

## The role of location in residential location choice models: a review of literature

Patrick M. Schirmer

Institute for Transport Planning & Systems, ETH Zurich <sup>a</sup>

Michael A.B. van Eggermond

Future Cities Laboratory, Singapore ETH Centre <sup>b</sup>

Kay W. Axhausen

Institute for Transport Planning & Systems, ETH Zurich <sup>c</sup>

**Abstract:** Geospatial data available to researchers has increased tremendously over the last several decades, opening up opportunities to define residential location in multiple ways. This has led to a myriad of variables to define “location” in residential location choice models. In this paper, we propose a common classification for location variables and categorize findings from a wide range of studies. We find similar preferences but different measurement methods and market segments for locations across different study regions. Recent studies consider the residential unit as choice alternative, making it possible to include a detailed description of the built environment. However, these studies are still limited in number and the inclusion of socioeconomic environment is more common. Transport land-use models can benefit from the inclusion of points of interest, such as schools, network distances, and the distance to previous locations. For the results of location choice models to be transferable to different disciplines, and avoid multi-collinearity, it is necessary to present different model specifications, including variables of interest in different disciplines.

**Keywords:** residential location choice; discrete choice; land use; location

### 1 Introduction

Residential household location is one of the driving forces of urban dynamics. It impacts employment, economic development, social structure, spatial segregation, and the transport system. Understanding and modeling residential location choice behavior is a primary concern for urban planners, policymakers, and researchers.

The roots of residential location modeling can be traced back to the first advances by von Thunen (1826) in land-use modeling; he explained transport costs' effect on activity locations and land market function through a single market in an agricultural region. Integral to modeling is the bid-rent concept; landowners are willing to rent their properties to the highest bidder. Alonso (1964) applies this bid-rent concept to residential locations and considers a mono-centric city with employment opportunities. Individuals and households choose their residential location by maximizing a utility function depending on expenditure in goods, size of the land lot, and distance to the city center.

Parallel to Alonso, Lowry (1964) applies the gravity model to residential location. Lowry assumes an initial set of basic employment centers per zone. Households are allocated to zones based on a deterrence function describing the number of workers employed and living in a certain zone. Residential attractiveness is measured by the amount of land available for development in a particular zone.

---

<sup>a</sup>patrick.schirmer@ivt.baug.ethz.ch

<sup>b</sup>eggermond@ivt.baug.ethz.ch

<sup>c</sup>axhausen@ivt.baug.ethz.ch

<sup>a</sup>patrick.schirmer@ivt.baug.ethz.ch

<sup>b</sup>eggermond@ivt.baug.ethz.ch

<sup>c</sup>axhausen@ivt.baug.ethz.ch

With the introduction of the discrete modeling framework (McFadden 1978), we find many studies describing residential location choice. The strong point of this framework is its ability to quantify the impact of different residential location characteristics and their interaction with household characteristics. With detailed specification of an alternative, one can avoid describing a residential location solely by its distance to the city center and the corresponding mono-centric perspective; residential alternatives and location with a wide range of attributes—a contributing factor for diversity found in cities—can then be used.

Early studies relied mainly on census data as aggregation level; they characterized a zone as alternative and used zonal attributes to describe possible moving destinations for households (Anas 1982; Weisbrod *et al.* 1980). In the last two decades, we have seen increasing availability of disaggregated data: census data is available down to a one-meter resolution. Cadastral information is not only digitalized, but it can be linked to buildings' or individual residential units' attributes. Geo-databases allow combinations of various datasets, as well as adding a spatial reference point. With the increasing availability of such data, ways to represent residential alternatives and their environment have changed; recent models consider buildings or units as choice alternatives and include building- and location-specific attributes (Habib and Miller 2009; Lee and Waddell 2010a).

The range of hypotheses and attributes tested in residential location choice models has evolved significantly and varies across studies. While early studies compared similar levels of granularity and employed similar methodology, more recent studies differ in granularity level and study focus. This has complicated comparisons of location preferences across different studies and limited ability to compare household preferences. One may assume that tastes are similar for comparable study areas and that there is a set of recommended variables for modeling residential location choice.

This paper investigates “location” within residential location models using a literature review that classifies different studies' reported variables, finds common relevant attributes between studies, defines various interaction terms, and summarizes the main literature elements.

## 2 Residential choice and operationalization of space

### 2.1 Discrete choice modeling

In this review, we consider residential location choice studies employing the discrete choice framework. Within this framework, a decision maker chooses a single alternative from a set of mutually exclusive alternatives, the choice set, thus obtaining a certain level of utility from each alternative and choosing the alternative providing most utility.

The decision maker obtains a certain level of utility from an alternative. However, the researcher cannot see all attributes of the alternatives faced by the decision maker, nor does he know the decision maker's attributes. Thus, an error term is introduced to capture the difference between true and observed utility (e.g., Train 2003).

If these discrete choice model characteristics are translated to residential location choice, the residential alternative is the dependent variable. As highlighted in the introduction, this residential alternative can be a certain zone, neighborhood, building, or residential unit. Independent variables could describe the residential alternative itself (number of rooms, price level, etc.). Residential location choice, by definition, is a spatial choice and, as such, description of the residential alternative consists partly of residential unit and alternative location attributes, discussed in Section 2.3. Section 2.2 discusses attributes of the decision maker in the residential alternative.

### 2.2 Characterizing the decision maker

Attributes (characteristics) describing the decision-making household include size, income, and social class or ethnic background. Within the choice set of a decision maker, values of these sociodemographic attributes are constant among different alternatives in the choice set. Sociodemographic attributes can only be included in the specification of an alternative if they vary from alternatives. Normally, this is accomplished by interacting a sociodemographic attribute with an alternative specific attribute. Examples include preference for schools in households with children or distance to an urban center for single member households. By introducing these interaction variables in the specification of an alternative, one can estimate different taste preferences for different market segments.

Characterizing households by their life cycle is one possible way to obtain market segmentation. Characteristics of household composition, such as number of children or number of employed persons, are aggregated to a life cycle. Examples include single

parent households, married couples without children, and households with teenagers. In Section 3, preferences of households in different life cycles will be highlighted.

A second differentiation of households is based on household lifestyle; this concept emerged during the 1980s in social sciences and helps models to better describe mode choice, residential location choice, and vehicle ownership. To determine household lifestyle, households are grouped based on daily life behavior and their cultural, social, and leisure preferences (Müller 1992).

Walker and Li (2007) differentiate between three types of lifestyles and find different preferences for urban density, retail and service density, and lot sizes, with car-oriented households preferring lower densities and bigger lot sizes. Krizek and Waddell (2003) differentiate between nine lifestyles, each with differences in location choice (e.g., preferences for urban density) and mobility behavior (e.g., preferences for travel distance).

### 2.3 Location

One sees a widely varied range of attributes describing location in residential location choice studies, and each of these studies categorize and operationalize spatial variables in a separate way. A common classification of location variables, to simplify assessing locational preferences across different studies, seems to be lacking. For our literature review, we propose the following classification of location variables:

- *Built environment*, defined by geometries and volumes of spatial objects, including buildings, parcels, blocks, and connecting networks (both road and public transport networks). These objects occupy a place in space; for instance, buildings contain firms and households. These spatial objects can also be considered independently, e.g., percentage of built-up environment or amount of open space.
- *Socioeconomic environment* is an umbrella term, characterized by variables describing various aspects of society: population size, income level, ethnic distribution, age, and education level. These variables are usually available on a certain aggregation level, such as neighborhood, postal district, census block, or grid cell. These aggregates are based on individual data points, located in the built environment.
- *Points of interest* provide functions relevant to the public. These points are distributed partly by market forces and partly by urban policy through zoning, land-use plans, and regulation. As “point of interest” implies, location is simplified to a single point. For instance, the city center is defined as a single point within the central business district, or a hospital to one center point. The simplification of larger spatial objects to single points can sometimes be useful, e.g., when counting the number of supermarkets within a certain distance. However, in other cases, simplification may be counterproductive: for example, the reduction of a park to a single point in space. Distance to the entrance might be more relevant, or a description of having a park view from a dwelling is helpful.
- *Accessibility* is the product of interest points and transport network. Generally speaking, accessibility is a special measurement used by people or firms to define spatial separation (Hansen 1959). As such, it contains a transport component and a land-use component.

