

The rebirth of urban subcenters: How subway expansion impacts the spatial structure and mix of amenities in European cities

EPB: Urban Analytics and City Science
2022, Vol. 49(4) 1266–1282
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23998083211056955
journals.sagepub.com/home/epb



Tabea S Sonnenschein 

Sustainable Urbanization Lab, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA; Department of Human Geography and Planning, Faculty of Geoscience, Utrecht University, Utrecht, Netherlands

Simon Scheider 

Department of Human Geography and Planning, Faculty of Geoscience, Utrecht University, Utrecht, Netherlands

Siqi Zheng 

Sustainable Urbanization Lab, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

Why do some neighborhoods thrive, and others do not? While the importance of the local amenity mix has been established as a key determinant of local livability, its link to urban transport infrastructure remains understudied, partially due to a lack of data. Using spatiotemporal social media data from Foursquare, we analyze the impact of metro stations which opened between 2014 and 2017 on the amenity mix of surrounding neighborhoods in nine European cities: Rome, Milan, Barcelona, Budapest, Warsaw, Sofia, Vienna, Helsinki, and Stuttgart. Thereby, we study three properties of the local amenity mix: its density, multifunctionality, and the heterogeneity between amenity types. For this purpose, we propose a new measurement of multifunctionality, which calculates the entropy of the locally present amenity set incorporating the degree of similarity between amenity types. For causal inference, we use Difference-in-Difference Regression based on Propensity Score Matching and Entropy Balancing. Our findings show that in most cities, subway expansion had a significant positive impact on the local amenity density and multifunctionality and that especially the social amenities—Arts & Entertainment, Restaurants and Nightlife—responded strongly. Moreover, considerable agglomeration forces seem to prevail, causing existing subcenters to benefit most from new metro stations.

Corresponding author:

Tabea Sonnenschein, Department of Human Geography and Planning, Faculty of Geoscience, Utrecht University, Heidelberglaan 8, Utrecht 3584 CS, Netherlands.
Email: t.s.sonnenschein@uu.nl

Keywords

Subway expansion, location theory, consumer amenities, multifunctionality, similarity-based diversity, spatiotemporal analysis

“When distance and convenience set in; the small, the various and the personal wither away.” (Jane Jacobs, 1961)

Introduction

Only recently consumer amenities and the cultural life they nurture have been recognized as a crucial mechanism of urban growth and human capital attraction, extending the traditional focus on the availability of jobs as the determinant of the success of cities (Clark, 2004; Florida, 2002). On the neighborhood level, the amenity mix of a neighborhood is considered a core component of its livability already since Jane Jacobs' (1961) observation that “only diverse surroundings have the practical power of inducing a natural, continuing flow of life and use” (p. 101). It is the mixture of work, cultural, residential, and commercial activity that produces vibrant environments with safety guaranteed through “Eyes on the Streets.” Since then, studies showed that the collocation of different public and private amenities in residential areas reduces the amount of crime in a neighborhood (Humphrey et al., 2020), incentivizes walking (Forsyth et al., 2008), decreases overall energy use (Zegras, 2005) and generates Marshallian agglomeration benefits for consumer amenities such as labor sharing, value chain linkages and knowledge spillovers (Duranton and Puga, 2004).

Transit-Oriented Development (TOD) projects combine this prioritization of walkable, mixed-use design with an effort to reverse the accelerating trend of motorized transport polluting and congesting our cities (Littman, 2017; UN, 2016). They aim at concentrating residential, work, and commercial spaces around a decentralized network of mass transit stations. Thereby, TOD projects tackle the spatial mismatch between job opportunities and potential workers, between quality-of-life amenities and people that demand them (Gobillon et al., 2007). To adjust to growing cities and improve public transit coverage, a core TOD policy is public transport expansion. A new transit station can have a substantial impact on neighborhoods. However, can it help developing the other element of TOD—mixed land use—by attracting jobs and consumer amenities? In other words, do transit stations and amenity centrality go hand in hand? There are a considerable number of studies showing how transport expansion impacts population growth, housing markets and job distributions as well as a large branch of literature studying the location decisions of different industry types. However, the relationship between urban transport infrastructure and neighborhood consumer amenity developments has been only sparsely addressed.

Spatial proximity or spatial accessibility has long been established as a key determinant of both human and physical interaction (Tobler, 1970). Taking away barriers to mobility should increase the number of people willing to travel the respective route and consequently, increase the number of consumer amenities that move to the origin or destination mostly due to increased market size. One of the fastest and therefore most popular modes of intracity public transport across distance is subway transit, which connects places in a sparse network structure. While cities are growing the fastest in Africa and Asia, European cities are interesting cases to study because they have a historically strong public transport culture, demonstrated by the continent having the oldest metro systems and the most subway stations per 100.000 residents on Earth (Gonzalez-Navarro and Turner, 2018). There has not been a large expansion project in any European big city during the last decade. However, together the small expansion projects happening mostly on the fringes of Europe's diverse cities are ideal cases for testing the relation between subway expansion—as part of TOD—and the local amenity mix in a cross-comparative study.

In this article, we analyze the impact of metro stations that opened between 2014 and 2017 on the amenity mix of surrounding neighborhoods in nine European cities. Rome, Milan, Barcelona, Budapest, Warsaw, Sofia, Vienna, Helsinki, and Stuttgart opened together 84 new metro stations. To measure the amenity mix, we exploit user-generated, location-based social media data. The place-based social media platform Foursquare offers high resolution, georeferenced data on consumer amenity openings differentiating by amenity type (Foursquare, 2020). This allows us to assess the mixture of amenities in a neighborhood along three dimensions: density, multifunctionality, and specialization on specific amenity types. To this end, we propose a new measurement of multifunctionality, which calculates the entropy of the amenity set incorporating the degree of similarity between amenity types. To examine the heterogeneity between amenity types, Foursquare's venue meta-categories, Shops & Consumer Services, Gastronomy, Nightlife, Arts & Entertainment, Professional Services, and Recreation are used. To test our hypotheses, we use Difference-in-Difference Regression based on Propensity Score Matching and Entropy Balancing.

We find that in most cities new subway stations positively impact the local amenity mix in terms of both, density and multifunctionality. Moreover, social amenities—such as Arts & Entertainment, Food venues, and Nightlife—respond most strongly, which could indicate both a flourishing of local economic life or a gentrification process. Furthermore, previous amenity subcenters benefit most of new metro stations due to the predominance of strong agglomeration forces that appear to affect consumer amenities. In the next chapter, we discuss related theories and derive three hypotheses to be tested. We then describe the operationalization of the research questions in terms of data and methods. Afterward, the results are discussed, followed by some concluding remarks.

Literature review and hypotheses

Consumer amenities as a development machine?

