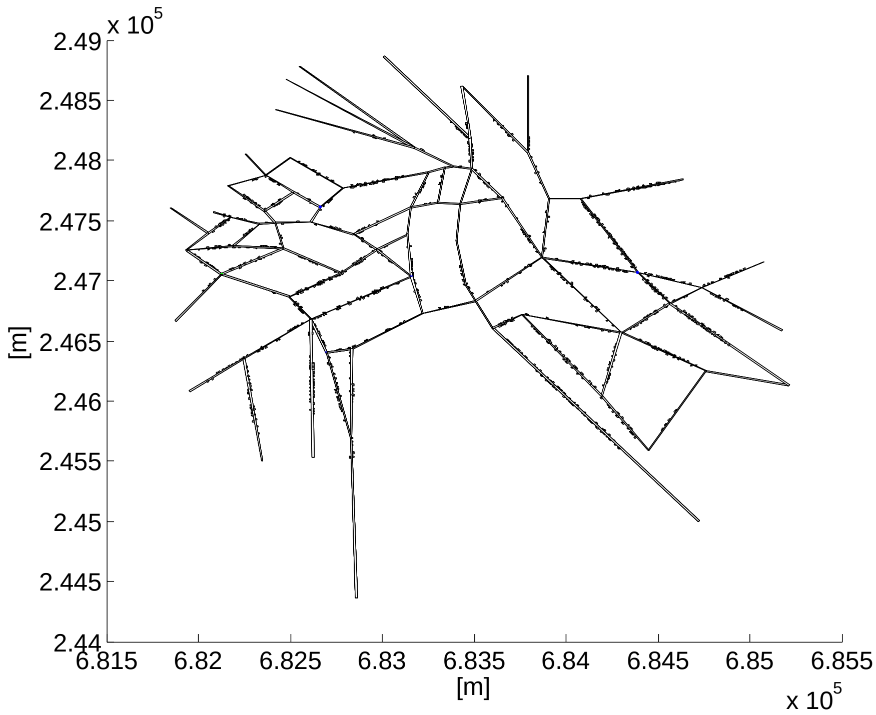


Figure 8.25: Search time distribution

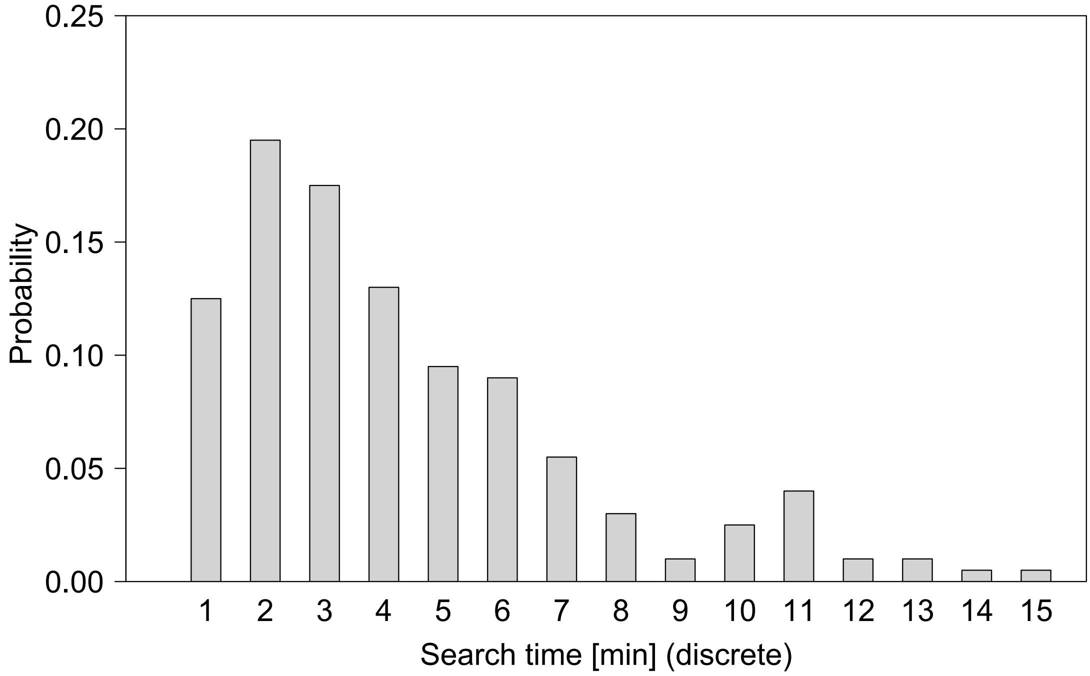
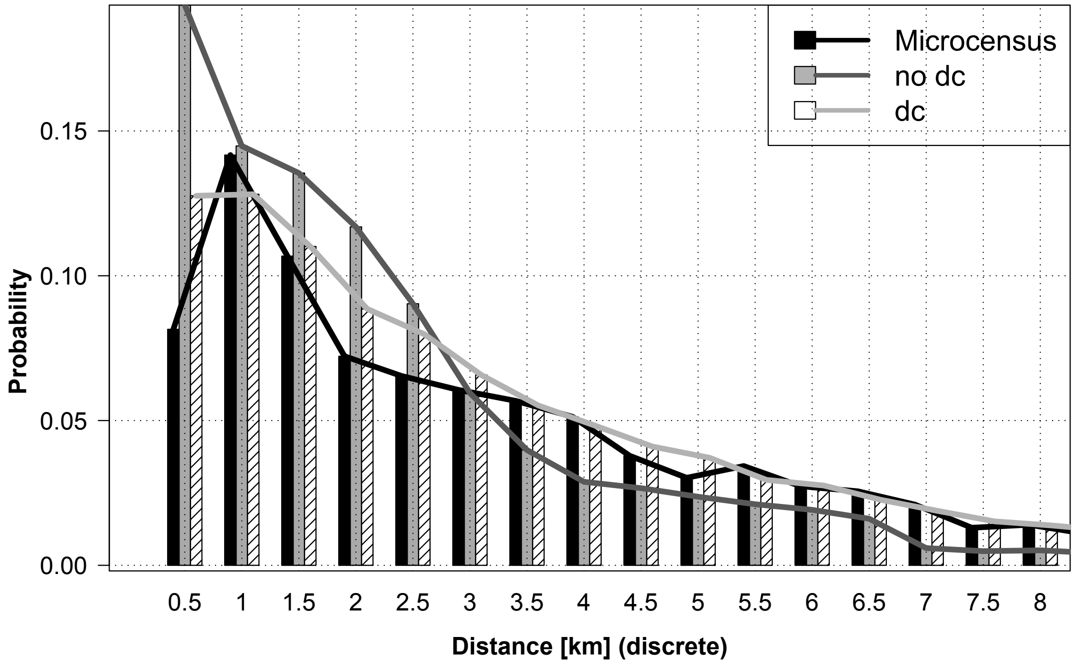


Figure 8.26: Zurich multi-modal scenario (relaxed state)

(a) Shopping trips



(b) Leisure trips

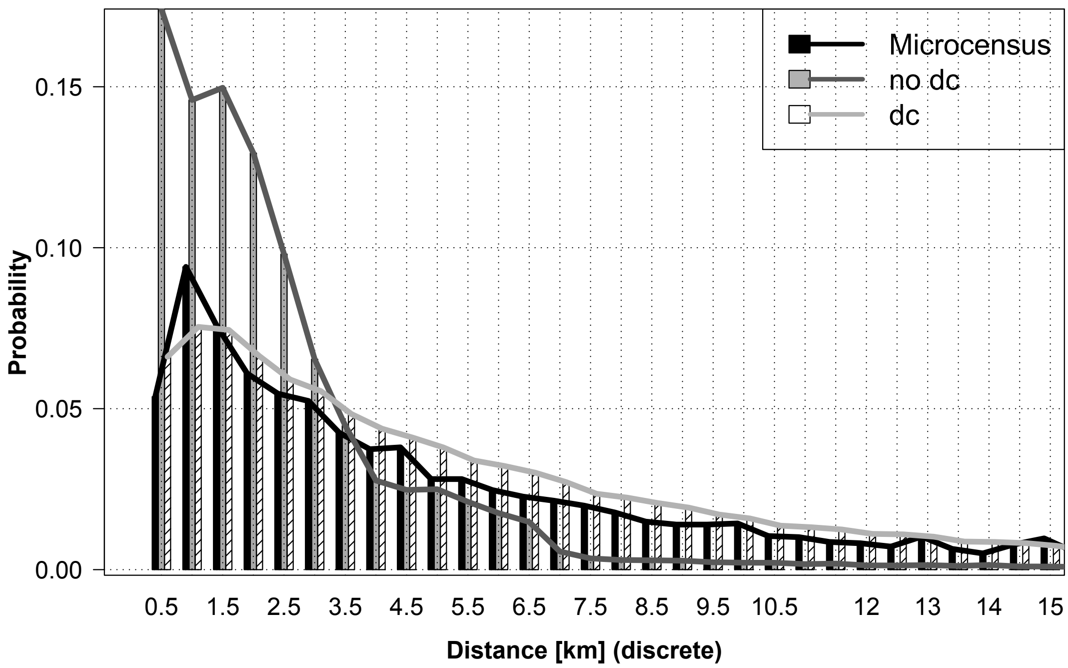


Figure 8.27: Daily traffic volumes for 123 links compared to traffic counts. Per link k the relative error is used, i.e, $(vol_{simulated,k} - vol_{counted,k})/vol_{counted,k}$.

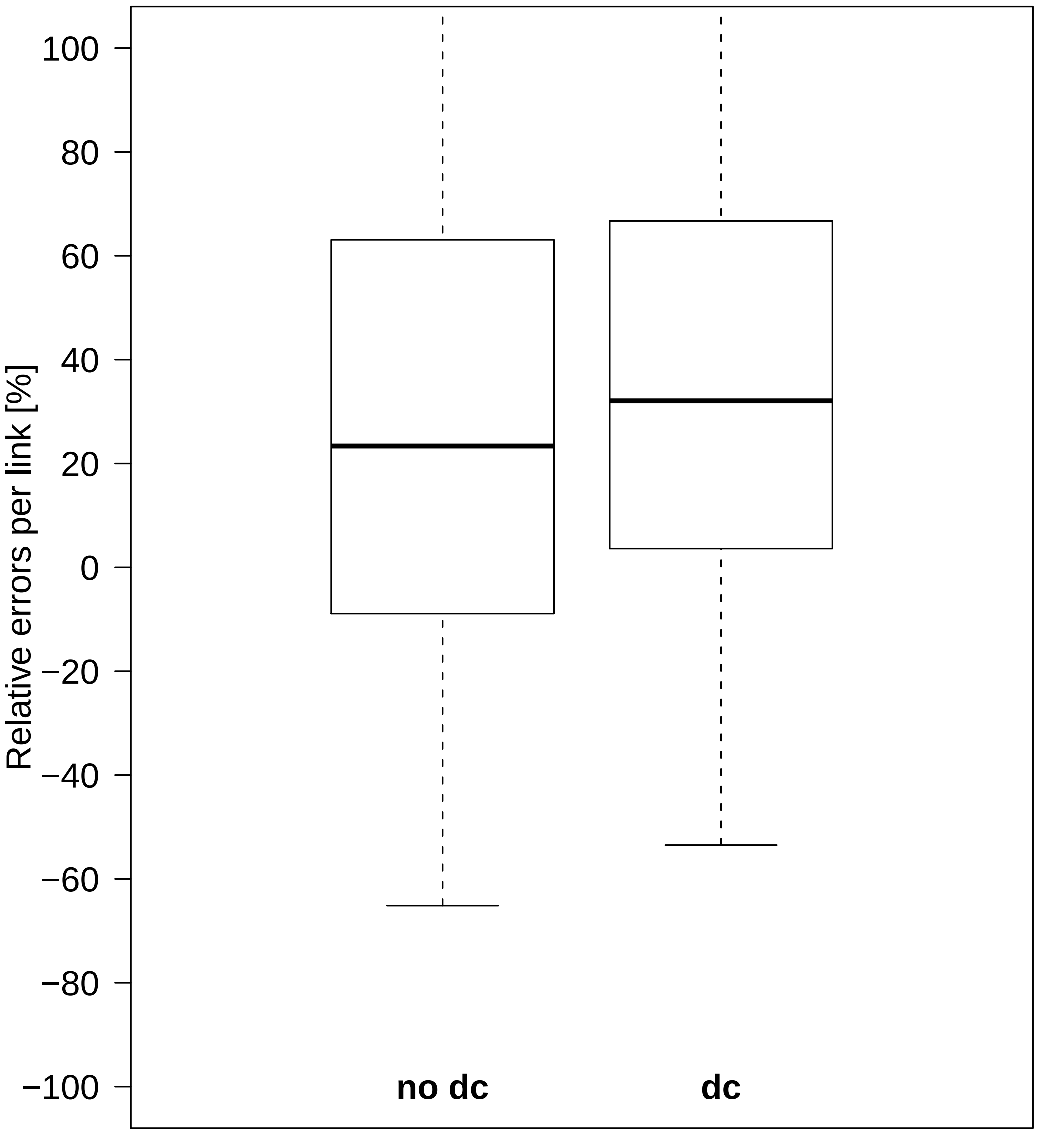
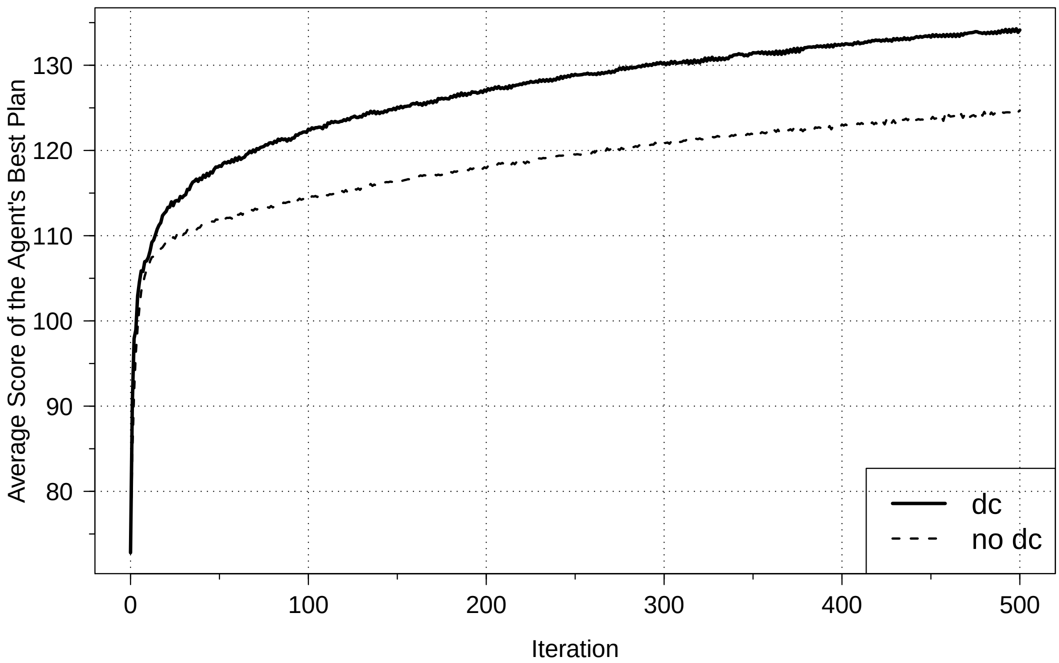


Figure 8.28: Average of the agent's best scores with and without destination choice (dc)



Chapter 9

Discussion

9.1 Discussion

9.1.1 Summary

This thesis solves the research problems formulated in Chapter 1 to the extent, that an operational MATSim destination choice module is provided being applicable in large-scale scenarios and conveniently adaptable for similar simulation frameworks. Main contribution is an approach to efficiently generate quenched randomness, being a standard requirement for many problems in iterative models. The thesis further contributes to the understanding of microsimulation variability and to destination choice set specification, where a map-based survey lays a basis for further approaching this problem. Besides providing data for our own analyses, the survey methodology has inspired personal communications and the source code has been passed to researchers developing similar surveys.

Furthermore, the main choice determinants and previous methods are identified, and the destination choice process as such is analyzed to broaden the methodological fundament of microsimulation destination choice modeling. Agent interaction modeling in the transport infrastructure, i.e., the competition for space-time slots) is extended in this work to the activities infrastructure, seldom investigated in previous models.

9.1.2 Current and Further Development

The two key transport model purposes are *understanding* simuland and *forecasting* the system's reaction to changes such as infrastructure and management measures. Relevant indicators measuring transport system behavior need to be defined in both cases. MATSim, and many similar models, are designed to be flexible tools, not a priori tailored to any specific purpose. Thus, specification of such indicators is non-trivial. It has become standard to compare simulated link volumes to measured road counts.

For MATSim, hourly link volumes for an average working day are computed as described in Horni (2007); Horni and Grether (2007). To a certain extent, these count data are an appropriate means to assess models and performed improvements. For example, count data clearly show that the model described in Section 4.2 strongly improved the microsimulation results; for the MATSim Zurich scenario the relative error in link volumes was strongly reduced.

However, the limits of model validation with count data are narrower than assumed at first sight. Count data, are a highly aggregated information, and they are hence structurally not a very appropriate validation means for microsimulations. Completely implausible microlevel processes could lead to plausible counts through the coaction of the high flexibility of microsimulations and capacity limits having the function of model constraints bringing the model to a reasonable domain. If a road is often operated at the capacity limit in simuland, then the inclusion of these limits can produce a very good match with count data in any case.

It is also difficult to show the benefits of disaggregate models over aggregate ones if assessment is exclusively based on aggregate validation data. Disaggregate data suitable for validation is provided by the Swiss Microcensus. However, large parts of it are used for model creation and calibration and are thus not available anymore for validation.

For this work, the assessment of model improvements by count data has reached its limits now. The improvements described in Chapters 5, 7 and Section 8.6, although based on obviously very important choice determinants, do not show a large effect in daily count data. Link volume sensitivity is relatively low. For the newer MATSim Zurich scenario, the situation is even more difficult. While the destination choice module corrects the distance distributions, count data comparison is made worse. The initial scenario with too short shopping and leisure distances is not well-calibrated. Clearly, as a next step, the analysis should be refined and spatio-temporally restrained to hours and areas relevant for shopping. However, count data remains problematic for assessing model improvements in this context. Thus, as a next big step before further developing the MATSim destination choice module, model validation data needs to be collected.