In spatial analysis, several methods measure attributes in each of these categories, including measurement of the distance between analysis points and one or more variables describing location. Distance can be calculated through Euclidean distance, network distance, travel time, travel cost, or other generalized cost functions. A second method is using a boundary from an administrative data set and aggregate variables in this region. A spatial buffer is a special kind of boundary, where a circular buffer is drawn around the point of analysis. This buffer then aggregates variables describing location. A third method involves estimating a kernel density, or calculating the density for a predefined measurement area (square kilometer, hectare, spatial boundary). Finally, one can calculate a ratio—an aggregation where a variable subset is divided by the total set within a spatial reference area.

With this classification—and by highlighting different spatial analysis methods—we show how to measure urban environment in different ways. For instance, population density can also be measured in terms of dwelling density. Characterizing urban form more broadly, one can move away from arbitrary terms such as “distance to central business district,” a term valid only in monocentric environments. Another way to characterize urban centers is to calculate population and firm density, or job and retail accessibility. Finally, different stakeholders are interested in different aspects of residential location models. An architect, urban

planner, or real estate developer will be more interested in estimating demand for a certain building types, while policymakers are interested in advantages of certain amenities for different households.

Ultimately, location attributes depend on choice set alternative resolution (e.g., zone, neighborhood, unit) and the disaggregation level data available. For instance, attributes describing zones do not provide details about local neighborhood characteristics.

The spatial extent to which attributes are perceived by different decision makers can differ. For instance, the proximity of coffee shops is only relevant to a decision maker when they are located within walking distance (e.g., 500 meters or 1000 meters). [Guo and Bhat \(2007\)](#) consider different spatial boundaries for several spatial attributes. They differentiate between fixed neighborhood boundaries, based on census tracts and blocks, against other possible neighborhood extents by using Euclidean distances and network distances. The results of their study show that different types of households consider different spatial boundaries in their location choice. Households tend to cluster close to other households of similar composition, race, and income; with spatial boundaries being different for each race and household group. Their findings also show that spatial effects differ between attributes describing the built environment and the socioeconomic distribution. The socioeconomic environment is relevant within the same block or census-tract, while land-use mix, measured within a 3.2-kilometer radius, has a positive effect.

## 2.4 Comparing literature

For this review, we selected a range of papers from searches on Google Scholar and individual publisher websites; the studies include peer-reviewed journal papers as well as working papers and theses, with attention paid to studies including location variables. While many studies consider only residential location choice, others cover several choice dimensions, such as vehicle ownership and residential location choice. When describing these studies' results, we included only results relating to residential location choice. Table 1 shows the studies reviewed.

**Table 1:** Overview of reviewed studies

| Study                   | Region                        | Sample             | Level of granularity          |
|-------------------------|-------------------------------|--------------------|-------------------------------|
| Andrew & Meen 2006      | London and South East England | 999 households     | city districts                |
| Axhausen et al. 2004    | Karlsruhe                     | 349 households     | housing unit                  |
| Belart 2011             | Kanton Zurich                 | 866 households     | households                    |
| Ben-Akiva & Bowman 1998 | Boston                        | 1,259 households   | traffic analysis zone         |
| Bergle et al 2006       | Greater Zurich Area           | 1,000 households   | households                    |
| Chen et al. 2008        | Puget Sound Region, Seattle   | 1,455 households   | zone                          |
| de Palma et al. 2005    | Greater Paris Region          | 589,355 households | city districts                |
| de Palma et al. 2007    | Greater Paris Region          | 11,786 movers      | municipality                  |
| Eliasson 2010           | Stockholm region              | 3,043 records      | dwelling type                 |
| Guo & Bhat 2007         | San Francisco Bay Area        | 4,791 residences   | census block, circular buffer |
| Habib & Miller 2009     | Greater Toronto Area          | 292 households     | dwelling unit                 |
| Kim et al. 2005         | Oxfordshire, UK               | 3,072 SP records   | residential unit              |
| Lee & Waddell 2010a     | Puget Sound Region, Seattle   | 1,677 households   | building                      |
| Lee & Waddell 2010b     | Puget Sound Region, Seattle   | 4,739 households   | building                      |
| Pinjari et al 2008      | San Francisco Bay Area        | 2,793 individuals  | zones                         |
| Pinjari et al. 2011     | San Francisco Bay Area        | 5,147 households   | traffic analysis zone         |
| Srour et al. 2002       | Dallas County                 | 1,215 households   | census block, parcel          |
| Vyvere et al 1998       | Louvain-la-Neuve, Belgium     | 120 households     | residential unit              |
| Waddell 2006            | Puget Sound Region            | 2,400 households   | cell of 150m x 150m           |
| Weisbrod et al. 1980    | Minneapolis, St. Paul         | 487 households     | zones                         |
| Zhou & Kockelman 2008   | Austin Texas                  | 533 households     | single family housing         |
| Zolfaghari et al. 2012  | Greater London                | 12,836 households  | traffic analysis zone         |
| Zondag & Pieters 2005   | Netherlands                   | 12,000 households  | regions and zones             |

[Ewing and Cervero \(2010\)](#) highlight several meta-analysis drawbacks. First, by making no distinction between strong and weak studies, contamination of stronger studies with a higher sample size is a danger. Second, meta-analyses inevitably compare apples and oranges because of differences in estimation methods, dependent and independent variables, and sampling units. Publication bias is almost inevitable, as only significant variables will be reported.

Second, when constructing residential location choice models, analysts face special challenges specifying attributes for the utility function and defining alternative sets. Spatially, similar variables may be available for the choice alternative. To avoid multicollinearity, the analyst cannot enter all possible variables in the utility function, e.g., number of dwellings in a specified area and population density, or house prices and income level. Dependent on data availability and policy questions, the analyst chooses the utility function specification, reducing comparability of residential location choice models and resulting in very different models with a wide range of variables. This can be seen in Figure 1, with the number of variables per category in each study.