The relevance of consumer amenities for economic development has been vigorously discussed in the last decades. One core thinker who sparked the debate in 2002 was Richard Florida, who started arguing that our economies are increasingly driven by creativity, which gave rise to a new social group: The Creative Class, consisting of entrepreneurs, technologists, artists, and innovators, who make our economies flourish. These “talents,” Florida believed, would be attracted by urban amenities, such as tolerance but also consumer amenities, particularly informal, almost bohemian types of amenities like Nightlife or Restaurants. Clark et al. joined the argument stating that “the city increasingly becomes an Entertainment Machine, leveraging culture to enhance its economic well-being” (2004: 292). Instead of a traditional focus on production, consumption to fulfill quality-of-life demands had become the central economic process of post-industrial cities. While Social and Entertainment amenities are the focal point of the Creative Class thesis, it is also the mere density and diversity of amenities of cities that is attractive. According to Glaeser (2012), the prism of lifestyles and interests they represent enables us to foster the kinds of collaborations that make cities the places of creativity and innovation

Due to the popularity of the Creative Class thesis, policies have been adapted to focus on attracting the creative workers with lifestyle and consumption opportunities. However, while some studies find that amenities are successful in attracting human capital (Wenting et al., 2011), not in all cases this led to job creation, especially not in peripheral places (Argent et al., 2013). What is more, the people-centered policies have been criticized for contributing to regional economic polarization, making central places even more central (Storper, 2018). Also within cities, the focus on attracting high-skill labor has tilted many neighborhoods toward the interests of the affluent, which gave rise to the phenomenon of gentrification and consequent displacement (Smith, 1996).

While these problems are real, their prevalence differs depending on the types of amenities (local or chain store) and the process of amenity creation (through bottom-up emergence or top-down planning). Consumer amenities do not only influence the economic geography through the human capital it attracts, but also through local value chains. Large shares of revenues of especially small, locally owned stores “reverberate back into the local economy via subcontracting from local providers, payments to local employees, and improvements made in the public infrastructure around stores, as well as through indirect investments into employee health and retirement benefits” (Sevtsuk, 2020: 4). Consumer amenities also benefit a neighborhood in many non-economic ways by increasing its livability (Jacobs, 1961) through, for example, a reduction of crime (Humphrey et al., 2020), improved walkability (Forsyth et al., 2008), or promoting the formation of social ties (Small and Adler, 2019). Hereby, small, local stores and restaurants again have a much more positive impact compared to large chains. Based on observational data from social media, this study tests whether subway expansion can trigger such an emergence of small, social, informal amenities, perhaps even in the rather struggling outskirts areas.

Transport systems and location theory

The impact of transport regimes on urban form and spatial dynamics could be studied in the past decades based on more data and better methods. Innumerable studies have investigated the impact of transport infrastructure on the built environment, population density and employment. The underlying rationale is that changes in transportation costs or accessibility shift incentives to move for both people and economic activities. In their comprehensive meta-study, Kasraian et al. (2016) showed that proximity to rail investments seems to induce population growth in surrounding areas. However, the few studies that investigate the impact of rail expansion on employment growth show contradictory results (Cervero and Landis, 1997; McMillen and Lester, 2003), revealing the need for a cross-comparative study of the consequences of transport systems for economic activities. However, it is important to keep in mind that the reversed causality might also apply. There is soft evidence for a possible selection bias of subway stations being placed in already amenity-rich areas (Schuetz, 2015).

For understanding what impacts amenity formations, Netz and Taylor’s (2002) application of location theory to retail stores becomes relevant. Their study concludes that consumer service providers must deal with offsetting incentives with regard to agglomeration and dispersion. Consumer services can take advantage of the market size effect of agglomeration and meanwhile avoid the price competition associated with the market share effect by sufficiently differentiating the quality of their services. Their theory hints at the possible correlation between density and diversity of amenities in space. Moreover, it shows that consumer amenities are driven mainly by consumer patronage.

To analyze how customers spatially choose shops they visit, a series of gravitation models have been developed (Huff and Jenks, 1968; Reilly, 1931). Studying the breaking points of retail influence, Reilly (1931) found that more populous cities attract more outside retail trade than smaller ones, and cities attract more trade from close by than from further away. The lineage of studies building upon his *Law of Retail Gravitation* found “that customers do not necessarily shop at the closest store, but patronize locations in proportion to the attractiveness of the retail centers and in inverse proportion to their distance” (Joseph and Kuby, 2011: p. 423). Sevtsuk and Kalvo (2017) prove that these models are still quite accurate in predicting retail patronage. Combined these theories give rise to Hypothesis 1:

Hypothesis 1: New subway stations increase the density of consumer amenities in surrounding neighborhoods, whereby areas that had many amenities before the station opened react stronger.

Based on the literature on travel and consumption, we propose to extend the concept of area attractiveness to incorporate multifunctionality. In his trip-chaining model of consumer shopping, [Anas \(2007\)](#) claims that “travel is intimately related to both consumption and the allocation of time among discretionary activities” (p. 163), whereby urbanites tend to choose their destinations among alternative substitutes by chaining stops so that it suits their income, time constraints and consumption demand. If people can take care of many things at the same place through diverse amenities, it provides them with an efficient and stimulating way of shopping. Taking moreover into account that the market share effect forces amenities to diversify ([Netz and Taylor, 2002](#)), the second hypothesis becomes:

Hypothesis 2: New subway stations increase the multifunctionality of the amenity mix in surrounding neighborhoods.

The consumption processes will be further analyzed in the next section to better understand amenity heterogeneity.

Amenity heterogeneity

Consumer amenities show crucial differences in the social groups that tend to visit them. For example, Nightlife venues are mostly visited by young adults, while Art Museums are mostly attended by educated people and tourists. This is relevant because some social groups tend to use subway transport more than others and may thus stimulate the local market, as studies finding higher pedestrian activity and staying time in metro station catchment areas compared to other urban areas suggest ([Kim et al., 2017](#)). Therefore, it is important to understand who travels with the subway.

A large body of literature on modal choice analysis aims at identifying the conditions under which specific groups of people choose their mode(s) of transport, whereby convenience, time and financial costs are the major factors ([Buehler, 2011](#)). Since car ownership is expensive but allows for fast and flexible traveling, high-income households tend to travel more by car instead of the subway than less affluent ones. Moreover, personal “environmental preferences increase the likelihood of choosing an environmentally friendly mode” ([Johansson et al., 2006](#): p. 517). Considering the interest of young people in Nightlife and their higher likelihood to use public transport because of budgetary and environmental reasons, one might expect Nightlife venues to pop up near new subway stations.

Amenities in subway neighborhoods are not only attracted by people that use the subway system but also by people that move close by. A substantial amount of literature has shown how urban rail transit investment capitalizes in neighboring home values ([Debrezion et al., 2007](#)). [Zheng et al. \(2016\)](#) proved that in Beijing an increase in the amount and diversity of restaurants close to new metro stations can partially explain the increase in home values. The increase in housing prices due to rising neighborhood attractiveness may induce a displacement of less affluent and an inflow of more affluent households that can afford and demand more and specific types of consumer amenities. We assume that Arts & Entertainment venues may indicate an underlying process of gentrification.

