NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Konstantin A. Kholodilin, Irina Krylova,
Darya Kryutchenko

FINDING THE CONSUMER CENTER OF ST. PETERSBURG?

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: ECONOMICS
WP BRP 165/EC/2017

Finding the consumer center of St. Petersburg?[☆]

Konstantin A. Kholodilin[a,b,*], Irina Krylova[b], Darya Kryutchenko[c]

[a] *DIW Berlin, Germany*
[b] *NRU HSE St. Petersburg, Russia*
[c] *St. Petersburg State University, Russia*

Abstract

In the urban economics, the distribution of people and real estate prices depends on the location of the central business district. As distance from the city center increases, both prices and population density diminish, for travel costs increase in terms of time and money. As manufacturing gradually leaves the cities, the importance of consumer amenities as attractors of population to the urban areas increases. The role of the business center is being taken over by the consumer center. This paper identifies the location of the consumer center of St. Petersburg — the second largest city in Russia and its former capital. For this purpose using data from open sources on the Internet regarding the location of different types of urban amenities, the indices of their spatial density are computed. Using weights based on coefficients of spatial variation and surveys, the individual indices are aggregated to two general centrality indices. Their unique maxima correspond to the city center of St. Petersburg, which is located on Nevsky prospekt, between Fontanka river and Liteinyi prospekt.

Keywords: St. Petersburg; urban amenities; consumer city center; 2D kernel density estimation.

JEL classification: R14; R15; C43.

[☆] The standard disclaimer applies. We thank Anna Smirnova for providing the geocoded and supplied with the city-center-related tags Instagram data. We are also grateful to the participants of the Leontieff Center seminar and scientific seminar of the economic department of the NRU-HSE that took place on April 25 and on April 26, 2017, respectively, in St. Petersburg.

[*] Corresponding author: kkholodilin@diw.de.

Contents

List of Tables

List of Figures

1. Introduction

In the urban economics, the distribution of people and real estate prices depends on the location of the central business district (CBD) of a city; see Alonso (1964), Mills (1967), and Muth (1969). As the distance from the city center increases, prices and population density diminish, reflecting increasing travel costs in terms of time and money.

Glaeser et al. (2001) show that as manufacturing gradually leaves the cities, the importance of consumer amenities as an attractor of population to the urban areas increases. The role of a business center is being replaced by that of the consumer center. While cities once relied on jobs to attract people, urban amenities (restaurants, shops, education opportunities, museums, etc.) are becoming critical. Using a model in which natural and urban amenities play a central role, Brueckner et al. (1999) explain the spatial distribution of different social classes across the city. If the city center has plenty of these amenities, then *ceteris paribus* the rich will be concentrated in the center, while the poor will live on the periphery. Otherwise, the central part of the city will be populated by the low-income families, whereas the high-income households will lodge in the suburbs. Moreover, as Clark (2003) establishes, different types of amenities attract different groups of the population, whose differences are more nuanced than simply income level. For example, college graduates tend to live in settlements with less natural and more urban amenities, while seniors favor more the natural amenities. Inventors are more likely to live in the places where both natural and urban amenities are in abundance.

In the applied literature, the proximity to different natural and urban amenities is considered to be a factor determining the real estate values: Luttik (2000) (green areas, water, and open spaces); Bourassa et al. (2004) (view); Rietveld et al. (2007) and Brandt and Maennig (2012) (railway stations); Ahlfeldt and Maennig (2010) (stadiums).

Relatively few studies apply the hedonic approach to Russian data. Most of them focus on Moscow. Магнус and Пересецкий (2010) analyze the determinants of the asking housing prices in Moscow. Two spatial variables are used: travel time to the next subway station and distance from the nearest subway station to the city center. The authors set the center of Moscow to be Red Square, based on the circular shape of the Russian capital city. Красильников and Щербакова (2011a) estimate an hedonic model using data on the asking prices of dwellings in St. Petersburg. This study uses the same two spatial variables. The coordinates of the city center are computed by averaging the coordinates of all dwellings in their sample. The estimated city center is located in the Peter and Paul Fortress. Красильников and Щербакова (2011b) employ similar methodology in order to identify city centers in their study on four Russian metropolises (Moscow, Novosibirsk, St. Petersburg, and Yekaterinburg). Similarly, Катышев and Хакимова (2012) use in their analysis of housing prices in Moscow two variables of proximity to the center and to the closest subway station. As the center of Moscow they take the subway station Okhotnyi ryad, which is about 0.5 km from Red Square. In addition, since the focus of their study is on the environmental quality, they also consider the distance to the nearest factory. Чугунов (2013) uses a much wider list of amenities to assess their impact on the housing prices in Moscow: 1) the distance to the nearest subway station; 2) the distance to the secondary schools and their quality (measured by the performance of the pupils); 3) the number of parks; 4) the number of sports facilities; 5) the number of health care institutions per 10,000 persons; and 6) the number of municipal police units per 10,000 persons. In contrast, this study does not employ any measure of proximity to the city center, capturing the spatial heterogeneity of prices by the district dummies. Носов and Цыпин (2015) investigate the determinants of asking prices for one-room apartments in a medium-size Russian city Orenburg. In order to capture spatial factors they take advantage of spatial clusters obtained by the k-means clustering technique and of the distance to the city center, which is defined as the central post office of the city. In Russia, post offices are typically used as departure points to measure the distances. Kholodilin and Ulbricht (2015), who estimate hedonic regressions for 48 large European cities, including seven Russian cities (Kazan, Moscow, Nizniy Novgorod, Rostov on Don, Samara, St. Petersburg, and Yekaterinburg), capture spatial effects only by district dummies.

As the amount and variety of information published on the Internet increase, the possibilities of exploiting it to measure the natural and urban amenities at the microlevel (individual parks, shops, restaurants, etc.) expand extraordinarily. For example, Ahlfeldt and Wendland (2016) suggest a method of computing the so-called potential spaces taking in account the geographical distribution of different natural and urban amenities objects.

The aim of this paper is to develop a simple and easily applicable method of delineating the consumer city center. As an example, it is used to identify the exact coordinates of the consumer center of St. Petersburg, Russia, i.e., the point with the highest density of consumer urban amenities in the whole city. This information can be used for different purposes. For example, it can be employed in the hedonic analysis of the housing prices and rents, where both a detailed evaluation of the impact of different urban amenities and a compact representation of all the relevant amenities by a single index are desirable. Furthermore, when assessing accessibility, it is critically important to know where the city center is located. In our case, the center of St. Petersburg identified in this paper can be used to construct the isochrones[1] (equal travel time curves), which require choosing the coordinates of the city center. Finally, the estimation of the population density gradient requires an exact knowledge of the central business district location. If the center coordinates are misspecified, then, as Alperovich and Deutsch (1992) demonstrated, this can lead to an underestimation of the gradient.

St. Petersburg is the second largest city in Russia and its former capital. During the 20th century, it underwent many dramatic changes related to wars and revolutions. In 1918, after having served as the capital of the vast Russian Empire for two centuries, it became a regionally important city. Since then, three times St. Petersburg has had its population drastically decrease: it lost more than half of its population during both the Russian Civil War of 1918–1920 and World War Two; and during the 1990s, as a result of radical socioeconomic and political transformations, St. Petersburg lost 500,000 citizens, ending the decade with 4.5 million residents. It is only during the early 2000s that the city managed to recover in terms of population, exceeding 5.2 million in 2016. In the 1930s, there was a plan to displace the political center of the city from the neighborhood of the Winter Palace[2], to the south by about 11 km, in the direction of Moscow. However, the entry of Russia into WWII made this plan obsolete. Overall, the central planning system that was in place in Russia between 1917 and 1990 tried to spatially distribute amenities in a planned manner in accordance with its non-market principles. Despite all these changes, the city kept many of the cultural values accumulated over the years in form of palaces, museums, and theaters. Its historic center is a UNESCO World Heritage Site. The transition to a market economy that started in the early 1990s led to a rapid increase in amenities, especially shops and restaurants.

[1]See, e.g., Kholodilin (2016).

[2]This was once the official residence of Russian monarchs. Today it is The Hermitage Museum.

The paper has following structure. Section 2 reviews the literature on delineating city centers. Section 3 introduces the method of finding the location of city center used in this paper and describes the underlying data. In section 4, the estimated coordinates of the consumer city center are contrasted with alternative estimates based on different techniques and data. Finally, section 6 concludes.