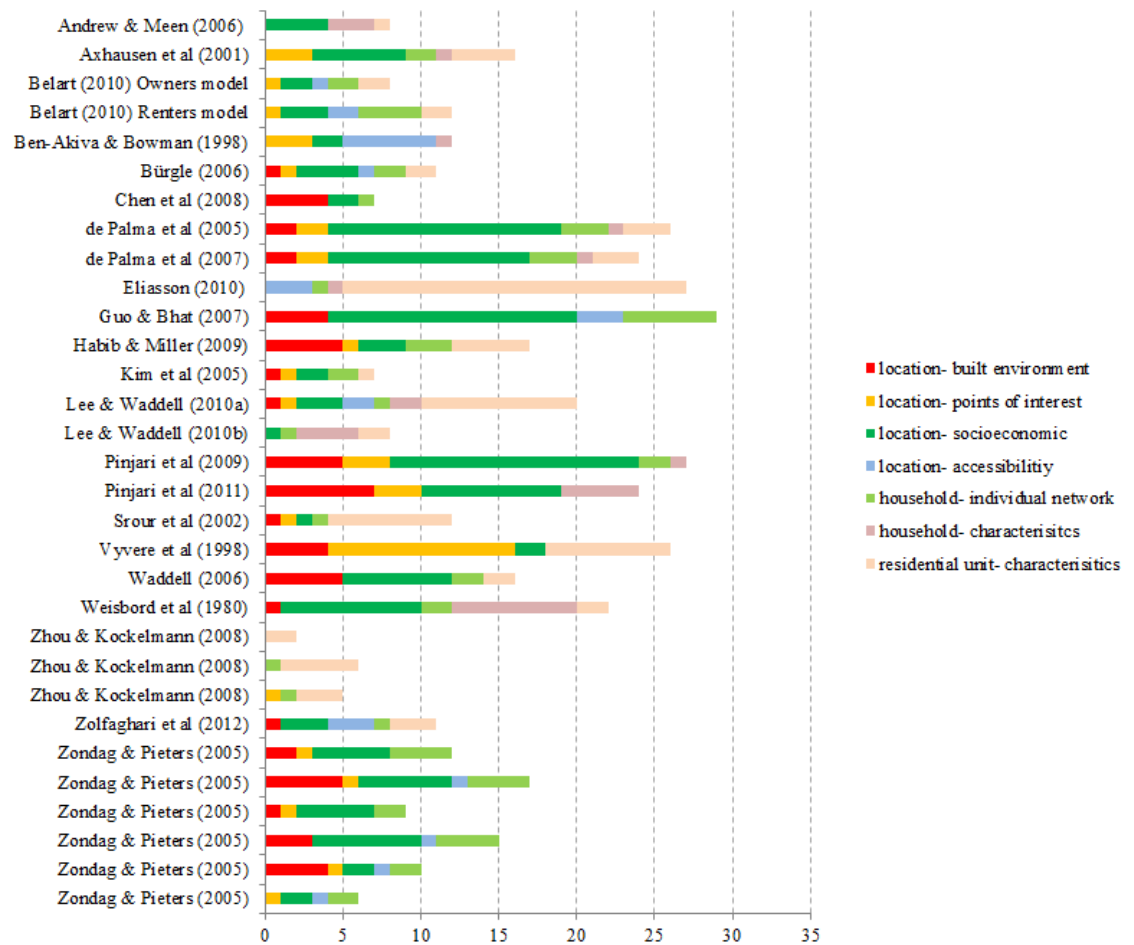


Figure 1: Variables per category in different residential location choice studies.

Third, with different scales and variable distributions, discrete choice models cannot be compared if they have not been already standardized. The same holds true when comparing different models: they cannot be directly compared unless standardized variables have been used, or elasticities have been calculated. As this is rarely seen in the literature reviewed, it becomes a limitation when comparing literature and results. Within this review, we can rely only on reported results; we must focus on significance level and positive or negative influence within the utility function, as well as the authors' primary conclusions.

Nonetheless, we believe it is essential to explore and group the array of variables in literature to define similarities and differences across different studies. We expect that peoples' behavior when facing comparable residential location choice characteristics and choices in can be modeled with a similar set of variables. With this review, we hope to provide a set of variables that can be used in residential location choice models and can capture main choice behavior elements.

### 3 Measuring residential location preferences

#### 3.1 Residential unit

As discussed in Section 2.1, the choice alternative has been modeled very differently in residential location choice models. In this section, we focus on disaggregated characteristics of residential units, which can be described as parcels, buildings, and dwellings. If dwelling is a choice alternative, it is possible to interact household attributes with residential units attributes, e.g., household size in relation to apartment size. Lee and Waddell (2010a) conclude that these dwelling characteristics tend to dominate accessibility indicators. Our own study in the Zurich area also showed that including dwelling characteristics strongly enhanced the models (Schirmer *et al.* 2013).

##### 3.1.1 Costs, price, and value

Because price is such an important determinant in choosing a residence, sales price or rental costs are often included in the model specification. Price captures various location characteristics as seen in hedonic regression models (i.e., models, like Löchl and Axhausen 2010), but any models that include property price mention its significant negative impact. While some studies implement price as untransformed value (Andrew and Meen 2006; Kim *et al.* 2005; Vyvere *et al.* 1998; Zolfaghari *et al.* 2012), others include a logarithmic price transformation (de Palma *et al.* 2005, 2007; Habib and Miller 2009; Lee and Waddell 2010a,b). Several studies interact price with household income (Habib and Miller 2009; Zolfaghari *et al.* 2012). Zolfaghari *et al.* (2012) see a positive price impact when constraining the choice set based on price, indicating that households choose the best available alternative within a certain range. Ratio of price to household income is also significant when it is the only price-related variable in the model specification (Belart 2011; Bürgle 2006; Lee and Waddell 2010a; Waddell 2006; Weisbrod *et al.* 1980; Zhou and Kockelman 2008). Walker and Li (2007) observe that price sensitivity decreases with rising income; integration of the ratio and, ultimately, the logarithmic transformation seems reasonable. Alternatively, interaction between price and income groups can further improve model estimates. Although Axhausen *et al.* (2004) state that owners and tenants of shared accommodation are more willing to pay for their location than others, a further differentiation of these groups is recommended.

Srouf *et al.* (2002) and Waddell (2006) also explore the effect of improvement on residence value. Both studies report that this variable is valued positively by all household types tested; this is not surprising, since a highly valued building should ensure a high quality standard. Curiously enough, Weisbrod *et al.* (1980) found no significant impact from this variable. One important issue with inclusion of price in residential models is price endogeneity: the unit's price is correlated with the model's error term, possibly because variables correlated with price are omitted. Guevara (2005) addresses this issue and introduces a two-step estimation method to overcome endogeneity.

##### 3.1.2 Unit size

Few studies model actual dwelling as choice alternative; we include other representations of unit size. When given, residential unit size can be included as an absolute value, or in interaction with household size, to capture space per person (Belart 2011; Bürgle 2006). Zhou and Kockelman (2008) model the location of single-family house owners with the lot as location and include lot size, as well as its ratio to household size, as an explanatory variable. Axhausen *et al.* (2004) subtract the observed mean space per person in a zone from the individual space per person to capture regional differences. These approaches yield similar results: households seek more space per person when relocating.

In addition to dwelling size, the number of rooms and bedrooms has been considered in several studies. Habib and Miller (2009) observe a positive impact for more rooms in their reference dependent model. Eliasson (2010) presents an approach similar to a two-step regression to measure the number of rooms per floor space. This formulation allows implementation of both variables and avoids multicollinearity between floor space and number of bedrooms. Their model estimates show that single-person households prefer a lower number of rooms with more floor space per room, while all non-single households favor additional rooms instead—the privacy element. The integration of room count into the residential choice model can thus enhance the model's explanatory power, but it may increase demands for corrections of correlation with floor space. Similar to the floor space variable, the number of rooms might then be integrated in relation to household size, although no study tried this approach.

### 3.1.3 Housing type

Various studies have documented preferences for a specific housing type. However, due to differences in cities' historic heritages and different structures in residential real estate markets (i.e., differences between public and private housing, tenant protection, and mortgage types), different preferences are to be expected. [Vyvere et al. \(1998\)](#) find a general preference for houses in Belgium; [Habib and Miller \(2009\)](#) find a negative preference for attached houses in Toronto. [Lee and Waddell \(2010a,b\)](#) differentiate between household types and find that single-person households and renters prefer multi-family houses; households with children favor single-family buildings. [Axhausen et al. \(2004\)](#) observe that households prefer a type of housing similar to their previous home and propose including previous location type as a variable. Eventually, this can capture self-selection effects; households with different lifestyles prefer a certain housing type. On the other hand, this ignores changes in household composition that might be a reason to relocate. In conclusion, the type of housing is relevant to residential location choice and differs per household type: apparently families favor single-family (detached) houses and single-person households are more attracted by multi-family houses.

### 3.1.4 Other features

[Vyvere et al. \(1998\)](#) find the number of garages important to households who own cars when relocating. The same study indicates that house owners generally dislike buildings constructed before 1960. Historic buildings' value has not been specifically researched, although [Srouf et al. \(2002\)](#) indicate that zones with an older average building age are preferred.

## 3.2 Location attributes

### 3.2.1 Built environment

As most location choice models are based on census data, using built space variables (mainly derived from cadastral information and volumetric appearance) is uncommon. In literature, we can differentiate variables for land-use mix, open space, structural densities, built densities, network buffers, and settlement areas.

#### *Built density*

[Waddell \(2006\)](#) finds that dwelling density has a negative impact on residential utility for all family households in the Puget Sound Region. Dwelling unit density as logarithmic expression, however, has a significant and positive influence for all household types tested.

