

Destination choice modeling with spatially distributed constraints

Basil J. Vitins

Alexander Erath

Transport and Spatial Planning

May 2019

STRC

19th Swiss Transport Research Conference
Monte Verità / Ascona, May 15 – 17, 2019

Transport and Spatial Planning

Destination choice modeling with spatially distributed constraints

Basil J. Vitins
ASE, ETH Zurich
Gartenhofstrasse 17,
Stefano-Franscini-Platz 5
CH-8004 Zurich, CH-8093
Zurich
phone:

Alexander Erath
FHNW
Hofackerstrasse 30
CH-4132 Muttenz
phone:
alexander.erath@fhnw.ch

basil.vitins@ase.ch,
basil.vitins@ivt.baug.ethz.ch

May 2019

Abstract

Destination choice models are a key component of any transport and land use model, for scenario evaluation and forecast. Destination choice models describe highly complex decisions between spatially distributed alternatives. Employed explanatory variables usually include generalized travel cost for available transport modes, personal attributes such as age, income and also situational attributes like trip purpose. New kind of destination choice models account for information on individuals and exploit agent-based modeling advantages. In existing literature, already applied methodologies often stem from econometrics, discrete choice theory and utility maximization; estimating parameters on revealed or stated individual choice behavior.

This paper contributes to a remaining gap of complementary data integration, focusing on empirically measured cross-sectional flows in destination choice models: While earlier work (Vitins et al., 2016) focused on destination capacity constraints, this paper presents a framework to additionally integrate cross-section flows between distinct geographic areas as additional constraints. This allows choice model optimization against cross-section transport flows which can be obtained from a cordon surveys or mobile phone data. Proposed optimization methodology – based on extended shadow price theory – accommodates these complementary data sources. The new generic and robust optimization methodology improves destination choice modeling and quality while maintaining econometric theory. As a proof of concept, suggested methodology is applied successfully in a real-case agent-based application covering the tri-national Basel region, an urban agglomeration with about 2 million residents and a large set of $2 \cdot 10^4$ destination alternatives. Improvements are up to 20% and more on a cross-section level and even higher on a choice alternative level, compared to calculations ignoring shadow prices.

1 Introduction

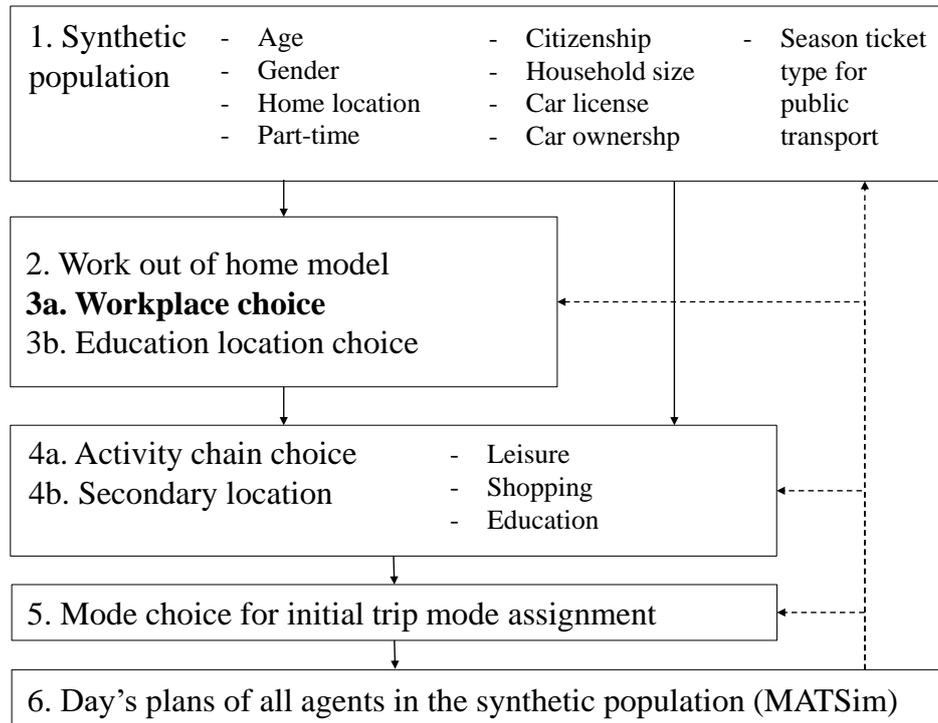
Destination choice assigns individuals with defined characteristics to destinations where they perform a planned activity. Destination choice therefore defines origin and destination pairs (OD) for activities such as home, work, education, shopping. In aggregated form those OD pairs sum up to massive flow between certain geographic areas. In microeconomics and econometrics, destination choice mirrors trip decisions of specific characteristics, and highly affects plans of individuals, households or firms. Destination choice therefore characterizes market behavior and potential for many mobility services and business locations. Destination choice also drives various economic processes and societal structures due to the massive sums of generated trips and flows. In macroeconomics, destination choice affects social welfare through accessibility and productivity changes (e.g. Venables, 2007). Accessibility to additional destinations increase productivity e.g. in agglomeration areas, and therefore increase welfare. Destination choice applications allow to account for short- and long-term revenue based on adapted or new trips as well as induced demand due to modified or new mobility services, transport infrastructure and spatial distribution of activity locations.

From a technical perspective, destination choice is often firmly embedded in travel demand modeling and transport modeling in general, and a key component for overall calculation of traffic flows and travel demand. It interferes with practically all other travel choices such as mode choice, departure time, or route choice, and therefore rises complexity on integration. Destination choice influences global and aggregated values such as distance distribution, vehicle-miles-travelled (VMT) and network flows. Therefore, it also affects practically all mobility scenario, forecast calculations and evaluations.

1.1 Embedded destination choice

Destination choice models are often key components of a larger travel demand model and transport simulation. In our case study, the proposed destination choice model is embedded in a large-scale activity- and agent-based travel demand model. An overview of the entire demand generation methodology is given in Figure 1 with arrows indicating some main dependencies of model elements, and dotted arrows indicating potential feedback loops. As a basis, all agents and their home locations are georeferenced on the basis of individual buildings and define a synthetic population with differing personal attributes (top of Figure 1). Proposed destination choice model assigns workplaces and school locations to corresponding working population and students, whereas this assignment is the main content of this paper. This paper focusses

Figure 1: Outline of travel demand calculation with embedded workplace choice model.



specifically on commuters, however, also students of three different age groups are assigned to school locations with basically the same methodological procedure. Afterwards, activity chains are assigned to all agents to determine additional, secondary activities, duration and time. Secondary location choice assigns activity to georeferenced locations, considering activity purpose. Mode choice model assigns an initially chosen transport mode to each individual trip, which is later optimized using MATSim.

From the supply side, above activity-based model is coupled with the high-performance, agent-based and multimodal transport simulation MATSim (Horni et al., 2016), including modules for simulations such as autonomous transport services. Additionally, feedback loops can return information to the activity-based model, such as travel time, for e.g. demand calculation refinement including destination choice, or scenario calculations.

1.2 Destination choice in travel demand modeling

This section highlights selected research in transportation related to destination choice. Destination choice builds up on the idea of individuals having a choice between different destination

alternatives, normally a finite set of distinct choice alternatives. This individual choice suits to micro-econometric choice theory (e.g. Ben-Akiva, 1973; McFadden, 1974). Utility maximization theory assumes different utilities among choice alternatives, and decisions depending on utility maximization of individuals. Those models are usually implemented using a logistic regression or logit model approach. As in other choice models, explanatory choice variables normally refer to categories of destination quality, personal variables, situational variables or a combination of them, e.g. car availability, car-only accessibility at a given destination can affect preferences for distinct destination and location of discretionary activities. Individual attributes might affect choice behavior and preference for a specific destination type, such as high income relates to specific economic sectors (Vitins et al., 2016). As in other choice models, outcomes are used for enhanced model evaluations e.g. for elasticities, or for superimposed transport models.

There are many model applications for destination choice. Activity-based models often include destination choice as part of activity chain definition, besides other choice models, such as out-of-home choice, mode choice, route or departure time choice. Vovsha et al. (2004) and Rasouli and Timmermans (2014) presented a comprehensive overview of activity-based modeling. Examples are DaySim (Bradley et al., 2010), ActivitySim (ActivitySim, 2019), SimMobility (Adnan et al., 2016) or CEMDAP (Bhat et al., 2004), all having a slightly different focus on e.g. required data sources, underlying activity theory, modeling framework, econometric foundation, and software applications. Many examples exist for activity-based theory. E.g. Bowman and Ben-Akiva (2001) proposed an activity-based disaggregate travel demand model integrating individual choices from a travel demand perspective. They applied a nested form of multiple tiers for multi-dimensional choice processes. Activity patterns, tours and time of day are estimated with choice models as well as primary and secondary destination and mode choice. Techniques are among others multinominal and nested logit models, and alternative sampling for destinations. Here, it is added that destination choice is a process of activity modeling and decoupled from supply modeling and simulation. Destination choice models, and also proposed methods, are therefore applicable for agent-based demand modeling, such as MATSim, or for aggregated models.

Above citations show how destination choice is embedded in a set of interfering individual travel choices. While the traditional four step model (e.g. Ortuzar and Willumsen, 2011) sequentially separate main travel choices and calculates all travel choices independently and recursively, one also argues that the choices should be seen as an entity of non-separable choices (e.g. Ben-Akiva, 1973). Given the complexity of the different, interfering choices and simultaneous calculation, it is clear that a complete inclusion and interaction of all factors is not achieved yet also when looking at the current research state. We are therefore at a stage where we implemented certain integrated models (e.g. Ben-Akiva, 1973; Adler and Ben-Akiva, 1976; Fotheringham, 1986,

and many more). Currently, it seems that computational power to estimate and to apply these models is still a constraint in practical applications and requests deliberate model decisions when applying theory in practice. Besides, also data availability is limited and limits model applications, e.g., census costs are very high for travel data, and personal data is often restricted for certain sources (e.g. mobile phone data).