Consumer amenities differ also in their market entry barriers which affect how quick and strong they react to subway stations. The two most important factors are the dominance of economies of scale, which makes it difficult for small firms to enter the market, and set-up costs with their attached financial risks ([Porter, 1980](#)). Restaurants tend to require lower set-up costs compared to, for example, Medical Offices (a professional venue) or a Spa (a recreation venue). Moreover, the former are usually quite decentralized and do not depend heavily on economies of scale. This should make them much more likely to respond quickly and numerously to location opportunities. The third hypothesis becomes:

Hypothesis 3: There is a heterogeneous reaction to subway connection of different amenity types, whereby Nightlife, Arts & Entertainment and Restaurant venues are expected to react stronger than other types.

The discussed literature is summarized in the following conceptual model (see [Figure 1](#)), which will help interpret the results.

Methodology

Case selection and spatial units

To be able to assess the treatment effect of new stations with a temporal lag and a possible selection bias, data on 1 year before and 2 years after the new metro station opening is required. Since Foursquare has reliable data since 2012 ([Foursquare, 2020](#)), the eligible cases are metro openings between 2014 and 2017. The selection is further limited to cities with more than 500.000 inhabitants as they rely more on metro transit ([Gonzalez-Navarro and Turner, 2018](#)). This leads us to select Rome, Milan, Barcelona, Budapest, Warsaw, Sofia, Vienna, Helsinki, and Stuttgart, which together opened 84 new metro stations during that period, as the best European cases for our study (see [Supplementary Table 1](#) for population and subway expansion data on these cities). It should be noted that most new stations are located in outskirts areas.

To define the cities' urban boundaries, Nightlight satellite imagery was used (see the section *Definition of Urban Areas* of the [Supplementary Material](#)). The spatial units of analysis are based on a regular point grid randomly generated over the urban area, where points are 100 m distant from each other and their properties are defined by a neighborhood of 300 m geodesic radius, a

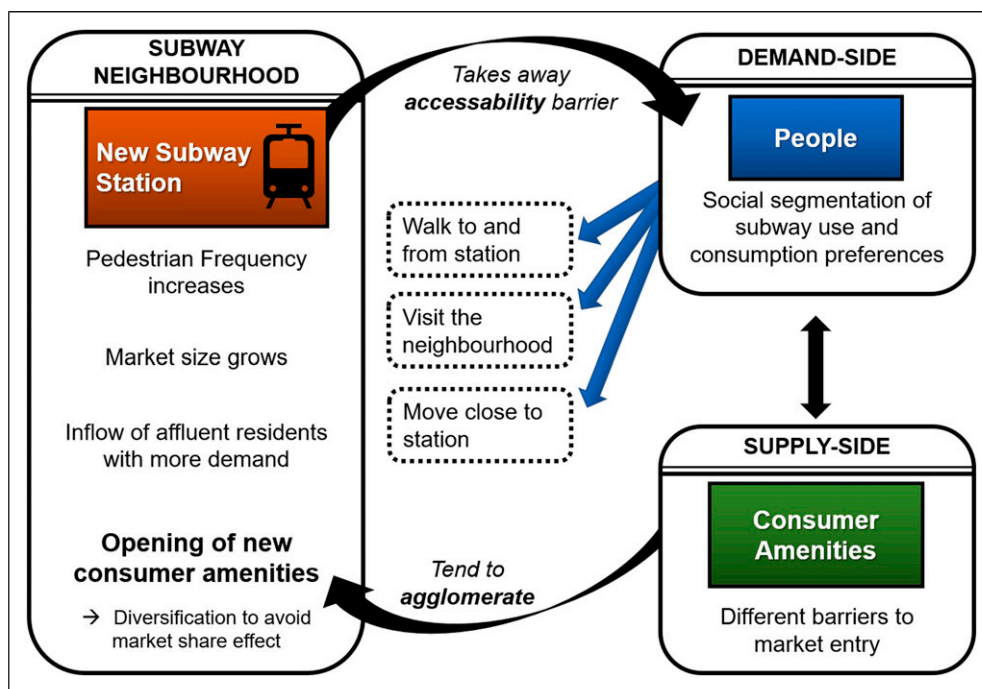


Figure 1. Conceptual framework: metro expansion and consumer amenities.

distance most urbanites are willing to walk (Yang and Diez-Roux, 2012). The grid represents an amenity density and multifunctionality field. The overlapping neighborhoods reduce the arbitrariness of the borders of the units of analysis (see [supplementary materials](#) to see all the figures and tables).

Data and variables

As the core data source on amenities, we chose *Foursquare*, one of the most extensive and complete social media datasets on places, which provides reliable historical data on venue creation dates and detailed categories of venues from 2012 onward (Foursquare, 2020; McKenzie et al., 2015; Weerdenburg et al., 2019).¹ The venues are classified into standardized categories by users, which reflects their subjective experience of a place. Due to the large-scale use and crowd sourcing of Foursquare, its amenity categorizations must have utility and therefore quality.¹ For example, in contrast to, for example, firm registries, it is able capture small and informal amenities. Moreover, Foursquare's venues have been shown to predict with high confidence the location of retail, hotels, restaurants and cafes (Spyratos et al., 2017). [Supplementary Figure 7](#) shows the number of Foursquare Arts & Entertainment, Retail and Food venues in comparison to the respective number of Orbis firms per city. The proportions of venues per type and the intercity ranking of Foursquare venues resemble the ones found in Orbis. The few exemptions can be explained through different definitions of urban areas and limitations in data quality of both Foursquare and Orbis (see [Supplementary Materials](#), p. 13 for more information).

Due to its standardized categorization and data generation, Foursquare data has the special property that it allows for comparability across cities. There are 937 categories of venues in total, organized in a tree-like structure of up to 4 subcategories (see [Supplementary Figure 2](#) for a graph representation of the taxonomy). The meta-categories we used are "Arts & Entertainment," "Food," "Shops & Services," "Nightlife Spots," "Professional Places," "Outdoors & Recreation." From the latter, only Recreation amenities, specifically sports and spa facilities, are selected. In contrast to that, "Arts & Entertainment" contains rather high-class entertainment like museums, theaters or theme parks and "Nightlife" captures the pubs and clubs that engage the young and informal. The final dataset contains 562 categories. [Supplementary Figure 1](#) shows the sample size per meta-category for each city and [Supplementary Table 11](#) lists the final Foursquare venue categories assigned to the meta-categories.

After collecting the data using Foursquare's Web API, the number and type of venue openings in the cities of interest can be regenerated for various points in time. While unfortunately only amenity openings that survived until now are contained in the data, it can be argued that those that survived are the most successful ones, which makes them more interesting anyway. Moreover, the same bias exists across all cities, allowing us to draw comparisons. We assume that this data limitation should not interfere with causal analysis, as there is little reason to suspect that new subway neighborhoods have more amenities closing shortly after opening compared to other places.