Similar results are found when using population density as a variable instead of dwelling density (Section 3.2.3). Population density is positive for young households; they prefer areas with a high residential population density. Buildings and dwellings provide space to allocate population and thus can be expected to be a proxy for population density.

#### *Structural density*

[Pinjari et al. \(2009\)](#) include structural variables based on building shape and urban form and introduce length of networks and number of blocks per square mile. All households prefer a bike lane option; more blocks are negative for households with high incomes. This is the only study to include such variables; they offer an intriguing approach to geometric information and urban characteristics and should be further researched. [link to Clifton and WSTLUR-review](#)

#### *Transport networks*

Networks of streets and public transportation are the links that define a location's accessibility (see also Section 3.2.4). Although basically positive, networks are also a source of emissions, i.e., noise and bad air quality along main roads. [Bürgele \(2006\)](#) finds that the proximity to major roads or railways has a negative effect on residential utility in the Zurich area. [Vyvere et al. \(1998\)](#) also observe this in their stated preference survey in Louvain-la-Neuve (Belgium). Strangely enough, [Waddell \(2006\)](#) documented a positive influence in Seattle and [de Palma et al. \(2005\)](#) in Paris, reporting a preference for arterial and highway proximity. [De Palma et al. \(2005\)](#) investigated an aggregated zonal level, and neither study included accessibility calculations. The "distance to highway" variable in their study thus owed more to accessibility effects than emissions.

### *Open space*

Several authors report on the use of open space and green area as variables in residential location choice, although it is unclear whether they mean recreational areas or un-built areas. [Habib and Miller \(2009\)](#) find that green area, measured as a percentage of a neighborhood, is positive for households in Toronto. [Chen et al. \(2008\)](#) reported similar results in Seattle and—for open space measured as an absolute value within a zone—[Zondag and Pieters \(2005\)](#) in the Netherlands, with both studies interacting the variable with various household types. The latter study also finds water surface in a zone positive, measured as an absolute value, on residential utility for employed persons, explaining the architectural attraction to waterfronts in the Netherlands' residential housing market. In an anomaly, [Guo and Bhat \(2007\)](#) find a negative impact of open space variables for couple-only households in San Francisco, but these results apply only to a single-household type in one study. We expect to see a preference for proximity to open space, independent of life cycles and household types.

### *Land use*

Land-use mix has been explored in different residential location models; however, the exact specification is not detailed in most studies. A common observation is the negative impact of industrial land use near a residential location ([Habib and Miller 2009](#); [Weisbrod et al. 1980](#)). Mixed land use often differentiates only between residential and commercial (office and retail) and does not include industrial land use. [Waddell \(2006\)](#) includes an indicator for mixed land use in his study on Seattle. His models show a positive impact for young households. [Guo and Bhat \(2007\)](#) also find a positive estimate for households without a car. Additionally, they report a negative preference for the percentage of residential land use in proximity to the residential location for several household types, which indicates a preference for mixed land use. Meanwhile [Pinjari et al. \(2009\)](#) report a different effect for the same study area, and observe that homogeneous regions are favored when using commercial fraction and land-use mix as variables in their residential location choice models. They do not differentiate between household types in their study, which leads us to expect that mixed land use is valued by households favoring urban areas, young households, and households not owning a car; other household types seem to prefer a more homogeneous neighborhood. A further differentiation between households with children has not yet been documented. In literature on residential location choice, interaction between built environment and choice behavior is still limited, at least compared to mode choice ([Cervero and Kockelman 1997](#)) and car ownership ([Chu 2002](#)).

### 3.2.2 Points of interest

Points of interest (POIs) are attractive locations for the public that can be created by urban planners and policy-makers or generated by a market reaction; they can be located in buildings or outside, and are commonly abstracted as points in space. Examples are educational facilities like schools or recreational facilities like sports fields. Certain types of POIs, i.e., administrative function centers or retail complexes, are indicators of center structures as well.

POIs reported in residential location choice literature can be generally grouped into categories: education, transportation, leisure, retail and service, or urban center. Variables on POIs are described in several ways in literature: as distance from a location in the form of network distances, Euclidean distance, travel costs, or the number of POIs within a certain predefined boundary.

### *Education*

One strong common denominator for all various study areas stood out: residents want proximity to educational facilities. [Pinjari et al. \(2009\)](#) mention density of schools in a zone as universally positive, although not highly significant. [Axhausen et al. \(2004\)](#) and [Vyvere et al. \(1998\)](#) observe a similar effect, reporting a negative reaction to longer distance from schools. None of the studies reviewed include other educational facilities, such as tuition centers, or a differentiation of catchment areas per school type.

### *Service and retail*

[Zondag and Pieters \(2005\)](#) explored service functions; they find service density in a zone increases residential utility for all household types tested. Only non-single households with employees are unaffected; however, this group reacts to accessibility for all travel modes, assumed to be partly correlated. [Lee and Waddell \(2010a\)](#) use the neighborhood jobs logarithm as a variable in their model and this has a positive effect, while [Guo and Bhat \(2007\)](#) find that lower-income households and single households are more



likely to be located near employment centers. We expect that service density and retail density have a common positive influence on residential choice.

Influence of proximity to retail facilities has been explored only within stated preference surveys in the literature we reviewed. All studies find them to have a positive effect on residential utility. [Vyvere et al. \(1998\)](#) show that the distance to grocery shops is positive when less than 500 meters and distance to a shopping center is appreciated when it is less than 5 kilometers. [Kim et al. \(2005\)](#) find the travel cost factor (which we expect to be correlated to distance) increases the probability of moving and reduces the utility of a residential location utility.

#### *Recreation and sport*

[Pinjari et al. \(2009, 2011\)](#) report on density of sport and recreational facilities variables for residential location choice and find the number of physical recreation centers significantly enhances location utility. The number of natural recreational centers has only a minor influence for households with bicycles. No other study we reviewed included measurements on sport and recreational facilities, but based on this report and our own observations, we expect proximity to sport and recreation facilities to be generally positive, as long as noise is not a problem.

#### *Transportation facilities*

[De Palma et al. \(2005\)](#) report that proximity to subways in Paris is valued, while households dislike proximity to railway stations; this is related to their noisy surroundings and the multiple retail services that group around stations. [Vyvere et al. \(1998\)](#) report a positive effect of proximity to bus stations in their stated preference survey in Louvain-la-Neuve. [Habib and Miller \(2009\)](#) find individual car traffic and proximity to highway exits to be positive. Apparently, proximity to local public transport stops (bus, tram, subway) is a boon for residents, while transport facilities for long-distance transport, such as railway stations and highway exits, are valued only by certain groups of persons dependent on them. Otherwise, noise and pollution make them undesirable.

#### *Urban characteristics and center*

Several studies include explicit variables describing urban characteristics in the residential location choice. [Andrew and Meen \(2006\)](#) find a relationship between life cycles in their study on London and show that households tend to move toward the city core when they are young and move away from the city later on. [De Palma et al. \(2005, 2007\)](#) show a significant negative value for a variable representing the center of Paris. [Kim et al. \(2005\)](#) includes a variable for city settlement in their stated preference survey of Oxfordshire. Their residential model shows a clear tendency to move out of the city as well. [Belart \(2011\)](#) and [Bürgle \(2006\)](#) define the central business district in Zurich as a spatial reference point and report a tendency for all households to move away from this spatial point. [Axhausen et al. \(2004\)](#) include the distance to Mittelzentrum and Oberzentrum. Their model shows the distance to a mayor center (Oberzentrum) to be valued by local households. Proximity to smaller center structures (Mittelzentrum) has a positive impact on residential utility [Axhausen et al. \(2004\)](#), leading to the conclusion that dense urban areas are generally disliked, while local center structures are valued within residential location choices. [Zondag and Pieters \(2005\)](#) performed a second study with a more differentiated approach, distinguishing four urban characteristics (urban centers, urban neighborhoods, local village centers, and local village green neighborhoods) and explore the impact on the residential location choice for different household types. Although several of their models report significant urban characteristics, it is not possible to distinguish general tendencies, which might rely on the household classifications. Other studies lead us to assume that urban characteristics have an effect on residential location, closely related to life cycles of households. Households tend to move away from the city core during this later phase. However, all these models lack urban characteristic variables. The extent and dimensions of the city core are not clearly defined, nor is the location of a single spatial reference point pinpointed. The definition of a point or zone as urban core thus seems less convincing as they are not reproducible or defined; this again makes studies difficult to compare. Instead, we believe that models should investigate capturing these characteristics using other spatial variables, e.g., built density, density of services, and public transport density. Also, a further differentiation of household types should be attempted, as young households tend to favor urban areas. As a side note, not only does location choice behavior differ, but also relocation probability changes over a life span and is additionally influenced by life-cycle events such as a change of marital status, a job change, or the birth of a child ([Andrew and Meen 2006](#); [Beige 2008](#); [Eluru et al. 2009](#); [Kim et al. 2005](#); [Lee and Waddell 2010a](#)).