1.3 Utility-based approach

In econometric choice theory, individuals decide independently about their destinations, based on alternatives' estimated utilities. In choice theory, a utility-based approach can account for generalized travel costs, destination quality, personal and situational variables. A range of literature exists on destination choice and probabilistic choice models. Exemplary, (1a) shows the general logit formula with utility u , coefficients β , alternative specific (transformed) attribute X . (1a) can be refined for generalized utilities, and other utility functions; (1b) describes multinomial logit assumptions.

$$p_i = \frac{e^{u_i}}{\sum_j e^{u_j}} \quad (1a)$$

$$\text{and for MNL with: } u_i = \beta_{0,i} + \sum_k \beta_k X_{i,k} \quad (1b)$$

Ben-Akiva and Lerman (1985) defined a multi-dimensional choice where individuals not only choose between workplaces but also between residential locations; and overall utility includes a sum of utilities of all dimensions, including the combination of dimensions (e.g. for generalized travel costs). Their approach is similar to a two-sided or doubly constraint approach (Section 1.4). Since a nested logit model is proposed to solve this multi-dimensional choice formulation, ordering of choice dimensions needs to be discussed for both location and workplace choice. For this reason, different papers have proposed different hierarchies, e.g. workplace choice as exogenous, or simultaneous decision (e.g. Waddell, 1993).

In the following, several aspects are listed especially relevant for destination choice, which complicate choice behavior:

1. Primary destination choice such as workplace choice is not only an individual's choice, but a two sided agreement between employers (and their locations) and employees, restaurants and customers or sport facilities being useful for certain activities. It is therefore not a pure and isolated individual's choice.
2. Potential capacity limitations at destination alternatives are unknown to individuals and cannot be considered in their decision process to a full extent. E.g. because of negative travel utility, individuals prefer closer locations with a certain probability. However, unevenly distributed origins (home locations) and destinations (e.g. workplaces, unevenly distributed professions) impede such trips to the shortest location due to potential overcrowdedness or availability and suitability of a location. This leads to market competition for attractive and popular workplaces.
3. Also, it is not a priori guaranteed that there is an equilibrium situation, on an individual level and on a firm level, respectively. On a firm level, one cannot, a priori, assume a market clearing situation for workplace location choice (e.g. McFadden, 1977; de Palma et al., 2007).

Technically, additional aspects arise with a utility-based choice model:

4. Destination choice models heavily depend on survey data. Whenever survey data is stratified, or biased on spatial attributes, it can be assumed that certain behaviors are not well captured for a given stratum within the model. It is possible to overcome this problem by estimating parameters for under-represented strata, but a certain parameter bias can be expected to persist during estimation, leading to a bias in destination choice model applications.
5. Recent, very elaborate, discrete choice models can deal with complex choice situations, e.g. spatial correlation (see also Section "Spatial correlation"). However, it is still possible that not all correlations are actually captured within the model, making systematic errors possible.
6. Perception of generalized travel costs might not be proportional to actual costs. One can assume that perception might be approximated with complex functions, but residual errors might still remain and bias modeling outcome (Bekhor and Prashker, 2008; Vovsha et al., 2012).
7. In case of commuting, contracts between employer and employee avoid overcrowdedness

at job locations. In case of school trips, assignments to school avoid overcrowdedness at school locations, in case of restaurant trips or recreational activities such as tennis, reservations and refusals prevent from overcrowdedness. Therefore, we need to implement such mechanisms also in our model where needed and meaningful.

8. For geographic areas that include borders, additional factors capturing the resistance to choose a location in another geographic area or language area might need to be included in the model.
9. Even if a model is fitted on a representative data set, parameter values might have to be changed in scenario applications. For scenarios with changing destination capacities, commuters might not be completely reassigned to all changing destinations due to influences of generalized travel costs on choice behavior.

1.4 Gravity models

The gravity model, widely and traditionally applied in transport modeling, mirrors gravitational attraction and repulsion, and is often seen as alternative to a logit based model formulation. Exemplary, (2) shows a basic gravity model where O_i are trips originating from i , D_j trips terminating in j , $f(d_{i,j})$ an impedance based on trip impedance such as distance (or time), and gravitational parameter G .

$$N_{i,j} = G \frac{O_i D_j}{f(d_{i,j})} \quad (2)$$

In traditional models, origin destination flows are fitted with a matrix-fitting method (e.g. Deming and Stephan, 1940). Marginal totals of all origins and destinations are exogenous inputs, for each considered trip purpose and person class. Examples are number of trip origins ('production') or workplaces at destinations ('attraction'). Origin-destination matrix elements are mutated such that matrix marginals fit with exogenous origin and destination marginals. An advantage of this approach is its robust procedure and forced convergence to the lowest overall marginal deviation. A difference to choice models is the underlying gravitational assumptions mirroring physical laws.

Anas (1983) showed that the logit model of joint origin-destination choice is consistent with

a doubly constrained gravity model. For comparison, the logit model implements a joint origin-destination zone pair choice and adds capacity constraints to both origins and destinations. It is remarkable that both approaches are equivalent views on the same problem, but with separate theoretical justification. Gravity models refer to information minimization and entropy maximization, while behavioral modeling refers to utility maximization and potentially explores individual utility in more advanced formulations. In the meanwhile, especially utility-based choice models have evolved in transportation theory and application; and developments have reproduced more detailed behavioral characteristics.

1.5 Market models

Market models are additional methods applied for destination choice. Especially market situations differ from a choice situations described in MNL (1b), also in the case of destination choice applications. There is difference between primary destination choice for residence, firms, commuters or scholars, and secondary destination choice for shopping and leisure activities. In the first case, individuals are not completely free to choose because of mutual agreement with e.g. a given employer or landlord, and their ties rely on a joined decision process between multiple parties, such as employee and employer. Travel utility might not be maximized, and individuals' choices potentially bias from their favorite destinations in case of access restrictions. Contrarily, and in the case of shopping and leisure, individuals are free to choose in most cases between shop alternatives, or visiting friends, though, in some cases with restrictions to capacity or accessibility, e.g. at restaurants, theatre, or club memberships. Because of this almost completely free choice, theory for secondary destination choice can potentially diverge from ideas of primary destination choice (e.g. Ben-Akiva, 1973; McFadden, 1974; Ben-Akiva and Bierlaire, 1999). Primary destination choice such as workplace choice therefore might resembles residential destination choice as applied in land use models.

Moving towards market-related methodologies, Hackney et al. (2013) and Hurtubia et al. (2010) summarized land market models and their implementation in land use models and were focusing 'bid' or 'auction' models, and hedonic pricing models. 'Bid' models mirror individuals maximizing their surplus, and adjust prices at destinations. Hedonic price models are often couples with utility-based choice models, and iterative procedures avoid over- and undercrowdedness (e.g. Zhou and Kockelman, 2011). Mentioned models are able to mirror market dynamics for large-scale models; it remains unclear if they are also applicable for transportation modeling since auction behavior possibly diverge from transport-related behavior.

1.6 Capacity constraints and shadow prices

Shadow prices serve as an impedance for attractive alternatives with limited capacity. Shadow prices reflect alternatives' constraints as explained in the following example based on micro-economics and productivity optimization. In an inefficient market situation, demand possibly exceeds supply due to distributional effects; then, shadow price implementation allows a solution of the optimization problem (Albouy, 2004). In economics, shadow prices are used to estimate unknown costs of a certain good or alternative. A price $p > 0$ can be assumed as well as a stock of X with sold units x where $0 \leq x \leq X$, and the objective $\max(px)$. Customers buying units x optimize their utility u , with respect to their time and budget limits, resulting in at least two dependent optimization problems. Any resource is considered a constraint (e.g. time, units), if the number that customers would like to use exceeds availability.

In linear programming, shadow price is associated with a constraint, and defines how much the optimal value of the objective would increase per unit increase in the amount of resources available, or how much individuals are willing to pay. They are in most cases equal to the solution of the dual variables of a given constrained linear program (Wagner, 1975). This idea of shadow prices is transferred and adjusted for destination choice in the following.

In destination choice, shadow prices can be assigned to designated destinations and regarded as additional impedance for persons choosing these destinations. Shadow prices can thus account for these alternatives' capacity constraints and scarcity, which can neither be captured with model parameters, nor with error terms (ϵ) of choice model distributions. Vitins et al. (2016) successfully applied a destination choice model on entire Singapore city-state, and described a destination choice model in detail, especially also in combination with generalized utility and $\log(\text{alternative capacity})$ (Lerman, 1977). Vitins et al. (2016) showed that destination choice coupled with shadow prices improves model quality considerably. Proposed methodology was successfully applied in an agent-based model of Singapore with 1'087 alternative destinations.

For more transportation application, literature provides information on shadow prices and some effects in behavioral choice models. However, detailed methodology and references are missing for destination choice models. Spiess (1996) applied shadow prices on parking choice. Gupta et al. (2014) used shadow prices for park-and-ride lots to identify the best parking lot, as did Davidson et al. (2011). Hammadou and Papaix (2014) applied shadow prices in a mode choice model for CO₂ pricing. de Palma et al. (2007) modeled housing markets, and defined that supply and prices in the housing market might clear demand, depending on the specific situation. They presented algorithms for constant and variable demand in the housing market context.

Literature has also focused on joint travel and residential costs based on mathematical programs. Beckmann and Wallace (1969) introduced shadow prices, similar to the proposed approach, for welfare maximization of home location changes with infrastructure changes, including housing rents. Los (1979) modified utility function, transportation and residential costs to improve residential choice and the usage and impact of transportation; travel behavior and residential location are coupled in mathematical programs.

1.7 Technical aspects in destination choice modeling

1.7.1 Spatial correlation

Spatial correlation between alternatives is problematic in choice model estimation due to unobserved spatial and demographic attributes. Bhat and Guo (2004) formulated a model for residential choice brought together random taste variation and correlation among choice alternatives, both key in choice modeling. They showed a dominant effect of travel times in their empirical test case, as well as significant spatial correlation among contiguous destinations, and differential responsiveness of households to exogenous variables. Sener et al. (2011) accommodates spatial correlation among any types of spatially distributed destinations (e.g. not restricted to contiguous destinations) including a case study in the San Francisco bay area. Spatial correlation can be tackled with considerable computational resources. It remains unclear how spatial correlation methods approach large-scale applications and agent-based model implementations.