Descriptive statistics and maps for the Foursquare data and the data sources and computation method for the rest of the variables can be found in the [Supplementary Materials](#).²

Similarity-based diversity

In previous studies, Multifunctional Amenity Mix (or Multifunctional Land Use) has been mainly assessed by counting the number of locally present amenity types (Vreeker et al., 2004) or adding up the proportions of each locally present amenity type to its total amount (Batty et al., 2004). Our new multifunctionality measure integrates the idea of amenity complementarity by quantifying the

similarity of venue types. Thereby, we assume that more similar amenities and their activities are more likely to be substitutes for each other, while dissimilar amenities are more complementary across a daily schedule. For example, people that attend a theater show are less likely to go to the cinema afterward, another form of Entertainment, but could be probably more willing to go to a pub or restaurant instead, which involves different activities. Based on Jane Jacobs' (1961) conceptualization of urban vibrancy, the core idea behind our measure is that for a diversity of people to use a given space over time, they need to be kept occupied with different kinds of activities. A simple entropy measure would fail to accurately capture the multifunctional quality of a neighborhood because it treats all categories on a par.

We, therefore, propose a *similarity-based diversity measure* as an indicator of multifunctionality. It assesses how uniform a population is distributed over a set of categories, taking into account that similar categories might be confused with each other or are more substitutive in nature. The idea is to *redistribute* a nominal distribution over categories (*cat*) according to some similarity measure between categories and then measure entropy over this new distribution. The diversity of this distribution should get higher, the higher the entropy of the original distribution is, as well as the more dissimilar the categories are from each other.

The frequencies of categories are given in terms of the following list

$$P = [P(1), \dots, P(k)] \tag{1}$$

where $k = |cat|$. Next, our measure requires a similarity measure between categories, where identity is 1, and all other similarities are between 0 and 1, such that the most dissimilar categories have a similarity of zero. Based on Foursquare's sensible venue category taxonomy, a pairwise similarity measure of amenities can be computed, whereby amenity types are more similar the shorter their taxonomic distance, that is, the shorter the shortest path within the Foursquare amenity taxonomy (McInnes et al., 2014)

$$\text{sim}(i,j) = 1 - \left(\frac{\min(\text{Path}_{i,j})}{\max(\min(\text{Path}_{n,n}))} \right) \tag{2}$$

with

$$\text{sim}(i,j) \in \{0, \dots, 1\}, \quad j, i \in \text{cat} \tag{3}$$

We now modify the frequency of some goal category (here: *j*) by letting it *profit* from other categories (e.g., *i*) in relation to their similarity. The new frequency of *j* (*Padd(j)*) is a *similarity-weighted sum* over all categories (including itself) that will be bigger the more similar *j* is to the other categories

$$\text{Padd}(j) = \sum_{i=1}^k P(i) * \text{sim}(i,j) \tag{4}$$

Note that in case similarity to all other categories is 0, $\text{Padd}(j)=P(j)$. In order to keep the sum (the total amount of instances) over all categories constant (this is what it means to “redistribute”), each profit that was shifted toward *j* subsequently needs to be *subtracted* from the respective category *i* from which it stems

$$\text{Psub}_j(i) = P(i) * (1 - \text{sim}(i,j)) \tag{5}$$

Note that also in this case, if the similarity is 0, $\text{Psub}_j(i) = P(i)$. The new distribution over categories *i*, where amounts were shifted toward one category *j*, thus looks like this

$$P_j^{\text{sim}}(i) = \begin{cases} P_{\text{add}}(j), & \text{if } i = j \\ P_{\text{sub}_j}(i), & \text{otherwise} \end{cases} \quad (6)$$

We can now measure Shannon entropy over this new distribution

$$\text{Entropy } H_{\text{sim}}(j) = - \sum_{i=1}^k P_j^{\text{sim}}(i) \log P_j^{\text{sim}} \quad (7)$$

It should be clear that this new entropy measure satisfies our goal: In the extreme case of all categories being equal, the entire amount is shifted toward a single category j , which results in an extremely non-uniform distribution with entropy 0. In case all categories are maximally distant from each other, then the entropy simply becomes the entropy of the original distribution. The maximal similarity between two different categories derived from the Foursquare taxonomy is 0.89, and the density of the similarities follows a Gaussian distribution.

Finally, in order to nullify the effects of shifting profits only into one direction, we compute Entropy $H_{\text{sim}}(j)$ for each possible goal category $j \in \text{cat}$ and average the effects

$$H_{\text{sim}} = \frac{\sum_{j=1}^k \text{Entropy } H_{\text{sim}}(j)}{k} \quad (8)$$

The presented way of measuring a similarity-based diversity will serve here as the indicator of the multifunctionality of a neighborhood. However, the formula can also be applied to other cases in which inter-category similarity is of relevance.

The relationship between amenity density and multifunctionality of amenities is non-linear but positive (see Supplementary Figure 3). An increase of density by several new amenities will increase multifunctionality less at already highly dense areas due to a greater likelihood of repetition and similarity to other amenities compared to the same number of amenities being added to a low-density area. [Supplementary Figure 4](#) shows that our new similarity-based diversity measure of multifunctionality differs considerably from the entropy of the amenity distributions.

Regression models

To test the hypotheses, Difference-in-Difference Regression is applied, whereby two different control-unit-matching methods are exploited to test for causal effects of treatments: Propensity Score Matching and Entropy Balancing. For both preprocessed datasets, we tested eight dependent variables: The change in the number of total amenities (H1), change in multifunctionality (H2), and change in the number of a specific amenity type from before the station opened to 2 years afterward, whereby the latter is differentiated between Food, Nightlife, Arts & Entertainment, Shops & Services, Recreation and Professional Places (H3). To account for the difference between existing subcenters and amenity poor areas receiving a new station (part of H1), the interaction effect between the metro neighborhood dummy and the density of amenities beforehand is tested.

As a first control matching method, nearest neighbor Propensity Score Matching (PSM) is applied to mimic a RCT, whereby treated units (up to 1200 m distance from a new station) are matched to control units (minimum 1500 m distant to any metro station) with a similar likelihood of receiving a new metro station ([Rosenbaum and Rubin, 1983](#)). We define this likelihood as being determined by the demand for a station, or accessibility need. Therefore, population density, the density of amenities before the subway expansion, the unemployment rate and the distance to the CBD are included in the matching algorithm. For validation, the distributions of these variables were compared between the treated and newly matched control units to ensure their similarity. As a

second method, we applied Entropy Balancing (EB), a data preprocessing method to reweight control units (minimum 1500 m distant to any metro station) so that they optimally represent the covariate distribution of the treated units (up to 1200 m distance from a new station) (Hainmueller, 2012). The covariates for which balance is achieved are the population in 2012, the change in population density between 2012 and 2018, the unemployment rate, the Euclidean distance to the CBD, the street accessibility to the nearest subcenter and the density of amenities before the subway expansion. [Supplementary Table 4](#) specifies the distribution of the most important control variable—amenity density before expansion—for the treated and the PSM and EB control units

$$\Delta \text{amen_dens}_{ijt} = f \left(\begin{array}{l} \text{metro} + \text{pop}_{2012} + \Delta \text{pop}_{2012-2018} + \text{unempl} + \text{dist}_{\text{CBD}} + \\ \text{access_subcenter} + \text{amen_dens_before} + \varepsilon_{ijt} \end{array} \right) \quad (9)$$