2. Approaches to delineating the city center

Despite the importance of the notion of CBD in urban and housing research, as a rule, in the literature, its location is arbitrarily chosen. Typically, the choice of its coordinates is not justified.

At the same time, there is an extensive literature devoted to determining the city center (see Table 1 for a concise overview), which can be divided in two unequal groups: a couple of studies by urban economists and many works by economic geographers.

Urban economics: One of the first urban economists addressing this issue is Alperovich (1982). He uses the population density gradient model in order to identify the location of the CBD of Tel Aviv-Yafo. Departing from the hypothesis of diminishing population density as the distance to the center becomes larger and using different functional forms modeling this relationship, he undertakes a grid search and chooses from all the candidates the point, for which the adjusted R^2 is maximized. He uses data on population at the level of census tracts, which produce a detailed picture of the geographical distribution of population density. However, such information is not always available. Moreover, the census tract boundaries are predetermined and do not reflect actual local housing market areas.

Alperovich and Deutsch (1994) suggest an approach, which estimates the coordinates of the CBD by including them as unknown parameters in the econometric model and applying the maximum likelihood method. This allows determining not only the region, to which the CBD belong, but also the CBD's precise point coordinates. In addition, it is possible to test various hypotheses about the location of the center. For example, one can test whether the CBD is shifting in space due to the changing structure of the city. This approach permits to flexibly model the potential nonlinearities using the Box-Cox transformations.

Economic geography: Economic geographers represent an independent and a very different strand of the literature. One of the first studies to address the question of delineating a city center is Murphy and Vance (1954), which employs land use data. In particular, two indicators are used: 1) the total space to ground floor ratio and 2) the "central business use" space to ground floor area at the block level. For many decades, the approach of Murphy and Vance (1954) dominated economic geography. After nearly 60 years, it was modified by Taubenböck et al. (2013), who use detailed data on the intensity of the land use taken from both open sources and satellite pictures, then applying morphological 3D modeling of the land use at the level of blocks with the object of delimiting the CBD of Paris. This method, with its objectivity and flexibility, is very data demanding and computation intensive. Furthermore, its applicability depends to large extent on the country and regional differences in the heights of buildings, which are determined, for instance, by the ground or by legal height restrictions. Buildings within the historical center of St. Petersburg are subject to legal restrictions respecting height, among other conditions.

Thurstain-Goodwin and Unwin (2000) suggest an innovative approach that subjects sectoral employment data attached to the centroids of the postal code districts to spatial smoothing using 2D kernel density estimation (KDE). The resulting empirical functions of spatial density are aggregated into a single index by computing their weighted average. It should be noted that the weights are determined arbitrarily, a weakness of the approach. This method is improved by Borruso and Porceddu (2009) and Lüscher and Weibel (2013) in terms of both the input data and the weighting scheme. Borruso and Porceddu (2009) collect microlevel data on different activities (clothing; arts and culture; banks and insurance companies; retail; etc.) from the Yellow pages and georeference each establishment. Then, a KDE of all these features taken together is done. Based on the resulting isolines the city center is delineated using three standard deviations as a threshold. We find, however, that mixing together different urban amenities is difficult to justify. Various amenities have different frequencies: for instance, there many more shops than theaters. At the same time, some amenities are more typical of a center than others. When mixed in a single data set, the amenities that are less typical of a center, but occurring more frequently overall, can have a larger impact on the estimated location of the city center, thus biasing the resulting coordinates. Lüscher and Weibel (2013) use point-of-interest data, that is, microlevel information on commercial establishments (accommodation;

eating and drinking; attractions; etc.) supplied by official UK bodies. The authors conduct an internet survey to identify whether features should be considered as typical or atypical for a city center. Based on the results, they determine weights for each feature. For each feature a 2D KDE is carried out. The resulting smoothed spatial distributions are aggregated using the survey-based weights. Finally, the boundaries of city centers are determined using as the area, for which the computed city center typicality exceeds 0.5. The computed centers are compared to the "comparative centers" based on alternative representations (tourist maps, Wikipedia, and Flickr).

Apart from the density of urban amenities, other density indicators have attracted the attention of researchers. Hollenstein and Purves (2010) take advantage of the tagged and georeferenced images from the photography website Flickr.com. The city center boundaries are obtained through a KDE of the locations of the pictures tagged as referring to the inner city (downtown, cbd, central, innercity, citycenter). Sun et al. (2016) use location-based social networking (LBSN) data. They take advantage of the fact that georeferenced and time-stamped "check-ins" (sometimes referred to as a type of volunteered geographic information) represent the displacements of the LBSN users and tend to be clustered in space, especially where commercial facilities (shops, restaurants, cinemas, etc.) abound. Therefore, these LBSN mobility data can serve as an indicator of the LBSN users' mobility. The data are collected from Gowalla, a LBSN. Clusters of point data are constructed and the boundaries of city center are defined as the boundaries of the Voronoi polygons around the points belonging to the largest cluster.

A very unusual approach employed by Montello et al. (2003) asked people on the street to draw the boundaries delimiting, on a paper map, where they are 50% and 100% confident downtown is located. The intersection of the hand-drawn maps can be considered as a conventionally defined city center.

To the best of our knowledge, there is only one study (Kolossov et al. 2002) that determines location of the central business district for a Russian city, namely Moscow. Using the addresses of businesses (the stores selling non-food products; the main offices and branches of banks; and the companies providing business services) compiled from reference/ informational and advertising publications the authors compute the density of these firms at the level of the nearly 500 post-office districts of Moscow. By informally aggregating information on the agglomeration of retail trade, finance as well as producer and business services the researchers identify the existing and potential business districts of the Russian capital as of 2000.

3. An index of urban amenities

3.1. Data

In order to identify the consumer center of St. Petersburg, Russia, it is necessary to identify the locations of each type of urban amenity. With these data, it is possible to compute the spatial density indices for each type. These indices can be interpreted as an accessibility measure for different amenities at any point within the city.

The sources of data and the number of observations of each type of amenities are reported in Table 2. Overall, 18 types of urban amenities are considered: banks, cinemas, fitness clubs, food stores, healthcare establishments (medical centers, hospitals, dental clinics, women's consultation clinics, early treatment centers, etc.), hairdressers, kindergartens, lawyers, museums, notaries, pharmacies, schools, shops (shoes, cloths, jewelry, etc.), restaurants, shopping malls, places of worship, and theaters. The largest number of observations is available for the shops (7139), while the smallest number is for the cinemas (90). The data were collected from various websites containing information (name, type of establishment, its geographical coordinates, and sometimes its price range as well as the rating based on the client votes) about different specialized individual establishments.

3.2. Spatial densities of urban amenities

In accordance with the literature, in order to estimate the spatial density of each amenity, we took advantage of the two-dimensional kernel density estimation method.[3] Following Borruso and Porceddu (2009), we split the city in the squares approximately 200 m per side and obtained a grid with 127,500 cells (375×340). The window size, BW, was determined separately for each coordinate using the following rule of thumb:[4]

$$BW = 4 \times 1.06 \times \min\{\sigma_x, h\} \times N^{-\frac{1}{5}} \tag{1}$$

[3]The spatial smoothing across both coordinates was done using the function *kde2d* from the package *MASS* of the statistical and graphical programming language **R**.

[4]See Venables and Ripley (2002), equation (5.5) on page 130.

where σ_x is the standard deviation of the variable x (in this case, the variable is either longitude or latitude of individual objects); N is the number of observations of variable x.

$$h = \frac{Q_3 - Q_1}{1.34} \tag{2}$$

where Q_1 and Q_3 are the first and third quartiles of the variable x, respectively.

The smoothed spatial distribution of selected urban amenities is depicted in Figures 1 and 2. In accordance with Christaller (1980), p. 28 and pp. 54–63, who distinguishes between the higher and lower order central goods, each amenity has a different degree of centrality. Figure 1 shows the four most decentralized amenities (schools, kindergartens, pharmacies, and hairdressers), while Figure 2 displays the four most centralized amenities (museums, notaries, restaurants, and theaters). For example, while the schools are widely scattered over the territory of the city, the museums are mainly concentrated in the historical districts of the city (Admiralteiskiy, Vasileostrovskiy, Petrogradskiy, and Tsentral'nyi) and have a clear cut center. Such differences are easy to understand when the nature of the services provided by, for example, the pharmacies and the theaters is taken into account. The former satisfy more basic needs and, hence, must be located close to the customers, while the latter are aimed at satisfying higher order requirements of a much more limited group of customers. Moreover, the geographical distribution of many theaters is determined by their history: prior to the October 1917 revolution, it was mainly higher income individuals attending theater performances; consequently these were built close to the neighborhoods where such persons lived. At that time, most nobles had their palaces close to the imperial palace.