Besides others, with GPS Montini et al. (2013); Rieser-Schüssler (2012); Hackney et al. (2007); Axhausen et al. (2012) and GSM Nijkamp (2009) surveys, peoples presence at activity locations can be validated with high temporal precision and completeness. Despite legal barriers, license plate surveys might be another productive means for generating validation data. Another rich validation data source are business volumes, customer frequencies and loyalty card information. However, getting these data directly is difficult as they are usually part of business secret. To get shopping destination load, a possibility might be the manual collection of infrastructure data for single stores such as number of cash points, parking lots, available shopping baskets, or the size of shelves. Even more precise would be direct supply information such as amounts of perishable products like vegetables or fresh bread.

In addition to validation data collection, following validation strategies might be productive in MATSim. For agent-based evaluation the Microcensus persons could be transformed into MATSim agents, included in the simulation population, and then compared to the empirical data. This idea is often formulated in personal communication and it is worth testing, but a problem might be, that best results are achieved if the simulation does not change anything for these agents. To overcome short of validation data, small-scale validation for a cut-out of the scenario or k-fold validation might be a solution. A particularly interesting undertaking could be to compare forecasts with the real system's answer to a substantial infrastructure modification. A natural example for MATSim would be the construction of the Zurich Westumfahrung (Balmer et al., 2009).

Having said that, for model development and validation, following two (non-obvious) boundaries must be constantly paid attention to. On the one hand, large-scale scenarios essentially represent a full-population investigation. Importantly, the testing of model improvements does not only hunt for and rely on statistical significance but for substantive significance (the "*oomph*" as it is called in in Ziliak and McCloskey (2006) (and one of its critical reviews Mayer (2012)). On the other hand, even in very large scenarios, small random fluctuations, create an aggregate limit for model improvement. Even if we know all the micro-motives, we also need to know *when* these are triggered. The small-scale fluctuations triggered by micro-motives might not average out, but potentially add up to a substantial bias. A very simple example taken from the author's daily life is shown in Figure 9.1(a). The grocery store, marked with a red circle, lies directly at an important distributor road with many commuters and provides competitive prices and free and nearby parking. It covers the common and popular Migros product range and additionally special local products are available, generating a plus in attractiveness. However, Figure 9.1(b) shows why the store and in particular its parking lots are much less frequented than supposed considering above advantages. Even

Figure 9.1: Micro-motives and macro-scale biases

(a) Easy turn (b) Uncomfortable turn

the author often passes the store although it is perfectly located at his work route. If the signal-light for the oncoming lane shows green, very dense traffic makes it very uncomfortable to turn left and cross the line. A traffic jam results on the own lane. If there were parking lots on both sides of the road, the store might enjoy up to 50% load increase. Elimination of these kind of biases in the model would require a too high detail level, meaning that the modeling error is irreducible at some point of the exercise, where this stopping point may come sooner than one usually expects.

Having said that, further development looks promising along the lines identified below (Horni and Axhausen, 2012a).

9.2 Further Research Avenues

9.2.1 Further Heterogeneity of Agents and Alternatives

9.2.1.1 Income

MATSim attribute range and thus agents' and alternatives' heterogeneity is relatively small. Most important attribute, central in any econometric model, income, and also the derived measures such as value-of-time (VOT) are not yet included in standard MATSim scenarios. Methodically, income is easy to survey. It is not latent, as for example, preference for fresh food. Practically, however, privacy issues render data collection nevertheless difficult and costly. Nevertheless, valuable sources exist for Switzerland (Swiss Federal Statistical Office (BFS), 2008b, 2007, 2006). In addition, rents could be considered, being a less sensitive proxy for income. Importantly, spatial income distribution, in other words, spatial separation of wealth (and other socio-demographic attributes, possibly) defines required spatial model resolution. Smaller geographical heterogeneity means that resolution must be higher to capture the income effects.

9.2.1.2 Shopping Activity Duration and Travel Time, Variation of β_{dur}

MATSim's first principle is minimization of travel time, leaving more time for performing activities. For the standard configuration applied in most studies, value-of-time is static and identical for all persons and all activities. While activity performing time maximization is reasonable for leisure and in-home time, it is highly questionable for shopping. Apart from leisure shopping, it can be assumed that shopping is mainly focused on pursuing a purpose as quickly as possible. It can be further assumed, that a mixed calculation of money and time expenses, both for shopping and traveling, is performed by the persons. For the next modeling extensions, β_{dur} and β_{trav} should be made dependent on activity type and person. Application of a mixed logit approach is natural here.

The mixed calculation of travel and shopping costs might also partly explain the spatial indifference hypothesis discussed in Section 8.3.2.1.

9.2.1.3 Activity Sub-Classification

Another way for incorporating observed heterogeneity is finer activity classification. The most frequently used Switzerland and Zurich scenarios contain only 5 activity types (home, work, education, leisure and shop). Improved versions differentiating shopping (e.g., grocery vs. non-grocery) and leisure activities are available, but only in an experimental manner (Horni et al., 2011e). The National Travel Surveys (Swiss Federal Statistical Office (BFS), 2006) and Swiss Federal Enterprise Census 2001 (Swiss Federal Statistical Office (BFS), 2008a) provide a relatively detailed classification of activities for demand and supply side, respectively. Most activities are in principle performed by anybody, e.g., everybody does grocery and non-grocery shopping. Thus, activity sub-classification is not expected to have the largest effect for socio-demographics-based assignment of activity chains to persons, but for chain-based destination choice. In particular, multi-purpose shopping activities and maybe also agglomeration effects are expected to be modeled much better with a finer activity-classification. Two consecutive shopping activities, now differentiated by sub-types, cannot be, a priori, performed in the same store anymore. This is an improvement as surveys likewise only report shopping activities made in different stores. Technically, the destination choice module is now ready to handle any number of activity types providing the improvements discussed above.

9.2.2 Destination Choice Equilibration

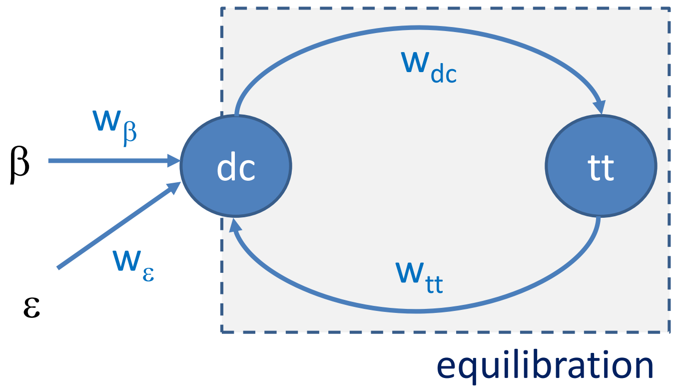
As shown earlier, equilibrium is a key concept in transport modeling. Traditional assignment procedures equilibrate static network flows given a fixed demand. For route choice, a strong influence of competitive travel time can be assumed beyond doubt. For destination choice, however, travel time differences between alternatives may be less important. In other words, in Figure 9.2, the choice weights w_β and w_ϵ compared to the weights of the equilibration variables w_{tt} w_{dc}, might be stronger for destination choice than for route choice. Hence, the equilibration assumption can be relaxed to a certain extent, allowing for stronger approximations, substantially reducing the computational burden. This thesis exploits this assumption by applying a best response approach. Further speed-ups might be possible.

Destination choice equilibration is also a future topic in terms of its behavioral base. The papers, investigating the empirical basis of the convenient assumption of equilibrium (e.g., Mahmassani and Chang, 1986) should be considered further and their applicability to destination choice should be assessed.

9.2.3 Longer Time Horizon

Travel time and money budgets and their stability and accurate observability are subject of intense and sometimes controversial discussion (see e.g., Goodwin (1981); Gunn (1981); Mokhtarian and Chen (2004), Kuhnimhof and Gringmuth (2009, p.182)). Nevertheless, in the first place, the budget concept is plausible, and it can easily be assumed that decisions are guided by budgets on different time-scales, meaning that budgets are balanced within different periods (Kuhnimhof and Gringmuth, 2009, p.179). Furthermore, when looking at the large weekend peaks of shopping and

Figure 9.2: MATSim destination choice equilibration: w_{tt} is a weight for travel time on the destination choice and w_{dc} is a weight for destination choice on the aggregate link travel times. w_{β} and w_{ϵ} are weights for the estimated parameters β and the error terms ϵ, respectively.



leisure activities, it can be assumed that time and money budgets for shopping and leisure activities are based (minimally) on a weekly horizon. Extending MATSim from an average working day to a weekly horizon is, thus, probably necessary for investigating such choices (see e.g., Ordóñez Medina et al. (2012)). In other words, applying a weekly horizon allows to explain time-variations instead of having them in the model as unobserved heterogeneity [1] . This is important as a large part of variability is usually intra-personal, i.e., temporal. Two thirds, respectively, more than half of variance being intra-personal is reported by Kuhnimhof and Gringmuth (2009) and Chikaraishi et al. (2010).

A week scenario—at the moment not yet including destination choice—was set up by Horni and Axhausen (2012b), implementing a warm-start mechanism to reduce the huge computational burden given by the enlarged scenario.

9.2.4 Speeding up the Destination Choice Module

9.2.4.1 Sampling

For speeding-up the destination choice module, a procedure that samples alternatives according to quenched randomness was successfully introduced in Section 8.6. A systematic analysis is outstanding.

9.2.4.2 Choice Dimension Dependencies

In MATSim to date, essentially all combinations of choices of different dimensions (i.e., time-route-mode-destination choices) are evaluated over the course of the iterations. However, replanning of agents' day plans is not done with consideration of choice dimensions' dependencies. Using the knowledge about choice combinations, e.g., its likelihood, might generate a substantial speed-up, relaxing the computational problems mentioned above. This is particularly true as it can be assumed that many choices regularly come as packets of choices, such as after-work shopping on the work route.

9.2.5 Incorporation of Spatial Correlations

In Section 8.6, inclusion of a term for spatial agglomerations τ_{agglo} in destination choice is described. Another possibility to account for corre-

[1] for a discussion of variability see e.g., Horni et al. (2011c) and for a discussion of MATSim modeling of longer-term decisions see e.g., Horni et al. (2012a)

lations between alternatives is estimation of models going beyond MNL. This is an important topic for the MATSim future.

9.2.6 Artificial Intelligence Approaches and Social Networks Models

Cognitive models of persons' spatial mental map (Axhausen, 2006; Chorus and Timmermans, 2009; Hannes et al., 2008; Mondschein et al., 2008; Arentze and Timmermans, 2004; Golledge and Timmermans, 1990; Bettman, 1979; Timmermans, 2008; Cadwallader, 1975) are promising in destination choice context to get under control its large choice sets. Apart from Dobler et al. (2009) cognitive models have not been applied yet in MATSim.

Furthermore, to capture complex destination choices, the MATSim agents could efficiently and consistently be extended toward complex decision making agents in artificial intelligence approaches such as Bazzan et al. (1999); Rossetti et al. (2002b,a); Rossetti and Liu (2005); Rindsfüser (2005); Dougherty et al. (1994); Wild (1994); Sadek (2007); Sadek et al. (2003); Anderson (2002); Bielli et al. (1994); Rao and Georgeff (1995); Gilbert (2006). The approach has a high parallelization capability, where, in the extreme, every agent can be simulated on its own CPU. However, similar as for the so-called within-day replanning, the decision making complexity must be adequately embedded in the equilibration process. At this point, equilibrium models and rule-based models overlap to some extent.

Finally, social networks are very promising, especially to capture leisure destination choices for leisure activities, which are very often social events. In MATSim, social coordination is investigated by Dubernet and Axhausen (2013); Hackney (2009).

9.2.7 General MATSim Topics: Convergence, Equilibrium and Volume-Delay Relationship

9.2.7.1 Convergence

MATSim as a research undertaking is subject to constant further development. For recent revisions, the author's experiments showed relatively week convergence above iteration 100 (e.g., Figure 8.28). This problem, occurring with and without destination choice and for different plan selectors (*SelectBest* and *SelectExpBeta*), requires future investigation.

Related to this issue, the initial investigations about MATSim equilibrium Meister (2011) require extension by analyses on existence, uniqueness and stability, maybe supported by the early mathematical approaches mentioned in Section 2.1.2.

Further analysis is also required for MATSim's volume-delay relationship. Horni and Montini (2013a) found that the relationship between travel time and link load is linear for MATSim. This might be a problem as traditionally for high loads, in particular when including intersections, non-linear relationships are assumed. The preliminary experiments need to be reproduced and, possibly, the queue simulation should be adapted, for example, by the back-traveling gaps, available in earlier mobility simulations (Charypar et al., 2009).

Bibliography

Abel, T. (1976) *Man is the measure: a cordial invitation to the central problems of philosophy*, Free Press, a division of Macmillan Publishing Company, London.

Adler, T. J. and M. E. Ben-Akiva (1976) Joint-choice model for frequency, destination and travel mode for shopping trips, *Transportation Research Record*, **569**, 136–150.

Adler, T. J. and M. E. Ben-Akiva (1979) A theoretical and empirical model of trip chaining behavior, *Transportation Research Part B: Methodological*, **13B** (3) 243–257.

AIMSUN (2013) AIMSUN, webpage, `http://www.aimsun.com`.

Alam, S. B. and K. G. Goulias (1999) Dynamic emergency evacuation management system using geographic information system and spatiotemporal models of behavior, *Transportation Research Record*, **1660**, 92–99.

Albrecht, B., T. Lütolf and T. Waltert (1998) Verkehrskennwerte von Einkaufszentren: Verkehrsaufkommen, Parkfelder und Verkehrsflächen, *Research Report*, **VSS 7/95**, Schweizerischer Verband der Strassen- und Verkehrsfachleute (VSS), Albrecht & Partner AG, Lucerne.

Albrecht, C.-M. (2009) *Einkaufsstress*, Schriftenreihe des Instituts für Marktorientierte Unternehmensführung (IMU), Universität Mannheim, Gabler Verlag, Wiesbaden.

Aldskogius, H. (1977) A conceptual framework and a Swedish case study of recreational behavior and environmental cognition, *Economic Geography*, **53** (2) 163–183.

Algers, S., E. Bernauer, M. Boero, L. Breheret, C. Di Taranto, M. Dougherty, K. Fox and J.-F. Gabard (1998) A review of microsimulation models, *Research Report*, Smartest: Transport RTD Programme of the 4th Framework Programme, European Commission, Institute for Transport Studies, University of Leeds, Leeds.

Ammoser, H. and M. Hoppe (2006) Glossar Verkehrswesen und Verkehrswissenschaften, *Schriftenreihe: Diskussionsbeiträge aus dem Institut für Wirtschaft und Verkehr*, **2**, Institut für Wirtschaft und Verkehr, Dresden.

Anderson, J. R. (2002) Spanning seven orders of magnitude: A challenge for cognitive modeling, *Cognitive Science*, **26**, 85–112.

Anderson, S. P. and A. de Palma (2004) The economics of pricing parking, *Journal of Urban Economics*, **55**, 1–20.

Ansah, J. A. (1977) Destination choice set definition in travel behaviour modelling, *Transportation Research*, **11** (2) 127–140.

Arentze, T. A., A. W. J. Borgers, F. Hofman, S. Fujii, C.-H. Joh, A. Kikuchi, R. Kitamura and P. Timmermans, Harry J. P.van der Waerden (2001) Rule-based versus utility-maximizing models of activity-travel patterns: A comparison of empirical performance, in D. A. Hensher (ed.) *Travel Behaviour Research: The Leading Edge*, 569–583, Pergamon, Oxford.

Arentze, T. A., A. W. J. Borgers and H. J. P. Timmermans (1994) Geographical information systems and the measurement of accessibility in the context of multipurpose travel: A new approach, *Journal of Geographical Systems*, **1** (2) 87–102.

Arentze, T. A., D. F. Ettema and H. J. P. Timmermans (2013) Location choice in the context of multi-day activity-travel patterns: model development and empirical results, *Transportmetrica A*, **9** (2) 107–123.

Arentze, T. A., F. Hofman, H. Mourik and H. J. P. Timmermans (2000) Albatross: A multi-agent rule-based model of activity pattern decisions, *Transportation Research Record*, **1706**, 136–144.

Arentze, T. A., M. Kowald and K. W. Axhausen (2011) A method to model population-wide social networks for large scale activity-travel micro-simulation, *Working Paper*, **698**, IVT, ETH Zurich, Zurich.

Arentze, T. A., H. Oppewal and H. J. P. Timmermans (2005) A multipurpose shopping trip model to assess retail agglomeration effects, *Journal of Marketing Research*, **42** (1) 109–115.

Arentze, T. A. and H. J. P. Timmermans (2000) *ALBATROSS: A Learning-Based Transportation Oriented Simulation*, EIRASS, Eindhoven.

Arentze, T. A. and H. J. P. Timmermans (2004) Capturing the role of awareness and information search processes on choice set formation in models of activity-travel behavior, paper presented at the *83rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2004.

Arentze, T. A. and H. J. P. Timmermans (2005) An analysis of context and constraints-dependent shopping behaviour using qualitative decision principles, *Urban Studies*, **42** (3) 435–448.

Arentze, T. A. and H. J. P. Timmermans (2007) Robust approach to modeling choice of locations in daily activity sequences, *Transportation Research Record*, **2003**, 59–63.

Arentze, T. A., H. J. P. Timmermans, D. Janssens and G. Wets (2006) Modeling short-term dynamics in activity-travel patterns: From Aurora to Feathers, paper presented at the *Innovations in Travel Demand Modeling (ITM'06)*, Austin, May 2006.

Arnott, R., A. de Palma and R. Lindsey (1991) A temporal and spatial equilibrium analysis of commuter parking, *Journal of Public Economics*, **45**, 301–335.

Arnott, R., A. de Palma and R. Lindsey (1993) A structural model of peak-period congestion: A traffic bottleneck with elastic demand, *The American Economic Review*, **83** (1) 161–179.

ASTRA (2006) Swiss federal roads authority, webpage, `http://www.astra.admin.ch/`.

Axhausen, K. W. (1988) Eine ereignisorientierte Simulation von Aktivitätenketten zur Parkstandswahl, Ph.D. Thesis, University of Karlsruhe, Karlsruhe.