Equation (9) shows the regression model for both PSM and EB. For testing different hypotheses, the dependent variable, change of amenity density before and after station opening ($\Delta \text{amen_dens}$), in equation (9) is replaced by change in multifunctionality ($\Delta \text{multifunct}$) and change in the density of the amenity types ($\Delta \text{type_dens}$), respectively. The control variables are the population in 2012, the change in population density between 2012 and 2018, the unemployment rate, the Euclidean distance to the CBD, the street accessibility to the nearest subcenter. Finally, we add a control for the selection bias that stations may be placed in already amenity-rich or multifunctional neighborhoods. For all regression models, all variables that are not dummies are standardized using Z-score normalization. More information on the methods can be found in the [Supplementary Materials](#).

Results

In Rome, Helsinki, Stuttgart, Sofia, Warsaw, and Milan, the opening of the metro station caused a significant increase in the density of amenities (H1) in the 1200 m around it compared to both their E-balanced as well as the PS-matched control units (see [Table 1](#)). Barcelona and Vienna on the other hand, only show significant effects in the EB-Regression.

Across all cities and methods, the number of venues before the station opened is the strongest predictor of amenity increases, proving that agglomeration forces are considerably strong for consumer amenities. An additional test of the interaction effect of the amenity density before the station opened and the new metro station dummy, in both the PS-matched as well as E-balanced datasets, shows that, as expected, previous amenity subcenters in Rome, Helsinki, Stuttgart, Warsaw, Sofia, and Milan benefited more from the new metro station. In Barcelona and Vienna, the interaction effect is significant in one of the two models.

Population Density is in most cities negatively related to amenity density increases, indicating that there is still a strong separation of residential and commercial uses in these cities. The analysis of the impact of the population increase between 2012 and 2018, on the other hand, shows very mixed results, making it very difficult to interpret. Areas with higher unemployment experience significantly stronger amenity increases than those with lower unemployment in Barcelona, Budapest, Stuttgart, and Milan, whereas the opposite relation only holds in Rome and Vienna. As expected, in line with the retail gravity models, the more walking distance to a subcenter, the fewer amenities opened, whereby this relationship is only significant in a few cities. Also, the distance to the CBD is negatively correlated to amenity density increases in most cities. Together, the entered variables can explain a large percentage of the variation indicated by high adjusted R-squared results.

Next, we analyzed the impact of subway expansion on the multifunctionality of the local Amenity Mix (H2) based on *similarity-based diversity*. The results demonstrate a significant increase in multifunctionality in new subway neighborhoods compared to both their PSM-matched

Table 1. Treatment effect of new metro stations on the density of local amenities.

	Rome		Barcelona		Vienna		Budapest		Helsinki		Stuttgart		Sofia		Warsaw		Milan	
	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p
Ebalance-based Regression																		
(Intercept)	-0.030	***	-0.273	***	0.000	***	-0.241	***	-0.270	***	-0.071	***	-0.123	***	-0.201	***	-4.458	***
Metroneighbourhood_dum	0.074	***	0.275	***	0.060	***	-0.242	***	0.588	***	0.458	***	0.202	***	0.423	***	1.108	***
Population 2012	-0.017	***	-0.171	***	-0.067	***	-0.111	***	-0.026	**	-0.143	***	0.102	***	0.026	***	-0.120	***
Pop change 2012-2018	0.025	***	0.031	***	0.151	***	-0.030	***	-0.007	***	0.204	***	0.037	***	-0.031	***	-0.963	***
Percentage unemployed	-0.015	***	0.306	***	-0.101	***	0.094	***	-0.010	***	0.249	***	-0.032	***	-0.032	***	0.096	***
Distance to center	-0.070	***	-0.268	***	-0.066	***	-0.263	***	0.091	***	-0.451	***	0.046	***	-0.019	***	-1.429	***
Nr venues before station	0.884	***	0.887	***	0.769	***	1.085	***	0.902	***	0.652	***	0.985	***	1.025	***	0.784	***
Street-level accessibility to nearest subcenter	-0.003	***	-0.057	**	-0.014	***	0.217	***	-0.041	***	-0.113	***	0.065	***	0.110	***	-0.236	***
Adjusted R-squared	0.792		0.860		0.764		0.968		0.591		0.606		0.880		0.907		0.873	
PSM-based regression																		
(Intercept)	-0.024	***	-0.010	***	-0.006	***	0.018	**	-0.087	***	-0.167	***	-0.059	***	-0.088	**	-0.053	***
Metroneighbourhood_dum	0.048	***	0.020	***	0.012	***	-0.036	***	0.174	***	0.334	***	0.117	***	0.176	***	0.105	***
Population 2012	-0.002	***	-0.036	***	-0.093	***	-0.038	***	-0.005	***	-0.071	**	0.034	***	0.071	*	-0.121	***
Pop change 2012-2018	0.023	***	0.034	**	0.118	***	-0.010	**	-0.010	***	0.163	***	0.021	***	0.042	***	-0.030	***
Percentage unemployed	-0.011	**	0.042	**	-0.067	***	0.028	***	0.007	***	0.160	***	0.083	***	0.083	***	0.023	**
Distance to center	-0.038	***	-0.115	***	0.009	***	-0.014	*	0.057	***	-0.265	***	0.024	***	-0.111	*	-0.014	***
Nr venues before station	0.861	***	0.859	***	0.810	***	0.975	***	0.785	***	0.560	***	0.968	***	0.913	***	0.986	***
Street-level accessibility to nearest subcenter	-0.009	***	-0.003	***	-0.050	**	-0.005	***	0.009	***	-0.092	***	0.032	***	0.057	*	-0.011	***
Adjusted R-squared	0.791		0.803		0.676		0.947		0.621		0.592		0.910		0.919		0.897	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 2. Treatment effect of new metro stations on the multifunctionality of the amenity mix.