### 3.2.3 Socioeconomic environment

Variables describing socioeconomic environment are considered in most studies and are usually available through census data and statistics of municipalities on an aggregated level, often consisting of administrative boundaries. Attributes used to characterize locations are aggregated household statistics (size, age, income, origin, children, workers) and employment rates. This group of variables represents the non-fixed urban landscape configuration, which is constantly in flux. Land-use simulation models focus on simulating these dynamics and use data models that can capture these measurements (Waddell *et al.* 2003). In their analysis of neighborhood perception, Guo and Bhat (2007) explain that these attributes are often used in interaction with household characteristics to capture similarities in neighborhoods.

#### *Population density*

One attribute used in most residential choice models is population density. De Palma *et al.* (2005, 2007) utilize three different formulations in their model specification: absolute density, log of density, and change of population density. Two studies produce the same result: absolute population density is negative when all residential alternatives are considered.

Other studies also find population density has a negative effect on residential choice (Kim *et al.* 2005; Lee and Waddell 2010a; Weisbrod *et al.* 1980). Zondag and Pieters (2005) differentiate between six household types in the Netherlands; all household types—except employed, one-person household—generally react negatively to population density. Employed, one-person households are attracted by densely populated areas. Guo and Bhat (2007) find a negative impact on residential utility for small families and high-income households, but a positive impact for all other households, including young households and single-person households, which supports the previous observations. Bürgle (2006) shows that young households are strongly attracted to high-density population areas in Zurich. As mentioned previously, this type of household generally prefers urban areas, mixed land use and high-density dwelling units. Pinjari *et al.* (2009, 2011) find a general positive influence using a logarithmic formulation of population density, but a negative impact for senior households and households with children in the untransformed formulation. Zolfaghari *et al.* (2012) report a positive influence of population density for all households. In conclusion, population density is generally viewed negatively by households, especially by families; however, certain types of households, namely young households and single-person households, value population density because they prefer urban areas.

#### *Household types*

Several studies include variables describing the proximity of similar households. This similarity can, for instance, be expressed by the percentage of households with similar ethnic backgrounds, the same household composition, or the same income level. De Palma *et al.* (2005, 2007) show that most households in Paris tend to search nearby areas for households of same size; only two-person households do not do this, albeit with low significance. Guo and Bhat (2007) find a positive influence and also use the difference between household size and average zonal household size as a variable—which has a negative impact—underlining the previous results. This approach also has been mentioned in other studies: Weisbrod *et al.* (1980), Pinjari *et al.* (2009), Pinjari *et al.* (2011), Zolfaghari *et al.* (2012); all find a negative impact when household sizes differ. Waddell (2006) introduces two variables: households of similar age and similar size. He is the only author who reports a significant effect for older households, defined as those where the head of the household is older than 40.

Lee and Waddell (2010a,b) demonstrate that families like being near other households with children; they use the percentage of similar families within 600 meters as an explanatory variable in their location choice models. Bürgle (2006) observes a similar effect when using density of children in the neighborhood as a variable.

Two studies report that young households tend to cluster in the Seattle and Paris study areas (de Palma *et al.* 2005, 2007; Lee and Waddell 2010a). The latter study additionally shows a negative influence of a high density of young households, which might represent correlation effects to the preference of young households for urban areas.

Pinjari *et al.* (2009, 2011) and Weisbrod *et al.* (1980) use the fraction of senior persons as explaining variables on residential choice of senior households and find a positive influence as well. Waddell (2006) uses density of same-age households instead of segments. His models show only significant positive influence for young single-person households. Similar to his observations on household size, we expect that the integration of a density—instead of a percentage—will capture the negative effects of population density.

De Palma *et al.* (2005, 2007) classified households according to the number of employed individuals and assumed that households segregate based on this characteristic. In a first study in 2005, they found a significant positive impact, an effect not observed again in a later study. These reports confirm that households prefer to locate around similar households (age, size, and presence of children). This fact is thus an essential variable for modeling residential location choice and needs to be integrated as an interaction term. The share of same-type households tends to be valued positively; difference between average size and a household's own size will probably be negative.

A question not explored in the literature—and still to be researched—is the following: to what point are socioeconomic variables causal explanatory variables at all? Is it that similar households tend to group or that they have similar housing, spatial, and locational preferences and tend to segregate?

#### *Household origin and race*

Various studies have observed segregation effects when defining ethnic groups or household place of origin as variables in residential location choice models. De Palma *et al.* (2005, 2007) differentiated between foreign-headed and French-headed households in the Paris region and demonstrated that households with a foreign head tend to group while households with a French head perceive the vicinity of foreign households negatively. Waddell (2006) finds the same effect for minority households and white households. While minorities favor minority neighborhoods, white households will avoid minority neighborhoods. Pinjari *et al.* (2009, 2011) and Guo and Bhat (2007) differentiate between various ethnic groups within their models and observe that all households tend to locate near households of the same group. We expect that households tend to group according to their ethnic group and origin and that this is an important variable in residential location models.

#### *Household income*

Several studies used income groups as explaining variables in their residential location choice models; this has proved significant. Weisbrod *et al.* (1980) observe a preference; high-income households tend to locate around other high-income households, while de Palma *et al.* (2005, 2007) observe a grouping of low-income households in their models. Zondag and Pieters (2005) also show that middle- or high-income households prefer neighborhoods in the same income group, but find that low-income households want to be near middle-income households. Their models also differentiate between household types in the Netherlands and show that retired seniors prefer middle-income neighborhoods and dislike high-income neighborhoods. Waddell (2006) includes interaction terms for low-, middle-, and high-income distribution according to household size and age. His models show a positive influence for all income groups, but are of significance only for some low- and high-income household types. Further studies by Guo and Bhat (2007) and Pinjari *et al.* (2009, 2011) include the difference between individual household income and average zonal income and find that this difference has a highly significant negative value. All these reports show that households have a tendency to relocate around household groups with similar incomes.

#### *Housing costs*

Spatially aggregated housing costs are closely related to sales prices or rental costs of individual units described in the previous section. Studies observe that these housing costs have a significant negative influence on location utility (Axhausen *et al.* 2004; Pinjari *et al.* 2009, 2011; Zondag and Pieters 2005). Srour *et al.* (2002) include the average lot value in a zone (land price) as an explaining variable and find a negative impact on the residential utility of a location. Guo and Bhat (2007) include the ratio of income to average housing price in a zone within their model, instead of the absolute formulation, which is negative. As these studies use a zone instead of an individual dwelling unit as choice alternative, this is not surprising. The average price of residential dwelling within a zone is not expected to have explanatory power unless it is used in relation to the individual costs of a residential unit.

#### *Employment*

Different approaches implement employment distribution into residential location choice models. Two studies integrate the unemployment rate in their models and show a negative effect on the utility of a residential location for London and Toronto (Andrew and Meen 2006; Habib and Miller 2009). Other studies use jobs density as an explaining variable; Zondag and Pieters (2005) find this significantly positive, but only for non-single, employed households of middle or high income, while Srour *et al.* (2002) ob-

serve a general attractive potential for high job density locations in Dallas. [Pinjari et al. \(2009, 2011\)](#) test this variable for different household types, but find no significant effects. [De Palma et al. \(2005, 2007\)](#) do not find any significant and consistent preferences, as this variable works both positively and negatively in their two studies. One explanation is that density of jobs has no consistent effect on residential location choice, but that households avoid locations with a high unemployment rate.