Figure 3 shows the centers of the smoothed spatial distribution of individual urban amenities. Most of them are clustered together in Tsentral'ny district. Three (education, sports, and food stores) are located more to the west, in Admiralteiskiy district.

The location of the consumer center of St. Petersburg is identified using a composite index of spatial density of the urban amenities, obtained by aggregating the smoothed spatial densities of individual amenities. The composite index is calculated as follows:

$$ADI_{ij} = \sum_{k=1}^{K} \widetilde{AD}_{ij}^{k} w_k \tag{3}$$

where

$$\widetilde{AD}_{ij}^{k} = \frac{AD_{ij}^{k} - \mu_k}{\sigma_k},$$

AD_{ij}^{k} is the spatial density of the k-th type of amenity in the cell ij; μ_k is the mean of smoothed spatial distribution of k-th amenity; σ_k is the standard error of amenity k; and w_k is the weight of this type. The weights are normalized in such a way that they sum to 1.

Variation-based weights: One way to determine the weights is to base them on the coefficient of variation of the smoothed spatial density. Indeed, the amenities that have clearly identified peaks are characterized by higher spatial density variance and, thus, higher coefficient of variation. By contrast, when amenities are evenly distributed across the surface of the city, then the coefficient of variation should be equal zero. Hence, the weights are computed as follows:

$$w_k = \frac{{}^{\sigma_k}\!/_{\mu_k}}{\sum_{k=1}^{K} {}^{\sigma_k}\!/_{\mu_k}} \tag{4}$$

where σ_k is the standard deviation of the spatial density of amenity k and μ_k is the mean of the spatial density of the amenity.

Table 3 reports coefficients of variation of individual amenities and corresponding weights.

The composite index of the spatial density of the urban amenities is shown in Figure 4. The darker shading corresponds to a higher density. The maximum of the index is attained at Nevsky prospekt, between Liteynyi and Ligovskii prospekts.

Survey-based weights: The second way to compute the weights is that of Lüscher and Weibel (2013). In order to obtain user defined weights, we conducted a survey in January – March 2017 in St. Petersburg. The survey consists of 10 questions, falling into two broad categories: 1) individual characteristics of the respondents (age, gender, educational level, size of the settlement of origin, and their nearest crossroads); 2) characteristics of the city center (what do the respondents associate with the city center as well as which amenities they find typical for the city center and which are not).

We received 140 correctly filled questionnaires. Since the survey was conducted mostly among the students of the National Research University – Higher School of Economics St. Petersburg, the share of young persons (aged between 18 and 24) exceeds 74%. In addition, females make up two-thirds of the respondents. Over 46% of respondents are university graduates. Finally, 35% of respondents come from St. Petersburg, almost 14% are from other cities with population exceeding 1 million, about 47% are from smaller cities, and slightly more than 4% are from the countryside.

Figure 5 shows the survey-based weights of 12 amenity groups. Cultural amenities and restaurants are perceived as the most typical for the city center, while the sports and health care amenities are thought to be the most atypical ones.

The general amenities index with survey-based weights is computed as:

$$ADI_{ij}^{S} = c_m \times \sum_{k=1}^{K} AD_{ij}^{k} w_k^{S} + c_a \tag{5}$$

where

$$c_m = \frac{1}{\sum_{k=1}^{K} |w_k^S|} \quad \text{and} \quad c_a = c_m \times \sum_{k=1}^{K^-} |w_k^S| \tag{6}$$

are normalization constants ensuring that $0 \leq ADI_{ij}^{S} \leq 1$; while $k = 1, 2, \ldots, K^-$ are the indices of the negative survey-based weights.

In cases when a category includes more than one individual amenity type, these types are aggregated to the category index using simple averaging. The correspondence between amenity types and categories is shown in Table 2.

Similar to Lüscher and Weibel (2013), the area-like amenities (open green spaces) are transformed into spatial densities by computing the share of the land devoted to the green areas within a circular window of 240 m radius around the center of each raster cell.

The resulting general amenities index is displayed in Figure 6. The distance between both estimates of consumer city center is 1.1 km. The survey-based center is located more to the north than the variation-based center; see black and green dots in Figure 7.

4. Validation of the results

In order to check the robustness of our results we use several alternative methods of finding the city center location and delineating the city center as a region.

4.1. Gradient method

First, we use the approach suggested in Alperovich and Deutsch (1994). For this purpose the following two regressions were estimated using the method of the maximum likelihood:

Exponential regression

$$p_i = \alpha e^{-\gamma d\Big((\theta_1,\theta_2),(c_{i1},c_{i2})\Big)+\varepsilon_i} \tag{7}$$

where p_i is the population or employment density in the i-th municipal district (MD); α is a parameter measuring the population or employment density in the CBD; γ is the so-called density gradient that describes the diminishing population or employment density as the distance from the CBD increases; $d\Big((\theta_1, \theta_2), (c_{i1}, c_{i2})\Big)$ is the distance in km between the CBD and the centroid of the district i (θ_1 and θ_2 are the longitude and latitude of the CBD, while c_{i1} and c_{i2} are the coordinates of the centroid of the i-th MD); ε_i is the error term.

Box-Cox regression

$$\frac{p_i^{\lambda} - 1}{\lambda} = \alpha + \gamma d\Big((\theta_1, \theta_2), (c_{i1}, c_{i2})\Big) + \varepsilon_i \tag{8}$$

where λ is a nonlinearity parameter determining the functional form of the regression. When $\lambda \to 0$, the model tends to an exponential form, while when $\lambda = 1$ it takes the linear form.

The data on the area, population, and employment in all 111 municipal districts of St. Petersburg, as of 2015, are taken from the *Database of the municipal district indicators* of the St. Petersburg's statistical office Petrostat.[5] The distribution of population and employment density by municipal districts is shown in the upper panels of Figure 8. This graph displays the population and employment density by the municipal districts of St. Petersburg. The darker the shading, the higher the density. The highest employment density is observed in the historical center of the city and gradually declines toward the city periphery. For the population, the picture is not that clear cut, for there are some municipal districts with high population density that are relatively far from the city center.

The results of estimation of models (7) and (8) for the population and employment are reported in Table 4. In all cases, the density gradient is negative and statistically significant, which implies that the population and employment densities decay as the distance from the CBD

[5]http://petrostat.gks.ru/wps/wcm/connect/rosstat_ts/petrostat/ru/statistics/Sant_Petersburg/db/

increases. The gradient varies between -0.12 and -0.19. The estimated nonlinearity parameter λ is 0.190 for the population and 0.105 for employment. This means that in the former case, the Box-Cox transformation model is more different from the exponential specification than in the latter case; see Figure 8, which depicts the dependence between population density (panel (c)) and employment density (panel (d)), on the one hand, and distance to the CBD, on the other. Both functions (exponential and Box-Cox) differ in their middle part. As the distance to the CBD increases from 0 to 10 km, the population density diminishes from over 200 to 60 persons/ha for the exponential specification and from 160 to 110 persons/ha for the Box-Cox one. Employment density decreases from over 60 to 10 persons/ha for the exponential and from 50 to 25 persons/ha for the Box-Cox specification.

The parameters θ_1 and θ_2 are the longitude and latitude of the CBD, respectively. The CBDs identified using the Alperovich and Deutsch (1994) method are shown together with the variation- and survey-based consumer centers in Figure 7. All these center estimates lie in Tsentral'nyi (Central) district. Moreover, all fall within the boundaries of the UNESCO World Heritage Site that makes up the historic center of St. Petersburg; the area denoted in orange on the map. The distance between the variation-based consumer center and CBDs is 1) approximately 0.4 and 0.8 km (exponential and Box-Cox specifications, correspondingly) for the population and 2) about 1.2–1.4 km for employment. The distance between the survey-based consumer center and CBDs is 1) approximately 1.3 and 1.9 km (exponential and Box-Cox specifications, correspondingly) for the population and 2) around 1.8 km for employment, similarly for both specifications.