	Rome		Barcelona		Vienna		Budapest		Helsinki		Stuttgart		Sofia		Warsaw		Milan	
	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p
Ebalance-based Regression	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p
(Intercept)	0.014	.	0.259	***	-0.236	***	-0.232	***	0.063	***	0.122	***	-0.074	***	-0.217	***	-1.067	***
Metroneighbourhood_dum	0.296	***	0.047	***	0.051	***	-0.196	***	0.356	***	0.045	.	0.162	***	0.033	***	0.101	***
Population 2012	-0.071	***	-0.391	***	-0.062	***	0.001	***	0.096	***	0.069	.	0.397	***	0.120	***	-0.002	**
Pop change 2012-2018	0.075	***	0.139	***	0.096	***	0.027	***	0.030	***	0.037	**	0.112	***	0.007	***	-0.068	***
Percentage unemployed	0.021	***	0.262	***	-0.082	***	-0.025	***	-0.023	***	0.110	***		***	0.092	***	-0.014	***
Distance to center	-0.028	***	-0.353	***	-0.235	***	-0.169	***	0.103	***	-0.106	***	-0.058	***	-0.397	***	-0.403	***
Nr venues before station	0.076	***	-0.328	***	-0.042	***	-0.111	***	-0.124	***	-0.249	***	-0.116	***	-0.321	***	-0.005	.
Street-level accessibility to nearest subcenter	-0.035	***	-0.149	***	-0.067	***	-0.148	***	-0.074	***	-0.013	***	-0.033	**	0.024	***	-0.035	***
<i>Adjusted R-squared</i>	0.038		0.352		0.055		0.148		0.054		0.087		0.051		0.082		0.105	
PSEM-based regression	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p	Beta.	p
(Intercept)	-0.140	***	-0.174	***	-0.004	***	0.086	***	-0.212	***	-0.069	.	-0.045	.	-0.062	.	0.007	
Metroneighbourhood_dum	0.281	***	0.348	***	0.008	***	-0.173	***	0.424	***	0.139	.	0.090	**	0.125	.	-0.014	
Population 2012	-0.082	***	-0.065	.	-0.023	.	0.060	***	0.104	***	0.211	***	0.279	***	0.322	.	0.002	
Pop change 2012-2018	0.057	***	0.005	***	0.065	**	0.093	***	0.038	**	0.042	***	0.100	***	0.195	.	-0.069	**
Percentage unemployed	0.019	.	0.083	**	-0.034	.	-0.054	***	-0.026	.	0.136	***		***	0.163	.	-0.030	***
Distance to center	-0.031	.	-0.189	***	-0.082	**	-0.183	***	0.077	***	-0.110	***	-0.056	**	-0.719	***	-0.243	***
Nr venues before station	0.079	***	-0.167	***	0.004	***	-0.337	***	-0.141	***	-0.337	***	-0.138	***	-0.497	***	-0.147	***
Street-level accessibility to nearest subcenter	-0.028	.	-0.041	.	-0.071	.	-0.114	***	-0.076	***	0.029	.	-0.044	.	0.176	.	-0.026	
<i>Adjusted R-squared</i>	0.030		0.072		0.055		0.097		0.050		0.075		0.059		0.114		0.039	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

and E-balanced counterparts in Helsinki, Rome, Sofia, Barcelona, and Stuttgart (see [Table 2](#)). All other cities, except for Budapest, show significant results of subway expansion for multifunctionality in one of the two preprocessing models

Due to the decreasing returns of amenity density for multifunctionality, one can observe that many cities have a negative relationship between the level of multifunctionality before the stations opened and the increase in multifunctionality 2 years after the station opened. Therefore, we have tested which cities had a significantly higher multifunctionality of areas around new metro stations compared to control units before the stations opened, which could explain why the increase is not substantial. It turns out that this selection bias applies to Sofia, Warsaw, Budapest, Vienna, Helsinki, and Rome for both the PSM and EB datasets. The results also show that the variables that predict amenity density very well are weak in predicting multifunctionality, as the consistently low adjusted R^2 values reveal. This leads us to question how much of the multifunctionality of an amenity mix is based on randomness or, instead, what explanatory variables might be omitted.

Moving on to the findings regarding the heterogeneity between amenity types, which can be found in [Table S9 of the Supplementary Materials](#), Arts & Entertainment venues increase significantly more in new metro neighborhoods compared to their controls in 8 out of 9 cities, Food venues in 7 and Nightlife in 4 of them, supporting H3. Shops and Services significantly increase in 3 out of 9 cities, Professional venues also in 3 and Recreational venues only in 2 of them. The average standardized coefficient is highest for Arts & Entertainment, followed by Food venues, then Nightlife and Shops and Services. This further highlights the possibility of an underlying gentrification process, where more affluent, educated, and young people move into the new neighborhoods. We additionally tested this by analyzing the interaction effect of the metro neighborhood dummy and population increase, which proved to be positive and significant in Rome, Vienna, and Helsinki for both Arts & Entertainment as well as Nightlife, and in Stuttgart only for Arts & Entertainment. The number of respective venues before the station opened is the strongest predictor for new venues across all cities and venue types, whereby this agglomeration effect is the strongest for Shops and Services, Gastronomy and Professional venues.

Discussion and conclusion

Taking advantage of the growing availability of georeferenced social media data for urban studies, we used Foursquare venues to evaluate the spatiotemporal effect of subway expansion on amenity density, multifunctionality of the local amenity mix and different venue types. The findings show that there is a significant impact of new subway stations on the local amenity mix, but that the effect varies across cities and venue types. Subway neighborhoods in Rome, Helsinki, Stuttgart, Sofia, Warsaw, and Milan seem to benefit much more in terms of amenity density from the new stations than the ones in Vienna, Budapest, and Barcelona (H1). It also becomes clear that previous amenity subcenters profit significantly more from new metro stations than amenity deserts, hinting at strong agglomeration forces. Helsinki, Rome, Sofia, Barcelona, and Stuttgart have significant increases in multifunctionality, which speaks to the area's increased ability to attract dissimilar amenities that involve activities that could not be found there previously (H2). Our new similarity-based diversity measurement for assessing the multifunctionality of an amenity mix can capture an important quality, which might approximate Jane Jacobs' idea of urban vibrancy (1961) and a diversity of lifestyle activities. It could be therefore an important factor for increasing local livability. We hope that it can be further used and tested in other studies.

Finally, our results show that indeed Arts & Entertainment, Food venues and Nightlife have been impacted most strongly by new subway stations, as was expected (H3). It is interesting that especially the social, cultural, and informal types of amenities have been affected, strongest of them the high-class Arts & Entertainment amenities. On the one hand, they are said to particularly attract

the Creative Class (Florida, 2002) and may play a role in creating the collaborative environment that makes urbanites innovate (Glaeser, 2012). On the other, they could be a symptom of gentrification and displacement, indicating that TOD projects may potentially not benefit those for whom they were originally intended. Future work should focus on better methods and data to distinguish amenity-led gentrification and displacement from gains in local economic life and livability, an analytical goal that goes beyond the current study.

Certainly, this study has a few more limitations. While the Foursquare amenity data has many advantages, we could not access past amenities that did not survive until our time of data collection. Moreover, we have not analyzed the obvious network effect that subway expansions have on the whole city. Due to the limited scope of the research and lack of data, we had to omit a few variables that could have helped to better explain why the impact of new stations varies in strength and composition between the cities. One factor is the path dependency of transportation and the built environment, as a city, once it is constructed, is hard to modify on a large scale. Some subway neighborhoods perhaps do not have enough vacant land available to create new amenities. In others, the amendments of zoning regulation and building codes, governmental investments or even the construction of new buildings could have enabled the development of new amenities. Next, subway culture varies between cities, whereby the level of stigmatization and therefore use may influence the impact a new station has on the surrounding areas. Finally, in different cities the subway expansion has been performed in neighborhoods of different socio-economic geographies for which our control of the local unemployment percentage cannot account. As the PSM and EB approach can only address the selection bias due to observed attributes, we still face some challenges to identify the causal relationships due to endogeneity coming from unobserved variables.