1.7.2 Sampling techniques and calculation costs

McFadden (1978) provided techniques for sampling alternatives from the full set of alternatives due to infeasible data requirements; they were applied for residential location choice, and with aggregation of alternatives and similarities of alternatives. Lee and Waddell (2010) presented a nested logit model for residential choice with sampling techniques including a correction procedure for sampling bias. Thill (1992), Kwan and Hong (1998) and others describe choice set formations for geographic areas. They consider behavioral realism such as space-time prisms and trip chaining, information acquisition for alternatives, and other constraint oriented approaches. Besides that, they describe data requirements, computational advances and spatial (GIS) calculations, or stochastic processes. Lee and Waddell (2010) modeled residential location choice in a nested logit (NL) two-tier approach, where the upper tier decides the binary

decision of moving at all. They introduced sampling techniques for very large alternative sets, and a correction procedure for the NL model. Ben-Akiva and Bowman (1998) introduced a residential choice model, embedded in a activity based model similar to Bowman and Ben-Akiva (2001). Beside other variables, their model includes individual activity patterns, conditioned to accessibility. Frejinger et al. (2009) and Nerella and Bhat (2004) and proposed sampling methods and found that a large number of observations are needed to achieve reasonable model parameter values. Nerella and Bhat (2004) suggested drawing $\frac{1}{8}$ of the full choice set size as a minimum and $\frac{1}{4}$ as a desirable sample share - in the case of their MNL models, and non-MNL models are even more demanding about required sample size.

High calculation costs are still ubiquitous in model estimation and application. Many applications in literature run on smaller census data set or smaller geographic regions. However, when it comes to application, computer resources scale quadratic or more, and available computer resources might be insufficient, or parallelization and online computer resources are still limited, resulting in e.g. infeasibly long simulation and convergence time. Then, most complex and ‘accurate’ model might be impractical in real-world applications, and consequently, model definition potentially trades off against precision. In destination choice, alternative sets can be considerably large, and depending on model definition, models are estimated either on sampled data only, or full and complete data sets.

1.8 Research question

Gap on complementary data usage such as cross-section flows:

Section 1.6 discusses shadow prices as a promising method for capacity constraint consideration. A generalization of mentioned shadow prices at destination also allows to potentially constrain origins. At cross-sections, flow counts from one section to another would be exogenous and presumed as known input constraint. Generally, any origin-destination relation and defined subset of all trips can be potentially assumed as (empirically) known exogenous input constraint. These situations are relevant especially when flows of given origin-destinations systematically deviate from expected, empirical cross-section flows. In case of deviating flows, individual’s behavior is therefore not fully accounted in the underlying choice model, and considering such cross-section data is relevant for improvements on destination choice models.

Potential applications are manifold for subdivided origin-destination section and corridors (cross-sections). Such sections can be defined e.g. on geographical, or political reasons. Examples of geographically defined sections are mountain ranges separating two regions, or rivers and

other water bodies. Examples of politically motivated sections are country borders. In both cases, counts can be available on specific cross-sections, e.g. on mountain passes, or customs. In transport models, specific geographically or politically defined sections can be mirrored within model perimeter, and modeled flows can be compared on cross-sections with empirical cross-section values. Knowing about a methodology to account for this additional data source potentially increases choice model quality.

Derived research questions:

Two research questions are derived from above discussion:

1. Given travel census data availability, can we increase model quality by adding complementary data and constraints such as destination capacity (Section 1.6) or empirical cross-section flows?
2. Is it possible to utilize estimated destination choice parameters in large-scale applications with ten thousands of potential destinations within reasonable calculation times (Section 1.7)?

While focusing on an encompassing choice model in Vitins et al. (2016), this paper focuses on generalized theory for spatial constraints in destination choice, and applicability in a large-scale case study (tri-national agglomeration of Basel) with $2 \cdot 10^4$ potential destinations, political borders and different language areas.

2 Data Sources for Basel Case Study Application

2.1 Spatial Data

The spatial extent of the model covers an area of about 5,460 km² (85 km long and 91 km wide) with 1.96 million inhabitants as of 2015. Basel as a main city lies in the corner of the three nations Switzerland, Germany, France; therefore the region is often referred as the trinational region as the Basel metropolitan area extends across all three countries.

Population distribution is displayed in Figure 7(a) exemplary for Basel city, and workplace distribution in Figure 7(b) (placed in the Result section for later outcome comparison). Data is

collected from BfS (2017) for Switzerland and complemented for Germany and France based on Bau- und Verkehrsdepartement Basel-Stadt (2015); BfS (2017). Population and workplaces are collected for 20,645 zones as indicated in Figure 2(a). In an inner core area around Basel, zones are generated based on inhabited hectare raster cells (with side length of 100m). The side length of the zones in the outer core area amounts to 200m, while the remaining area is covered by zones that represent the municipal borders.

2.2 Survey data

2.2.1 Cross-border flows

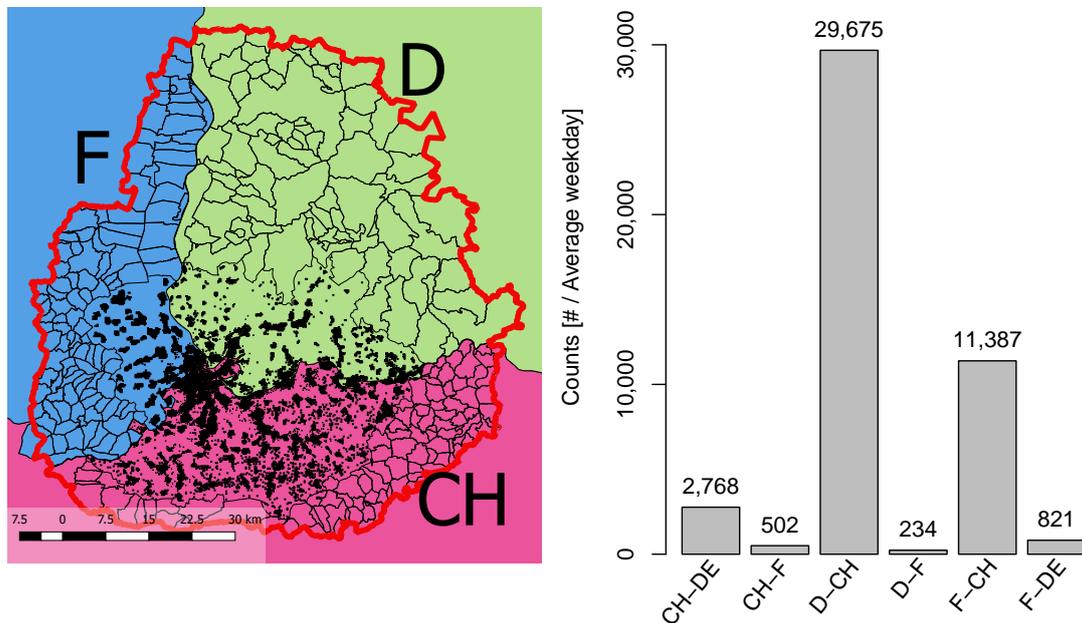
Data on cross-border flows are derived from a cordon line survey which was conducted in 2010, available through Bau- und Verkehrsdepartement Basel-Stadt (2015) for this paper. They are available as average weekday counts for specific trip purposes, and visualized in Figure 2(b) for commuters. Values include home-workplace trips, excluding return trips. Flow differences are considerable and are accounted for in the proposed choice model approach. To fit an agent-based destination choice model in such conditions was the main motivation to develop the model presented in this paper. Cross-border flows are only available for a reduced region compared to the entire model perimeter, indicated in Figure 2(a) with small rasters. Identical behavior is assumed for remaining flows (also in line with proposed method below).

2.2.2 Micro census 2015

To compare travel distance distributions derived from the model with data from observed behaviour, we employ the Swiss Mobility and Transport Microcensus (MTMC) (Swiss Federal Statistical Office (BFS), 2017) which was conducted in 2015 by the Federal Office for Statistics of Switzerland (further referred to as "MZMV 2015"). Data of 57,090 interviewed persons and their households were collected in 2015, providing a substantial sample of the population and travel behavior. Selected households were drawn systematically within the country of Switzerland, with higher densities in larger cities and agglomeration regions. MZMV 2015 is used for outcome comparison and reference distributions; in this paper, only data points within perimeter area are selected from census data.

Figure 2: Model overview and considered commuters cross-section flows.

(a) Map of model region with perimeter border, (b) Counted cross-section flows between regions zones, and country borders.



Source: Bau- und Verkehrsdepartement Basel-Stadt (2015)

2.2.3 Commuter flows 2014

Results of the proposed shadow price methodology are compared against a commuter flow evaluation conducted originally for entire Switzerland for 2014 (BfS, 2017). For this study, data is cut at perimeter borders, and is further referred to as "Pendlermatrix". Commuting flows and distances of "Pendlermatrix" are compared with flows and distances calculated by the proposed methodology.

2.3 Existing destination choice model of Basel region

Logsum parameter values are used in the following from the existing destination choice model of the identical region of Basel (Bau- und Verkehrsdepartement Basel-Stadt, 2015). The existing destination choice model is based on a logsum utility calculation for each potential destination, including car, public transport (pt), bicycle and pedestrian mode. Mode utilities are estimated with specific parameters and corresponding variables. For car, *travel times*, *distance*, *estimated park search time*, *gas costs*, *parking and car costs* are considered in the model. For pt, *in-vehicle*

time, access time, ticket costs, vehicle interchange, service frequency, season card availability card were used. *Travel time* is used for bicycle and pedestrian trips, calculated with average travel speeds. Model parameters are estimated for individual trip purposes, namely commuting, educational, shopping, leisure and service trips. Separate models are implemented for each trip purpose, with separate parameter values.

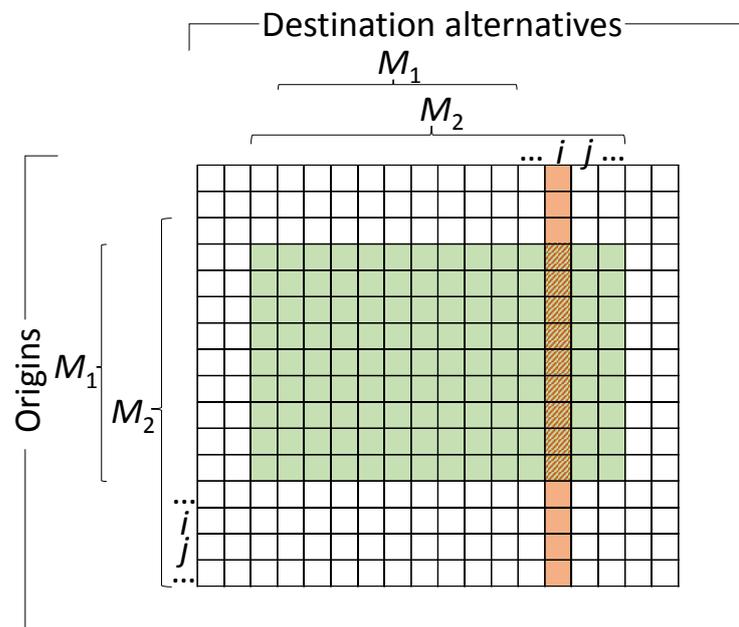
In the existing destination choice model (Bau- und Verkehrsdepartement Basel-Stadt, 2015), destination choice calculation is based on a logsum term as described above and a gravitation approach with gravitation parameter optimization. For this paper, logsum model parameters are identical with the existing model. Compared to the existing model, this paper uses an MNL based destination assignment approach with logsum calculation, above described variables and identical parameter values, and adaption on the global logsum parameter values due to missing gravitational parameter (2).