Axhausen, K. W. (1997) Data needs of activity scheduling models, in D. F. Ettema and H. J. P. Timmermans (eds.) *Activity-Based Approaches to Travel Analysis*, 229–241, Pergamon, Oxford.

Axhausen, K. W. (2006) Neue Modellansätze der Verkehrsnachfragesimulation: Entwicklungslinien, Stand der Forschung, Forschungsperspektiven, *Stadt Region Land*, **81**, 149–164.

Axhausen, K. W. (2007) Activity spaces, biographies, social networks and their welfare gains and externalities: Some hypotheses and empirical results, *Mobilities*, **2**, 15–36.

Axhausen, K. W. (2012) Social networks and travel: Recent international results, presentation, ITRN, Belfast, August 2012.

Axhausen, K. W. and R. Herz (1989) Simulating activity chains: German approach, *Journal of Transportation Engineering*, **115** (3) 316–325.

Axhausen, K. W. and J. W. Polak (1991) Choice of parking: Stated preference approach, *Transportation*, **18** (1) 59–81.

Axhausen, K. W., J. W. Polak, M. Boltze and J. Puzicha (1994) Effectiveness of the parking guidance information system in Frankfurt am Main, *Traffic Engineering and Control*, **35** (5) 304–309.

Axhausen, K. W., N. Rieser-Schüssler and L. Montini (2012) Zwischen Tagebuch und GPS Beobachtung: Wie weiter mit der Messung des Verkehrsverhaltens?, presentation, Technical University Graz, Graz, June 2012.

Axhausen, K. W. and C. Weis (2010) Predicting response rate: A natural experiment, *Survey Practice*, **3** (2).

Axhausen, K. W., A. Zimmermann, S. Schönfelder, G. Rindsfüser and T. Haupt (2002) Observing the rhythms of daily life: A six-week travel diary, *Transportation*, **29** (2) 95–124.

Baier, R., C. Hebel, C. Peter and K. H. Schäfer (2000) Gesamtwirkungsanalyse zur Parkraumbewirtschaftung, *Research Report*, Bundesamt für Strassenwesen (BAST), Büro für Stadt- und Verkehrsplanung Dr.Ing. Reinhold Baier GmbH, Aachen.

Baker, J., D. Grewel and A. Parasuraman (1994) The influence of store environment on quality inferences and store image, *Journal of the Academy of Marketing Science*, **22** (4) 328–339.

Balmer, M. (2007) Travel demand modeling for multi-agent traffic simulations: Algorithms and systems, Ph.D. Thesis, ETH Zurich, Zurich, May 2007.

Balmer, M., K. W. Axhausen and K. Nagel (2006) An agent-based demand-modeling framework for large scale micro-simulations, *Transportation Research Record*, **1985**, 125–134.

Balmer, M., A. Horni, K. Meister, F. Ciari, D. Charypar and K. W. Axhausen (2009) Wirkungen der Westumfahrung Zürich: Eine Analyse mit einer Agenten-basierten Mikrosimulation, *Final Report*, Baudirektion Kanton Zurich, IVT, ETH Zurich, Zurich, February 2009.

Balmer, M., K. Meister, R. A. Waraich, A. Horni, F. Ciari and K. W. Axhausen (2010) Agenten-basierte Simulation für location based services, *Final Report*, F&E Förderung: Science to Market, **KTI 8443.1 ESPP-ES**, Datapuls AG, IVT, ETH Zurich, Zurich, February 2010.

Balmer, M., B. Raney and K. Nagel (2005) Adjustments of activity timing and duration in an agent-based traffic flow simulation, in H. J. P. Timmermans (ed.) *Progress in Activity-Based Analysis*, 91–114, Elsevier, Oxford.

Balmer, M. and M. Rieser (2004) Generating daily activity chains from origin-destination matrices, *Working Paper*, **243**, IVT, ETH Zurich, Zurich.

Baltas, G. and P. Papastathopoulou (2003) Shopper characteristics, product and store choice criteria: A survey in the greek grocery sector, *International Journal of Retail & Distribution Management*, **31** (10) 498–507.

Banks, J. (2001) *Discrete-event system simulation*, Prentice Hall, Upper Saddle River.

Barnard, P. O. (1987) Modelling shopping destination choice behaviour using the basic multinomial logit model and some of its extensions, *Transport Reviews*, **7** (1) 17–51.

Barton-Aschman Associates, Inc.and Cambridge Systematics, Inc. (1997) *Model Validation and Reasonableness Checking Manual*, Travel model improvement program (TMIP), Federal Highway Administration, Arlington, Texas, February 1997, http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/mvrcm/.

Bates, J. (2000) History of demand modelling, in D. A. Hensher and K. J. Button (eds.) *Handbook of Transport Modelling*, chap. 2, 11–33, Pergamon, Amsterdam.

Batt, P. J. (2009) Factors influencing the consumer's choice of retail store, paper presented at the *19th IAMA Annual World Forum and Symposium*, Budapest, June 2009.

Baudirektion Kanton Zurich (2008) Bauprojekt Westumfahrung Zürich, webpage, `http://www.westumfahrung.ch/`.

Bawa, K. and A. Ghosh (1999) A model of household grocery shopping behavior, *Marketing Letters*, **10** (2) 149–160.

Bazzan, A. L. C., J. Wahle and F. Klügl (1999) Agents in traffic modelling - from reactive to social behaviour, in W. Burgard, T. Christaller and A. B. Cremers (eds.) *Lecture Notes in Computer Science: Proceedings of the 23rd Annual German Conference on Artificial Intelligence: Advances in Artificial Intelligence 1999*, vol. 1701, 303–306, Springer.

Beckmann, M., C. B. McGuire and C. B. Winston (1956) *Studies in the Economics of Transportation*, Yale University Press, New Haven.

Bekhor, S., C. Dobler and K. W. Axhausen (2011) Integration of activity-based with agent-based models: an example from the Tel Aviv model and MATSim, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Bekhor, S. and J. N. Prashker (2008) GEV-based destination choice models that account for unobserved similarities among alternatives, *Transportation Research Part B: Methodological*, **42** (3) 243–262.

Bell, D. R., T.-H. Ho and C. S. Tang (1998) Determining where to shop: Fixed and variable costs of shopping, *Journal of Marketing Research*, **35** (3) 352–369.

Bell, S. J. (1999) Image and consumer attraction to intraurban retail areas: An environmental psychology approach, *Journal of Retailing and Consumer Services*, **6** (2) 67–78.

Ben-Akiva, M. E. (1974) Structure of passenger travel demand models, *Transportation Research Record*, **526**, 26–42.

Ben-Akiva, M. E. and B. Boccara (1995) Discrete choice models with latent choice sets, *International Journal of Research in Marketing*, **12** (1) 9–24.

Ben-Akiva, M. E., J. L. Bowman and D. Gopinath (1996) Travel demand model system for the information era, *Transportation*, **23** (3) 241–266.

Ben-Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.

Benekohal, R. F. and G. Abu-Lebdeh (1994) Variability analysis of traffic simulation outputs: Practical approach for TRAF-NETSIM, *Transportation Research Record*, **1457**, 198–207.

Benenson, I., K. Martens and S. Birfir (2008) PARKAGENT: an agent-based model of parking in the city, *Computers, Environment and Urban Systems*, **32** (6) 431–439.

Berg, W. D., P. A. Koushki, C. L. Krueger and W. L. Bittner (1976) Development of a simulation model for regional recreational travel, *Transportation Research Record*, **569**, 96–106.

Bernardin, V. L., F. S. Koppelman and D. Boyce (2009) Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Berry, B. J. L., H. G. Barnum and R. J. Tennant (1962) Retail location and consumer behavior, *Papers of the Regional Science Association*, **9**, 65–106.

Bettman, J. R. (1979) *An Information Processing Theory of Consumer Choice*, Addison-Wesley, Massachusetts.

BfS (2011) Swiss federal statistical office (bfs), webpage, `http://www.bfs.admin.ch`.

Bhat, C. R. (1998) Activity-based travel demand modeling, presentation, Oregon Symposium, September.

Bhat, C. R. (2005) A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions, *Transportation Research Part B: Methodological*, **39** (8) 679–707.

Bhat, C. R., J. Y. Guo, S. Srinivasan and A. Sivakumar (2004) A comprehensive econometric microsimulator for daily activity-travel patterns (CEMDAP), *Transportation Research Record*, **1894**, 57–66.

Bhat, C. R. and F. S. Koppelman (2003) Activity-based modeling of travel demand, in R. W. Hall (ed.) *Handbook of Transportation Science*, 39–65, Springer, New York.

Bhat, C. R. and I. N. Sener (2009) A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units, *Journal of Geographical Systems*, **11** (3) 243–272.

Bhat, C. R. and S. K. Singh (2000) A comprehensive daily activity-travel generation model system for workers, *Transportation Research Part A: Policy and Practice*, **34** (1) 1–22.

Bielli, M., G. Ambrosino and M. Boero (eds.) (1994) *Artificial Intelligence Applications to Traffic Engineering*, VSP, Utrecht.

Bierlaire, M. (2013) *An introduction to BIOGEME (Version 2.2)*, `http://biogeme.epfl.ch/`.

Bierlaire, M. and F. Crittin (2006) Solving noisy, large-scale fixed-point problems and systems of nonlinear equations, *Transportation Science*, **40** (1) 44–63.

Bierlaire, M., R. Hurtubia and G. Flötteröd (2009) A comparative analysis of implicit and explicit methods to model choice set generation, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.

Bierlaire, M., T. Lotan and P. L. Toint (1997) On the overspecification of multinomial and nested logit models due to alternative specific constants, *Transportation Science*, **31** (4) 363–371.

Blum, J. (1954) Multidimensional stochastic approximation methods, *Annals of Mathematical Statistics*, **25** (4) 737–744.

Bodenmann, B. R. (2005) Modelle zur Standortwahl von Unternehmen, *Working Paper*, **336**, IVT, ETH Zurich, Zurich.

Bonabeau, E. (2002a) Agent-based modeling: Methods and techniques for simulating human systems, *Proceedings of the National Academy of Sciences of the United States of America*, **99** (3) 7280–7287.

Bonabeau, E. (2002b) Predicting the unpredictable, *Harvard Business Review*, **80** (3) 5–11.

Bonsall, P. W. and I. A. Palmer (2004) Modelling drivers' car parking behaviour using data from a travel choice simulator, *Transportation Research Part C: Emerging Technologies*, **12** (5) 321–347.

Borgers, A. W. J. and H. J. P. Timmermans (1986) City centre entry points, store location patterns and pedestrian route choice behaviour: A microlevel simulation model, *Socio-Economic Planning Sciences*, **20** (1) 25–31.

Bossel, H. (2004) *Systeme, Dynamik, Simulation: Modellbildung, Analyse und Simulation komplexer Systeme*, Books on Demand.

Bourahla, M. and M. Benmohamed (2005) Formal specification and verification of multi-agent systems, *Electronic Notes in Theoretical Computer Science*, **123**, 5–17.

Bowman, J. L. (2009a) Historical development of activity based model theory and practice (part 1), *Traffic Engineering and Control*, **50** (2) 59–62.

Bowman, J. L. (2009b) Historical development of activity based model theory and practice (part 2), *Traffic Engineering and Control*, **50** (7) 314–318.

Bowman, J. L. and M. E. Ben-Akiva (1996) Activity-based travel forecasting, paper presented at the *Activity-Based Travel Forecasting Conference*, New Orleans, June 1996.

Bowman, J. L. and M. E. Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A: Policy and Practice*, **35** (1) 1–28.

Bowman, J. L., M. A. Bradley, Y. Shiftan, T. K. Lawton and M. E. Ben-Akiva (1999) Demonstration of an activity based model system for Portland, in H. Meersman, E. van de Voorde and W. Winkelmans (eds.) *World Transport Research*, vol. 3, 171–184, Pergamon, Oxford.

Boyce, D. and H. C. W. L. Williams (2003) Urban travel forecasting in the USA and UK, paper presented at the *43rd Congress of the European Regional Science Association*, Jyväskylä, August 2003.

Boyce, D. and I. Williams (eds.) (forthcoming) *Forecasting Urban Travel: Past, Present and Future*, Edward Elgar, Northampton.

Boyce, D. E., L. J. LeBlanc and K. S. Chon (1988) Network equilibrium models of urban location and travel choices: A retrospective survey, *Journal of Regional Science*, **28** (2) 159–183.

Bradley, M. A. (2005) Activity-based models in practice: Portland & San Francisco, presentation, TRB 2005 Workshop, Activity-Based Approaches: Theory, Methods, Data & Applications, Washington, D.C.

Bradley, M. A. and J. L. Bowman (2006) Design features of activity-based microsimulation models for U.S. metropolitan planning organizations, paper presented at the *Innovations in Travel Demand Modeling (ITM'06)*, Austin, May 2006.

Bradley, M. A. and J. L. Bowman (2009) SCAG project 09-012: Strategy for activity-based travel demand model development with travel survey - final report, *Research Report*, Southern California Association of Governments - SCAG, Mark Bradley Research & Consulting, Transportation Systems and Decision Sciences, Santa Barbara.

Bradley, M. A., J. L. Bowman and B. Griesenbeck (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution, *Journal of Choice Modelling*, **3** (1) 5–31.

Braess, D. (1969) Über ein Paradoxon aus der Verkehrsplanung, *Unternehmensforschung*, **12**, 258–268.

Brown, S. (1994) Retail location at the micro-scale: Inventory and prospect, *The Service Industries Journal*, **14** (4) 542–576.

Brunner, J. A. and J. L. Mason (1968) The influence of driving time upon shopping center preference, *Journal of Marketing*, **32** (2) 57–61.

Buliung, R. N. and T. Hernández (2009) Places to shop and places to grow, *Research Report*, NEPTIS Studies on the Toronto Metropolitan Region, NEPTIS, University of Toronto and Ryerson University, Toronto.

Buliung, R. N., T. Hernández and J. E. Mitchell (2007) Exploring the relationship between big-box retailing and consumer travel demand in the Greater Toronto area, *Canadian Journal of Transportation*, **1** (2) 121–141.

Buliung, R. N., M. J. Roorda and T. K. Remmel (2008) Exploring spatial variety in pattterns of activity-travel behaviour: Initial results from the Toronto Travel-Activity Panel Survey (TTAPS), *Transportation*, **35** (6) 697–722.

Burmeister, B., A. Haddadi and G. Matylis (1997) Application of multi-agent systems in traffic and transportation, *IEE Proceedings Software Engineering*, **144** (1) 51–60.

Burnett, K. P. (1980) Spatial constraints-oriented modelling as an alternative approach to movement: Microeconomic theory and urban policy, *Urban Geography*, **1** (1) 53–67.

Burnett, K. P. and S. Hanson (1979) Rationale for an alternative mathematical approach to movement as complex behavior, *Transportation Research Record*, **723**, 11–24.

Burnett, P. (1977) Tests of a linear learning model of destination choice: Applications to shopping travel by heterogeneous population groups, *Geografiska Annaler*, **59** (2) 95–108.

Cadwallader, M. (1975) A behavioral model of consumer spatial decision making, *Economic Geography*, **51** (4) 339–349.

Cadwallader, M. (1995) Interaction effects in models of consumer spatial behaviour, *Applied Geography*, **15** (2) 135–145.

Caliper (2013) Transmodeler, webpage, `http://www.caliper.com/transmodeler/`.

Cambridge Systematics Inc. (2008) Tel Aviv activity schedule travel demand model system: A tour-based approach, *Research Report*, Ministry of Transport, Cambridge Systematics Inc., December 2008.

Cambridge Systematics, Inc. (2008) *Travel Model Validation Practices Peer Exchange White Paper*, Travel model improvement program (TMIP), Federal Highway Administration, Arlington, Texas, December 2008.

Cameron, G. D. B. and G. I. D. Duncan (1996) PARAMICS-parallel microscopic simulation of road traffic, *Journal of Supercomputing*, **10** (1) 25–53.

Cantillo, V. and J. d. D. Ortúzar (2006) Implications of thresholds in discrete choice modelling, *Transport Reviews*, **26** (6) 667–691.

Carrasco, J. A., B. Hogan, B. Wellman and E. J. Miller (2008) Collecting social network data to study social activity-travel behavior: An ego-centered approach, *Environment and Planning B*, **36** (6) 961–980.

Carrasco, N. (2008) Deciding where to shop: Disaggregate random utility destination choice modeling of grocery shopping in canton Zurich, Master Thesis, IVT, ETH Zurich, Zurich.

Cascetta, E. and A. Papola (2009) Dominance among alternatives in random utility models, *Transportation Research Part A*, **43** (2) 170–179.

Casey, H. J. (1955) Applications to traffic engineering of the law of retail gravitation, *Traffic Quarterly*, **IX** (1) 23–35.

Castiglione, J., J. Freedman and M. A. Bradley (2003) Systematic investigation of variability due to random simulation error in an activity-based microsimulation forecasting model, *Transportation Research Record*, **1831**, 76–88.

Cetin, N. (2005) Large-scale parallel graph-based simulations, Ph.D. Thesis, ETH Zurich, Zurich.

Chamhuri, N. and P. J. Batt (2009) Factors influencing the consumer's choice of retail food store, *Stewart Postharvest Review*, **5** (3) 1–7.

Chapin, F. S. (1974) *Human Activity Patterns in the City*, John Wiley & Sons, Hoboken.

Chapman, R. G. and R. Staelin (1982) Exploiting rank ordered choice set data within the stochastic utility model, *Journal of Marketing Research*, **19** (3) 288–301.

Charnes, A. and W. W. Cooper (1961) Multicopy traffic network models, in R. Herman (ed.) *Theory of Traffic Flow*, 85–96, Elsevier, Amsterdam.

Charypar, D., K. W. Axhausen and K. Nagel (2006) Implementing activity-based models: Accelerating the replanning process of agents using an evolution strategy, paper presented at the *11th International Conference on Travel Behaviour Research (IATBR)*, Kyoto, August 2006.

Charypar, D., K. W. Axhausen and K. Nagel (2007) Event-driven queue-based traffic flow microsimulation, *Transportation Research Record*, **2003**, 35–40.

Charypar, D., M. Balmer and K. W. Axhausen (2009) High-performance traffic flow microsimulation for large problems, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Charypar, D. and K. Nagel (2003) Generating complete all-day activity plans with genetic algorithms, paper presented at the *10th International*

Conference on Travel Behaviour Research (IATBR), Lucerne, August 2003.

Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with genetic algorithms, *Transportation*, **32** (4) 369–397.

Chikaraishi, M., J. Zhang, A. Fujiwara and K. W. Axhausen (2010) Exploring variation properties of time use behavior based on a multilevel multiple discrete-continuous extreme value model, *Transportation Research Record*, **1** (2156) 101–110.

Chiu, Y.-C., J. Bottom, M. Mahut, A. Paz, R. Balakrishna, S. T. Waller and J. Hicks (2010) A primer for Dynamic Traffic Assignment, *Technical Report*, ADB30 Transportation Network Modeling Committee, Transportation Research Board, Washington, D.C.

Chorus, C. G. and H. J. P. Timmermans (2009) An empirical study into the influence of travel behavior on stated and revealed mental maps, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Ciari, F. and K. W. Axhausen (2011) Modeling location decisions of retailers with an agent-based approach, *Arbeitsberichte Verkehrs- und Raumplanung*, **671**, IVT, ETH Zurich, Zurich.

Ciari, F., M. Löchl and K. W. Axhausen (2008) Location decisions of retailers: An agent-based approach, paper presented at the *International Conference on Recent Advances in Retailing and Services Science*, Zagreb, July 2008.

Cirillo, C., E. Cornélis, L. Legrain and P. L. Toint (2003) Combining spatial and temporal dimensions in destination choice models, paper presented at the *European Transport Conference*, Strasbourg, October 2003.

Colome, R. and D. Serra (2000) Supermarket key attributes and location decisions: A comparative study between British and Spanish consumers, *Economic and Business Working Paper*, **469**, Pompeu Fabra University, Barcelona.

Comparis (2011) Remis zwischen Coop und Migros (komplette Vergleichsliste), February 2011, `http://www.comparis.ch/~/media/files/mediencorner/konsumentenstimme/ausgabe%202011_06/preisvergleich_detailhandel.pdf`.

Converse, P. D. (1949) New laws of retail gravitation, *Journal of Marketing*, **14** (3) 379–384.

Cools, M., B. Kochan, T. Bellemans, D. Janssens and G. Wets (2011) Assessment of the effect of micro-simulation error on key travel indices: Evidence from the activity-based model FEATHERS, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Correa, J. and N. E. Stier-Moses (2010) Wardrop equilibria, in J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh and J. C. Smith (eds.) *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, Hoboken.

Crompton, J. L. (1992) Structure of vacation destination choice sets, *Annals of Tourism Research*, **19** (3) 420–434.

Dafermos, S. (1971) An extended traffic assignment model with applications to two-way traffic, *Transportation Science*, **5** (4) 367–389.

Dafermos, S. (1980) Traffic equilibrium and variational inequalities, *Transportation Science*, **14** (1) 42–54.

Dafermos, S. (1983) An iterative scheme for variational inequalities, *Mathematical Programming*, **26** (1) 40–47.

Daganzo, C. F. (1983) Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost jacobians, *Transportation Science*, **17** (3) 282–300.

Daganzo, C. F. and Y. Sheffi (1977) On stochastic models of traffic assignment, *Transportation Science*, **11** (3) 253–274.

Daly, A. J. (2013) Forecasting travel demand using large-scale models, keynote, 13th Swiss Transport Research Conference, Ascona, April 2013.

Darbha, S., K. R. Rajagopal and V. Tyagi (2008) A review of mathematical models for the flow of traffic and some recent results, *Nonlinear Analysis*, **69** (3) 950–970.

Datye, S. S. (2012) Variability analysis of transport microsimulations with MATSim, *Term Report*, IVT, ETH Zurich.

David, N. (2009) Validation and verification in social simulation: Patterns and clarification of terminology, *Epistemological Aspects of Computer Simulation in the Social Sciences*, **5466**, 117–129.

Davies, F. M., M. M. H. Goode, L. A. Moutinho and E. Ogbonna (2001) Critical factors in consumer supermarket shopping behaviour: A neural network approach, *Journal of Consumer Behaviour*, **1** (1) 35–49.

de Palma, A., N. Picard and P. A. Waddell (2007) Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region, *Journal of Urban Economics*, **62** (2) 204–230.

Delleart, B. G. C., T. A. Arentze, M. Bierlaire, A. W. J. Borgers and H. J. P. Timmermans (1998) Investigating consumers' tendency to combine multiple shopping purposes and destinations, *Journal of Marketing Research*, **35** (2) 177–188.

Delleart, B. G. C., A. W. J. Borgers and H. J. P. Timmermans (1997) Development and test of stage-dependent conjoint choice experiments, *Journal of Retailing and Consumer Services*, **4** (1) 25–37.

DemoSCOPE und Planungsbüro Jud (2007) Rennweg mit und ohne Parkplätze: Erhebungen und Befragungen, *Final Report*, Stadt Zürich, DemoSCOPE und Planungsbüro Jud, Zurich.

Dial, R. B. (1971) A probabilistic multipath traffic assignment model which obviates path enumeration, *Transportation Research*, **5** (2) 83–111.

Dieussaert, K., K. Aerts, T. Steenberghen, S. Maerivoet and K. Spitaels (2009) SUSTAPARK: an agent-based model for simulating parking search, paper presented at the *AGILE International Conference on Geographic Information Science*, Hannover.

Dijst, M. and V. Vidakovic (1997) Individual action space in the city, in D. F. Ettema and H. J. P. Timmermans (eds.) *Activity-Based Approaches to Travel Analysis*, 117–134, Pergamon, Oxford.

Dobler, C. (2013) Travel behaviour modelling for scenarios with unpredictable events - methods and implementation, Ph.D. Thesis, ETH Zurich, Zurich.

Dobler, C., M. Balmer and K. W. Axhausen (2009) Simulation of information oriented knowledge models, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.

Dobler, C. and G. Lämmel (2012) Integration of a multi-modal simulation module into a framework for large-scale transport systems simulation, paper presented at the *6th International Conference on Pedestrian and Evacuation Dynamics*, Zurich, June 2012.

Doherty, S. T. and K. W. Axhausen (1998) The development of a unified modeling framework for the household activity-travel scheduling process, paper presented at the *4th NECTAR Conference*, Tel Aviv, April 1998.

Domencich, T. and D. McFadden (1975) *Urban Travel Demand: A Behavioral Analysis*, North-Holland, Amsterdam.

dos Santos, M. I. R. and A. M. O. Porta Nova (1999) The main issues in nonlinear simulation metamodel estimation, paper presented at the *Winter Simulation Conference*, Phoenix, December 1999.

Dougherty, M. S., H. R. Kirby and R. D. Boyle (1994) Using neural networks to recognise, predict and model traffic, in M. Bielli, G. Ambrosino and M. Boero (eds.) *Artificial Intelligence Applications to Traffic Engineering*, 233–250, VSP, Utrecht.

Dowling, R., J. H. Holland and A. Huang (2002) *California Department of Transportation: Guidelines for Applying Traffic Microsimulation Modeling Software*, Downling Associates, Inc., September 2002.

Downs, A. (1962) The law of peak-hour expressway congestion, *Traffic Quarterly*, **16** (3) 393–409.

Dubernet, T. and K. W. Axhausen (2013) A Framework to Represent Joint Decisions in a Multi-Agent Transport Simulation, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Dupuis, P. and A. Nagurney (1993) Dynamical systems and variational inequalities, *Annals of Operations Research*, **44** (1) 7–42.

DYNASMART (2013) DYNASMART, webpage, `http://www.its.uci.edu/~paramics/sim_models/dynasmart.html`.

East, R., P. Harris, W. Lomax, G. Wilson and K. Hammond (1998) Customer defection from supermarkets, *Advances in Consumer Research*, **25** (1) 507–512.

Eberhard, L. K. (2002) A 24-h household-level activity based travel demand model for the GTA, Master Thesis, University of Toronto, Toronto.

Eggenberger, M. (2001) Mobilitätsverhalten: Einkaufs- und Freizeitverkehr - Glattal, *Technical Report*, **80**, RZU - Regionalplanung Zürich und Umgebung, Zurich.

Ehreke, I. (2008) Analyse der Zielwahl im motorisierten Individualverkehr - eine Betrachtung der Auftretungshäufigkeiten routinisierter Fahrten, Master Thesis, Freie Universität Berlin, Berlin.

Eiben, A. E. and J. E. Smith (eds.) (2003) *Introduction to Evolutionary Computing*, Springer, Berlin.

Epstein, J. M. (2008) Why model?, *Journal of Artificial Societies and Social Simulation*, **11** (4) 12.

Erath, A. (2005) Zeitkosten im Einkaufsverkehr, Master Thesis, IVT, ETH Zurich, Zurich.

Erath, A., A. Chakirov, P. J. Fourie, S. A. Ordóñez Medina, M. Shah, M. A. B. van Eggermond and K. W. Axhausen (2012) A large-scale agent-based transport travel demand model for Singapore: The implementation of MATSim, *Working Paper*, Future Cities Laboratory, Singapore-ETH Centre (SEC), Singapore.

Erath, A., N. Frank, R. Lademann and K. W. Axhausen (2007) The impact of travel time savings on shopping location choice or how far do people go to shop cheaply?, paper presented at the *14th EIRASS Conference in Advances in Activity Based Approaches*, San Francisco, June 2007.

Eroglu, S. and G. D. Harrell (1986) Retail crowding: Theoretical and strategic implications, *Journal of Retailing*, **62** (4) 346–363.

Eroglu, S. and K. A. Machleit (1990) An empirical study of retail crowding: Antecedents and consequences, *Journal of Retailing*, **66** (2) 201–221.

Eroglu, S., K. A. Machleit and T. Feldman Barr (2005) Perceived retail crowding and shopping satisfaction: The role of shopping values, *Journal of Business Research*, **58** (8) 1146–1153.

Esser, J. and K. Nagel (2001) Iterative demand generation for transportation simulations, in D. A. Hensher (ed.) *Travel Behaviour Research: The Leading Edge*, 689–709, Pergamon, Oxford.

Ettema, D. F., A. W. J. Borgers and H. J. P. Timmermans (1995) Competing risk hazard model of activity choice, timing, sequencing, and duration, *Transportation Research Record*, **1493**, 101–109.

Ettema, D. F., A. W. J. Borgers and H. J. P. Timmermans (1996) SMASH (simulation model of activity scheduling heuristics): Some simulations, *Transportation Research Record*, **1551**, 88–94.

Ettema, D. F., A. W. J. Borgers and H. J. P. Timmermans (1997a) A simulation model of activity scheduling behaviour, *Transportation Research Record*, **1413**, 1–11.

Ettema, D. F., A. J. Daly, G. de Jong and E. Kroes (1997b) Towards an applied activity-based travel demand model, paper presented at the *8th International Conference on Travel Behaviour Research (IATBR)*, Austin, September 1997.

Ettema, D. F. and H. J. P. Timmermans (eds.) (1997) *Activity-Based Approaches to Travel Analysis*, Pergamon, Oxford.

Feeney, B. P. (1989) A review of the impact of parking policy measures on travel demand, *Transportation Planning and Technology*, **13** (4) 229–234.

Feil, M., M. Balmer and K. W. Axhausen (2009a) Enhancement and estimation of MATSim's utility function, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.

Feil, M., M. Balmer and K. W. Axhausen (2009b) Generating comprehensive all-day schedules: Expanding activity-based travel demand modelling, paper presented at the *European Transport Conference*, Leeuwenhorst, October 2009.

Fellendorf, M., T. Haupt, U. Heidl and W. Scherr (2000) VISEM - an activity chain based traffic demand model, *Technical Report*, PTV, Karlsruhe.

FHWA (2013) TRANSIMS Background, webpage, http://www.fhwa.dot.gov/planning/tmip/resources/transims/.

filialsuche.ch (2013) filialsuche.ch, http://www.filialsuche.ch/.

Finkel, R. A. and J. L. Bentley (1974) Quadtrees a data structure for retrieval on composite keys, *Acta Informatica*, **4** (1) 1–9.

Fisher, M. and M. Wooldridge (1997) On the formal specification and verification of multi-agent systems, *International Journal of Cooperative Information Systems*, **6** (1) 37–65.

Flötteröd, G. (2012) MATSim as a Monte Carlo engine, presentation, 1st MATSim Conceptual Meeting, Zurich, September.

Flötteröd, G., M. Bierlaire and K. Nagel (2008) Bayesian calibration of dynamic traffic simulations, *Technical Report*, **081028**, TRANSP-OR, EPF Lausanne, Lausanne.

Flötteröd, G. and K. Nagel (2006) Modeling and estimation of combined route and activity location choice, paper presented at the *Intelligent Transportation Systems Conference (ITSC)*, Toronto, September 2006.

Foscht, T. and B. Swoboda (2007) *Käuferverhalten*, 3. edn., Gabler Verlag, Wiesbaden.

Fosgerau, M. (2001) PETRA - an activity-based approach to travel demand analysis, in L. Lundqvist and L.-G. Mattsson (eds.) *National Transport Models: Recent Developments and Prospects*, Advances in Spatial Science, 134–146, Springer, Stockholm.

Fotheringham, A. S. (1983a) A new set of spatial interaction models: The theory of competing destinations, *Environment and Planning A*, **15** (1) 15–36.

Fotheringham, A. S. (1983b) Some theoretical aspects of destination choice and their relevance to production-constrained gravity models, *Environment and Planning A*, **15** (8) 1121–1132.

Fotheringham, A. S. (1985) Spatial competition and agglomeration in urban modelling, *Environment and Planning A*, **17** (2) 213–230.

Fotheringham, A. S. (1986) Modelling hierarchical destination choice, *Environment and Planning A*, **18** (3) 401–418.

Fotheringham, A. S. (1988) Consumer store choice and choice set definition, *Marketing Science*, **7** (3) 299–310.

Fotheringham, A. S., T. Nakaya, K. Yano, S. Openshaw and Y. Ishikawa (2001) Hierarchical destination choice and spatial interaction modelling: A simulation experiment, *Environment and Planning A*, **33** (5) 901–920.

Fox, E. J., S. Postrel and A. McLaughlin (2007) The impact of retail location on retailer revenues: An empirical investigation, *Working Paper*, Cox School of Business, Southern Methodist University, Dallas.

Frank, M. and P. Wolfe (1956) An algorithm for quadratic programming, *Naval Research Logistics Quarterly*, **3**, 95–110.

Frei, A. and K. W. Axhausen (2007) Size and structure of social network geographies, *Working Paper*, **444**, IVT, ETH Zurich, Zurich.

Frei, A., K. W. Axhausen and T. Ohnmacht (2009) Mobilities and social network geography: Size and spatial dispersion - the Zurich case study results, in T. Ohnmacht, H. Maksim and M. Bergmann (eds.) *Mobilities and Inequalities*, 99–120, Ashgate, Farnham.

Frejinger, E., M. Bierlaire and M. E. Ben-Akiva (2009) Sampling of alternatives for route choice modeling, *Transportation Research Part B: Methodological*, **43** (10) 984–994.

Fried, M., J. Havens and M. Thall (1977) Travel behavior - a synthesized theory, *Final Report*, National Cooperative Highway Research Program, Transportation Research Board, Boston College, Washington, D.C.

Friesz, T. L. (1996) Dynamic systems, variational inequalities, and control-theoretic models for predicting time-varying urban network flows, *Transportation Science*, **30** (1) 14–31.

Friesz, T. L. (2010) Dynamic user equilibrium, in F. S. Hillier (ed.) *Dynamic Optimization and Differential Games*, chap. 9, 411–456, Springer, New York.

Friesz, T. L., D. Bernstein, T. Smith and B. W. Wie (1993) A variational inequality formulation of the dynamic network user equilibrium problem, *Operations Research*, **41** (1) 80–91.

Friesz, T. L. and D. H. Bernstein (2000) Analytical dynamic traffic assignment models, in D. A. Hensher and K. J. Button (eds.) *Handbook of Transport Modelling*, chap. 11, 181–195, Pergamon, Amsterdam.

Friesz, T. L., D. H. Bernstein, N. J. Metha, R. L. Tobin and S. Ganjalizadeh (1994) Day-to-day dynamic network disequilibrium and idealized traveler information systems, *Operations Research*, **42** (6) 1120–1136.

Fröhlich, P., K. W. Axhausen, M. Vrtic, C. Weis and A. Erath (2012) SP-Befragung 2010 zum Verkehrsverhalten im Personenverkehr, *Research Report*, Swiss Federal Office for Spatial Development (ARE), IVT, ETH Zurich, Berne.

Gallo, M., L. D'Acierno and B. Montella (2011) A multilayer model to simulate cruising for parking in urban areas, *Transport Policy*, **18** (5) 735–744.

Gao, W., M. Balmer and E. J. Miller (2010) Comparison of MATSim and EMME/2 on Greater Toronto and Hamilton Area network, Canada, *Transportation Research Record*, **2197**, 118–128.

Gärling, T., K. Brännäs, J. Garvill, R. G. Golledge, S. Gopal, E. Holm and E. Lindberg (1989) Household activity scheduling, paper presented at the *Transport Policy, Management & Technology Towards 2001*, vol. IV, 231–248, Ventura.

Gärling, T., M.-P. Kwan and R. G. Golledge (1994) Computational-process modeling of household activity scheduling, *Transportation Research Part B: Methodological*, **28B** (5) 355–364.

Gautschi, D. A. (1981) Specification of patronage models for retail center choice, *Journal of Marketing Research*, **18** (2) 162–174.

Gawron, C. (1998) An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model, *International Journal of Modern Physics C*, **9** (3) 393–408.

Gehlke, C. E. and K. Biehl (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material, *Journal of the American Statistical Association*, **29** (185) 169–170.

Gerrard, B., B. Still and A. Jopson (2001) The impact of road pricing and workplace parking levies on the urban economy: results from a survey of business attitudes, *Environment and Planning A*, **33**, 1985–2002.

Gibb, J. and J. L. Bowman (2007) Convergence of an activity-based travel model system to equilibrium: Experimental designs and findings, paper presented at the *11th National Transportation Planning Applications Conference*, Daytona Beach, Florida, May 2007.

Gilbert, N. (2006) When does social simulation need cognitive models?, in D. Sun (ed.) *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, chap. 19, 428–432, Cambridge University Press, Cambridge.