Despite that, our study does show that the spatiotemporal dynamics of different cities display surprisingly many parallels. The evidence supports the relevance of subway expansion as a TOD policy tool that can attract amenities even to outskirts, thereby transforming them into more multifunctional, mixed-use neighborhoods. However, the results also show that subway connection alone cannot miraculously vitalize amenity deserts. A minimum number of pre-existing amenities are necessary for new subway stations to have an amenity-attracting effect on neighborhoods. More research is needed to understand how subway expansion interacts with other TOD policies and which TOD policies packages are the most effective in achieving a decentralized structure of mixed-use around mass transit stations. In an age of urbanization, where an ever-larger share of people has to live outside central areas, TOD may become part of a larger strategy to tackle the geographical inequality in accessing amenities and jobs.

Acknowledgment

The research was supported by the MIT Sustainable Urbanization Lab, which funded the stay of Tabea Sonnenschein as a visiting student in the lab. We moreover thank Pierre-Alexandre Balland and Adriano Borges Ferreira Da Costa for the support and mentorship during the process.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Tabea S Sonnenschein  <https://orcid.org/0000-0001-6592-9548>

Simon Scheider  <https://orcid.org/0000-0002-2267-4810>

Siqi Zheng  <https://orcid.org/0000-0002-4467-8505>

Supplementary material

Supplementary material for this article is available online.

Notes

1. Foursquare employs so called “Superusers” to keep the venue data up to date (Foursquare, 2020).
2. Moreover, all scripts, including the similarity-based diversity function, can be found within the following GitHub folder: <https://github.com/TabeaSonnenschein/SubwayExpansionAndAmenityMix>

References

- Anas A (2007) A unified theory of consumption, travel and trip chaining. *Journal of Urban Economics* 62(2): 162–186. doi: [10.1016/j.jue.2006.05.002](https://doi.org/10.1016/j.jue.2006.05.002).
- Argent N, Tonts M, Jones R, et al. (2013) A creativity-led rural renaissance? Amenity-led migration, the creative turn and the uneven development of rural Australia. *Applied Geography* 44: 88–98. doi: [10.1016/j.apgeog.2013.07.018](https://doi.org/10.1016/j.apgeog.2013.07.018).
- Batty M, Besussi E, Maat K, et al. (2004) Representing multifunctional cities: density and diversity in space and time. *Built Environment* 30(4): 324–337. doi: [10.2148/benv.30.4.324.57156](https://doi.org/10.2148/benv.30.4.324.57156).
- Buehler R (2011) Determinants of transport mode choice: a comparison of Germany and the USA. *Journal of Transport Geography* 19: 644–657. doi: [10.1016/j.jtrangeo.2010.07.005](https://doi.org/10.1016/j.jtrangeo.2010.07.005).
- Cervero R and Landis J (1997) Twenty years of the Bay area rapid transit system: land use and development impacts. *Transportation Research Part A* 31(4): 309–333. doi: [10.1016/S0965-8564\(96\)00027-4](https://doi.org/10.1016/S0965-8564(96)00027-4).
- Clark NT (2004) Urban amenities: Lakes, Opera, and juice bars: do they drive development? In: Clark NT (ed) *The City as an Entertainment Machine*. Bingley: Research in Urban Policy, Emerald Group Publishing Limited, Vol. 9, 103–140.
- Debrezion G, Pels E and Rietveld P (2007) The impact of railway stations on residential and commercial property value: A meta-analysis. *Journal of Real Estate Finance and Economics* 35(2): 161–180. doi: [10.1007/s11146-007-9032-z](https://doi.org/10.1007/s11146-007-9032-z).
- Duranton G and Puga D (2004) Micro-foundations of urban agglomeration economies. In: Henderson JV and Thisse JF (eds) *Handbook of Regional and Urban Economics*. Amsterdam, NL: Elsevier, Vol. 4, pp. 2063–2117.
- Florida R (2002) *The Rise of the Creative Class: And How its Transforming Work, Leisure, Community and Everyday Life*. New York, USA: Basic Books.
- Forsyth A, Hearst M, Oakes JM, et al. (2008) Design and destinations: factors influencing walking and total physical activity. *Urban Studies* 45(9): 1973–1996. doi: [10.1177/0042098008093386](https://doi.org/10.1177/0042098008093386).
- Foursquare (2020) Places Data. Foursquare Developer. Retrieved on April 04, 2020, from: <https://developer.foursquare.com/places>
- Glaeser E (2012) *Triumph of the City. How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. New York: Penguin.
- Gobillon L, Selod H and Zenou Y (2007) The mechanisms of spatial mismatch. *Urban Studies* 44(12): 2401–2427. doi: [10.1080/00420980701540937](https://doi.org/10.1080/00420980701540937).
- Gonzalez-Navarro M and Turner MA (2018) Subways and urban growth: evidence from earth. *Journal of Urban Economics* 108(C): 85–106. doi: [10.1016/j.jue.2018.09.002](https://doi.org/10.1016/j.jue.2018.09.002).