Methodological similarities between the two approaches, MNL and gravity, are described in Anas (1983) in detail with all formulae transformation (see also Section 1.3 and 1.4). This paper excludes further model re-estimation and focuses on spatially distributed constraints. Still, proposed methodology of this paper also allows replacement of given utility formula and to substitute it with a more comprehensive generalized utility approach. (If budget would allow, generalized utility estimation is recommended as examined in Vitins et al. (2016)).

3 Methodology

Proposed methodology focuses on destination choice embedded in transport models, and therefore refers to agents travelling to destinations for activity performance. Two generic constraints are added in the following, namely for destinations and cross-section flows. Destinations potentially have certain capacity constraints; e.g. at workplaces, we assume that workplace numbers are met either exactly, or serve as upper limit. Destination constraints are relevant in destination choice models, because capacity is potentially exceeded when applying a choice model on each agent without knowing decisions of other agents. Technically, workplace constraints are defined as one-dimensional constraint (array), whereas each array value refers to only one specific destination. Another dimension is added for cross-section flows. In the two-dimensional case, we refer to exogenous information on cross-sections, such as counts, between defined geographic regions. Counts can be available between two regions separated by a mountain range or political border, where we potentially know the flow counts. In the two-dimensional approach, cross-section counts serve as constraint when assigning destinations. One- or two-dimensional

Figure 3: One-dimensional destination constraint (orange) i and two-dimensional cross-section constraint M_1 – M_2 (green).



constraints are therefore independent of each other, and can be considered alone or jointly in destination choice calculations. Figure 3 visualizes both categories of spatially distributed constraints; x and y axes are destinations and origins, such as workplaces and home locations.

Proposed methodology offers diverse application potential for the one- and / or two-dimensional constraints as showed in Figure 3, such as cross-section flow estimation, capacity constraint alternatives (e.g. vehicle size in route choice, destination capacity on destination choice) on various scales. The methodology described in the following is a generalization of the one-dimensional case described in Vitins et al. (2016), and additionally includes the two-dimensional case.

3.0.1 Methodological prerequisites

Two relevant requirements are discussed for the constrained destination choice. (1) Proposed methodology should be applicable to scenario and forecast applications. Therefore, constraints need to be adaptive e.g. in the case of future population growth. (2) Incomplete data might still be considered in proposed methodology, especially at available flow counts which cover only part of the model area. Following list describes requirements for spatial constraints:

1. One-dimensional case:
 - a) Capacity constraints at single destinations.
 - b) Capacity constraints at a defined subset of specified destinations.
2. Two-dimensional case:
 - a) Cross-section constraint between defined geographic regions.
 - b) Cross-section constraint between subset of defined geographic regions.
3. Different approximation options:
 - a) Ceiling option for a given upper limit (e.g. capacity constraints at workplaces).
 - b) Bottom option for a given lower limit (e.g. known lower bound at workplaces).
 - c) Exact approximation for a given value (e.g. precise workplace number or cross-section flow counts).
4. Scenario and forecast capabilities with two alternative options:
 - a) Identical parameter weights for scenario and forecast as in calibrated scenario (see below explanation).
 - b) Newly estimated parameter weights with new constraints.

As another requirement, proposed methodology should be applied in the following environments:

1. Applicable in a reasonable geographic size for a transport model (reference case-study includes 2 Mio. inhabitants).
2. Calculation time and computational memory requirements needs to be reasonable for this size (case-study includes 20,645 destination as potential alternatives).

3. Algorithm can be applied in both agent-based applications, and macroscopic flow models (case-study has both options).

Spatial correlation describes potential correlation between specific workplace with neighboring workplaces. In this paper, it is assumed that spatial correlation is less relevant and therefore excluded based on the following reasons:

1. Today, job search (workplace assignment) often takes place on internet search machines and might exclude spatial correlation since the internet search is not affected by spatial distances.
2. Job search based on individual contacts might be affected by spatial correlation, because of the distribution of individual contacts in space. However, number of individual contacts decrease with distance (Kowald et al., 2013), and therefore in relevance.
3. Industry sector consideration can further improve choice models. This behavior is covered with the industry – profession type match described in generalized utility, and potentially reduce correlation effects.

3.1 Shadow price calculations for destination choice models

Following methodology aims at refining existing model approaches for destination choice in transport models. Proposed methodology starts at the same premises as MNL choice models (Section 1.3). Moreover, certain modification are added in model theory, described below:

1. Shadow prices are defined and calculated as dis-utility added to destinations due to capacity restrictions. So, shadow prices are positive, and are negatively perceived by choice makers.
2. It is assumed impossible to assign more workers at designated workplaces than indicated in the workplace survey data. On the other side, it might be possible to have certain vacant workplaces. Therefore, an assignment balance is assumed with a certain upper restriction determined by the maximum number of workplaces, which should not be violated.
3. Cross-section flows are exogenous values; it is assumed that they match exactly with the modeled flows.
4. Constraint sums need to always allow a feasible solution on destination assignments.

5. Choice model parameters for utility calculations are given as input parameters, and therefore estimated beforehand (Bierlaire, 2019).

Starting with utility-maximizing theory (1a), probability $p_{i,j}$ and flow $g_{i,j}$ of choosing alternative destination j are defined as:

$$g_{i,j} = P_i \cdot p_{i,j} = P_i \cdot \frac{e^{u_{i,j}}}{\sum_j e^{u_{i,j}}} \quad (3)$$

whereas $u_{i,j}$ equals deterministic utility of alternative j . P_i is the number of individuals commuting from origin i (a building or a zone). Individual attributes and situational attributes are omitted in (3), by omitting corresponding indices. However, individual and situational attributes can be easily re-added for completeness.

It can be shown that the following convex minimization problem (4a) is equivalent to (3), by applying the Kuhn-Tucker conditions, feasible because considered function is partially differentiable (Horst, 1979):

$$\text{Min} \sum_{i,j} g_{i,j} (\ln(g_{i,j}) - 1 + u_{i,j}) \quad (4a)$$

subject to:

$$\sum_j g_{i,j} = P_i \quad (4b)$$

Capacity restrictions of destination zones j (4c) are added as addition constraints to the optimization problem above ((4a),(4b)). Moreover, two-dimensional matrix restrictions are also added (4d) where we define all geographic subregions for origin – destination cross-section flows with $M = \{M_1, M_2, \dots, M_m\}$ where $\{i, j\} \in M_m$ for both origin and destination subregions (see Figure 3 for a schema visualization). Known cross-section flow values are defined in $B = \{B_{M_1, M_1}, B_{M_1, M_2}, \dots\}$ and added as additional constraints in (4d). In-/equality of these constraints depends on available data and are due to known upper limit, lower limit, or exact approximation of a given cross-section flow.

$$\sum_i g_{i,j} \leq C_j \quad (4c)$$

$$\sum_{i \in M_1, j \in M_2} g_{i,j} = B_{M_1, M_2} \quad (4d)$$

Total number of employees should not exceed total number of workplaces available (4e), or the problem above becomes unfeasible. This requirement is assumed as given beforehand. However, proposed methodology is also applicable if constraint (4e) is violated, but only with resulting oversaturated destinations. Initialization of B -values can be incomplete, meaning that not all possible cross-section flows of M are needed for proposed methodology. Still, defined set of cross section flow constraints should be compatible and inclusive (e.g. for M_1 see (4f)).

$$\sum_k C_k \geq \sum_k P_k \quad (4e)$$

$$\sum_m B_{M_1, M_m} \leq \sum_{i \in M_1} P_i \quad (4f)$$

In practical model applications, inequation of capacity constraints as in (4e) means that the sum of workers is lower or equal to the sum of workplaces available. Here, two aspects complicate practical applications: First, workers traveling from outside the model perimeter are not considered in the model, and their workplaces should be subtracted. This is actually often the case in any destination choice model. Second, part-time workers can be considered with appropriate adaptations.

We introduce the Lagrange function defined as the following with vectors λ_1 , λ_2 , and matrix λ_3 added as Lagrange multipliers, to reduce equality and inequality constraints. λ_1 has length i , λ_2 has length j , λ_3 has length $\leq i \times j$. λ_2 is constraint to ≥ 0 due to the inequality of (4c):

$$L(g_{i,j}, \lambda_1, \lambda_2, \lambda_3) = \sum_{i,j} (g_{i,j} \cdot \ln(g_{i,j}) - g_{i,j} + u_{i,j} \cdot g_{i,j}) + \lambda_1 \left(\sum_j g_{i,j} - P \right) + \lambda_2 \left(\sum_i g_{i,j} - C \right) + \lambda_3 \left(\sum_{i \in M_1, j \in M_2} g_{i,j} - B \right) \quad (5)$$

Now, optimality conditions of (5) are ($\lambda_2 \geq 0$):

$$g_{i,j} = e^{-\lambda_{1,i}-\lambda_{2,j}-\lambda_{3,i \in M_1, j \in M_2}-u_{i,j}} \quad \text{for all } i, j. \quad (6)$$

and

$$\sum_j e^{-\lambda_{1,i}-\lambda_{2,j}-\lambda_{3,i,j}-u_{i,j}} = P_i \quad \text{for all } i. \quad (7a)$$

$$\sum_i e^{-\lambda_{1,i}-\lambda_{2,j}-\lambda_{3,i,j}-u_{i,j}} \leq C_j \quad \text{for all } j. \quad (7b)$$

$$\sum_{i \in M_1, j \in M_2} e^{-\lambda_{1,i}-\lambda_{2,j}-\lambda_{3,i,j}-u_{i,j}} = B_{M_1, M_2} \quad \text{for all defined } M_n, M_o \text{ pairs.} \quad (7c)$$

Unlike (4a) – (4c), dual problem (7a) – (7c) comes without constraints; λ_1 , λ_2 and λ_3 can be determined by solving (7a) – (7c).