Thus, the centers of St. Petersburg determined using different techniques are relatively close to each other. The differences between them are related, firstly, to the fact that the CBD reflects the concentration of the productive activities, while consumer center characterizes the consumption opportunities. Secondly, they are based on different information (population/employment vs. urban amenities) at different aggregation levels (municipal districts vs. individual amenity objects). The variation-based consumer center has the shortest average distance to all other center estimates, unlike the survey-based one.

4.2. *Place of living method*

Our second alternative approach is based on the survey that we conducted in St. Petersburg in January – March 2017. We asked the respondents to judge whether they live in the city center, between the center and the periphery, or in the periphery. They also had to name the closest crossroads to their place of living. The purpose of this question was to determine, more or less, the respondent's home coordinates without making the respondents reveal potentially sensitive information regarding their personal address. Figure 9 shows the corresponding crossroads classified into the three categories according to the proximity to the city center. The city center category denoted by red letter C represents the union of individual estimates of the center and, thus, the largest estimate of central area.

4.3. *Drawing method*

The third approach follows that of Montello et al. (2003). During the survey, the same respondents were asked to draw on the map they were provided with two regions delineating what is in their opinion the city center with probability of 100% and 50%. The rest of the map was coded as having the zero probability of being the city center. The survey respondents supplied 51 maps with two central regions. The individual maps were geocoded[6], rasterized using a 1000×1000 grid, and averaged. As a result we obtain a consensus region that most of survey participants who drew the map perceive as being the St. Petersburg city center with the highest probability; see Figure 10. It is surrounded by areas with gradually declining probability of centrality. Both estimates of the consumer center belong to the region with the highest probability of being central. The point with maximum probability of being city center assigned by the survey participants is located between both consumer center estimates, slightly to the west. Consumer center estimates do not belong to the area of the highest probability of being central, but are located very close to it.

4.4. *User-generated content method*

The fourth center delineation method we use follows Hollenstein and Purves (2010). Here, we take advantage of the geocoded and tagged pictures from Instagram, the most widely used

[6] Geocoding was done using QGIS.

photography website in Russia. From the over 8 million photographs taken by persons whose city of origin is denoted as St. Petersburg we selected those having geographical coordinates and such tags as "center", "centre", or "центр" (which is "city center" in Russian). Over two-thirds of the persons who posted their photos are females and 81% of those whose age is known are younger than 30.[7] The smoothed spatial distribution of these photographs is shown Figure 11. It is very concentrated along the historical avenue of the city, Nevsky prospekt. The peak of the distribution is located at the subway station "Kanal Griboyedova", being about 1 km to the west of the variation-based consumer center. However, the latter belongs to a local peak of the spatial distribution of Instagram center-related pictures.

5. The relevance of this study

The knowledge of exact location of the city center is important due to various reasons. First of all, the results of both the empirical and theoretical urban literature depend crucially upon where exactly the city center is. To name just a few topics: intra-urban wage gradients (Eberts 1981 and Ihlanfeldt 1992), gradients of built-up density (Guérois and Pumain 2008), and urban business locations (Stahl 1987 and Egan and Nield 2000).

Furthermore, information about the city center can be useful for the urban planners. For example, during the transition from the centrally planned to the market economy, the spatial distribution of consumer amenities changed a lot in Russian cities. This was a consequence of the abandonment of the Soviet system of the decentralized distribution of jobs, a surge in the service sector, a tremendous rise in the amount of privately owned real estate, and a rapid increase in the motorization of the population. Even in an established market economy, the location of a city center is not fixed: it can move over time across the space. The shifts of the city center bring about changes in the spatial distribution of employment, which affects traffic flows (with an increased traffic volume along the roads leading to the new city center) that need to be accommodated by the city planners; see Ingram (1998).

The information on the spatial distribution of urban amenities collected during this study can also be used to identify the areas of the city that are undersupplied by the amenities (Páez

[7] The demographic characteristics of the Instagram users are obtained from their VK (the largest Russian social network, an analog of Facebook) account linked to their Instagram account.

et al. 2010). This can be helpful in designing the place-based policies, that is, government efforts to enhance the performance of disadvantaged areas, including individual city neighborhoods (Neumark and Simpson 2014).

This information about the way amenities are distributed across the city can also be valuable to other market participants like the owners of restaurants, retailers, hoteliers, etc., since they need data on spatial distribution of urban amenities in order to optimally locate their own establishments.

6. Conclusion

In this study, using microdata about various objects of urban consumer amenities collected on the Internet, we constructed individual indices of the spatial density of urban amenities in St. Petersburg, Russia. These indices are aggregated to two alternative composite indices of centrality depending on the way the weights are computed. Given that the centrality index with variation-based weights is, on average, the closest to the alternative center estimates, we took it as a proxy for the location of the consumer center of St. Petersburg. This consumer center is located on Nevsky prospekt, between Fontanka river and Liteinyi prospekt.

The results of the amenities-based approach were cross-checked using four alternative techniques: population/employment density gradients, place of living approach, map drawing method, and user-generated content method. All in all, the method employed in this paper produces plausible results that are confirmed by other approaches using different techniques and different data.

The findings of this paper can be used, for example, in the hedonic regressions of housing prices and rents as well as in the analysis of the determinants of the spatial distribution of employment and economic activities in St. Petersburg. Moreover, it is also applicable in an historical analysis of the urban economy in general and housing market in particular.

References

Ahlfeldt, G. M. and W. Maennig (2010). Impact of sports arenas on land values: Evidence from Berlin. *Annals of regional science 44*(2), 205–227.

Ahlfeldt, G. M. and N. Wendland (2016). The micro-cosmos of German cities: New insights into the supply of and demand for urban amenities. In T. Just and W. Maennig (Eds.), *Understanding German Real Estate Markets*, pp. 79–88. Cham: Springer International Publishing Switzerland.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Cambridge: Harvard University Press.

Alperovich, G. (1982). Density gradients and the identification of the central business district. *Urban Studies 19*, 313–320.

Alperovich, G. and J. Deutsch (1992). Population density gradients and urbanization measurement. *Urban Studies 29*, 1323–1328.

Alperovich, G. and J. Deutsch (1994). Joint estimation of population density functions and the location of the central business district. *Journal of Urban Economics 36*(3), 239–248.

Borruso, G. and A. Porceddu (2009). A tale of two cities: Density analysis of CBD on two midsize urban areas in Northeastern Italy. In B. Murgante, G. Borruso, and A. Lapucci (Eds.), *Geocomputation and Urban Planning*, pp. 37–56. Berlin, Heidelberg: Springer-Verlag.

Bourassa, S. C., M. Hoesli, and J. Sun (2004). What's in a view? *Environment and Planning A 36*(8), 1427–1450.

Brandt, S. and W. Maennig (2012). The impact of rail access on condominium prices in Hamburg. *Transportation 39*(5), 997–1017.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999). Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review 43*, 91–107.

Christaller, W. (1980). *Die zentralen Orte in Süddeutschland. Eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Clark, T. N. (2003). Urban amenities: Lakes, opera, and juice bars do they drive development? In T. N. Clark (Ed.), *The City as an Entertainment Machine*, pp. 103–140. New York: JAI Press/Elsevier.

Eberts, R. W. (1981). An empirical investigation of intraurban wage gradients. *Journal of Urban Economics 10*(1), 50–60.

Egan, D. J. and K. Nield (2000). Towards a theory of intraurban hotel location. *Urban Studies 37*(3), 611–621.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of Economic Geography 1*(1), 27–50.

Guérois, M. and D. Pumain (2008). Built-up encroachment and the urban field: a comparison of forty European cities. *Environment and Planning A 40*(9), 2186–2203.

Hollenstein, L. and R. Purves (2010). Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science 1*, 21–48.

Ihlanfeldt, K. R. (1992). Intraurban wage gradients: Evidence by race, gender, occupational class, and sector. *Journal of Urban Economics 32*(1), 70–91.

Ingram, G. K. (1998). Patterns of metropolitan development: what have we learned? *Urban studies 35*(7), 1019–1035.

Kholodilin, K. (2016). War, housing rents, and free market: A case of Berlin's rental housing market during World War I. *European Review of Economic History 20*, 322–344.

Kholodilin, K. A. and D. Ulbricht (2015). Urban house prices: A tale of 48 cities. *Economics – The Open-Access, Open-Assessment E-Journal 9*, 1–43.