### *School quality*

[Kim et al. \(2005\)](#) and [Zhou and Kockelman \(2008\)](#) find school quality to have a positive influence; as it is a non-interaction variable, this holds for all households. [Chen et al. \(2008\)](#) include school quality as interaction variable for households with children and without children and find it to be slightly stronger for households with children, but positive and significant for both. [Andrew and Meen \(2006\)](#) include the GCSE levels and A-level scores (percentage of children obtaining 5+ GCSE grades; percentage of children obtaining 3+ A/AS levels) in their model and find a positive impact of the first variable and a negative impact for the latter. They conclude that households would send their children to other schools (in an area with poor local neighborhood schools) instead of moving, but also see a correlation to a deprivation index in these situations. This index combines a “number of indicators covering income, employment, health deprivation and disability, education skills and training, housing quality, and geographical access to services into a single score for each area” ([Andrew and Meen 2006](#)). [Weisbrod et al. \(1980\)](#) include the ratio of teachers to students as a quality indicator. This has a positive influence on the residential utility of a location for households with children.

We assume that school quality is most important for households with children or who plan to have children, even if some studies find the variable positive for almost all households. Further research is needed on the expected correlation with other sociodemographic variables and how/if school quality can be measured in a land-use simulation.

### *Other variables*

Some additional variables are used in only a few studies, so that no assumptions can be made on their influence and use in residential choice models. [Weisbrod et al. \(1980\)](#) include property tax rate per household as a variable in the model specification, which does have a slightly significant negative impact. [Bürgle \(2006\)](#) also finds a negative impact for the study area of the greater Zurich region, though with a high significance, which reflects the local divergence of tax rates and the strong competition between municipalities in Switzerland.

Two studies mention the use of rental vacancy within residential choice models. While [Zondag and Pieters \(2005\)](#) report the positive influence of vacant housing in neighborhoods in the Netherlands on all household types tested, [Bürgle \(2006\)](#) finds a negative impact using the municipal rental vacancy. [Andrew and Meen \(2006\)](#) and [Weisbrod et al. \(1980\)](#) observe that crime rate has a significant negative impact on residential location choice. In the reviewed studies, [Guo and Bhat \(2007\)](#) are the only authors to include the share of owner-occupied housing for residential location choice models and find this variable to be positive for owners' location utility.

## 3.2.4 Access and accessibility

### *Access*

Access to work, or commuting time, is often included when the workplace of the decision maker is known and the commuting time can be computed for all alternatives. Generally, it is obtained from a transport model and needs to be calculated for every choice set alternative. Commuting time is usually a negative influence on the residential utility of a location ([Axhausen et al. 2004](#); [Guo and Bhat 2007](#); [Habib and Miller 2009](#); [Lee and Waddell 2010a](#); [Zhou and Kockelman 2008](#); [Zolfaghari et al. 2012](#)). If commuting time is not available, studies use commuting distance as network-based distance or Euclidean distance instead ([Belart 2011](#); [Bürgle 2006](#); [Chen et al. 2008](#); [Srouf et al. 2002](#)), which has also always been negative but has a lower significance. Few studies differentiate between commuting time by car or public transportation, or factor in commuting costs ([Kim et al. 2005](#); [Pinjari et al. 2009, 2011](#)), although all have a negative impact on the residential utility when significant. Public transit commuting times are more important than commute times by private transport ([de Palma et al. 2007](#)). [Belart \(2011\)](#) and [Pinjari et al. \(2009\)](#) include average commuting time/commuting distance of all workers in a household, which further enhances model performance. A weighted approach based on the percentage of employees in a household has not seemed to enhance the model ([Belart 2011](#)). To capture the reduced effect of long-distance variations, [Bürgle \(2006\)](#) uses a formulation within an exponential function that is also implemented by [Belart \(2011\)](#).

### Accessibility

Accessibility is a measure of spatial distribution of activities around a point, adjusted for people and firms' ability and desire to overcome this spatial separation or "the potential of opportunities of interaction" (Hansen 1959). Accessibility calculations are applied to represent local and regional differences within the urban landscape. Five main types of accessibility measures appear in literature (Bhat *et al.* 2002): spatial separation, cumulative opportunities, gravity measures, utility measures, and time/space measures.

Guo and Bhat (2007) and Zolfaghari *et al.* (2012) include employment accessibility in their location choice models. A low accessibility to employment is a deterrent for by households in San Francisco and London. The latter study also shows that single households are more likely to reside near employment centers; this preference is in line with the preference for a higher population density and ultimately reflects a preference for urban areas.

Srouf *et al.* (2002) calculate different types of accessibility measures. They conclude that the cumulative opportunity accessibility index—in which accessibility is determined by opportunities available within 30 minutes of travel time—provides the best results for location choice, based on statistical significance and easy behavioral interpretation. In their study, they find a preference for accessibility to employment. Belart (2011) differentiates between accessibility by transit and car. Employment accessibility by the transit network has a positive significant effect for the location choice of households having no car. Employment accessibility by private transport, however, has a negative influence on residential utility for households owning a car. The latter can indicate negative effects associated with accessibility by car, such as noise and particle emissions. In the same region, Bürgle (2006) notes that households without a car prefer to live in high population accessibility locations.

Zondag and Pieters (2005) calculate a log-sum measure for all travel purposes and a log-sum per travel purpose (work, education, other). In their move-stay model, they find that fewer households are willing to move away from easily accessible locations than from less accessible locations. In their location choice model, households with medium to high income do not see accessibility for all travel purposes as an asset. On the other hand, employed single households and non-single households older than 65 view accessibility for purposes other than work and education positively.

Chen *et al.* (2008) measure accessibility to open space and find it positive for households with children. However, accessibility is perceived as less important than other factors, such as income and other household characteristics

Ben-Akiva and Bowman (1998) consider a log-sum measure based on an individual's activity schedule and do not find that this measure improves their model, like Eliasson (2010), who considers direct utility from households' optimal activity patterns. As well as accessibility to employment, he considers access to services and shops and decides that these accessibility measures should be included in location choice models.

Zondag and Pieters (2005) state that the effect of accessibility is marginal in comparison to attributes describing the residential unit. Lee and Waddell (2010a) include more detailed accessibility measures. They include a time/space prism measure to calculate shop accessibility and a highly disaggregated commute travel time measure. These two measures are highly significant and have a relatively large influence. They argue that these factors are being considered by decision makers and should be included in residential location models.

### 3.3 The role of previous location and social networks

#### Previous location

is this a value of the residential unit or of a location? Daily activities make up a household's individual "mind map." This map can strongly influence location choice. A limited number of number of studies includes households' previous location. Axhausen *et al.* (2004) and de Palma *et al.* (2005, 2007) find that households tend to stay close to their previous location or remain in the same district. Zondag and Pieters (2005) combine the Euclidean distance between previous and current location and find it to be the most dominant variable in their model for various household types tested. Proximity to the previous location is preferred for all households, except for households over age 65.

Habib and Miller (2009) introduce a reference dependent model and assume that new location is dependent on the current location. Households consider a gain in bedrooms positive and are concerned about a loss; they are positive to gains in open area and have a loss aversion attitude to losses in open areas.

### *Social network*

Gordan (1992) mentions the desire to maintain social networks, which is relevant for residential relocation. Vyvere *et al.* (1998) introduce distance to social contacts in their model specification; moving farther away from social contacts is perceived negative. Belart (2011) further explores this effect and observes a preference for proximity to social contacts when using distance to social contacts weighted with by number of meetings per month. As shown in the complementary modeling study of this paper (Schirmer *et al.* 2013), this variable correlates with previous location; it is unclear whether location is influenced by social contacts or whether the social network is formed by the location.