Gilbert, N. and K. G. Troitzsch (2005) *Simulation for the Social Scientist*, 2. edn., Open University Press, Maidenhead.

Gillen, D. W. (1977) Estimation and specification of the effects of parking costs on urban transport mode choice, *Journal of Urban Economics*, **4** (2) 186–199.

Gillen, D. W. (1978) Parking policy, parking location decisions and the distribution of congestion, *Transportation*, **7** (1) 69–85.

Glazer, A. and E. Niskanen (1992) Parking fees and congestion, *Regional Science and Urban Economics*, **22** (1) 123–132.

Golias, J., G. Yannis and M. Harvatis (2002) Off-street parking choice sensitivity, *Transportation Planning and Technology*, **25** (4) 333–348.

Golle, P. and K. Partridge (2009) On the anonymity of home/work location pairs, paper presented at the *7th International Conference on Pervasive Computing*, 390–397, Nara, May 2009.

Golledge, R. G. (1970) Some equilibrium models of consumer behavior, *Economic Geography*, **46** (2) 417–424.

Golledge, R. G. and H. J. P. Timmermans (1990) Applications of behavioural research on spatial problems: Cognition, *Progress in Human Geography*, **14** (1) 57–99.

Goodwin, P. B. (1981) The usefulness of travel time budgets, *Transportation Research Part A*, **15A** (1) 97–106.

Google (2010a) Google Maps, webpage, `http://maps.google.com/`.

Google (2010b) Google Maps API Family, webpage, `http://code.google.com/apis/maps/index.html`.

Gopal, S., R. L. Klatzky and T. R. Smith (1989) Navigator: A psychologically based model of environmental learning through navigation, *Journal of Environmental Psychology*, **9** (4) 309–331.

Gottardi, G. and S. Bürgler (1999) Güterverkehrsmodell Kanton Zürich, *Strasse und Verkehr*, **2**.

Goulias, K. G. and R. Kitamura (1992) Travel demand forecasting with dynamic microsimulation, *Transportation Research Record*, **1357**, 8–17.

Gunn, H. (1981) Travel budgets—a review of evidence and modelling implications, *Transportation Research Part A*, **15A** (1) 7–23.

Hackney, J. K. (2009) Integration of social networks in a large-scale travel behavior microsimulation, Ph.D. Thesis, ETH Zurich, Zurich.

Hackney, J. K. and K. W. Axhausen (2006) An agent model of social network and travel behavior interdependence, paper presented at the *11th International Conference on Travel Behaviour Research (IATBR)*, Kyoto, August 2006.

Hackney, J. K., M. Bernard, S. Bindra and K. W. Axhausen (2007) Predicting road system speeds using spatial structure variables and network characteristics, *Journal of Geographical Systems*, **9** (4) 397–417.

Hägerstrand, T. (1970) What about people in regional science?, *Papers of the Regional Science Association*, **24** (1) 7–21.

Hale, D. (1997) How many NETSIM runs are enough?, *McTrans Newsletter*, **11** (3) 4–5.

Hale, T. S. and C. R. Moberg (2003) Location science research: A review, *Annals of Operations Research*, **123** (1–4) 21–35.

Han, Q., T. A. Arentze, H. J. P. Timmermans, D. Janssens and G. Wets (2009) A multi-agent modeling approach to simulate dynamic activity-travel patterns, in A. L. C. Bazzan and F. Klügl (eds.) *Multi-Agent Systems for Traffic and Transportation Engineering*, 36–56, Information Science Reference, Hershey.

Handy, S. L. and K. J. Clifton (2001) Local shopping as a strategy for reducing automobile travel, *Transportation*, **28** (4) 317–346.

Hannes, E., D. Janssens and G. Wets (2008) Destination choice in daily activity travel: The mental map's repertoire, paper presented at the *87th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2008.

Hanson, S. and J. Huff (1988a) Repetition and day-to-day variability in individual travel patterns: Implications for classification, in R. G. Golledge and H. J. P. Timmermans (eds.) *Behavioural Modelling in Geography and Planning*, 368–398, Croom Helm, London.

Hanson, S. and J. O. Huff (1982) Assessing day-to-day variability in complex travel patterns, *Transportation Research Record*, **891**, 18–24.

Hanson, S. and J. O. Huff (1988b) Systematic variability in repetitious travel, *Transportation*, **15** (1–2) 111–135.

Harker, P. T. and J.-S. Pang (1990) Finite-dimensional variational inequality and nonlinear complementary problems: A survey of theory, algorithms and applications, *Mathematical Programming*, **48** (1-3) 161–220.

Harrell, G. D., M. D. Hutt and J. C. Anderson (1980) Path analysis of buyer behavior under conditions of crowding, *Journal of Marketing Research*, **17** (1) 45–51.

Hautzinger, H. (ed.) (2003) *Freizeitmobilitätsforschung - Theoretische und methodische Ansätze*, Verlag MetaGIS Infosysteme, Mannheim.

Hawas, Y. E. and M. A. Hameed (2009) A multi-stage procedure for validating microscopic traffic simulation models, *Transportation Planning and Technology*, **32** (1) 71–91.

Hensher, D. A. and J. King (2001) Parking demand and responsiveness to supply, pricing and location in the Sidney central business district, *Transportation Research Part A: Policy and Practice*, **35** (3) 177–196.

Henson, K. M. and K. G. Goulias (2006) Preliminary assessment of activity analysis and modeling for homeland security applications, *Transportation Research Record*, **1942**, 23–30.

Hernández, T. and D. Bennison (2000) The art and sciencce of retail location decisions, *International Journal of Retail & Distribution Management*, **28** (8) 357–367.

Hernández, T., D. Bennison and S. Cornelius (1998) The organisational context of retail locational planning, *GeoJournal*, **45** (4) 299–308.

Hobeika, A. G. (2005) TRANSIMS fundamentals, *Technical Report*, Virginia Polytechnic University, Virginia, July 2005.

Hoh, B., M. Gruteser, H. Xiong and A. Alrabady (2010) Achieving guaranteed anonymity in GPS traces via uncertainty-aware path cloaking, *IEEE Transactions on Mobile Computing*, **9** (8) 1089–1107.

Holden, D. J. (1989) Wardrop's third principle. urban traffic congestion and traffic policy, *Journal of Transport Economics and Policy*, **23** (3) 239–262.

Hollander, Y. and R. Liu (2007) The principles of calibrating traffic microsimualtion models, *Transportation*, **35** (3) 347–362.

Hong, S.-K., J.-h. Kim, H. Jang and S. Lee (2006) The roles of categorization, affective image and constraints on destination choice: An application of the NMNL model, *Tourism Management*, **27** (5) 750–761.

Hoogendoorn, S. P. and P. H. L. Bovy (2001) State-of-the-art of vehicular traffic flow modelling, *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, **215** (4) 283–303.

Hope, M., Y. S. Park and H. Ley (2009) Convergence control and equilibration (part 7), presentation, TRANSIMS Training Course at Transportation Research and Analysis Computing Center (TRACC), Chicago, December.

Horni, A. (2007) Vergleich: Simulation - Zählungen (Wie gut ist unsere Simulation?), presentation, Treffpunkt Science City, ETH Zurich, Zurich, November.

Horni, A. and K. W. Axhausen (2012a) How to improve MATSim destination choice for discretionary activities?, paper presented at the *12th Swiss Transport Research Conference*, Ascona, May 2012.

Horni, A. and K. W. Axhausen (2012b) MATSim agent heterogeneity and week scenario, *Arbeitsberichte Verkehrs- und Raumplanung*, **836**, IVT, ETH Zurich, Zurich.

Horni, A., D. Charypar and K. W. Axhausen (2011a) Empirically approaching destination choice set formation, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Horni, A., D. Charypar and K. W. Axhausen (2011b) Variability in transport microsimulations investigated for MATSim: Preliminary results, paper presented at the *11th Swiss Transport Research Conference*, Ascona, May 2011.

Horni, A., D. Charypar and K. W. Axhausen (2011c) Variability in transport microsimulations investigated with the multi-agent transport simulation MATSim, *Working Paper*, **692**, IVT, ETH Zurich, Zurich.

Horni, A. and F. Ciari (2009) Modeling shopping customer and retailers with the activity-based Multi-agent Transport Simulation MATSim, presentation, Interdisciplinary Seminar "Modeling Complex Socio-Economic Systems and Crises", Zurich, April 2009.

Horni, A. and F. Ciari (2011) Microsimulating choices of different agent types: Shoppers and retailers, poster presented at Rational Choice Seminar, San Servolo (Venice), November 2011.

Horni, A., F. Ciari and K. W. Axhausen (2012a) Coupling customers' destination choice and retailers' location choice in MATSim, *Arbeitsberichte Verkehrs- und Raumplanung*, **808**, IVT, ETH Zurich, Zurich.

Horni, A. and D. Grether (2007) Counts, internal presentation, MATSim-T Workshop, IVT, ETH Zurich, Castasegna, October.

Horni, A. and L. Montini (2013a) A Glimpse at Emergence in Agent-Based Microsimulations, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Horni, A. and L. Montini (2013b) ApplauSim: Rhythmischer Applaus als emergentes Phänomen?, *Arbeitsberichte Verkehrs- und Raumplanung*, **853**, IVT, ETH Zurich, Zurich.

Horni, A., L. Montini, R. A. Waraich and K. W. Axhausen (2012b) An agent-based cellular automaton cruising-for-parking simulation, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Horni, A., K. Nagel and K. W. Axhausen (2011d) High-resolution destination choice in agent-based demand models, *Working Paper*, **682**, IVT, ETH Zurich, Zurich.

Horni, A., K. Nagel and K. W. Axhausen (2012c) High-resolution destination choice in agent-based demand models, paper presented at the *91st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2012.

Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (2009a) Location choice modeling for shopping and leisure activities with MATSim: Combining micro-simulation and time geography, *Transportation Research Record*, **2135**, 87–95.

Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (2009b) Location choice modeling for shopping and leisure activities with MATSim: Utility function extension and validation results, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.

Horni, A., B. J. Vitins and K. W. Axhausen (2011e) The Zurich scenario: A technical overview, *Working Paper*, **687**, IVT, ETH Zurich, Zurich.

Horowitz, A. J. and J. J. Louviere (1995) What is the role of consideration sets in choice modeling?, *International Journal of Research in Marketing*, **12** (1) 39–54.

Horowitz, J. L. (1984) The stability of stochastic equilibrium in a two-link transportation network, *Transportation Research Part B: Methodological*, **18** (1) 13–28.

Horowitz, J. L. (1985) Travel and location behavior: State of the art and research opportunities, *Transportation Research Part A: Policy and Practice*, **19A** (5/6) 441–453.

Howard, J. A. and J. N. Sheth (1969) *The Theory of Buyer Behavior*, John Wiley, New York.

Huang, A. and D. Levinson (2011) Why retailers cluster: An agent model of location choice of supply chains, *Environment and Planning B*, **38** (1) 82–94.

Hubbard, B. (1978) A review of selected factors conditioning consumer travel behavior, *Journal of Consumer Research*, **5** (1) 1–21.

Huff, D. L. (1960) A topographical model of consumer space preferences, *Papers in Regional Science*, **6** (1) 159–173.

Huff, D. L. (1963) A probabilistic analysis of shopping center trade areas, *Land Economics*, **39**, 81–90.

Huff, J. O. and S. Hanson (1986) Repetition and variability in urban travel, *Geographical Analysis*, **18** (2) 97–114.

Hui, M. K. and J. E. G. Bateson (1991) Perceived control and the effects of crowding and consumer choice on the service experience, *Journal of Consumer Research*, **18** (2) 174–184.

Huisman, O. and P. Forer (2005) The complexities of everyday life: Balancing practical and realistic approaches to modelling probable presence in space-time, paper presented at the *17th Annual Colloquium of the Spatial Information Research Centre*, Dunedin (New Zealand), November 2005.

Hunt, J. D. (2006) Levels of disaggregation and degrees of aggregate constraint in transportation system modeling, paper presented at the *Innovations in Travel Demand Modeling (ITM'06)*, Austin, May 2006.

Hutchinson, T. P. (1993) *Essentials of Statistical Methods in 41 Pages*, Rumsby Scientific Publishing, Adelaide.

Ibrahim, M. F. (2002) Disaggregating the travel components in shopping centre choice, *Journal of Property Investment & Finance*, **20** (3) 277–294.

Illenberger, J., M. Kowald, K. W. Axhausen and K. Nagel (2010) Insights into a spatially embedded social network from a large-scale snowball-sample, *Working Paper*, **10–10**, Transport Systems Planning and Transport Telematics (VSP), Technical University Berlin, Berlin.

Innes, J. D., M. C. Ircha and D. A. Badoe (1990) Factors affecting automobile shopping trip destinations, *Journal of Urban Planning and Development*, **116** (3) 126–136.

Institut für Mobilitätsforschung (ed.) (2000) *Freizeitverkehr*, Springer, Berlin.

Iragael, J. (2007) The role of travel time budgets - representation of a demand derived from activity participation, paper presented at the *11th World Conference on Transportation Research*, Berkeley, June 2007.

IVT and ETH Zurich (2013) ETH Travel Data Archive, webpage, `http://tda.ethz.ch`.

Janssens, D., G. Wets, H. J. P. Timmermans and T. A. Arentze (2007) Modeling short-term dynamics in activity-travel patterns: the Feathers model, paper presented at the *11th World Conference on Transportation Research*, Berkeley, June 2007.

Jin, W.-L. (2005) A dynamical system model of the traffic assignment problem, *Transportation Research Part B: Methodological*, **41** (1) 32–48.

Joh, C.-H., T. A. Arentze and H. J. P. Timmermans (2004) Activity-travel scheduling and rescheduling decision processes: Empirical estimation of aurora model, *Transportation Research Record*, **1898**, 10–18.

Jones, P. M. (1979) New approaches to understand travel behaviour: the human activity approach, in D. A. Hensher and P. R. Stopher (eds.) *Behavioural Travel Modelling*, 55–80, Croom Helm Ltd, Kent.

Jones, P. M. and M. I. Clarke (1988) The significance and measurement of variability in travel behaviour, *Transportation*, **15** (1) 65–87.

Jones, P. M., M. C. Dix, M. I. Clarke and I. G. Heggie (1983) *Understanding Travel Behaviour*, Gower, Aldershot.

Jonnalagadda, J., N. Freedman, W. A. Davidson and J. D. Hunt (2001) Development of microsimulation activity-based model for San Francisco: Destination and mode choice models, *Transportation Research Record*, **1777**, 25–35.

Joubert, J. W., P. J. Fourie and K. W. Axhausen (2010) Large-scale agent-based combined traffic simulation of private cars and commercial vehicles, *Transportation Research Record*, **2168**, 24–32.

Jovicic, G. (2001) Activity based travel demand modelling: A literature study, *Danmarks TransportForskning*, **Note 8**, 1–59.

Kahn, B. E. and A. W. Marshall (1953) Methods of reducing sample size in Monte Carlo computations, *Journal of the Operational Research Society*, **1** (5) 263–278.

Kahn, B. E. and D. C. Schmittlein (1989) Shopping trip behavior: An empirical investigation, *Marketing Letters*, **1** (1) 55–69.

Kaplan, S. and S. Bekhor (2011) Exploring en-route parking type and parking-search route choice: Decision making framework and survey design, paper presented at the *2nd International Choice Modelling Conference*, Leeds, July 2011.

Kaplan, S., S. Bekhor and Y. Shiftan (2009) A two-stage model jointly revealing the determinants of non-compensatory conjunctive choice set formation and compensatory choice, paper presented at the *88th*

Annual Meeting of the Transportation Research Board, Washington, D.C., January 2009.

Kaufman, D. E., R. L. Smith and K. E. Wunderlich (1998) User-equilibrium properties of fixed points in dynamic traffic assignment, *Transportation Research Part C: Emerging Technologies*, **6** (1–2) 1–16.

Kawasaki, T. and K. W. Axhausen (2009) Choice set generation from GPS data set for grocery shopping location choice modelling in canton Zurich: The comparison to Swiss Microcensus 2005, *Working Paper*, **595**, IVT, ETH Zurich, Zurich.

Kelly, T. and K. Nagel (1998) Relaxation criteria for iterated traffic simulations, *International Journal of Modern Physics C*, **9** (1) 113–132.

Kemperman, A., A. W. J. Borgers and H. J. P. Timmermans (2002) Incorporating variety seeking and seasonality in stated preference modeling of leisure trip destination choice, *Transportation Research Record*, **1807**, 67–76.

Kickhöfer, B. (2009) Die Methodik der ökonomischen Bewertung von Verkehrsmaßnahmen in Multiagentensimulationen, Master Thesis, Technical University Berlin, Berlin.

Kim, B.-D. and K. Park (1997) Studying patterns of consumer's grocery shopping trip, *Journal of Retailing*, **73** (4) 501–517.

Kim, S.-J., W. Kim and L. R. Rilett (2005) Calibration of microsimulation models using nonparametric statistical techniques, *Transportation Research Record*, **1935**, 111–119.

Kinderlehrer, D. and G. Stampacchia (1980) *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York.

Kitamura, R. (1984) Incorporating trip chaining into analysis of destination choice, *Transportation Research Part B: Methodological*, **18** (1) 67–81.

Kitamura, R. (1996) Applications of models of activity behavior for activity based demand forecasting, paper presented at the *Activity-Based Travel Forecasting Conference*, New Orleans, June 1996.

Kitamura, R., C. Chen and R. Narayanan (1998) Traveler destination choice behavior: Effects of time of day, activity duration, and home location, *Transportation Research Record*, **1645**, 76–81.

Kitamura, R. and S. Fujii (1998) Two computational process models of activity-travel behavior, in T. Gärling, T. Laitila and K. Westin (eds.) *Theoretical Foundations of Travel Choice Modeling*, 251–279, Pergamon, Oxford.

Kitamura, R., A. Kikuchi and S. Fujii (2005) An overview of PCATS/DEBNetS micro-simulation system: Its development, extension, and application to demand forecasting, in R. Kitamura and M. Kuwahara (eds.) *Simulation Approaches in Transportation Analysis: Recent Advances and Challenges*, 371–399, Springer, New York.

Kitamura, R., E. I. Pas, C. V. Lula, T. K. Lawton and P. E. Benson (1996) The sequenced activity mobility simulator (SAMS): An integrated approach to modeling transportation, land use and air quality, *Transportation*, **23** (3) 267–291.