Kolossov, V., O. Vendina, and J. O'Loughlin (2002). Moscow as an emergent world city: international links, business developments, and the entrepreneurial city. *Eurasian Geography and Economics 43*(3), 170–196.

Lüscher, P. and R. Weibel (2013). Exploiting empirical knowledge for automatic delineation of city centres from large-scale topographic databases. *Computers, Environment and Urban Systems 37*, 18–34.

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Landscape and Urban Planning 48*(3–4), 161–167.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review 57*(2), 197–210.

Montello, D. R., M. F. Goodchild, J. Gottsegen, and P. Fohl (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation 3*(2 and 3), 185–204.

Murphy, R. E. and J. E. Vance (1954). Delimiting the CBD. *Economic Geography 30*(3), 189–222.

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.

Neumark, D. and H. Simpson (2014). Place-based policies. Technical report, National Bureau of Economic Research.

Катышев, П. К. and Ю. А. Хакимова (2012). Экологические факторы и ценообразование на рынке недвижимости (на примере г. Москвы). *Прикладная эконометрика 4*(28), 113–123.

Красильников, А. А. and А. А. Щербакова (2011a). Детерминанты цены на вторичном рынке недвижимости Санкт-Петербурга. *Экономические науки 11*(84), 93–99.

Красильников, А. А. and А. А. Щербакова (2011b). Ценообразование на вторичном рынке жилья в мегаполисах России. *Экономические науки 12*(85), 103–106.

Магнус, Я. Р. and А. А. Пересецкий (2010). Цены квартир в Москве. *Прикладная эконометрика 1*, 89–105.

Носов, В. В. and А. П. Цыпин (2015). Эконометрическое моделирование цены однокомнатной квартиры методом географически взвешенной регрессии. *Известия Саратовского университета. Новая серия. Серия Экономика. Управление. Право 15*(4), 381–387.

Чугунов, Д. Ю. (2013). Влияние факторов качества образования и социального окружения на стоимость жилья в Москве. *Журнал Новой экономической ассоциации 1*(17), 87–112.

Páez, A., R. Gertes Mercado, S. Farber, C. Morency, and M. Roorda (2010). Relative accessibility deprivation indicators for urban settings: definitions and application to food deserts in Montreal. *Urban Studies 47*(7), 1415–1438.

Rietveld, P., G. Debrezion, and E. Pels (2007). The impact of railway stations on residential and commercial property value: A meta-analysis. *The Journal of Real Estate Finance and Economics 35*(2), 161–180.

Stahl, K. (1987). Therories of urban business location. *Handbook of regional and urban economics 2*, 759–820.

Sun, Y., H. Fan, M. Li, and A. Zipf (2016). Identifying the city center using human travel flows generated from location-based social networking data. *Environment and Planning B: Planning and Design 43*(3), 480–498.

Taubenböck, H., M. Klotz, M. Wurm, J. Schmieder, B. Wagner, and T. Esch (2013). Delimiting central business districts — a physical approach using remote sensing. In *Proceedings of the JURSE 2013*, pp. 17–20.

Thurstain-Goodwin, M. and D. Unwin (2000). Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations. *Transactions in GIS 4*(4), 305–317.

Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S.* New York: Springer Verlag.

Appendix

Table 1: Approaches to delineating the city center

Paper	Data	Method	Number of cities (country)
Murphy and Vance (1954)	land use by blocks	indices of intensity and nature of land use	9 (USA)
Alperovich (1982)	population by census tracts	density gradient model, OLS	1 (Israel)
Alperovich and Deutsch (1994)	population by census tracts	density gradient model, exponential and Box-Cox	1 (Israel)
Thurstain-Goodwin and Unwin (2000)	sectoral employment by postal code districts from official statistics	2-d kernel density estimation and aggregation of the empirical density functions using arbitrary weights	1 (UK)
Montello et al. (2003)	street survey	drawing of city center by the respondents	1 (USA)
Borruso and Porceddu (2009)	urban amenities objects from Yellow pages	2-d kernel density estimation for all amenities simultaneously	2 (Italy)
Hollenstein and Purves (2010)	geocoded and tagged pictures from Flickr	2-d kernel density estimation	1 (Switzerland), 2 (UK), 2 (USA), and 1 (Australia)
Lüscher and Weibel (2013)	urban amenities objects from official statistics	2-d kernel density estimation and aggregation using survey-based weights	10 (UK)
Taubenböck et al. (2013)	building footprints and street networks from the OpenStreetMap as well as 2- and 3-d satellite pictures	3-d city model and blockwise classification using fuzzy logic	1 (France)
Sun et al. (2016)	check-ins in social network Gowalla	cluster analysis and Voronoi polygons	3 (Germany)

Table 2: Data sources

Amenity	Number of observations	Website	Category
Banks	874	www.banki.ru	Offices, notaries, banks
Cinemas	90	www.afisha.ru	Culture
Fitness clubs	224	http://sportgyms.ru/	Fitness, pools
Food stores	1655	www.yp.ru	Food stores
Hairdressers	1994	www.yp.ru	Beauty, hairdressers
Health care	2082	www.spbmed.info	Health care
Kindergartens	1028	http://detsadi-spb.ru	Education
Lawyers	425	www.yp.ru	Offices, notaries, banks
Museums	221	www.afisha.ru	Culture
Notaries	314	sanktpeterburg.tradeis.ru	Offices, notaries, banks
Open spaces	6975	http://data.nextgis.com/	Green spaces
Pharmacies	1026	www.spbmed.info	Health care
Restaurants	3138	www.restoclub.ru	Restaurants
Schools	635	http://apeterburg.com	Education
Shopping malls	176	http://peterburg2.ru	Shopping malls
Shops	7139	www.shopping-spb.su	Cloth and shoe stores
Temples	351	http://temples.ru/	Temples
Theaters	349	www.afisha.ru	Culture

Table 3: Variation-based weights of individual amenities

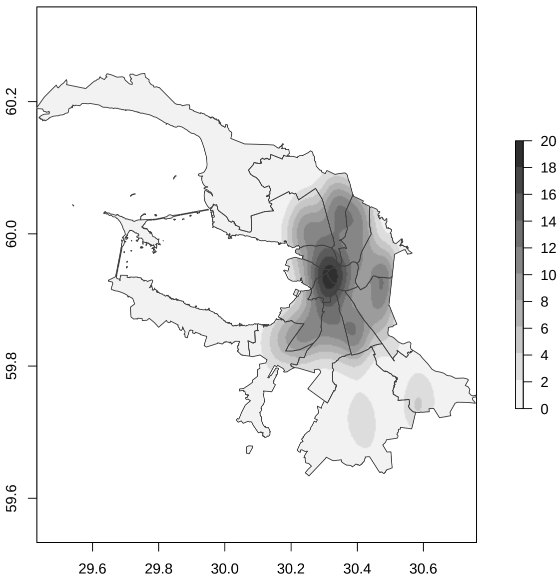
Amenity	Variation coefficient	Weight
Schools	2.421	0.034
Kindergartens	2.524	0.035
Pharmacies	2.573	0.036
Hairdressers	2.678	0.037
Temples	2.699	0.037
Cinemas	2.759	0.038
Food stores	2.832	0.039
Shopping malls	2.878	0.040
Fitness clubs	2.986	0.041
Health care	3.475	0.048
Lawyers	4.550	0.063
Banks	4.990	0.069
Shops	5.540	0.077
Museums	6.421	0.089
Notaries	6.430	0.089
Restaurants	7.774	0.108
Theaters	8.532	0.118