For efficiency, all variables in (7a) – (7c) are transformed: $\alpha = e^{-\lambda_1}$, $\beta = e^{-\lambda_2}$, $\gamma = e^{-\lambda_3}$ $U_{i,j} = e^{-u_{i,j}}$, where $\alpha, \beta, \gamma, U > 0$ and $\beta < 1$:

$$\alpha_i \cdot \sum_j \beta_j \gamma_{i,j} U_{i,j} = P_i \quad (8a)$$

$$\beta_j \cdot \sum_i \alpha_i \gamma_{i,j} U_{i,j} \leq C_j \quad (8b)$$

$$\gamma_{i,j} \cdot \sum_{i \in M_1, j \in M_2} \alpha_i \beta_j U_{i,j} = B_{M_1, M_2} \quad (8c)$$

3.2 Efficient algorithm to determine shadow prices on one and two dimensional constraints

Given above formulae, following set of unknown parameter values are given: utility u , λ_2 and λ_3 , whereas utility decomposes into its β parameters. The problem is a non-linear equation system with linear equality and inequality constraints and parameters needed to be fitted with empirical data. Normally, utility function fitting is based on a maximum likelihood method. Assuming a generalized logit model, it can be estimated with known methods (McFadden, 1974; Ben-Akiva and Lerman, 1985). Ignoring constraints, the maximum likelihood method can be applied with a various algorithms for unconstrained nonlinear optimization since it is strictly concave; and algorithm mainly differ in convergence speed and memory.

After estimating utility, β and γ values need to be found in an iterative approach. Algorithm 1 describes an iterative procedure to determine shadow prices. Threshold t defines how much capacity should not be exceeded by a given destination. E.g. $t = -2$ means that capacity can be exceeded by maximum 2. Algorithm 1 approximates a balanced situation within an adequate number of iterations, where all commuters are assigned to a workplace and where defined cross-section flows are approximated to empirical values. Shadow prices λ_2 , λ_3 can be viewed as additional utilities for each individual, to respect capacity constraints and cross-section flows. λ_2 reflect spatial development potential of underdeveloped, or even missing, but valuable alternatives. λ_3 is a shadow prices on a specific relation between M_1 and M_2 . Overall, it might be possible that similar results could be obtained by randomly assigning weights to locations compared to results achieved with Algorithm 1; however, Algorithm 1 efficiently approximates a balanced situation, where all the commuters find a designated working location under given cross-section constraints.

Proposed Algorithm 1 is based on a gradient projection methodology, similar to Haftka and Gürdar (1992, Chap. 5). It starts with a solution where β and cross-section shadow prices are set to 1 (both shadow prices are then 0). Generally, the algorithm assumes that active constraints at point g^n in the n^{th} iteration are also active at g^{n+1} . Subsequently, vectors β and cross-section shadow prices are redefined after intermediate results g^n when necessary. Equality condition of the cross-section impedance (4d) leads to a continuous recalculation of the cross-section shadow prices in each iteration n when values are over- and underestimated compared to the given constraints. In case of the destination capacities, β are only recalculated when the inequality is violated and destinations are oversaturated in iteration n . Similar optimization methods are applied widely and successfully in other fields.

Algorithm 1 Shadow price calculation for one and two dimensional constraints.

$n \leftarrow 0$

$\beta_n \leftarrow 1$

$\gamma_n \leftarrow 1$

while $C - \sum_i g_i < t$ **do**

1. Calculate the demand $g_{i,j,n}$ for each pair i, j based on (6) and (8a):

$$g_{i,j,n} \leftarrow \frac{P_i \cdot \beta_n \cdot \gamma_n \cdot U_{i,j}}{\sum_j \beta_n \gamma_n U_{i,j}} \quad (9)$$

2a. One dimensional case: Recalculate β parameters based on (6) and (8b):

$$\beta_{n+1,j} \leftarrow \min\left(\frac{C_j \cdot \beta_n}{\sum_i g_{i,j,n}}, 1\right) \quad (10)$$

2b. Two dimensional case: Recalculate γ parameters based on (6) and (8c):

$$\gamma_{n+1,i,j} \leftarrow \frac{B_{M_1, M_2} \cdot \gamma_n}{\sum_{i \in M_1, j \in M_2} g_{i,j,n}} \quad (11)$$

2c. One and two dimensional case: Recalculate β and γ parameters based on (6),(8b) and (8c):

3. $n \leftarrow n + 1$

end while

Shadow prices for destination alternatives: $\lambda_2 \leftarrow -\log(\beta)$

Shadow prices for cross section flows: $\lambda_3 \leftarrow -\log(\gamma)$

Terminate.

3.3 Data quality issues

Data quality and approximation of spatial data is often an issues for larger models. It might be that certain workplace numbers are incomplete for given areas and need to be approximated, or the number of part-time workers are not precisely known. This can lead to unequal sums between global destination capacity and agents assigned to a trip. Proposed algorithm handles these issues, and still approximates destination choice. If capacity is scarce, destinations are overfilled where it uses least utility. If capacity is more than enough destinations are filled up

where individual utility can be minimized. However, proposed methodology differs in origin and destination balancing as common matrix fitting approaches (e.g. Deming and Stephan, 1940). Algorithm 1 only fits destination assignment, and trip number remains as initially defined.

4 Results

This section contains three parts: First, Section 4.1 shows Algorithm 1 convergence related to capacities at destinations, and cross-section flows. Second, Section 4.2 discusses spatial shadow price distribution as a result of Algorithm 1. Third, Section 4.3 demonstrates further Algorithm 1 applications.

4.1 Algorithm 1 convergence characteristics

4.1.1 Global convergence

Algorithm 1 convergence is essential for destination saturation evaluation. This section discusses global convergence behavior, for two reasons. First, convergence is expected when applying Algorithm 1 as shown in above theory, and therefore results should conform accordingly. Second, convergence behavior is essential for determination of the number of iterations conducted with Algorithm 1, for final cut-off criteria definition.

Figure 4(a) shows median and mean of all destination workplace difference relative to available workplaces at destinations, for each iteration $n = 0 \dots 14$. Figure 4(a) depicts that over- and undersaturation slowly decrease towards a global workplace saturation at growing iteration number, which is the main purpose of Algorithm 1. Second, cross-section flows also converge towards a minimum. Figure 4(a) also shows cross-section mean values (median is ignored because of the low numbers of 6 examined cross section flows).

Figure 4(b) refers to the specific cross section flows between the countries and therefore contains more detail than global mean value evaluation in Figure 4(a). It shows larger differences between calculated cross-section flows and counts. More specifically, Figure 4(b) depicts unbalanced flows between both directions of each cross-section considered in the model. In the Basel case study, this is due to more directional flows of commuters in one direction. Figure 4(b) depicts under-estimation of flows in one direction, and over-estimation of flows in the other direction; this

holds for the cross-section Switzerland - Germany and Switzerland - France through all iterations, and France - Germany at some iterations. This is in line with directional flows observed in the counts.

Definition of cut-off criteria for Algorithm 1 is subject to interpretation. Major convergence takes place before iteration 10 (Figure 4(a)), however, small assignment changes are calculated after iteration 10 especially at specific cross-section flows (Figure 4(b)). It is assumed that convergence behavior also varies between case studies, and therefore needs to be an evaluation element of any application.

Algorithm 1 takes about 2.5 hours to calculate 15 iterations on a 2.6 GHz Intel Core i7 processor with 15.8 GB memory, with considerably optimized code structure. Iteration 0 uses about 1.5 hours of total calculation time due to encompassing utility calculation.

4.1.2 Workplace saturation

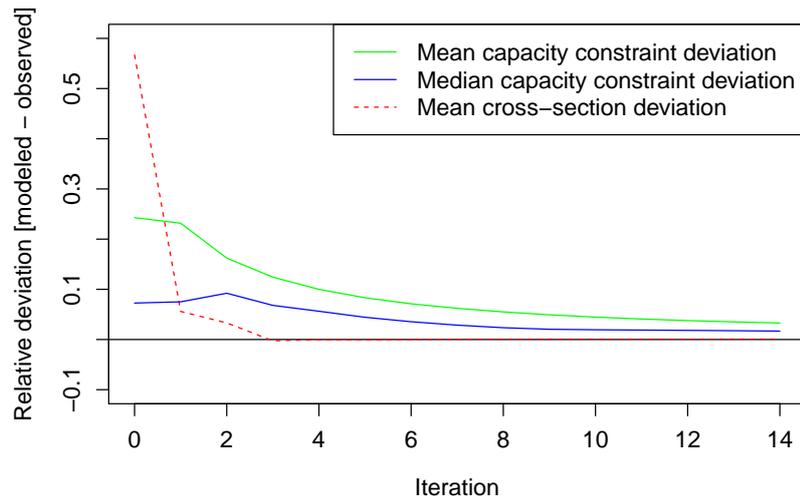
Figures 5(a) and 5(b) depict workplace saturation at iteration 0 and 15, respectively. Over- and under-saturated workplaces can be found in iteration 0, Figure 5(a), whereas workplaces are almost completely saturated in iteration 15 (Figures 5(b)). Especially oversaturation is minimized as a result of penalties to overcrowded destinations (ceiling condition in (4c)).

4.1.3 Trip length distribution

Figure 6 shows cumulative distance distribution of all modeled commuting trips, and potential distributional changes due to Algorithm 1 application. A priori, it is unclear how much Algorithm 1 affects trip length distribution. For comparison, Figure 6 includes two reference distance distributions. First reference distance distribution includes trips from the national micro census, selected for the Swiss perimeter area; second reference distance distribution represents the selected and encompassing commuting trip distribution (see Section 2.2.3 for data source description). Figure 6(a) shows the distance distribution of all commuters before applying Algorithm 1. Due to reference data availability, only commuting trips are displayed starting and ending in the Swiss perimeter area. At this stage, destination alternatives are potentially overcrowded, and cross-section flows potentially differ from observed counts. Figure 6(b) shows cumulative distance distribution after applying Algorithm 1 for 15 iterations, including cross-section flow balancing, and shadow prices calculations at destinations. Figures 6(a) and 6(b) both show that before and after applying Algorithm 1, overall population travel distributions

Figure 4: Algorithm 1 convergence behavior.