Kitamura, R., T. Yamamoto, Y. O. Susilo and K. W. Axhausen (2006) How routine is a routine? an analysis of the day-to-day variability in prism vertex location, *Transportation Research Part A: Policy and Practice*, **40** (3) 259–279.

Kleijnen, J. P. C. (1995) Verification and validation of simulation models, *European Journal of Operational Research*, **82** (1) 145–162.

Knight, F. H. (1924) Some fallacies in the interpretation of social cost, *Quarterly Journal of Economics*, **38** (4) 582–606.

Knox, S. D. and T. J. Denison (2000) Store loyalty: Its impact on retail revenue. an empirical study of purchasing behaviour in the UK, *Journal of Retailing and Consumer Services*, **7** (1) 33–45.

Koppelman, F. S. and J. R. Hauser (1978) Destination choice behavior for non-grocery-shopping trips, *Transportation Research Record*, **673**, 157–165.

Kowald, M. and K. W. Axhausen (2012) Focusing on connected personal leisure networks: Selected results from a snowball sample, *Environment and Planning A*, **44** (5) 1085–1100.

Kreibich, V. (1979) Modeling car availability, modal split and trip distribution by Monte-Carlo simulation, *Transportation*, **2** (8) 153–166.

Kroeber-Riel, W. and P. Weinberg (2003) *Konsumentenverhalten*, 8. edn., Verlag Vahlen, Munich.

Krumme, C., M. Cebrian and A. Pentland (2010) Patterns of individual shopping behavior, *Working Paper*, **arXiv:1008.2556 [cs.CY]**, The Media Laboratory, Massachusetts Institute of Technology, Cambridge.

Kuhnau, J. and K. G. Goulias (2002) Centre SIM: First-generation model design, pragmatic implementation, and scenarios, in K. G. Goulias (ed.) *Transportation Systems Planning: Methods and Applications*, 16–1–16–14, CRC Press, New York.

Kuhnimhof, T. and C. Gringmuth (2009) Multiday multiagent model of travel behavior with activity scheduling, *Transportation Research Record*, **2134**, 178–185.

Kuipers, B. (1978) Modeling spatial knowledge, *Cognitive Science*, **2** (2) 129–153.

Kurth, D. L., S. Childress, E. E. Sabina and T. F. Rossi (2006) Proposed validation and sensitivity testing of Denver region activity-based models, paper presented at the *Innovations in Travel Demand Modeling (ITM'06)*, Austin, May 2006.

Kwan, M.-P. (1997) GISICAS: An activity-based travel decision support system using a GIS-interfaced computational-process model, in D. F. Ettema and H. J. P. Timmermans (eds.) *Activity-Based Approaches to Travel Analysis*, 263–282, Pergamon, Oxford.

Lademann, R. (2007) Zum Einfluss von Verkaufsfläche und Standort auf die Einkaufswahrscheinlichkeit, in M. Schuckel and W. Toporowski (eds.) *Theoretische Fundierung und praktische Relevanz der Handelsforschung*, 143–162, Springer, Heidelberg.

LaHowara & Commander Spock (2013) An Agent-Based Cellular Automaton Cruising-For-Parking Simulation published on sourceforge.net, software, `https://sourceforge.net/projects/lahowara/`.

Lämmel, G. (2011) Escaping the tsunami: Evacuation strategies for large urban areas. concepts and implementation of a multi-agent based approach, Ph.D. Thesis, Technical University Berlin, Berlin.

Landau, U., J. N. Prashker and B. Alpern (1982a) Evaluation of activity constrained choice sets to shopping destination choice modeling, *Transportation Research Part A: Policy and Practice*, **16** (3) 199–207.

Landau, U., J. N. Prashker and B. Alpern (1982b) Evaluation of activity constrained choice sets to shopping destination choice modeling, *Transportation Research Part A*, **16** (3) 199–207.

Landsverk, O. B., D. Hughes and O. B. Landsverk (2003) Shopper loyalty and store choice: Insights from a study of Norwegian supermarkets, *European Retail Digest*, **38** (9) 1–8.

Lawe, S., J. Lobb, A. W. Sadek, S. Huang and C. Xie (2009) TRANSIMS implementation in Chittenden County, Vermont: Development, calibration, and preliminary sensitivity analysis, *Transportation Research Record*, **2132**, 113–121.

LeBlanc, L. J. (1973) Mathematical programming algorithms for large-scale network equilibrium and network design problems, Ph.D. Thesis, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston.

LeBlanc, L. J., E. K. Morlok and W. P. Pierskalla (1975) An efficient approach to solving the road network equilibrium traffic assignment problem, *Transportation Research*, **9** (5) 309–318.

Lee, K.-S., A. G. Hobeika, H. B. Zhang and H. Jung (2010) Travelers' response to value pricing: Application of departure time choices to TRANSIMS, *Journal of Transportation Engineering*, **136** (9) 811–817.

Lefebvre, N. and M. Balmer (2007) Fast shortest path computation in time-dependent traffic networks, paper presented at the *7th Swiss Transport Research Conference*, Ascona, September 2007.

Lemp, J. D. and K. Kockelman (2012) Strategic sampling for large choice sets in estimation and application, *Transportation Research Part A: Policy and Practice*, **46** (3) 603–613.

Lemp, J. D., L. B. McWethy and K. Kockelman (2007) From aggregate methods to microsimulation: Assessing benefits of microscopic activity-based models of travel demand, *Transportation Research Record*, **1994**, 80–88.

Lenntorp, B. (1976) Paths in space-time environments, Ph.D. Thesis, The Royal University of Lund, Department of Geography.

Ley, H. (2008) Introduction to the methodologies and implementation of TRANSIMS, presentation, TRANSIMS Training Course at Transportation Research and Analysis Computing Center (TRACC), Chicago, April.

Lima & Associates (2006) Lincoln MPO travel demand model: Draft model documentation, *Technical Report*, Lincoln Metropolitan Planning Organization, Lincoln, Nebraska.

Lin, D.-Y., N. Eluru, S. T. Waller and C. R. Bhat (2008) Integration of activity-based modeling and dynamic traffic assignment, *Transportation Research Record*, **2076**, 52–61.

Liu, R., D. van Vliet and D. P. Watling (2006) Microsimulation models incorporating both demand and supply dynamcis, *Transportation Research Part A: Policy and Practice*, **40** (2) 125–150.

Löchl, M. (2008) Standortplanung im Detail-/Einzelhandel: Auswertung von Interviews mit Unternehmen in Deutschland und der Schweiz, *Working Paper*, **492**, IVT, ETH Zurich, Zurich.

Lotzmann, U. (2009) TRASS: A multi-purpose agent-based simulation framework for complex traffic simulation applications, in A. L. C. Bazzan and F. Klügl (eds.) *Multi-Agent Systems for Traffic and Transportation Engineering*, 79–107, Information Science Reference, Hershey.

Lu, M. (2002) Mode Choice Methodology in TRANSIMS, Master Thesis, Virginia Polytechnic Institute and State University.

Ma, J. (1997) An activity-based and micro-simulated travel forecasting system: A pragmatic synthetic scheduling approach, Ph.D. Thesis, Department of Civil and Environmental Engineering, Pennsylvania State University, University Park.

Mahmassani, H. S. (1988) Some comments on activity-based approaches to the analysis and prediction of travel behavior, *Transportation*, **15**, 35–40.

Mahmassani, H. S. (1989) Dynamic models of commuter behavior: Experimental investigation and application to the analysis of planned traffic disruptions, presentation, World Conference on Travel Behaviour, Kyoto, Japan.

Mahmassani, H. S. and G.-L. Chang (1986) Experiments with departure time choice dynamics of urban commuters, *Transportation Research Part B: Methodological*, **20** (4) 297–320.

Mahmassani, H. S., T. Hu and R. Jayakrishnan (1995) Dynamic traffic assignment and simulation for advanced network informatics (DYNASMART), in N. H. Gartner and G. Improta (eds.) *Urban traffic networks: dynamic flow modeling and control*, Springer, Berlin.

Maier, G. and F. Toedtling (2006) *Regional- und Stadtökonomik 1: Standorttheorie und Raumstruktur*, Springer, Vienna, New York.

Maley, D. W. and R. R. Weinberger (2011) Food shopping in the urban environment: Parking supply, destination choice, and mode choice, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Manrai, A. K. and R. L. Andrews (1998) Two-stage discrete choice models for scanner panel data: An assessment of process and assumptions, *European Journal of Operational Research*, **111** (2) 193–215.

Manski, C. F. (1977) The structure of random utility models, *Theory and Decision*, **8** (3) 229–254.

Marchal, F. and K. Nagel (2005a) Computation of location choice of secondary activities in transportation models with cooperative agents, in F. Klügl, A. L. C. Bazzan and S. Ossowski (eds.) *Applications of Agent Technology in Traffic and Trasportation*, 153–164, Birkhäuser, Basel.

Marchal, F. and K. Nagel (2005b) Modeling location choice of secondary activities with a social network of cooperative agents, *Transportation Research Record*, **1935**, 141–146.

Marnay, C. and T. Strauss (1991) Effectiveness of antithetic sampling and stratified sampling in Monte Carlo chronological production cost modeling, *IEEE Transactions on Power Systems*, **6** (2) 669–675.

Marsden, G. (2006) The evidence base for parking policies - a review, *Transport Policy*, **13** (6) 447–457.

Martínez, F., F. Aguila and R. Hurtubia (2009) The constrained multinomial logit: A semi-compensatory choice model, *Transportation Research Part B*, **43** (3) 365–377.

MATSim (2013) Multi-Agent Transportation Simulation, webpage, `http://www.matsim.org`.

MATSim-T (2013) MATSim scenarios, webpage, `http://www.matsim.org/scenarios`.

Matsoukis, E. C. (1986) Road traffic assignment - a review, part I: Non-equilibrium methods, *Transportation Planning and Technology*, **11** (1) 69–79.

Matsoukis, E. C. and P. C. Michalopoulos (1986) Road traffic assignment - a review, part II: Equilibrium methods, *Transportation Planning and Technology*, **11** (2) 117–135.

Mayer, T. (2012) Ziliak and McCloskey's criticisms of significance tests: An assessment, *Econ Journal Watch*, **9** (3) 256–297.

McCarthy, P. S. (1980) A study of the importance of generalized attributes in shopping choice behaviour, *Environment and Planning A*, **12** (11) 1269–1286.

McFadden, D. (1978) Modeling the choice of residential location, in A. Karlqvist (ed.) *Spatial Interaction Theory and Residential Location*, 75–96, North-Holland, Amsterdam.

McNally, M. G. (2000) The four step model, in D. A. Hensher and K. J. Button (eds.) *Handbook of Transport Modelling*, chap. 3, 35–52, Pergamon, Amsterdam.

McNally, M. G. and A. Kulkarni (2001) A microsimulation of daily activity patterns, paper presented at the *80th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2001.

McNally, M. G. and C. R. Rindt (2008) The activity-based approach, *Working Paper*, **UCI-ITS-AS-WP-07-1**, Institute of Transportation Studies, University of California, Irvine.

McTrans (2013) Traffic Software Integrated System - Corridor Simulation, webpage, `http://mctrans.ce.ufl.edu/featured/tsis/`.

Meister, K. (2008) Erstellung von MATSim Facilities für das Schweiz-Szenario, *Working Paper*, **541**, IVT, ETH Zurich.

Meister, K. (2011) Contribution to agent-based demand optimization in a multi-agent transport simulation, Ph.D. Thesis, ETH Zurich, Zurich.

Meister, K., M. Balmer, F. Ciari, A. Horni, M. Rieser, R. A. Waraich and K. W. Axhausen (2010) Large-scale agent-based travel demand optimization applied to Switzerland, including mode choice, paper presented at the *12th World Conference on Transportation Research*, Lisbon, July 2010.

Meister, K., M. Frick and K. W. Axhausen (2005) A GA-based household scheduler, *Transportation*, **32** (5) 473–494.

Messinger, P. R. and C. Narasimhan (1997) A model of retail formats based on consumers' economizing on shopping time, *Marketing Science*, **16** (1) 1–23.

Milam, R. T. and F. Chao (2001) Recommended guidelines for the calibration and validation of traffic simulation models, paper presented at the *8th TRB Conference on the Application of Transportation Planning Methods*, Chorpus Christi, Texas, April 2001.

Miller, E. J. (1996) Microsimulation and activity-based forecasting, paper presented at the *Activity-Based Travel Forecasting Conference*, New Orleans, June 1996.

Miller, E. J. and M. E. O'Kelly (1983) Estimating shopping destination choice models from travel diary data, *The Professional Geographer*, **35** (4) 440–449.

Miller, E. J. and M. J. Roorda (2003) A prototype model of 24-hour household activity scheduling for the Toronto area, *Transportation Research Record*, **1831**, 114–121.

Miller, G. K. and C. T. Everett (1982) Raising commuter parking prices - An empirical study, *Transportation*, **11** (2) 105–129.

Mitchell, R. and C. Rapkin (1954) *Urban Traffic: A Function of Land Use*, Columbia University Press, New York.

mobiTopp (2013) mobiTopp Simulation Program Homepage, webpage, `http://www.mobitopp.net/`.

Mokhtarian, P. L. and C. Chen (2004) TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets, *Transportation Research Part A: Policy and Practice*, **38** (9-10) 643–675.

Mondschein, A., E. Blumenberg and B. D. Taylor (2008) Accessibility and cognition: Effect of transportation mode on spatial knowledge, paper presented at the *87th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2008.

Montini, L., A. Horni, N. Rieser-Schüssler and K. W. Axhausen (2012) Searching for parking in GPS data, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013) Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Nagel, K. (2012) Transport simulation as a complex adaptive system, presentation, LATSIS: 1st European Symposium on Quantitative Methods in Transportation Systems, Lausanne, September.

Nagel, K. and K. W. Axhausen (2001) Workshop report: Microsimulation, *Working Paper*, **49**, IVT, ETH Zurich, Zurich.

Nagel, K. and C. L. Barrett (1997) Using microsimulation feedback for trip adaptation for realistic traffic in Dallas, *International Journal of Modern Physics C*, **8** (3) 505–525.

Nagel, K. and G. Flötteröd (2009) Agent-based traffic assignment: Going from trips to behavioral travelers, paper presented at the *12th International Conference on Travel Behaviour Research (IATBR)*, Jaipur, December 2009.

Nagel, K. and M. Rickert (2001) Parallel implementation of the TRANSIMS micro-simulation, *Parallel Computing*, **58** (2) 1611–1639.

Nagel, K. and M. Schreckenberg (1992) A cellular automaton for freeway traffic, *Journal de Physique I*, **2** (1992) 2221–2229.

Nagurney, A. (1993) *Network Economics: A Variational Inequality Approach*, Kluwer, Dordrecht.

Nagurney, A. (2001a) Variational inequalities: Geometric interpretation, existence and uniqueness, in C. A. Floudas and P. M. Pardalos (eds.) *Encyclopedia of Optimization*, 2676–2679, Springer, New York.

Nagurney, A. (2001b) Variational inequalities: Projected dynamical system PDS, in C. A. Floudas and P. M. Pardalos (eds.) *Encyclopedia of Optimization*, 4002–4007, Springer, New York.

Nagurney, A. and D. Zhang (1996) *Projected Dynamical Systems and Variational Inequalities With Applications*, Kluwer, Boston.

Narayana, C. L. and R. J. Markin (1975) Consumer behavior and product performance: An alternative conceptualization, *Journal of Marketing*, **39** (4) 1–6.

Nash, J. (1950) Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United States of America*, **36** (1) 48–49.

Nash, J. (1951) Non-cooperative games, *The Annals of Mathematics*, **54** (2) 286–295.

NAVTEQ (2011) NAVTEQ Maps and Traffic, webpage, `http://www.navteq.com`.

Nerella, S. and C. R. Bhat (2004) Numerical analysis of effect of sampling of alternatives in discrete choice models, *Transportation Research Record*, **1894**, 11–19.

Newman, M. E. J. and G. T. Barkema (1999) *Monte Carlo Methods in Statistical Physics*, Oxford University Press, Oxford.

Nguyen, S. (1974) An algorithm for the traffic assignment problem, *Transportation Science*, **8** (3) 203–216.

Nicolai, T. and K. Nagel (2011) Investigating accessibility indicators for feedback from a travel to a land use model, paper presented at the *51st Congress of the European Regional Science Association*, Barcelona, September 2011.

Nicolai, T., L. Wang, K. Nagel and P. A. Waddell (2011) Coupling an urban simulation model with a travel model - a first sensitivity test, *Working Paper*, **11-07**, Transport Systems Planning and Transport Telematics (VSP), Technical University Berlin, Berlin.

Nijkamp, P. (2009) Mobile telephone data and traffic management: Exploratory research, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.

Noor, M. A., K. I. Noor and T. M. Rassias (1993) Some aspects of variational inequalities, *Journal of Computational and Applied Mathematics*, **47** (3) 285–312.

Odland, J. (1981) A household production approach to destination choice, *Economic Geography*, **57** (3) 257–269.

O'Kelly, M. E. (1983) Multipurpose shopping trips and the size of retail facilities, *Annals of the Association of American Geographers*, **73** (2) 231–239.

Oliveira, G. C., M. V. F. Pereira and S. H. F. Cunha (1989) A technique for reducing computational effort in Monte-Carlo based composite evaluation, *IEEE Transactions on Power Systems*, **4** (4) 1309–1315.

OpenAMOS (2011) OpenAMOS modeling framework, webpage, `http://simtravel.wikispaces.asu.edu/OpenAMOS+Modeling+Framework`.

Openshaw, S. (1984) *The Modifiable Areal Unit Problem*, Geo Books, Norwich.

Oppewal, H. and B. Hoyoake (2004) Bundling and retail agglomeration effects on shopping behavior, *Journal of Retailing and Consumer Services*, **11** (2) 61–74.

Oppewal, H., H. J. P. Timmermans and J. J. Louviere (1997) Modelling the effects of shopping centre size and store variety on consumer choice behavior, *Environment and Planning A*, **29**, 1073–1090.

Orcutt, G. H. (1957) A new type of socio-economic system, *Review of Economics and Statistics*, **39** (2) 116–123.

Ordóñez Medina, S. A., A. Erath and K. W. Axhausen (2012) Simulating urban transport for a week time horizon, paper presented at the *3rd Workshop on Time Use Observatory*, Santiago de Chile, March 2012.