Table 4: Employment and population density estimation results

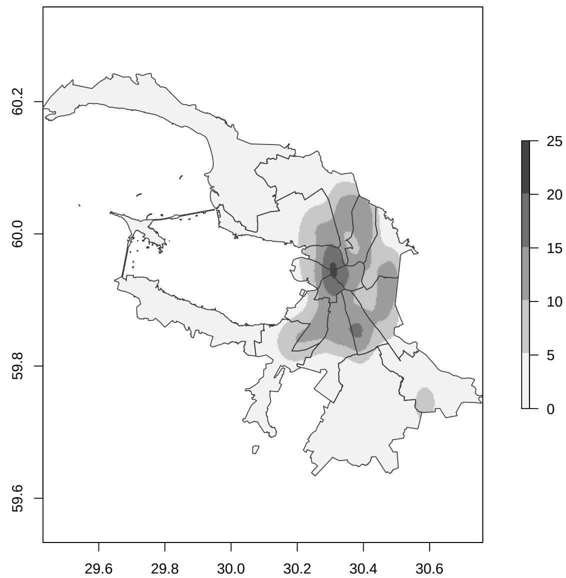
Coefficient	Dependent variable: population density		Dependent variable: employment density	
	Specification: exponential	Specification: Box-Cox	Specification: exponential	Specification: Box-Cox
	(1)	(2)	(3)	(4)
α	5.280^{***}	8.531^{***}	4.045^{***}	4.879^{***}
	(0.154)	(1.391)	(0.189)	(0.35)
γ	-0.115^{***}	-0.191^{***}	-0.140^{***}	-0.163^{***}
	(0.008)	(0.036)	(0.01)	(0.014)
σ	1.051^{***}	1.991^{***}	1.278^{***}	1.491^{***}
	(0.071)	(0.449)	(0.086)	(0.136)
θ_1	30.348^{***}	30.352^{***}	30.324^{***}	30.321^{***}
	(0.029)	(0.031)	(0.016)	(0.016)
θ_2	59.926^{***}	59.933^{***}	59.934^{***}	59.935^{***}
	(0.013)	(0.014)	(0.008)	(0.008)
λ		0.190^{***}		0.105^{***}
		(0.059)		(0.029)
Observations	111			

Note 1: *p<0.1; **p<0.05; ***p<0.01
Note 2: Numbers in parentheses are standard errors of the estimated parameters.

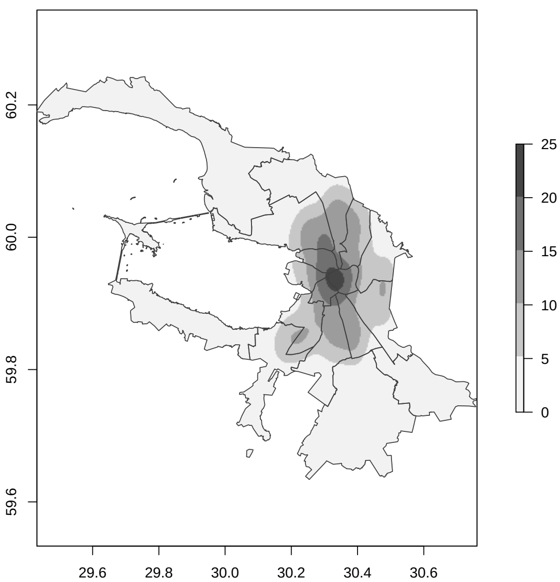
Figure 1: Amenities with lowest spatial concentration



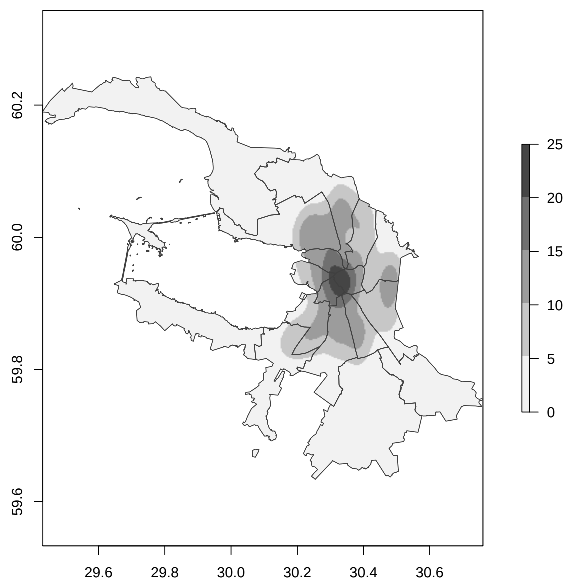
(a) schools



(b) kindergartens

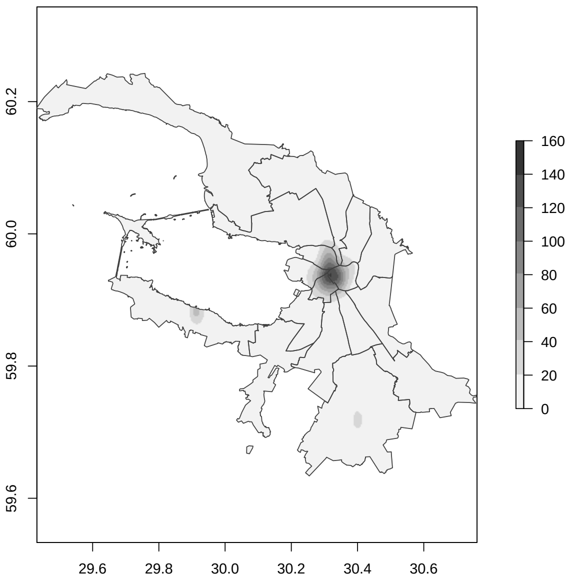


(c) pharmacies

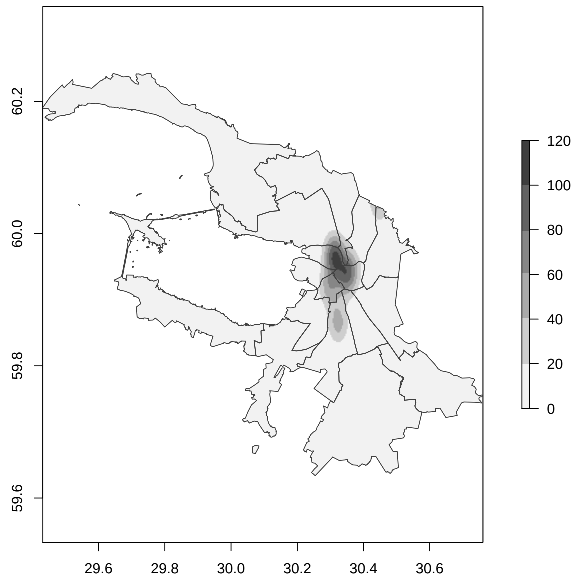


(d) hairdressers

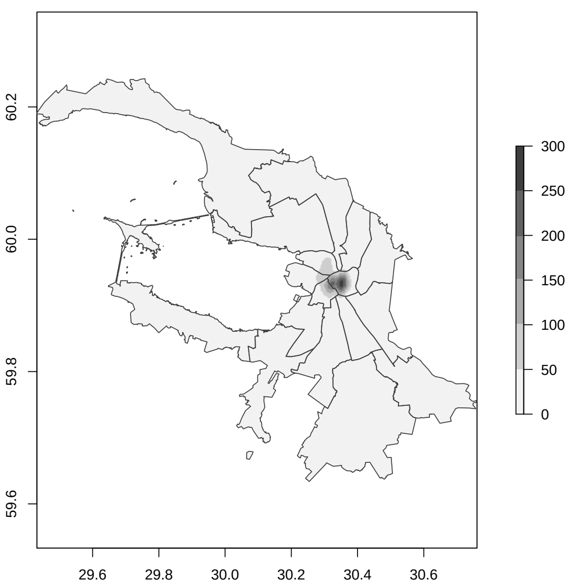
Figure 2: Amenities with highest spatial concentration



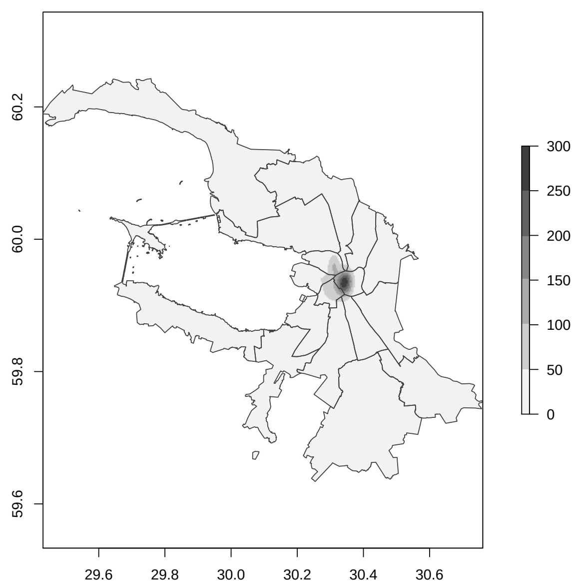
(a) museums



(b) notaries



(c) restaurants



(d) theaters

Figure 3: Centers of spatial distribution of individual amenities

Figure 4: Consumer center of St. Petersburg (variation-based weights)