(a) Destination constraints and cross-section flow constraint optimization.



(b) Convergence behavior of Algorithm 1 for cross-section flows, compared to counts.

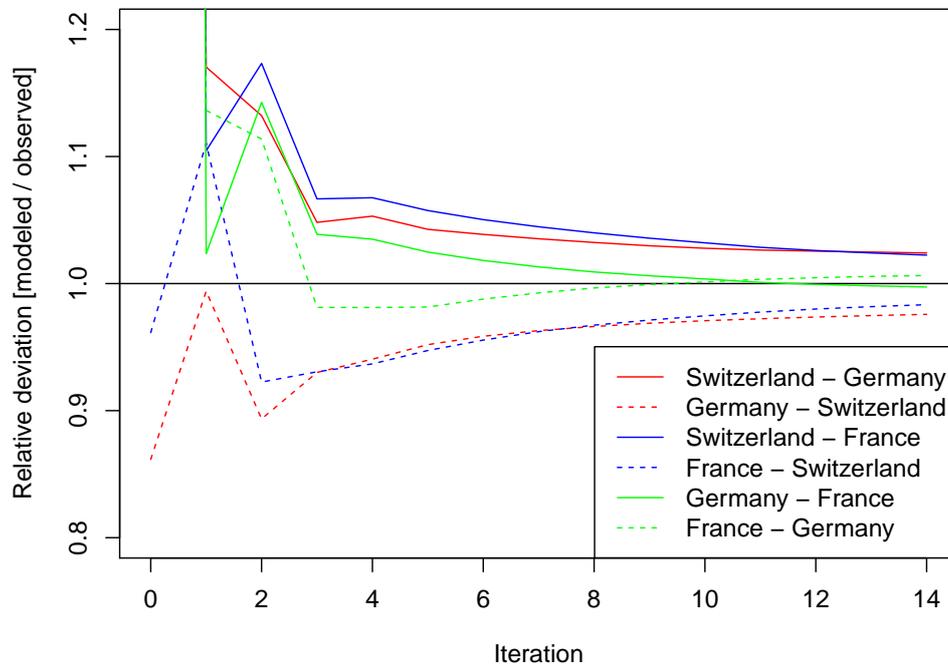
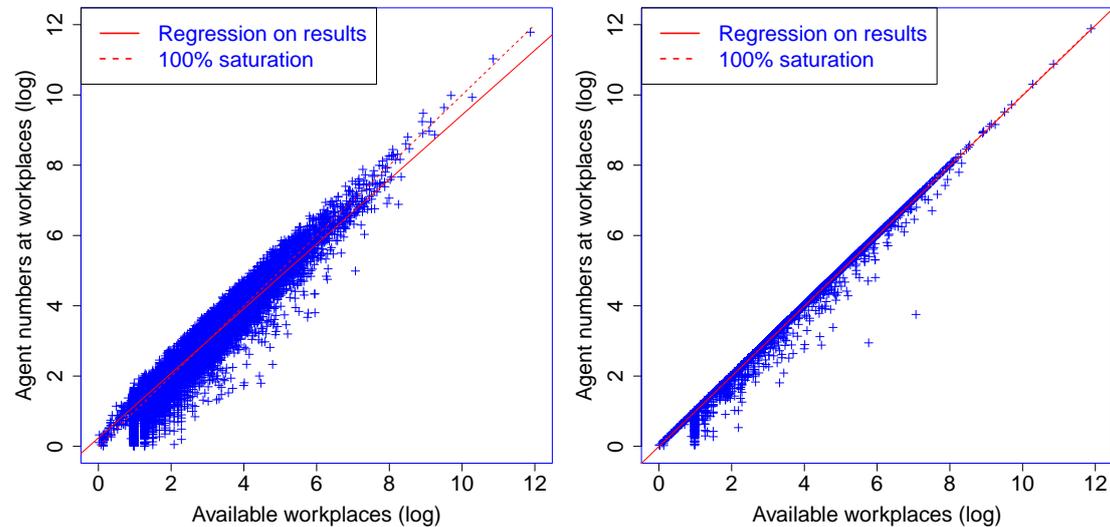


Figure 5: Regression with workplace capacities and workplace saturation.

- (a) Saturation after choice model application (iteration 0). (b) Number of workplaces and saturation (iteration 15).

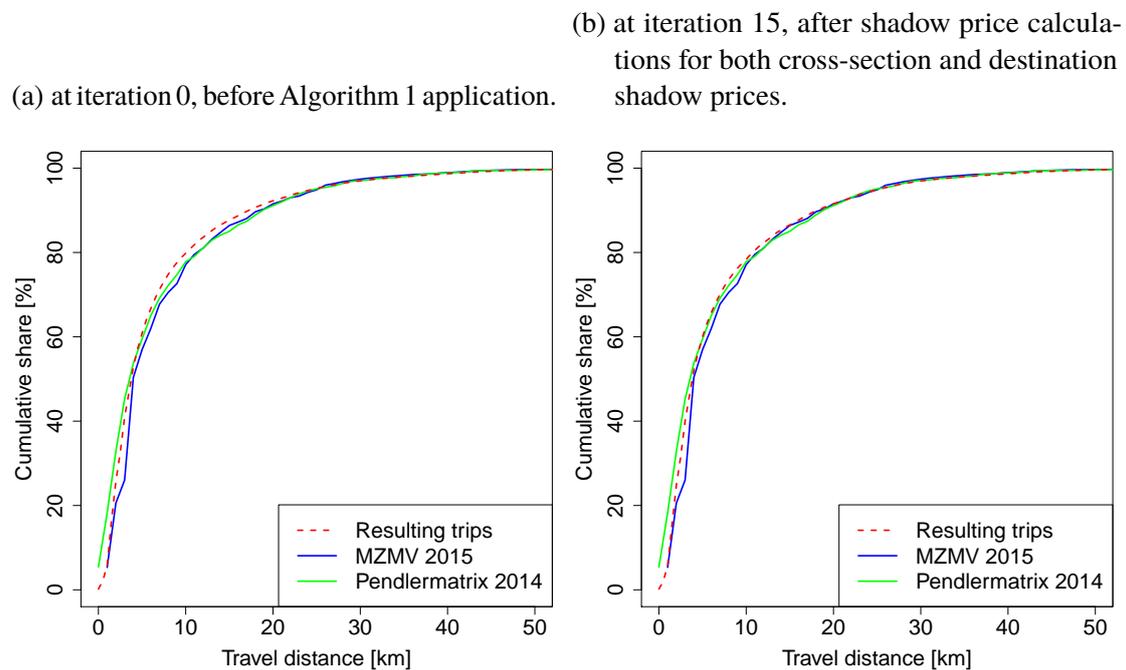


mirrors the reference data provided from alternative sources. Distance distribution only slightly increase after applying Algorithm 1 especially between 5km and 20km travel distance, because of reassigning agents to different locations. However, effects on travel distance distribution are minor, and overall fit most likely satisfies modeling needs.

4.2 Shadow price distribution and interpretation

This section describes spatial shadow price distribution, in comparison with population and workplace distribution. Insights improve model understanding, but also future land use development and planning. Figure 7(a) show spatial distribution of the Basel population, Figure 7(b) workplaces and 7(c) shadow price ($\lambda_2 = -\log(\beta)$) distribution. Shadow price distribution in Figure 7(c) depicts a centralistic pattern. Prices are higher at outskirts and generally lower in the city center. This distribution derives from a high concentration of jobs centralized near the city center, and a comparably higher population density in the outskirts. Accessibility might be a main reason for Basel's centralized development, and land development. While firms compete for very accessible locations, land prices rise and population are spread towards agglomeration with lower living costs, however, also lower workplace densities. For city planning, it is therefore key to know about workplace assignment.

Figure 6: Cumulative car distance distribution for commuters of the entire synthetic population of the Swiss perimeter area, in comparison with reference distributions (Section 2.2).



4.3 Applications, forecast and scenarios

Multiple applications are evaluated for this section. First, Algorithm 1 is applied independently for capacity constraint and cross-section flow optimization. Second, infrastructure scenario and third, forecast capabilities are described in the following.

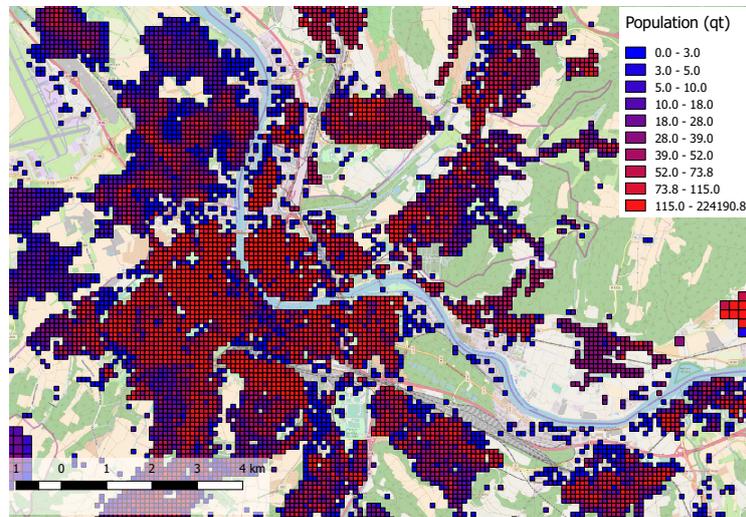
4.3.1 Independent capacity constraint and cross-section flow optimization

Algorithm 1 is modifiable for optimization of only (a) cross-section flow, and (b) destination capacity constraints.