## 4 Summary and discussion

### 4.1 Summary

This study systematically compared “location” in residential location choice models. A wide range of variables has been included in the different residential location models reviewed. Geospatial data available to researchers has increased tremendously over the last several decades, opening up opportunities to define residential location in multiple ways. Table 2 provides an overview of the different studies’ main findings.

Recent studies consider the residential unit or building as choice alternatives. This offers the possibility to include the units’ and buildings’ characteristics and, according to these studies, greatly improves model performance. Also, this offers the opportunity to include a detailed description of the buildings’ built environment and the calculation of detailed travel times and accessibility measures.

The role of built environment has been explored in a limited number of studies. Households prefer different land-use mixes in different life cycles. Density of the road and rail network captures emissions as noise and pollution and is perceived as negative for all household types.

Variables describing the socioeconomic environment are still important in residential location choice models. Household clustering can be observed in a wide range of studies, with households preferring to live near other households with a similar composition or income. Single households and young persons gravitate toward population-dense areas. No household types could tolerate high unemployment or crime rates.

Points of interest are included in several studies; similar behavioral preferences can be observed across different study regions. Education, service, retail, and local transport facilities are valued by all household types; both proximity to and the density of these amenities is appreciated. Preferences for stations and highway ramps vary, depending on household car availability and the noise level caused by these transportation facilities. Young households tend to favor proximity to the central business district, while other household types would rather locate away from the urban core.

Longer commuting distance is negative. Preferences for accessibility to employment differ between different studies. This can be due to different calculation methods; some studies calculate accessibility to employment opportunities by car, while other studies differentiate by travel mode and household car ownership. This can indicate that decision makers either search for a job close to their residence or, when relocating, take into account the location of their job and consider accessibility to jobs less important. The value of accessibility to shops is measured across different studies and is considered positive by decision makers.

Household location choice should be seen within a relocation context: the limited number of studies that include distance to previous residence or distance to social networks find that households want to stay close to their previous location.

### 4.2 Discussion

To an outsider, results found in residential location choice studies can seem obvious. This should be viewed positively for residential location choice studies, which can quantify attributes of importance to location choice. Despite criticism of the ability of discrete choice models to quantify an important choice not frequently made by households (Timmermans 2006), different residential location studies reveal similar preference structures.

The proposed classification of location into built environment, socioeconomic environment, points of interest, and accessibility can be used in several ways. First, it can serve as a classification to structure data available to researchers. Second, it can support discussions and decisions about variables to include in location choice models by showing which trade-offs exist when specifying

---

different models. Further quantitative research is required to determine which combinations of variables can be used to compare different model specifications.

Table 2: Main findings in reviewed studies

| Group / Subgroup                               | Variable   | Function | Radius | Sign |
|--|--|----------|--------|------|
| <b>Residential unit</b>                        |  |          |        |      |
| size   | floorspace per person * (non single hh & single hh)                            | value    |        | +    |
|  | rooms per person * non single hh   | value    |        | +    |
|  | room per person * single hh  | value    |        | -    |
| costs  | (price/income) * hh is owner   | value    |        | -    |
|  | (price/income) * hh is renter  | value    |        | -    |
| house type                                     | detached single family * hh with children                                      | value    |        | +    |
|  | multifamily housing * single person hh   | value    |        | +    |
| age  | building age   | value    |        | -    |
| <b>Built environment</b>                       |  |          |        |      |
| built density                                  | log (density of dwelling units in 600m)  | density  | 500    | +    |
|  | log(density of dwelling units) * hh has children                               | density  | 500    | -    |
|  | log(density of dwelling units) * (young hh or single hh)                       | density  | 500    | +    |
| open space                                     | share of open space or unbuilt space   | ratio    | 500    | +    |
|  | share of water   | ratio    | 500    | +    |
| land use                                       | share of commercial land use * (young hh or single hh or hh without car)       | ratio    | 500    | +    |
|  | share of commercial land use * hh has children                                 | ratio    | 500    | -    |
|  | share of industrial land   | ratio    | 500    | -    |
| network/noise                                  | buffer to arterials and railways (noise)                                       | boolean  |        | -    |
|  | structural density (amount of network links per sqkm)                          |          | 1000   |      |
| <b>Socio-economic environment</b>              |  |          |        |      |
| population density                             | population density   | density  | 1000   | -    |
|  | populations density * (young hh or single hh)                                  | density  | 1000   | +    |
| household type                                 | share of hh with same size   | ratio    | 1000   | +    |
|  | share of hh with same age  | ratio    | 1000   | +    |
|  | share of hh with children * hh has children                                    | ratio    | 1000   | +    |
|  | share of hh with same income cat   | ratio    | 1000   | +    |
|  | share of hh of same ethnic group/origin  | ratio    | 1000   | +    |
| employment                                     | unemployment rate  | value    |        | -    |
| crime  | crime rate   | value    |        | -    |
| <b>Points of interest</b>                      |  |          |        |      |
| education                                      | distance to school   | distance |        | -    |
| urban character                                | distance to urban center (CBD) * (nonsingle hh or hh has children)             | distance |        | +    |
|  | distance to urban center (CBD) * (young hh or single hh)                       | distance |        | -    |
|  | distance to local center   | distance |        | -    |
| service and retail                             | density of retail  | density  | 500    | +    |
|  | density of service   | density  | 500    | +    |
| sport and recreation                           | density of sport activity centers  | density  | 500    | +    |
|  | density of natural recreation centers  | density  | 500    | +    |
| transport                                      | distance to local transport  | distance |        | -    |
|  | distance to station * no car owner   | distance |        | -    |
|  | Very close proximity to station (noisy environment)                            | value    |        | -    |
|  | distance to highway exit * car owner   | distance |        | -    |
| <b>Access &amp; Accessibility</b>              |  |          |        |      |
| commuting time                                 | commuting time/distance/costs by car * car available                           | distance |        | -    |
|  | commuting time/distance/costs by pt * no car available                         |          |        | -    |
| socio-demographic                              | accessibility (cumulative opportunities) of jobs * (no car or young or single) | acc      |        | +    |
|  | accessibility (cumulative opportunities) of jobs * car                         | acc      |        | -    |
| POIs   | accessibility (cumulative opportunities) of shops                              | acc      |        | +    |
| <b>Previous location &amp; Social networks</b> |  |          |        |      |
| individual distances                           | distance to previous location  | distance |        | -    |
|  | distance to social network   |          |        |      |



#### 4.2.1 Inclusion of built environment

Socioeconomic variables are often used in location choice models and offer valuable insights into residential preferences. However, for architects and urban planners and the evaluation of their planning proposals, it is necessary to include a description of the built environment. While most studies do not include such a direct description, conclusions can be drawn from other preferences. Population density and dwelling density reveal similar preference structures for different household types. This can provide urban planners with guidelines concerning density and housing type.

Similar household types tend to cluster in close proximity (less than 1000 meters); similarity is expressed by households of similar composition, ethnic background, or income level. The spatial extent of these clusters differ by household type. This preference for clustering in a smaller area provides opportunities for a mix of housing types in larger areas, such as neighborhoods. Also, proximity to amenities is preferred, which is confirmed by the preference for mixed-use areas. Together, these results confirm that residential location choice behavior is more complex than a single function describing travel time to the central business district.

Clifton *et al.* (2008) provide an overview of urban form quantitative representation in different disciplines, which might guide the researcher in this point. We would also expect that the availability of the 3D models, i.e., the volumetric appearance, would allow for representation of spatial configurations and spatial qualities. Visibility, sun-prone and shadow-rich areas, and street sections are examples of such qualities, which have not been explored in the reviewed studies.

#### 4.2.2 Transport land-use models

The research in this paper is driven partly by possible applications of location choice models. One possible application is using choice models in transport land-use simulations. State-of-the-art transport land-use simulation models (e.g., UrbanSim, Waddell *et al.* 2003) include discrete choice models to quantify households' and individuals' decisions. UrbanSim's most detailed version, the "parcel" version, includes persons, households, jobs, residential units, buildings, and parcels in its data specification, explaining an important part of residential location choice. This review has shown that in addition points of interest (e.g., schools, school quality, retail, transport elements) and previous residential locations are attributes that should be included. Preferably, network distances and travel times to points of interest should be calculated, which makes it necessary to describe different transportation networks. In which way the location classification presented in this paper can be implemented remains open for further research.