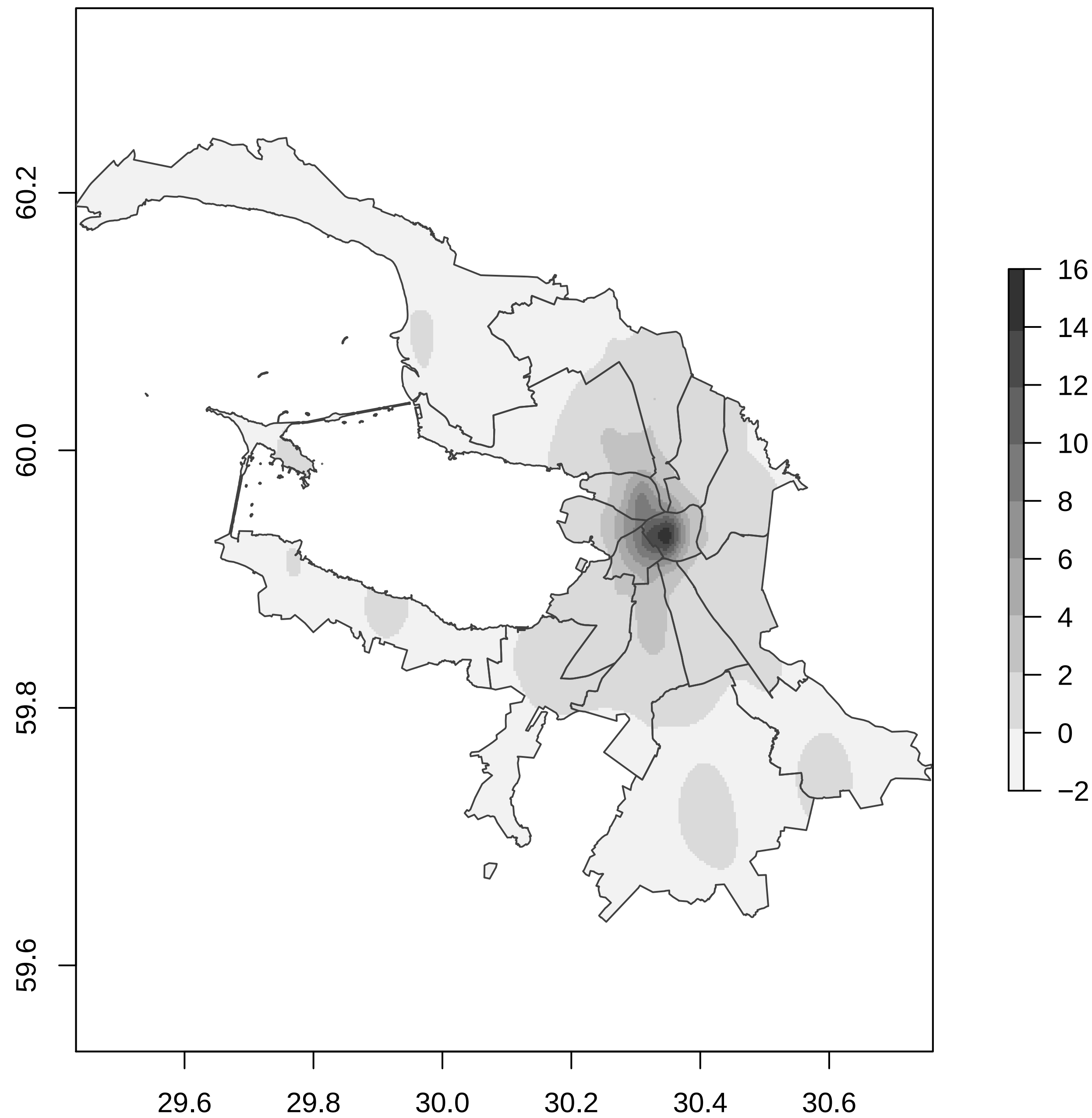


Figure 5: The survey-based weights of amenity groups

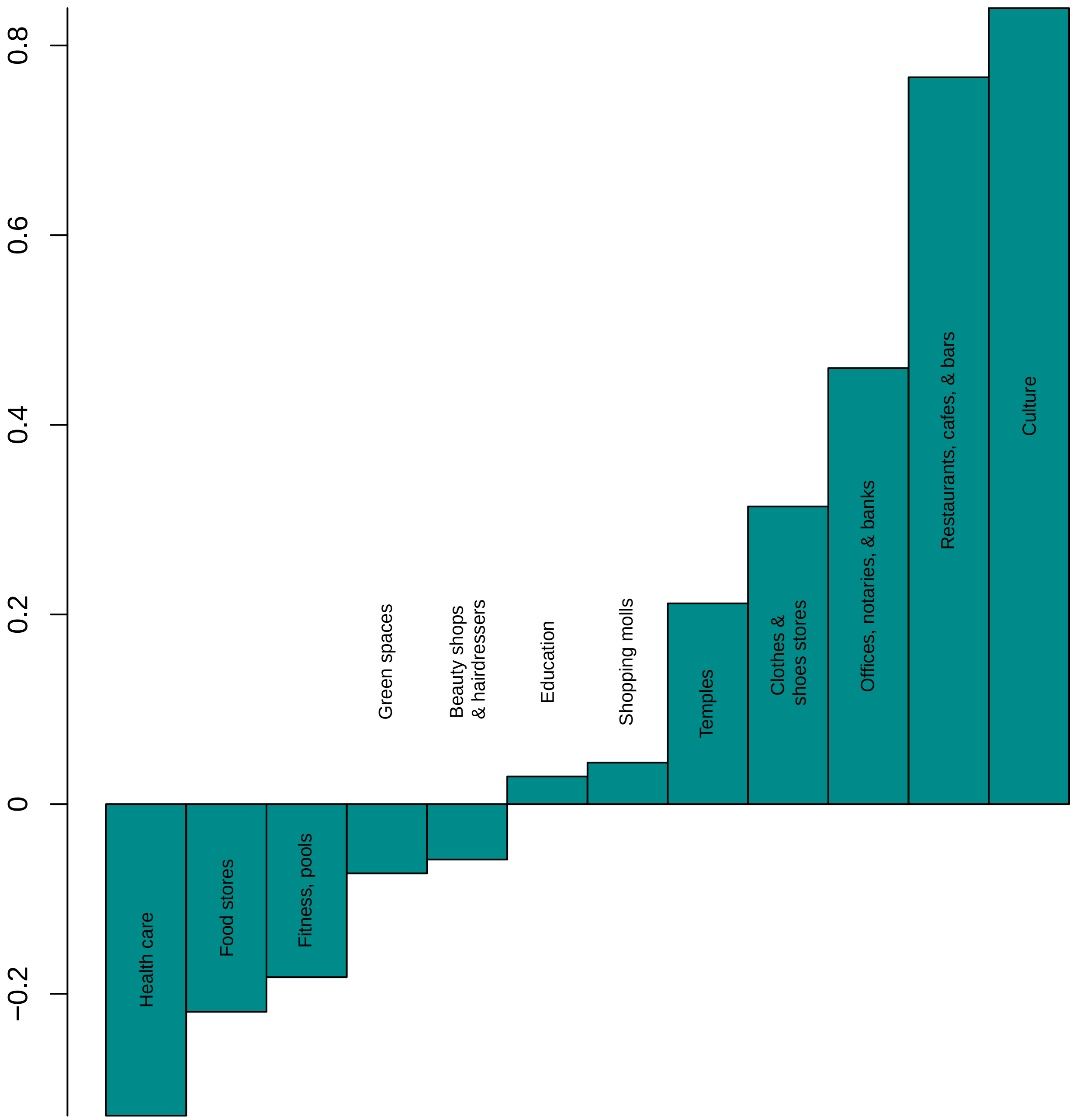


Figure 6: Consumer center of St. Petersburg (survey-based weights)