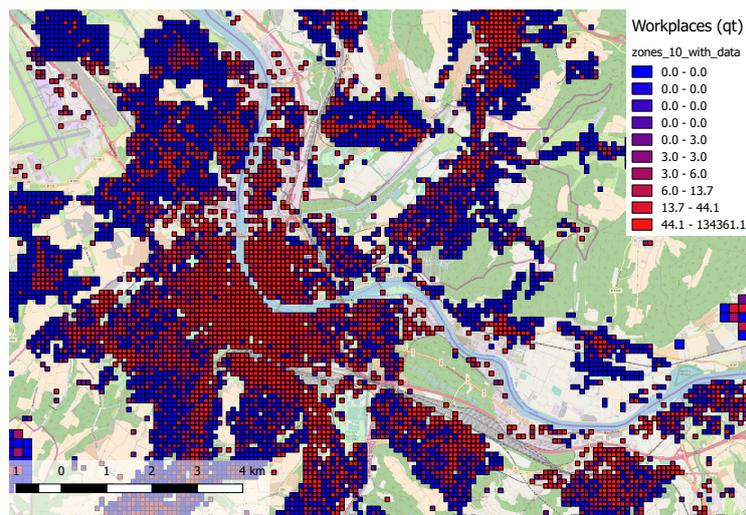
- (a) Cross-section counts are approximated in Algorithm 1, and capacity constraints at destinations are excluded during calculations. Step 2a and (10) are ignored in Algorithm 1. Figure 8(a) shows cross-section flow optimization; as expected, capacity constraints are violated after Algorithm 1 calculations.
- (b) Only capacity constraints at destination alternatives are considered in Algorithm 1, and cross-section counts are unnecessary and excluded during calculations. Step 2b and

Figure 7: Basel population, workplace and shadow prices distributions.

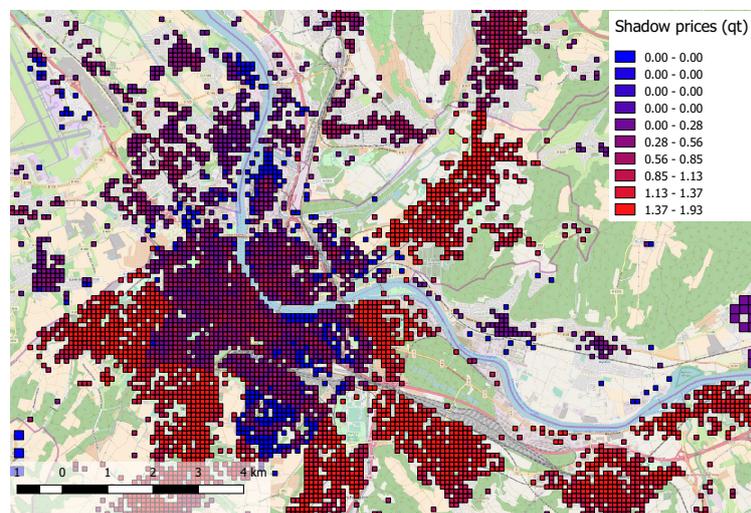
(a) Population distribution.



(b) Workplace distribution.



(c) Shadow prices λ_2 distribution after iteration 15.



(11) are ignored in Algorithm 1. Figure 8(b) shows capacity constraint optimization; as expected, cross-section counts are not approximated after Algorithm 1 calculations. This case is similar to calculation in Vitins et al. (2016).

4.3.2 Forecast capabilities

Growth in population and number of workplaces is considered in forecast scenarios. For destination choice, alternative destinations or destinations with changing capacities affect individual destination choice behavior. Often and for transport models, detailed population and workplace growth is calculated for forecast scenarios, and destination choice is recalculated; this procedure can be applied similarly with adapted alternatives' capacities and Algorithm 1. Differently, cross-section counts are missing for forecast. For this reason, two solutions (c) and (d) are proposed for improved forecast and destination choice calculation.

- (c) Calculated and optimized γ values for available cross-sections are reapplied for forecast destination choice calculation. In this solution, cross-section impedance values γ remains as in original, present destination choice calculation. Most probably, cross-section flows deviate from empirical flows due to implemented growth conditions in forecast population and workplaces. Figure 8(c) depicts average deviation of workplace capacities and cross-section flows, of a forecast scenario with global 20% population and workplace growth. Figure 8(c) shows deviation fore cross-section flows as supposed above, however, capacity constraints are met similar to above evaluations. (Content of Figure 8 is comparable to Figure 4(a).)
- (d) Cross-section counts are assumed to be known, which means that values either remain or grow accordingly (e.g. by 20%). Figure 8(d) shows convergence behavior with remaining cross-section counts as in original data. As expected, Algorithm 1 optimizes destination choice behavior and matches choice behavior with counts.

4.3.3 Scenario application

Figure 9 shows shadow price changes between two scenarios differing in a planned infrastructure, namely a fictitious city link connecting two city neighborhoods. Figure 9 depicts city neighborhood clusters changing in their β values depending on the affect of the link. It shows a systematic change in values. (Here, any further case-study interpretation is skipped because of possibly unfeasible assumption on underlying data.)

Figure 8: Global convergence behavior of Algorithm 1 over 15 iterations, differing in considered spatial constraints.

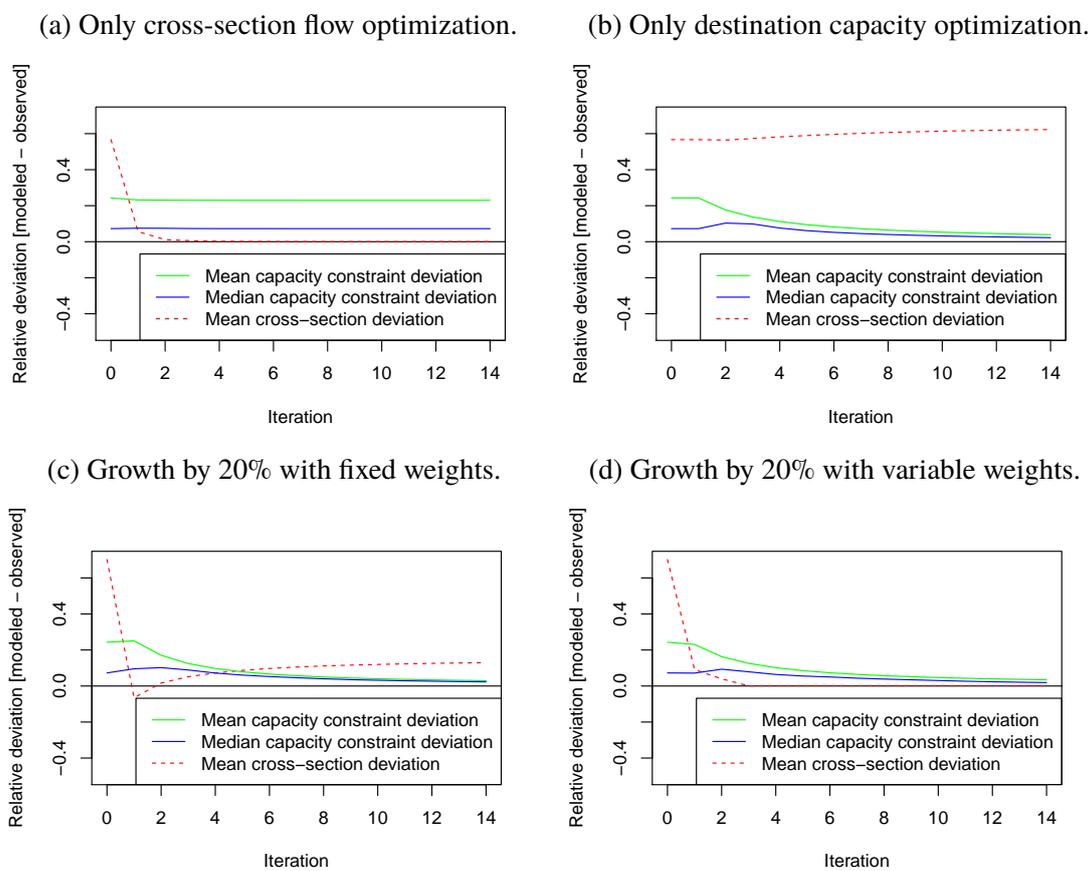
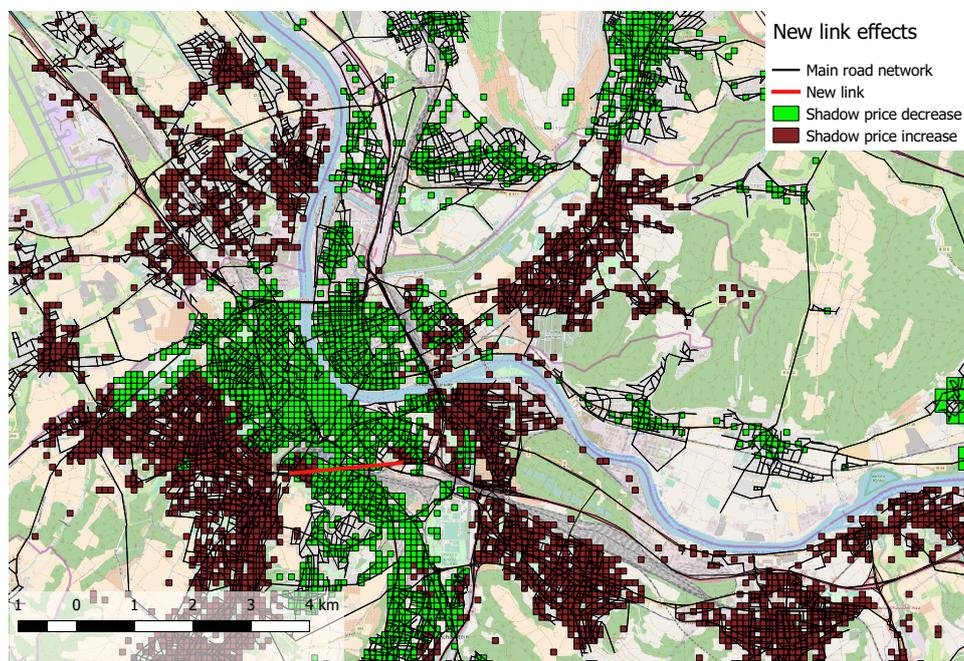


Figure 9: Clustering of β changes when adding new infrastructure, here road link.



Source: Assumptions differing from and uncomparable with any other planned study.

5 Discussion and Outlook

Two generic types of spatial constraints are successfully integrated in a new and robust algorithm to improve destination choice precision. Proposed methodology is applied in a detailed, large-scale model with $2 \cdot 10^4$ destination alternatives. Results show that applied methodology improves destination choice on a global, cross-section level and on a micro-level on choice alternatives. On cross-sections, precision improvements are as high as 20% and more compared to without Algorithm 1. Regarding choice alternatives and destination capacities, proposed methodology is able to avoid over-saturations at popular locations. Both methodologies for cross-section flows and destination capacities are applicable separately or simultaneously in destination choice calculations. Moreover, proposed methodology allow forecast and scenario calculation with necessary parameter choice (constant or variable impedance for cross-section flows). And for land use planning purposes, shadow prices help predict future spatial development potential of underdeveloped - or unused, valuable - destination alternatives.