Orgel, D. (1997) Consuming issues: Responses to a SN consumer survey show preferences about supermarket shopping and reveal merchandising oportunities, *Supermarket News*, **47** (34) 1–4.

Ortúzar, J. d. D. and L. G. Willumsen (2001) *Modelling Transport*, 3rd edn., John Wiley & Sons, Chichester.

Pagliara, F. and H. J. P. Timmermans (2009) Choice set generation in spatial contexts: A review, *Transportation Letters*, **1** (1) 181–196.

Pas, E. I. (1988) Weekly travel-activity behavior, *Transportation*, **15** (1) 89–109.

Pas, E. I. and F. S. Koppelman (1987) An examination of the determinants of day-to-day variability in individuals' urban travel behavior, *Transportation*, **14** (2) 3–20.

Patriksson, M. (1994) *The Traffic Assignment Problem, Models and Methods*, VSP, Utrecht.

PCATS (2011) Activity-travel simulator PCATS, webpage, `http://trans.kuciv.kyoto-u.ac.jp/tba/kikuchi/PCATS_DEBNetS/pcats.html`.

PCATS (2013) Activity-travel simulator PCATS, webpage, `http://trans.kuciv.kyoto-u.ac.jp/tba/kikuchi/PCATS_DEBNetS/pcats.html`.

Peeta, S. and A. K. Ziliaskopoulos (2001) Foundations of dynamic traffic assignment: The past, the present and the future, *Networks and Spatial Economics*, **1**, 233–265.

Pellegrini, P. A., A. S. Fotheringham and G. Lin (2005) An empirical evaluation of parameter sensitivity to choice set definition in shopping destination choice models, *Papers in Regional Science*, **76** (2) 257–284.

Pendyala, R. M. and C. R. Bhat (2006) Validation and assessment of activity-based travel demand modeling systems, paper presented at the *Innovations in Travel Demand Modeling (ITM'06)*, Austin, May 2006.

Pendyala, R. M., R. Kitamura, A. Chen and E. I. Pas (1997) An activity-based microsimulation analysis of transport control measures, *Transport Policy*, **4** (3) 183–192.

Pendyala, R. M., R. Kitamura and A. Kikuchi (2004) FAMOS: The Florida activity mobility simulator, paper presented at the *Conference on Progress in Activity-Based Analysis*, Maastricht, May 2004.

Pendyala, R. M., R. Kitamura, A. Kikuchi, T. Yamamoto and S. Fujii (2005) Florida activity mobility simulator: Overview and preliminary validation results, *Transportation Research Record*, **1921**, 123–130.

Pendyala, R. M., R. Kitamura, P. D. Reddy and Q. C. Chen (1995) Activity-based modeling system for travel demand forecasting, *Final Report*, Metropolitan Washington Council of Governments and Department for Transport, RDC Inc., San Francisco, September 1995.

Perrone, M. P. (1993) Improving regression estimation: Averaging methods for variance reduction with extensions to general convex measure optimization, Ph.D. Thesis, Brown University, Providence, Rhode Islands.

Petty, M. D. (2010) Verification, validation, and accreditation, in J. A. Sokolowski and C. M. Banks (eds.) *Verification, Validation, and Accreditation, in Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*, chap. 10, 325–372, John Wiley & Sons, Hoboken.

Pigou, A. C. (1920) *The Economics of Welfare*, Macmillan and Co., London.

Pinjari, A. R., N. Eluru, R. B. Copperman, I. N. Sener, J. Y. Guo, S. Srinivasan and C. R. Bhat (2006) Activity-based travel-demand analysis for metropolitan areas in Texas: CEMDAP models, framework, software architecture and application results, *Research Report*, **4080–8**, Texas Department of Transportation, Department of Civil, Architectural and Environmental Engineering, University of Texas Austin, Austin, October 2006.

Planungsbüro Jud (1990) Parksuchverkehr Zürich, *Research Report*, Stadtplanungsamt der Stadt Zürich, Planungsbüro Jud.

Planungsbüro Jud (2010) Belegung und Verkehrsaufkommen von Parkfeldern in der Stadt Zürich, *Research Report*, Tiefbauamt der Stadt Zürich, Planungsbüro Jud.

Polak, J. W. and K. W. Axhausen (1990) Parking search behaviour: A review of current research and future prospects, *Working Paper*, **540**, Transport Studies Unit, University of Oxford, Oxford.

Pons, F., M. Laroche and M. Mourali (2006) Consumer reactions to crowded retail settings: Cross-cultural differences between North America and the Middle East, *Psychology & Marketing*, **23** (7) 555–572.

Popkowski Leszczyc, P. T., A. Sinha and A. Sahgal (2004) The effect of multi-purpose shopping on pricing and location strategy for grocery stores, *Journal of Retailing*, **80** (2) 85–99.

Popkowski Leszczyc, P. T., A. Sinha and H. J. P. Timmermans (2000) Consumer store choice dynamics: An analysis of the competitive market structure for grocery stores, *Journal of Retailing*, **76** (3) 323–345.

Powell, W. B. and Y. Sheffi (1982) The convergence of equilibrium algorithms with predetermined step sizes, *Transportation Science*, **16** (1) 45–55.

Pozsgay, M. A. and C. R. Bhat (2001) Destination choice modeling for home-based recreational trips: Analysis and implications for land use, transportation, and air quality planning, *Transportation Research Record*, **1777**, 47–54.

Prayag, G. (2009) Tourists' evaluations of destination image, satisfaction, and future behavioral intentions—the case of Mauritius, *Journal of Travel & Tourism Marketing*, **26** (8) 836–853.

Rajagopal, K. R. (2006) Leisure shopping behavior and recreational retailing: A symbiotic analysis of marketplace strategy and consumer response, *Working Paper*, **06/2006**, Graduate School of Administration and Management (EGADE), Monterrey Institute of Technology and Higher Education (ITESM), Mexico City.

Ramadurai, G. and S. Ukkusuri (2008) Dynamic user equilibrium model for combined activity-travel choices using activity-travel supernetwork representation, *Working Paper*, Rensselaer Polytechnic Institute, Troy, NY, August 2008.

Ramadurai, G. and S. Ukkusuri (2010) Dynamic user equilibrium model for combined activity-travel choices using activity-travel supernetwork representation, *Networks and Spatial Economics*, **10** (2) 273–292.

Raney, B., A. Völlmy, N. Cetin, M. Vrtic and K. Nagel (2002) Towards a microscopic traffic simulation of all of Switzerland, *Lecture Notes in Computer Science*, **2329**, 371–380.

Rao, A. S. and M. P. Georgeff (1995) BDI agents: From theory to practice, paper presented at the *1st International Conference on Multi-Agent Systems*, San Francisco, June 1995.

Rauh, J., T. A. Schenk and P. Ulrich (2007) Einzelhandel und Verkehr - Ergebnisse einer Multiagentensimulation von Konsumentenentscheidungen, *Zeitschrift für Verkehrswissenschaft*, **78** (3) 176–191.

Recker, W. W. (1995) The household activity pattern problem: General formulation and solution, *Transportation Research Part B: Methodological*, **29** (1) 61–77.

Recker, W. W. and L. P. Kostyniuk (1978) Factors influencing destination choice for the urban grocery shopping trip, *Transportation*, **7** (1) 19–33.

Recker, W. W., M. G. McNally and G. S. Root (1986a) A model of complex travel behavior: Part I - theoretical development, *Transportation Research Part A: Policy and Practice*, **20** (4) 307–318.

Recker, W. W., M. G. McNally and G. S. Root (1986b) A model of complex travel behavior: Part II - an operational model, *Transportation Research Part A: Policy and Practice*, **20** (4) 319–330.

Recker, W. W. and H. J. Schuler (1981) Destination choice and processing spatial information: Some empirical test with alternative constructs, *Economic Geography*, **57** (4) 373–383.

Reilly, W. J. (1931) *The Law of Retail Gravitation*, Knickerbocker Press, New York.

Rhee, H. and D. R. Bell (2002) The inter-store mobility of supermarket shoppers, *Journal of Retailing*, **78** (4) 225–237.

Rieser, M. (2010) Adding transit to an agent-based transportation simulation, Ph.D. Thesis, Technical University Berlin, Berlin.

Rieser-Schüssler, N. (2012) Capitalising modern data sources for observing and modelling transport behaviour, *Transportation Letters*, **4** (2) 115–128.

Rieser-Schüssler, N. and K. W. Axhausen (2012) Development of psychometric scales to evaluate the attitude towards risk, environmentalism and variety seeking of public transport users, *Transportation Research Record*.

Rindsfüser, G. (2005) Simulation des Aktivitätenplanungsprozesses mit Methoden der Künstlichen Intelligenz, *Straßenverkehrstechnik*, **49** (3) 145–153.

Robbins, H. and S. Monro (1951) A stochastic approximation method, *Annals of Mathematical Statistics*, **22** (3) 400–407.

Robinson, R. V. F. and R. W. Vickerman (1976) The demand for shopping travel: A theoretical and empirical study, *Applied Economics*, **8** (4) 267–281.

Rogers, D. S. (2007) Retail location analysis in practice, *Research Review*, **14** (2) 73–78.

Roorda, M. J., E. J. Miller and K. M. N. Habib (2008) Validation of TASHA: A 24-h activity scheduling microsimulation model, *Transportation Research Part A: Policy and Practice*, **42** (2) 360–375.

Rossetti, R. J. F., R. H. Bordini, A. L. C. Bazzan, S. Bampi, L. Liu and D. van Vliet (2002a) Using BDI agents to improve driver modelling in a commuter scenario, *Transportation Research Part C: Emerging Technologies*, **10**, 373–398.

Rossetti, R. J. F. and L. Liu (2005) Activity-based analysis of travel demand using cognitive agents, in H. J. P. Timmermans (ed.) *Progress in Activity-Based Analysis*, 139–160, Elsevier, Oxford.

Rossetti, R. J. F., R. Liu, H. B. B. Cybis and S. Bampi (2002b) A multi-agent demand model, paper presented at the *13th Mini-Euro Conference and 9th Meeting of the Euro Working Group Transportation*, Bari, June 2002.

Rutter, C. M., D. L. Miglioretti and J. E. Savarino (2009) Bayesian calibration of microsimulation models, *Journal of the American Statistical Association*, **104** (488) 1338–1350.

Sadek, A. W. (2007) Artificial intelligence applications in transportation, in A. W. Sadek (ed.) *Artificial Intelligence in Transportation: Information for Application*, no. E-C113 in Transportation Research Circular, 1–6, Transportation Research Board, Washington, D.C.

Sadek, A. W., T. Springer and B. L. Smith (2003) Toward more effective transportation applications of computational intelligence paradigms, *Transportation Research Record*, **1836**, 57–63.

Saldo (2002) Detailhandel - Beim Wocheneinkauf fast 50 Franken gespart, May 2002, `http://www.saldo.ch/themen/beitrag/1014712/Detailhandel_-_Beim_Wocheneinkauf_fast_50_Franken_gespart`.

Salomon, I. and M. E. Ben-Akiva (1983) The use of the life-style concept in travel demand models, *Environment and Planning A*, **15** (5) 623–683.

Salvini, P. A. and E. J. Miller (2005) ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems, *Networks and Spatial Economics*, **5** (2) 217–234.

Sands, S., H. Oppewal and M. Beverland (2009) The effects of in-store themed events on consumer store choice decisions, *Journal of Retailing and Consumer Services*, **16** (5) 386–395.

Sanford Bernhardt, K. L. (2007) Agent-based modeling in transportation, in A. W. Sadek (ed.) *Artificial Intelligence in Transportation: Information for Application*, no. E-C113 in Transportation Research Circular, 72–80, Transportation Research Board, Washington, D.C.

Sbayti, H. and D. Roden (2010) Best practices in the use of micro simulation models, *Research Report*, American Association of State Highway and Transportation Officials, AECOM, Arlington.

Scherer, R. and C. Derungs (2008) Standortwahl von Unternehmen: Ein Entscheidungsprozess zwischen Rationalität und Emotionalität, *Working Paper*, Institut for Public Services and Tourism (IDT), University of St.Gallen (HSG), St. Gallen.

Schiffer, R. G. and T. F. Rossi (2009) New calibration and validation standards for travel demand modeling, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Schirmer, P. M., C. Zöllig, K. Müller, B. R. Bodenmann and K. W. Axhausen (2011) The Zurich case study of UrbanSim, paper presented at the *51st Congress of the European Regional Science Association*, Barcelona, September 2011.

Schlich, R. (2001) Analysing intrapersonal variability of travel behaviour using the Sequence Alignment Method, paper presented at the *European Transport Conference*, Cambridge.

Schlich, R. and K. W. Axhausen (2003a) Habitual travel behaviour: Evidence from a six-week travel diary, *Transportation*, **30** (1) 13–36.

Schlich, R. and K. W. Axhausen (2003b) Wohnumfeld und Freizeitverkehr - eine Untersuchung zur Fluchttheorie, *Working Paper*, **155**, IVT, ETH Zurich, Zurich.

Schlich, R., S. Schönfelder, S. Hanson and K. W. Axhausen (2002) Leisure travel in a historical perspective - changes in the structures of time and space use, *Working Paper*, **107**, IVT, ETH Zurich, Zurich.

Schlich, R., S. Schönfelder, S. Hanson and K. W. Axhausen (2004) Structures of leisure travel: Temporal and spatial variability, *Transport Reviews*, **24** (2) 219–237.

Schlich, R., A. Simma and K. W. Axhausen (2003a) Determinanten des Freizeitverkehrs - Modellierung und empirische Befunde, *Working Paper*, **190**, IVT, ETH Zurich, Zurich.

Schlich, R., A. Simma and K. W. Axhausen (2003b) Zielwahl im Freizeitverkehr, *Working Paper*, **181**, IVT, ETH Zurich, Zurich.

Schmidt, C. G. (1980) Location decision-making within a retail corporation, *The Journal of Regional Analysis & Policy*, **13**, 60–71.

Schnittger, S. and D. Zumkeller (2004) Longitudinal microsimulation as a tool to merge transport planning and traffic engineering models - the MobiTopp model, paper presented at the *European Transport Conference*, Strasbourg, October 2004.

Schüssler, N. (2006) Initial ideas on accounting for similarities between alternatives in route, mode and destination choice, presentation, 2nd Workshop on Discrete Choice Models, EPF Lausanne, Lausanne, September 2006.

Schüssler, N. (2010) Accounting for similarities between alternatives in discrete choice models based on high-resolution observations of transport behaviour, Ph.D. Thesis, ETH Zurich, Zurich.

Schüssler, N. and K. W. Axhausen (2007) Recent developments regarding similarities in transport modelling, paper presented at the *7th Swiss Transport Research Conference*, Ascona, September 2007.

Schweizerischer Verband der Lebensmittel-Detaillisten (2013) veledes.ch, `http://www.veledes.ch/documents/Februar_2012_deutsch.pdf`.

Scott, D. M. (2006) Constrained destination choice set generation: A comparison of GIS-based approaches, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.

Scott, D. M. and S. He (2012) Modeling constrained destination choice for shopping: a GIS-based, time-geographic approach, *Journal of Transport Geography*, **23**, 60–71.

Scrogin, D., R. Hofler, K. Boyle and J. W. Milon (2004) An efficiency approach to choice set generation, *Working Paper*, Department of Economics, University of Central Florida.

Sheffi, Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice Hall, Englewood Cliffs.

Sheffi, Y. and W. B. Powell (1981) A comparison of stochastic and deterministic traffic assignment over congested networks, *Transportation Research Part B: Methodological*, **15** (1) 53–64.

Shim, S. and M. A. Eastlick (1998) The hierarchical influence of personal values on mall shopping attitude and behavior, *Journal of Retailing*, **74** (1) 139–160.

Shocker, A. D., M. E. Ben-Akiva, B. Boccara and P. Nedungadi (1991) Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions, *Marketing Letters*, **2** (3) 181–197.

Shoup, D. (2005) *The High Cost of Free Parking*, Planners Press, Chicago.

Simma, A., P. Cattaneo, M. Baumeler and K. W. Axhausen (2004) Factors influencing the individual shopping behaviour: The case of Switzerland, *Working Paper*, **247**, IVT, ETH Zurich, Zurich.

Simma, A., R. Schlich and K. W. Axhausen (2002) Destination choice modelling for different leisure activities, *Working Paper*, **99**, IVT, ETH Zurich, Zurich.

SimTRAVEL (2013) TRANSIMS Application and Deployment, webpage, `http://simtravel.wikispaces.asu.edu/TRANSIMS+Application+and+Deployment`.

Sivakumar, A. and C. R. Bhat (2007) Comprehensive, unified framework for analyzing spatial location choice, *Transportation Research Record*, **2003**, 103–111.

Smith, M. C., A. W. Sadek and S. Huang (2008) Large-scale microscopic simulation: Toward an increased resolution of transportation models, *Journal of Transportation Engineering*, **134** (7) 273–281.

Smith, M. J. (1979) The existence, uniqueness and stability of traffic equilibria, *Transportation Research Part B: Methodological*, **13** (4) 295–304.

Smith, M. J. (1983) The existence and calculation of traffic equilibria, *Transportation Research Part B: Methodological*, **17** (4) 291–303.

Smith, M. J. (1993) Existence and calculation of dynamic user equilibria on congested capacity-constrained road networks, *Transportation Research Part B: Methodological*, **27** (1) 49–63.

Smock, R. B. (1962) An iterative assignment approach to capacity restraint on arterial networks, *Highway Research Board Bulletin*, **347**, 6–66.

Sobel, K. L. (1980) Travel demand forecasting by using the nested multinomial logit model, *Transportation Research Record*, **775**, 48–55.

Solomon, M. R. (2009) *Consumer Behavior*, Pearson Prentice Hall, Upper Saddle River.

SourceForge (2013) Open source software development web site, webpage, `http://sourceforge.net`.

Southworth, F. (1981) Calibration of multinomial logit models of mode and destination choice, *Transportation Research Part A*, **15** (4) 315–325.

Sparmann, U. and W. Leutzbach (1980) ORIENT: Ein verhaltensorientiertes Simulationsmodell zur Verkehrsprognose, *Working Paper*, **20**, Institut für Verkehrswesen, University of Karlsruhe.