- Hainmueller J (2012) Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20: 25–46. doi: [10.1093/pan/mpr025](https://doi.org/10.1093/pan/mpr025).
- Huff D and Jenks G (1968) A graphic interpretation of the friction of distance in gravity models. *Annals of the Association of American Geographers* 58: 814–824. doi: [10.1111/j.1467-8306.1968.tb01670.x](https://doi.org/10.1111/j.1467-8306.1968.tb01670.x).
- Humphrey C, Jensen ST, Small DS, et al. (2020) Urban vibrancy and safety in Philadelphia. *Environment and Planning B: Urban Analytics and City Science*. doi: [10.1177/2399808319830403](https://doi.org/10.1177/2399808319830403)
- Jacobs J (1961) *The Death and Life of Great American Cities*. New York: Random House.
- Johansson MV, Heldt T and Johansson P (2006) The effects of attitudes and personality traits on mode choice. *Transportation Research Part A* 40: 507–525. doi: [10.1016/j.tra.2005.09.001](https://doi.org/10.1016/j.tra.2005.09.001).
- Joseph L and Kuby M (2011) Gravity modelling and its impacts on location analysis. In: Eisel HA and Marianov V (eds) *Foundations of Location Analysis*. Boston, MA: Springer, 423–443.
- Kasraian D, Maat K, Stead D, et al. (2016) Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies. *Transport Reviews* 36(6): 772–792. doi: [10.1080/01441647.2016.1168887](https://doi.org/10.1080/01441647.2016.1168887).
- Kim T, Sohn DW and Choo S (2017) An analysis of the relationship between pedestrian traffic volumes and built environment around metro stations in Seoul. *KSCE Journal of Civil Engineering* 21(4): 1443–1452. doi: [10.1007/s12205-016-0915-5](https://doi.org/10.1007/s12205-016-0915-5).
- Littman T (2017) *Evaluating Accessibility for Transport Planning: Measuring People's Ability to Reach Desired Goods and Activities*. Victoria, CA: Victoria Transport Policy Institute.
- McInnes BT, Pedersen T, Li Y, et al. (2014) U-path: An undirected path-based measure of semantic similarity. *AMIA—Annual Symposium Proceedings*. Symposium: AMIA, 882–891.
- McKenzie G, Janowicz K, Gao S, et al. (2015) Poi pulse: a multi-granular, semantic signature-based information observatory for the interactive visualization of big geosocial data. *Cartographica* 50: 71–85. doi: [10.3138/cart.50.2.2662](https://doi.org/10.3138/cart.50.2.2662).
- McMillen DP and Lester TW (2003) Evolving subcenters: Employment and population densities in Chicago, 1970–2020. *Journal of Housing Economics* 12(1): 60–81. doi: [10.1016/S1051-1377\(03\)00005-6](https://doi.org/10.1016/S1051-1377(03)00005-6).
- Netz JS and Taylor BA (2002) Maximum or minimum differentiation? location patterns of retail outlets: the review of economics and statistics. *MIT Press* 84(1): 162–175.
- Porter M (1980) *Competitive Strategy*. New York: Free Press.
- Reilly WJ (1931) *The Law of Retail Gravitation*. New York: W.J. Reilly.
- Rosenbaum PR and Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41–55. doi: [10.1093/biomet/70.1.41](https://doi.org/10.1093/biomet/70.1.41).
- Schuetz J (2015) Do rail transit stations encourage neighbourhood retail activity? *Urban Studies* 52(14): 2699–2723. doi: [10.1177/0042098014549128](https://doi.org/10.1177/0042098014549128).
- Sevtsuk A (2020) *Street Commerce Creating Vibrant Urban Sidewalks*. Pennsylvania: University of Pennsylvania Press.
- Sevtsuk A and Kalvo R (2017) Patronage of urban commercial clusters: a network-based extension of the Huff model for balancing location and size. *Environment and Planning B: Planning and Design* 45(3): 508–528. doi: [10.1177/2399808317721930](https://doi.org/10.1177/2399808317721930).
- Small ML and Adler L (2019) The role of space in the formation of social ties. *Annual Review of Sociology* 45(1): 111–132. doi: [10.1146/annurev-soc-073018-022707](https://doi.org/10.1146/annurev-soc-073018-022707).
- Smith N (1996) *The New Urban Frontier: Gentrification and the Revanchist City*. London: Routledge.
- Spyratos S, Stathakis D, Lutz M, et al. (2017) Using Foursquare place data for estimating building block use. *Environment and Planning B: Urban Analytics and City Science* 44(4): 693–717. doi: [10.1177/0265813516637607](https://doi.org/10.1177/0265813516637607).
- Storper M (2018) Separate Worlds? Explaining the current wave of regional economic polarization. *Journal of Economic Geography* 18(2): 247–270. doi: [10.1093/jeg/lby011](https://doi.org/10.1093/jeg/lby011).
- Tobler WR (1970) A Computer movie simulating urban growth in the detroit region. *Economic Geography* 46: 234–240. doi: [10.2307/143141](https://doi.org/10.2307/143141).

- UN (2016) *Transport for Sustainable Development*. United Nations High-level Advisory Group on Sustainable Transport.
- Vreeker R, de Groot HLF, Verhoef ET, et al. (2004) Urban multifunctional land use: theoretical and empirical insights on economies of scale, scope and diversity. *Built Environment* 30(4): 289–306. doi: [10.2148/benv.30.4.289.57157](https://doi.org/10.2148/benv.30.4.289.57157).
- Weerdenburg D, Scheider S, Adams B, et al. (2019) Where to go and what to do: Extracting leisure activity potentials from Web data on urban space. *Computers, Environment and Urban Systems* 73: 143–156. doi: [10.1016/j.compenvurbsys.2018.09.005](https://doi.org/10.1016/j.compenvurbsys.2018.09.005).
- Wenting R, Atzema O and Frenken K (2011) Urban amenities and agglomeration economies? The locational behaviour and economic success of Dutch fashion design entrepreneurs. *Urban Studies* 48(7): 1333–1352. doi: [10.1177/0042098010375992](https://doi.org/10.1177/0042098010375992).
- Yang Y and Diez-Roux AV (2012) Walking distance by trip purpose and population subgroups. *American Journal of Preventive Medicine* 43(1): 11–19. doi: [10.1016/j.amepre.2012.03.015](https://doi.org/10.1016/j.amepre.2012.03.015).
- Zegras C (2005) *Sustainable Urban Mobility: Exploring the Role of the Built Environment*. Ph.D. Thesis, Massachusetts Institute of Technology, USA.
- Zheng S, Xu Y, Zhang X, et al. (2016) Transit development, consumer amenities and home values: evidence from Beijing's subway neighborhoods. *Journal of Housing Economics* 33: 22–33. doi: [10.1016/j.jhe.2016.05.003](https://doi.org/10.1016/j.jhe.2016.05.003).

Author Biographies

Tabea Sonnenschein is a PhD candidate at the Department of Human Geography and Spatial Planning and the Institute for Risk Assessment Sciences of Utrecht University. She is a former visiting student of Siqi Zheng at the MIT Sustainable Urbanization Lab. Tabea is interested in using computational tools to understand and address urban problems related to social inequality, climate change or public health. To that end, she is exploring Complex Systems Sciences, Agent-based modeling, GIS and other methods. Her PhD work focuses on building Spatial Agent-based Models of Urban Interventions to prevent Cardio-metabolic, pulmonary diseases. Her work contributes to the EXPANSE project.

Simon Scheider is an assistant professor of Geographic Information Science at the Department of Human Geography and Spatial Planning in Utrecht. His research lies at the interface between conceptual modeling, geographic data analysis and knowledge extraction. He is particularly interested in understanding the concepts underlying geodata, including reference systems, fields, places, spatial objects, events, trajectories and their relationship to human activities, and in practical methods for modeling and handling them with GIS and semantic technology. Simon holds a PhD from the Institute of Geoinformatics in Münster, Germany.

Siqi Zheng is the STL Champion Professor of Urban and Real Estate Sustainability at the Department of Urban Studies and Planning, and the Center for Real Estate of the Massachusetts Institute of Technology (MIT). In 2019 she established the MIT Sustainable Urbanization Lab, of which she is also the faculty director. Prof. Zheng's field of specialization is urban and environmental economics with a special focus on China and developing countries. Her research covers topics ranging from environmental sustainability and place-based policies (new planned cities and transportation infrastructure) to self-sustaining urban growth. Her co-authored book "Blue Skies over Beijing: Economic Growth and the Environment in China" was published in 2016.