Therefore, proposed methodology serves as robust solution for transport models and simulations, still maintaining choice heterogeneity especially relevant for agent-based simulations. It therefore can be efficiently implemented in large-scale agent-based models as in applied Basel case. In transport simulations, improvements are probably even higher than 20% on single routes after route assignment. According to our experiments and results, travel time distribution of all trips slightly changes when applying proposed methodology, compared to empirical reference sample distributions.

Still, additional research might be of interest, such as the following two topics: (1) Even though it is not calculated in this paper, one can assume that statistical model fit improves during Algorithm 1 iterations as in Vitins et al. (2016), where ρ^2 increases from $\rho^2 = 0.115$ to $\rho^2 = 0.168$ in contrast to ignoring capacity constraints. It therefore seems that Algorithm 1 is capable of further improving overall model prediction; this needs to be verified on cross-section flow calculations. (2) Geographic border effects might still occur in model application, similar to 'border effects' occurring in other transport models. Effects of persons from areas outside model perimeter might affect all calculations; pertaining to areas easily accessible from outside of the model perimeter. These external influences are difficult to capture quantitatively, e.g. at route choice or departure time choice. Additional studies can reveal quantitative extents of these effects.

Referring to Anas (1983), it is clear that proposed methodology extends the utility maximization approach, and further explores its broad application potential for large numbers of individuals with various characteristics. Compared to the choice model applied in this paper, it is also

possible to include additional data to estimate parameter values directly with a choice model methodology. Improvements can account for different economic sections and job types (industry, chemistry, etc.). Assignment should be in combination with a detailed destination choice model including generalized utilities to consider economic sections, as proposed in Vitins et al. (2016). Then, improvements account at least partially for specialization and economical benefits for agglomerations (e.g. as in Venables, 2007).

Various other planning tasks are potentially solvable with applied methodology: Vehicle occupancy will be additionally interesting for research and suitable for shadow prices, coupled with mode and route choice, for dense cities with frequently overloaded public transport lines and services. Battery charging stations for e-mobility or general parking choice might profit from the proposed methodology, along with transportation modeling. Land use related, primary location choice can be considered with proposed methodology for households and firm location. Spatial economics (as mentioned above) can profit, and probably any constraint resource distribution, in a wider sense.

6 Acknowledgements

We thank Alexei Korn, Artem Chakirov, Julia Wysling, Michael van Eggermond, and the city of Basel for providing the regional transport model (Bau- und Verkehrsdepartement Basel-Stadt, 2015).

7 References

- ActivitySim (2019) ActivitySim, webpage, <https://activitysim.github.io/>.
- 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.
- Adnan, M., F. Pereira, C. Lima Azevedo, K. Basak, M. Lovric, S. Raveau, Y. Zhu, J. Ferreira, C. Zegras and M. Ben-Akiva (2016) SimMobility: A multi-scale integrated agent-based simulation platform, paper presented at the *95st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2016.
- Albouy, D. (2004) Constrained optimization, shadow prices, inefficient markets, and government projects, *Working Paper*, **Economics 131**, University of California, Berkeley, Berkeley.
- Anas, A. (1983) Discrete choice theory, information-theory and the multinomial logit and gravity models, *Transportation Research Part B: Methodological*, **17** (1) 13–23.
- Bau- und Verkehrsdepartement Basel-Stadt (2015) Gesamtverkehrsmodell der Region Basel 2012, *Final Report*, Bau- und Verkehrsdepartement Basel-Stadt, Mobilität / Mobilitätsstrategie, Singapore, March 2015.
- Beckmann, M. J. and J. P. Wallace (1969) Evaluation of user benefits arising from changes in transportation systems, *Transportation Science*, **3** (4) 344–351.
- 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.
- Ben-Akiva, M. E. (1973) Structure of passenger travel demand models, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.
- Ben-Akiva, M. E. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.) *Handbook of Transportation Science*, chap. 2, 5–34, Kluwer, Dordrecht.
- Ben-Akiva, M. E. and J. L. Bowman (1998) Integration of an activity-based model system and a residential location model, *Urban Studies*, **35** (7) 1231–1253.
- Ben-Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.
- BfS (2017) Swiss federal statistical office (bfs), webpage, <http://www.bfs.admin.ch>.

- Bhat, C. R. and J. Y. Guo (2004) A mixed spatially correlated logit model: Formulation and application to residential choice modeling, *Transportation Research Part B: Methodological*, **38** (2) 147–168.
- 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.
- Bierlaire, M. (2019) *PandasBiogeme: a short introduction*, <http://transp-or.epfl.ch/documents/technicalReports/Bier18.pdf>.
- 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.
- 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.
- Davidson, B., P. Vovsha and J. Freedman (2011) New advancements in activity-based models, paper presented at the *Australasian Transport Research Forum 2011*, Adelaide, September 2011.
- 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.
- Deming, W. E. and F. F. Stephan (1940) On the least squares adjustment of a sampled frequency table when the expected marginal totals are known, *Annals of Mathematical Statistics*, **11** (4) 427–444.
- Fotheringham, A. S. (1986) Modelling hierarchical destination choice, *Environment and Planning A*, **18** (3) 401–418.
- 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.
- Gupta, S., P. Vovsha, R. Kumar and A. Subhani (2014) Incorporating cycling in ottawa-gatineau travel forecasting model, paper presented at the *5th Conference on Innovations in Travel Modeling*, Baltimore, April 2014.
- Hackney, J. K., B. J. Vitins and B. R. Bodenmann (2013) Market-Clearing Models in FaLC, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

- Haftka, R. T. and Z. Gürdar (1992) *Elements of Structural Optimization*, Springer, Heidelberg.
- Hammadou, H. and C. Papaix (2014) Estimating shadow-prices of urban transport mode choice for a second best pricing of CO₂ emissions, paper presented at the *93rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2014.
- Horni, A., K. Nagel and K. W. Axhausen (eds.) (2016) *The Multi-Agent Transport Simulation MATSim*, ubiquity press.
- Horst, R. (1979) *Nichtlineare Optimierung*, Carl Hanser Verlag, Munich.
- Hurtubia, R., F. Martánez, G. Flötteröd and M. Bierlaire (2010) Comparative analysis of hedonic rents and maximum bids in a land-use simulation context, paper presented at the *10th Swiss Transport Research Conference*, Ascona, September 2010.
- Kowald, M., P. van den Berg, A. Frei, J. A. Carrasco, T. A. Arentze, K. W. Axhausen, D. Mok, H. J. P. Timmermans and B. Wellman (2013) Distance patterns of personal networks in four countries: a comparative study, *Journal of Transport Geography*, **31**, 236–248.
- Kwan, M.-P. and X.-D. Hong (1998) Network-based constraints-oriented choice set formation using GIS, *Journal of Geographical Systems*, **5**, 139–162.
- Lee, B. H. Y. and P. A. Waddell (2010) Residential mobility and location choice: A nested logit model with sampling of alternatives, *Transportation*, **37** (4) 587–601.
- Lerman, S. R. (1977) Location, housing, automobile ownership, and mode to work: A joint choice model, *Transportation Research Record*, **610**, 6–11.
- Los, M. (1979) Combined residential-location and transportation models, *Environment and Planning A*, **11** (11) 1241–1265.
- McFadden, D. (1974) Conditional logit analysis of qualitative choice-behaviour, in P. Zarembka (ed.) *Frontiers in Econometrics*, 105–142, Academic Press, New York.
- McFadden, D. (1977) Modeling the choice of residential location, *Working Paper*, **477**, University of California, Berkeley, Yale University, Berkeley, New Haven.
- McFadden, D. (1978) Modeling the choice of residential location, in A. Karlqvist (ed.) *Spatial Interaction Theory and Residential Location*, 75–96, North-Holland, Amsterdam.
- 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.
- Ortuzar, J. d. D. and L. G. Willumsen (2011) *Modelling Transport*, 4th edn., John Wiley & Sons, Chichester.

- Rasouli, S. and H. J. P. Timmermans (2014) Activity-based models of travel demand: Promises, progress and prospects, *The International Journal of Urban Sciences*, **18** (1) 31–60.
- Sener, I. N., R. M. Pendyala and C. R. Bhat (2011) Accommodating spatial correlation across choice alternatives in discrete choice models: An application to modeling residential location choice behavior, *Journal of Transport Geography*, **19** (2) 294–303.
- Spiess, H. (1996) A logit parking choice model with explicit capacities, *Research Report*, EMME2 Support Center, INRO, Aegerten.
- Swiss Federal Statistical Office (BFS) (2017) *Verkehrsverhalten der Bevölkerung - Ergebnisse des Mikrozensus Mobilität und Verkehr 2015*, Swiss Federal Statistical Office (BFS), Neuchatel.
- Thill, J.-C. (1992) Choice set formation for destination choice modelling, *Progress in Human Geography*, **16** (3) 361–382.
- Venables, A. J. (2007) Evaluating urban transport improvements, *Journal of Transport Economics and Policy*, **41** (2) 173–188.
- Vitins, B. J., A. L. Erath and K. W. Axhausen (2016) Integration of a capacity constrained workplace choice model: Recent developments and applications with an agent-based simulation in Singapore, *Transportation Research Record*, **2564**, 1–13.
- Vovsha, P., M. A. Bradley and J. L. Bowman (2004) Activity-based travel forecasting models in the united states: Progress since 1995 and prospects for the future, paper presented at the *EIRASS Conference in Advances in Activity Based Approaches*, Maastricht, May 2004.
- Vovsha, P., S. Gupta, J. Freedman, W. Sun and V. Livshits (2012) Workplace choice model: Comparison of spatial patterns of commuting in four metropolitan regions, paper presented at the *91st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2012.
- Waddell, P. A. (1993) Exogenous workplace choice in residential location models: Is the assumption valid?, *Geographical Analysis*, **25** (1) 65–82.
- Wagner, H. M. (1975) *Principles of operations research: with applications to managerial decisions*, Prentice-Hall, Englewood Cliff.
- Zhou, B. B. and K. Kockelman (2011) Land use change through microsimulation of market dynamics: An agent-based model of land development and locator bidding in austin, texas, *Transportation Research Record*, **2255**, 125–136.