Spiggle, S. and M. A. Sewall (1987) A choice sets model of retail selection, *Journal of Marketing*, **51** (2) 97–111.

Stahel, A. (2012) Agent interactions in the activity infrastructure of transport microsimulations, Master Thesis, IVT, ETH Zurich, Zurich.

Stauffacher, M., R. Schlich, K. W. Axhausen and R. W. Scholz (2005) The diversity of travel behavior: Motives and social interactions in leisure time activities, *Working Paper*, **328**, IVT, ETH Zurich, Zurich.

Steenbergen, M. R. (2003) *Maximum Likelihood Programming in Stata*, Department of Political Science, University of North Carolina, Chapel Hill, `www.unc.edu/~monogan/computing/r/MLE_in_Stata.pdf`.

Stopher, P. R., D. T. Hartgen and Y. Li (1996) SMART: Simulation model for activities, resources and travel, *Transportation*, **23** (3) 293–312.

Sun, D. (ed.) (2006) *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge University Press, Cambridge.

Susilo, Y. O. and R. Kitamura (2005) Analysis of day-to-day variability in an individual's action space, *Transportation Research Record*, **1902**, 124–133.

Swait, J. D. (2001) Choice set generation within the generalized extreme value family of discrete choice models, *Transportation Research Part B*, **35** (7) 643–666.

Swait, J. D. and M. E. Ben-Akiva (1987) Incorporating random constraints in discrete models of choice set generation, *Transportation Research Part B*, **21** (2) 91–102.

Swiss Federal Statistical Office (BFS) (2000) Eidgenössische Volkszählung 2000, `http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/vz/uebersicht.html`.

Swiss Federal Statistical Office (BFS) (2001) *Eidgenössische Betriebszählung 2001 - Sektoren 2 und 3, GEOSTAT Datenbeschreibung*, Swiss Federal Statistical Office (BFS), GEOSTAT, Neuchatel, `http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/geostat/datenbeschreibung/betriebszaehlung05.Document.111424.pdf`.

Swiss Federal Statistical Office (BFS) (2006) *Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten*, Swiss Federal Statistical Office (BFS), Neuchatel.

Swiss Federal Statistical Office (BFS) (2007) *Einkommens-und Verbrauchserhebung 2004 (EVE 2004)*, Swiss Federal Statistical Office (BFS), Neuchatel, `http://www.bfs.admin.ch/bfs/portal/de/index/infothek/publ.Document.89847.pdf`.

Swiss Federal Statistical Office (BFS) (2008a) *Eidgenössische Betriebszählung 2008 - Sektoren 2 und 3, GEOSTAT Datenbeschreibung*, Swiss Federal Statistical Office (BFS), GEOSTAT, Neuchatel.

Swiss Federal Statistical Office (BFS) (2008b) Haushaltsbudgeterhebung, `http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/habe/intro.html`.

Swiss Federal Statistical Office (BFS) (2012) *Mobilität in der Schweiz - Ergebnisse des Mikrozensus Mobilität und Verkehr 2010*, Swiss Federal Statistical Office (BFS), Neuchatel.

Teller, C. (2008) Shopping streets versus shopping malls—determinants of agglomeration format attractiveness from the consumers' point of view, *The International Review of Retail, Distribution and Consumer Research*, **18** (4) 381–403.

Teller, C. and T. Reutterer (2008) The evolving concept of retail attractiveness: What makes retail agglomerations attractive when customers shop at them?, *Journal of Retailing and Consumer Services*, **15** (3) 127–143.

Thill, J.-C. (1992) Choice set formation for destination choice modelling, *Progress in Human Geography*, **16** (3) 361–382.

Thill, J.-C. and A. J. Horowitz (1997) Travel-time constraints on destination-choice sets, *Geographical Analysis*, **29** (2) 108–123.

Thompson, R. G. and A. J. Richardson (1998) A parking search model, *Transportation Research Part A: Policy and Practice*, **32** (3) 159–170.

Timmermans, H. J. P. (1980) Consumer spatial choice strategies: A comparative study of some alternative behavioural spatial shopping models, *Geoforum*, **11**, 123–131.

Timmermans, H. J. P. (1983) Non-compensatory decision rules and consumer spatial choice behavior: A test of predictive ability, *The Professional Geographer*, **35** (4) 449–455.

Timmermans, H. J. P. (1996) A stated choice model of sequential mode and destination choice behaviour for shopping trips, *Environment and Planning A*, **28** (1) 173–184.

Timmermans, H. J. P. (2001) Models of activity scheduling behavior, *Stadt Region Land*, **71**, 33–47.

Timmermans, H. J. P. (2008) Consumer choice of shopping centre: an information integration approach, *Regional Studies*, **16** (3) 171–182.

Timmermans, H. J. P., T. A. Arentze, M. Dijst, E. R. Dugundji, C.-H. Joh, L. Kapoen, S. Krijgsman, K. Maat and J. Veldhuisen (2002) Amadeus: A framework for developing a dynamic multiagent, multiperiod, activity-based micro-simulation model of travel demand, paper presented at the *81st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2002.

Timmermans, H. J. P., T. A. Arentze and C.-H. Joh (2000) Modelling learning and evolutionary adaptation processes in activity settings, *Transportation Research Record*, **1718**, 27–33.

Timmermans, H. J. P., T. A. Arentze and C.-H. Joh (2001) Modeling effects of anticipated time pressure on execution of activity programs, *Transportation Research Record*, **1752**, 8–15.

Timmermans, H. J. P., A. W. J. Borgers and P. van der Waerden (1992) Mother logit analysis of substitution effects in consumer shopping destination choice, *Journal of Business Research*, **24** (4) 177–189.

Timmermans, H. J. P., R. van der Heijden and H. Westerveld (1982) The identification of factors influencing destination choice: An application of the repertory grid methodology, *Transportation*, **11** (2) 189–203.

Timmermans, H. J. P. and P. van der Waerden (1992) Store performance, pedestrian movement, and parking facilities, in G. Heinritz (ed.) *The Attraction of Retail Locations: IGU-Symposium*, Verlag Michael Lassleben, Kallmünz.

TomTom MultiNet (2011) MultiNet: The most powerful map database ever built, webpage, `http://licensing.tomtom.com/OurProducts/MapData/Multinet/index.htm`. Accessed 25/05/2011.

Topp, H. H. (1991) Parking policies in large cities in Germany, *Transportation*, **18** (1) 3–21.

TRANSIMS (2009) TRansportation ANalysis and SIMulation System, webpage, May 2009, `http://transims-opensource.org/`.

TRANSIMS Open Source (2013) Getting Started with TRANSIMS, webpage, `http://code.google.com/p/transims/wiki/GettingStarted`.

TRB (2007) Metropolitan travel forecastig: Current practice and future direction, *Special Report*, Committee for Determination of the State of Practice in Metropolitan Area Travel Forecasting, Transportation Research Board, Washington, D.C.

Trucano, T. G., L. P. Swiler, T. Igusa, W. L. Oberkampf and M. Pilch (2006) Calibration, validation, and sensitivity analysis: What's what, *Reliability Engineering and System Safety*, **91** (10-11) 1331–1357.

Tversky, A. (1972) Elimination by aspects: A theory of choice, *Psychological Review*, **79** (4) 281–299.

Uncles, M. D. (1996) Classifying shoppers by their shopping-trip behaviour: a polythetic-divisive method, *Marketing Intelligence & Planning*, **14** (1) 35–44.

Underhill, P. (1999) *Why we buy*, Simon & Schuster Paperbacks, New York.

Underhill, P. (2004) *Call of the Mall*, Simon & Schuster Paperbacks, New York.

UrbanSim (2011) Open Platform for Urban Simulation, webpage, `http://www.urbansim.org`.

U.S. Bureau of Public Roads (1964) *Traffic Assignment Manual*, U.S. Department of Commerce, Washington, D.C.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (1998) The impact of the parking situation in shopping centres on store choice behaviour, *GeoJournal*, **45** (4) 309–315.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (2006) Attitudes and behavioral responses to parking measures, *European Journal of Transport and Infrastructure Research*, **6** (4) 301–312.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (2009) Consumer response to the introduction of paid parking in a regional shopping center, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

van der Waerden, P., H. Oppewal and H. J. P. Timmermans (1993) Adaptive choice behaviour of motorists in congested shopping centre parking lots, *Transportation*, **20**, 395–408.

van der Waerden, P., H. J. P. Timmermans and A. W. J. Borgers (2002) PAMELA: Parking analysis model for predicting effects in local areas, *Transportation Research Record*, **1781**, 10–18.

van Kenhove, P., K. de Wulf and W. van Waterschoot (1999) The impact of task definition on store-attribute saliences and store choice, *Journal of Retailing*, **75** (1) 125–137.

van Middelkoop, M., A. W. J. Borgers and H. J. P. Timmermans (2004) Merlin: Microsimulation system for predicting leisure activity-travel patterns, *Transportation Research Record*, **1894** (1) 20–27.

van Vliet, D. (1977) An application of mathematical programming to network assignment, in P. W. Bonsall, Q. Dalvi and P. J. Hills (eds.) *Urban transportation planning: current themes and future prospects*, 147–158, Abacus Press, Kent.

Vause, M. (1997) A rule-based model of activity scheduling behavior, in D. F. Ettema and H. J. P. Timmermans (eds.) *Activity-Based Approaches to Travel Analysis*, 73–88, Pergamon, Oxford.

Veldhuisen, J., H. J. P. Timmermans and L. Kapoen (2000a) Microsimulation model of activiy patterns and traffic flows: Specification, validation tests, and Monte Carlo error, *Transportation Research Record*, **1706**, 126–135.

Veldhuisen, J., H. J. P. Timmermans and L. Kapoen (2000b) RAMBLAS: a regional planning model based on the microsimulation of daily activity travel patterns, *Environment and Planning A*, **32** (3) 427–443.

Venter, C. and M. Hansen (1998) Flexibility and time dependence in activity scheduling models, *Transportation Research Record*, **1645**, 120–126.

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

Vitins, B. J., T. Dubernet, A. Horni, N. Rieser-Schüssler and K. W. Axhausen (2012) MATSim Moduswahl Untersuchungen, *Final Report*, VW AG, Wolfsburg, IVT, ETH Zurich, Zurich.

Vogt, W. (2001) Zur Numerik nichtlinearer Gleichungssysteme (Teil 1), *Working Paper*, **Preprint No. M 12/01**, IfMath, TU Ilmenau, Ilmenau.

Vovsha, P., R. Donnelly and S. Gupta (2008) Network equilibrium with activity-based microsimulation models, *Transportation Research Record*, **2054**, 102–109.

Vovsha, P., E. Petersen and R. Donnelly (2002) Microsimulation in travel demand modeling: Lessons learned from the New York best practice model, *Transportation Research Record*, **1805**, 68–77.

Vovsha, P., E. Petersen and R. Donnelly (2003) Explicit modeling of joint travel by household members: Statistical evidence and applied approach, *Transportation Research Record*, **1831**, 1–10.

Vrtic, M. (2005) Simultanes Routen- und Verkehrsmittelwahlmodell, *Straßenverkehrstechnik*, **49** (8) 393–401.

Vrtic, M., P. Fröhlich and K. W. Axhausen (2003) Schweizerische Netzmodelle für Strassen- und Schienenverkehr, in T. Bieger, C. Lässer and R. Maggi (eds.) *Jahrbuch 2002/2003 Schweizerische Verkehrswirtschaft*, 119–140, Schweizerische Verkehrswissenschaftliche Gesellschaft (SVWG), St. Gallen.

Vrtic, M., P. Fröhlich, N. Schüssler, S. Dasen, S. Erne, B. Singer, K. W. Axhausen and D. Lohse (2005) Erstellung des nationalen Personenverkehrsmodells für den öffentlichen und privaten Verkehr, *Research Report*, Swiss Federal Department for Environment, Transport, Energy and Communication, Swiss Federal Office for Spatial Development (ARE), Swiss Federal Roads Authority and Swiss Federal Office of Transport, IVT, ETH Zurich, Emch und Berger and Institute for Transportation Planning and Traffic, Technical University Dresden, Zurich.

Vrtic, M., N. Schüssler, A. Erath, K. Meister and K. W. Axhausen (2007) Tageszeitliche Fahrtenmatrizen im Personenverkehr an Werktagen im Jahr 2000, *Research Report*, Swiss Federal Department for Environment, Transport, Energy and Communication, Swiss Federal Office for Spatial Development (ARE), Swiss Federal Roads Authority and Swiss Federal Office of Transport, IVT, ETH Zurich, Zurich.

Vucetic, S., T. Fiez and Z. Obradovic (2000) Examination of the influence of data aggregation and sampling density on spatial estimation, *Water Resources Research*, **36** (12) 3721–3730.

Waddell, P. A. (2010) Overview of UrbanSim and the Open Platform for Urban Simulation, presentation, UrbanSim Tutorial, Zurich.

Walker, J. L. (2005) Making household microsimulation of travel and activities accessible to planners, *Transportation Research Record*, **1931**, 38–48.

Wang, D. and H. J. P. Timmermans (2000) Conjoint-based model of activity engagement, timing, scheduling, and stop pattern formation, *Transportation Research Record*, **1718**, 10–17.

Wang, J., L. Zhang, Q. Rao and W. Yang (2013) Large-scale agent-based transport simulation in shanghai, china, paper presented at the *92nd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2013.

Wang, J.-Y. (2008) Variance reduction for multivariate Monte Carlo simulation, *The Journal of Derivatives*, **16** (1) 7–28.

Waraich, R. A. and K. W. Axhausen (2012a) Agent-Based Parking Choice Model, *Transportation Research Record*, **2319**, 39–46.

Waraich, R. A. and K. W. Axhausen (2012b) An agent-based parking choice model, *Arbeitsberichte Verkehrs- und Raumplanung*, **696**, IVT, ETH Zurich, Zurich, August 2012.

Waraich, R. A., C. Dobler and K. W. Axhausen (2012a) Modelling Parking Search Behaviour with an Agent-Based Approach, presentation, 13th International Conference on Travel Behaviour Research, Toronto, July 2012.

Waraich, R. A., C. Dobler, C. Weis and K. W. Axhausen (2012b) Optimizing parking prices using an agent based approach, *Arbeitsberichte Verkehrs- und Raumplanung*, **794**, IVT, ETH Zurich, Zurich.

Waraich, R. A., C. Dobler, C. Weis and K. W. Axhausen (2013) Optimizing Parking Prices Using an Agent Based Approach, poster presentation, 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., January 2013.

Wardrop, J. G. (1952) Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers*, **1** (3) 325–362.

Wegmann, F. and J. Everett (2008) Minimum travel demand model calibration and validation guidelines for state of Tennessee, *Technical Report*, Center for Transportation Research, University of Tennessee, Knowxville, Tennessee.

Weidmann, U. (1992) Transporttechnik der Fussgänger - Transporttechnische Eigenschaften des Fussgängerverkehrs, Literaturauswertung, *Schriftenreihe*, **90**, IVT, ETH Zurich, Zurich.

Weiner, E. (2008) *Urban Transportation Planning in the United States*, 3. edn., Springer, New York.

Weis, C., M. Vrtic, J.-P. Widmer and K. W. Axhausen (2011) Influence of parking on location and mode choice: A stated choice survey, *Working Paper*, **684**, IVT, ETH Zurich, Zurich.

Weisbrod, G., R. J. Parcells and C. Kern (1984) A disaggregate model for predicting shopping area market attraction, *Journal of Retailing*, **60** (1) 65–83.

Wen, C.-H. and F. S. Koppelman (2000) A conceptual and methodological framework for the generation of activity-travel patterns, *Transportation (Netherlands)*, **27** (1) 5–23.

Widmer, J.-P. and M. Vrtic (2004) Einfluss von Änderungen des Parkierungsangebots auf das Verkehrsverhalten, *Research Report*, **VSS 1997/46**, Schweizerischer Verband der Strassen- und Verkehrsfachleute (VSS), Büro Widmer, Frauenfeld.

Wild, B. (1994) Using artificial intelligence in traffic engineering - perspectives and potential applications, in M. Bielli, G. Ambrosino and M. Boero (eds.) *Artificial Intelligence Applications to Traffic Engineering*, 29–42, VSP, Utrecht.

Willumsen, L. G. (2000) Travel networks, in D. A. Hensher and K. J. Button (eds.) *Handbook of Transport Modelling*, chap. 10, 165–180, Pergamon, Amsterdam.

Wolf, D. A. (2001) The role of microsimulation in longitudinal data analysis, *Canadian Studies in Population*, **28** (2) 313–339.

Wooldridge, M. (2009) *An Introduction to MultiAgent Systems*, John Wiley & Sons, West Sussex.

Wu, N. and W. Brilon (1997) Cellular automata for highway traffic flow simulation, paper presented at the *14th International Symposium on Transportation and Traffic Theory (ISTTT)*, Jerusalem, July 1997.

Yang, M., W. Wang, X. Chen, W. Wang, R. Xu and T. Gu (2009) Influence of spatial factors, individual sociodemographics and travel mode on destination choice in China, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Young, W. (1986) PARKSIM/1: a simulation model of driver behaviour in parking lots, *Traffic Engineering and Control*, **27** (12) 606–613.

Young, W. and R. G. Thompson (1987) PARKSIM/1: a computer graphics approach for parking lot layouts, *Traffic Engineering and Control*, **28** (3) 120–123.

Young, W., R. G. Thompson and M. A. P. Taylor (1991) A review of urban car parking models, *Transport Reviews*, **11** (1) 63–84.

Zängler, T. W. (2000) *Mikroanalyse des Mobilitätsverhaltens in Alltag und Freizeit*, Springer, Berlin.

Zhang, M. and J. Ma (2008) Developing calibration tools for microscopic traffic simulation final report part 1: Overview methods and guidelines on project scoping and data collection, *Research Report*, **UCB-ITS-PWP-2008-3**, California Partners for Advanced Transit and Highways (PATH), University of California, Davis, Davis.

Zheng, J. and J. Y. Guo (2008) Destination choice model incorporating choice set formation, paper presented at the *87th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2008.

Ziems, S. E., S. Bhargava, J. Plotz and R. M. Pendyala (2011) Stochastic variability in microsimulation modeling results and convergence of corridor-level characteristics, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Ziliak, S. T. and D. N. McCloskey (2006) *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*, The University of Michigan Press, Ann Arbor.