## References

- Alonso, W. 1964. *Location and Land Use*. Cambridge: Harvard University Press.
- Anas, A. 1982. *Residential Location Markets and Urban Transportation*. New York: Academic Press.
- Andrew, M. and G. Meen. 2006. Population structure and location choice: A study of London and south east England\*. *Papers in Regional Science*, 85(3):401–419.
- Axhausen, K., D. Scott, A. König, and C. Jürgens. 2004. Locations, commitments and activity spaces. In M. Schreckenberg and R. Selten, eds., *Human Behaviour and Traffic Networks*, pp. 205–230. Berlin: Springer.
- Beige, S. 2008. *Long-term and mid-term mobility decisions during the life course*. Ph.D. thesis, ETH Zurich, Zurich.
- Belart, B. C. 2011. *Wohnstandortwahl im Grossraum Zürich*. Master's thesis, IVT, ETH Zurich, Zurich.
- 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.
- Bhat, C. R., S. L. Handy, K. Kockelman, H. S. Mahmassani, A. Gopal, I. M. Srour, and L. Weston. 2002. Development of an Urban Accessibility Index: Formulations, Aggregation, and Application. Working Paper Report 7-4938-4, University of Texas Austin, Austin.
- Bürgle, M. 2006. Residential location choice model of the Greater Zurich area. In STRC, ed., *6th Swiss Transport Research Conference*. Ascona. URL [http://www.strc.ch/conferences/2006/Buergle\\_STRC\\_2006.pdf](http://www.strc.ch/conferences/2006/Buergle_STRC_2006.pdf).
- Cervero, R. and K. Kockelman. 1997. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3):199–219.
- Chen, J., C. Chen, and H. J. P. Timmermans. 2008. Accessibility trade-offs in household residential location decisions. *Transportation Research Record*, 2077:71–79.

- Chu, Y.-L. 2002. Automobile ownership analysis using ordered probit models. *Transportation Research Record*, 1805:60–67.
- Clifton, K., R. Ewing, G. Knaap, and Y. Song. 2008. Quantitative analysis of urban form: a multidisciplinary review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 1(1):17–45.
- de Palma, A., K. Motamedi, N. Picard, and P. Waddell. 2005. A model of residential location choice with endogenous housing prices and traffic for the paris region. *European Transport\Trasporti Europei*, 31:67–82.
- 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.
- Eliasson, J. 2010. The Influence of Accessibility on Residential Location. In F. Pagliara, J. Preston, and D. Simmonds, eds., *Residential Location Choice: Models and Applications*, Advances in Spatial Science, pp. 137–164. Berlin: Springer.
- Eluru, N., I. N. Sener, C. R. Bhat, R. M. Pendyala, and K. W. Axhausen. 2009. Understanding residential mobility: A joint model of the reason for residential relocation and stay duration. *Transportation Research Record*, 2(2133):64–74.
- Ewing, R. and R. Cervero. 2010. Travel and the built environment. *Journal of the American Planning Association*, 76(3):265–294.
- Gordan, I. 1992. Modelling approaches to migration and the labour market. In A. Champion and A. Fielding, eds., *Migration processes and patterns 1: Research Progress and Prospects*. London: Belhaven.
- Guevara, C. A. 2005. *Endogeneity and Sampling of Alternatives in Spatial Choice Models*. Ph.D. thesis, Massachusetts Institute of Technology, Boston.
- Guo, J. Y. and C. R. Bhat. 2007. Operationalizing the concept of neighborhood: Application to residential location choice analysis. *Journal of Transport Geography*, 15(1):31–45.
- Habib, K. M. N. and E. J. Miller. 2009. Reference-dependent residential location choice model within a relocation context. *Transportation Research Record*, 2133:92–99.
- Hansen, W. 1959. How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, 25(2):73–76.
- Kim, J., F. Pagliara, and J. Preston. 2005. The intention to move and residential location choice behaviour. *Urban Studies*, 42(9):1621–1636.
- Krizek, K. J. and P. A. Waddell. 2003. Analysis of lifestyles choices: neighborhood type, travel patterns, and activity participation. *Transportation Research Part B: Methodological*, 1807:119–128.
- Lee, B. H. Y. and P. A. Waddell. 2010a. Reexamining the influence of work and nonwork accessibility on residential location choices with a microanalytic framework. *Environment and Planning A*, 42(4):913–930.
- Lee, B. H. Y. and P. A. Waddell. 2010b. Residential mobility and location choice: A nested logit model with sampling of alternatives. *Transportation*, 37(4):587–601.
- Löchl, M. and K. W. Axhausen. 2010. Modelling hedonic residential rents for land use and transport simulation while considering spatial effects. *Journal of Transport and Land Use*, 3(2):39–63.
- Lowry, I. S. 1964. A model of metropolis. Technical report, RAND Europe.
- McFadden, D. 1978. Modeling the choice of residential location. In A. Karlqvist, ed., *Spatial Interaction Theory and Residential Location*, pp. 75–96. Amsterdam: North-Holland.
- Müller, H.-P. 1992. *Sozialstruktur und Lebensstile- Der neue theoretische Diskurs über soziale Ungleichheit*. Frankfurt am Main: Suhrkamp.
- Pinjari, A. R., C. R. Bhat, and D. A. Hensher. 2009. Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B: Methodological*, 43(7):729–748.
- Pinjari, A. R., R. M. Pendyala, C. R. Bhat, and P. A. Waddell. 2011. Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. *Transportation*, 38(6):933–958.
- Schirmer, P., M. Van Eggermond, and K. Axhausen. 2013. Measuring location in residential location choice - an empirical study in the Canton of Zurich. Utrecht.
- Srouf, I. M., K. Kockelman, and T. P. Dunn. 2002. Accessibility Indices: Connection to Residential Land Prices and Location Choices. *Transportation Research Record*, 1805:25–34.
- Timmermans, H. J. P. 2006. The saga of integrated land use and transport modelling: How many more dreams before we wake up? In K. W. Axhausen, ed., *Moving Through Nets: The Physical and Social Dimensions of Travel*, pp. 219–248. Oxford: Elsevier.
- Train, K. E. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

- von Thunen, J. H. 1826. *Der isolierte Staat in beziehung auf Landwirtschaft und Nationalökonomie*, Gustav Fisher, Stuttgart. English edition: *The isolated state (trans: Wartenburg CM (1966), edited by Hall P)*. Oxford: Pergamon.
- Vyvere, Y., H. Oppewal, and H. Timmermans. 1998. The validity of hierarchical information integration choice experiments to model residential preference and choice. *Geographical Analysis*, 30(3):254–272.
- Waddell, P. 2006. Reconciling household residential location choices and neighborhood dynamics. *Working Paper*.
- Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. 2003. Microsimulation of urban development and location choices: Design and implementation of UrbanSim. *Networks and Spatial Economics*, 3(1):43–67. ISSN 1566-113X.
- Walker, J. L. and J. Li. 2007. Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems*, 9(1):77–101.
- Weisbrod, G., M. E. Ben-Akiva, and S. R. Lerman. 1980. Tradeoffs in residential location decisions: transportation versus other factors. *Transport Policy and Decision Making*, 1(1).
- Zhou, B. and K. Kockelman. 2008. Microsimulation of residential land development and household location choices: bidding for land in austin, texas. *Transportation Research Record*, 2077:106–112.
- Zolfaghari, A., A. Sivakumar, and J. W. Polak. 2012. Choice Set Pruning in Residential Location Choice Modelling: A Comparison of Sampling and Choice Set Generation Approaches in Greater London. *Transportation Planning and Technology*, 35(1):87–106.
- Zondag, B. and M. Pieters. 2005. Influence of accessibility on residential location choice. *Transportation Research Record*, 1902:63–70.