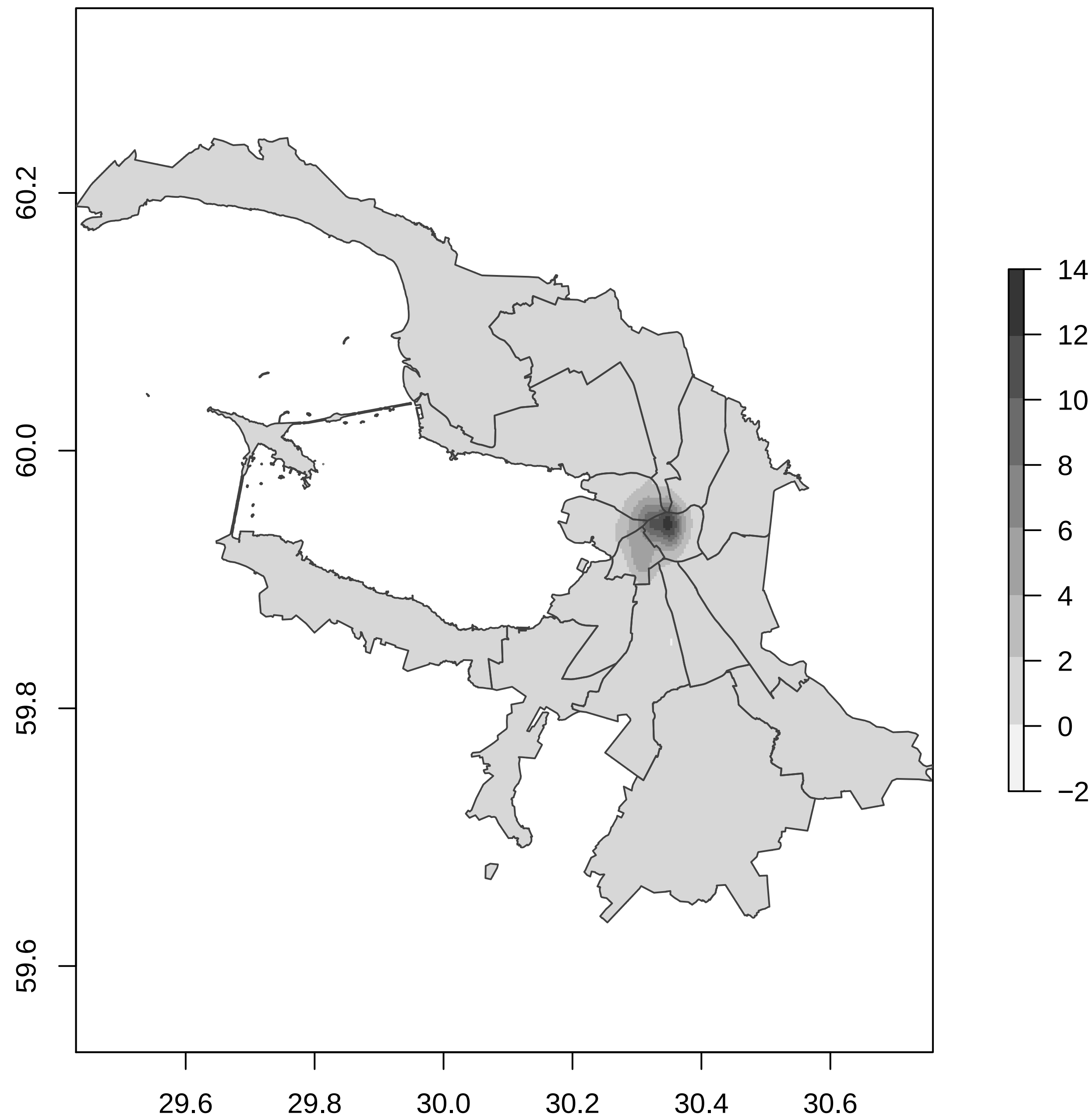


Figure 7: Alternative point estimates of city center

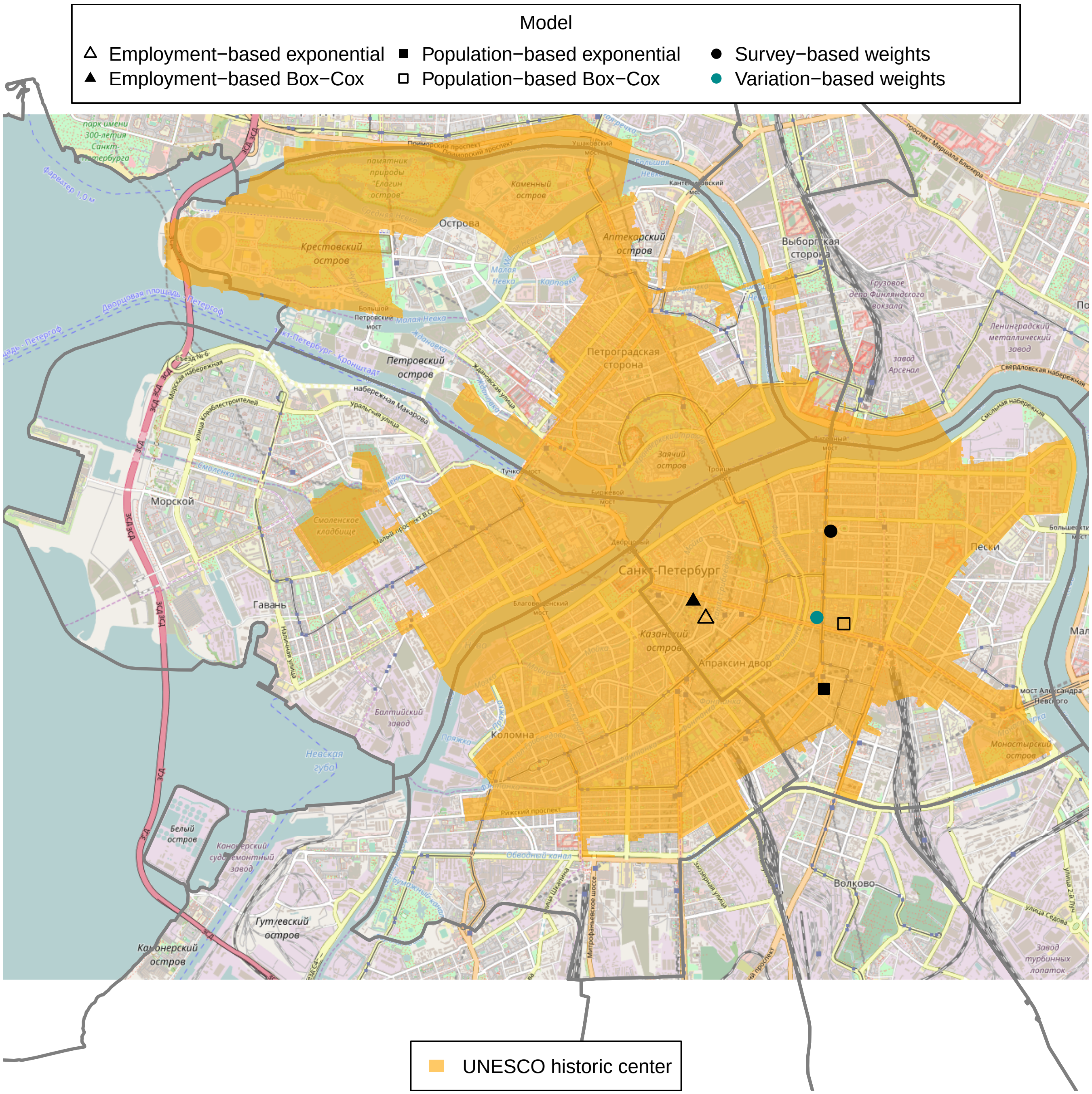
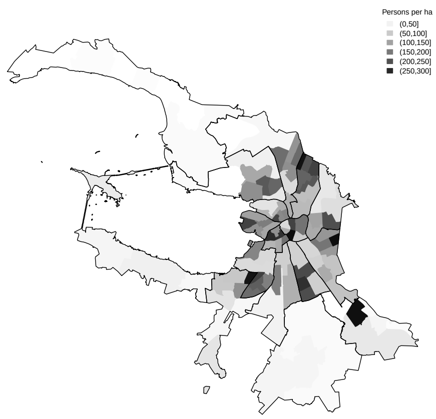
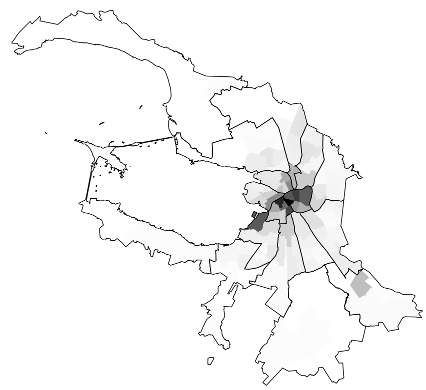


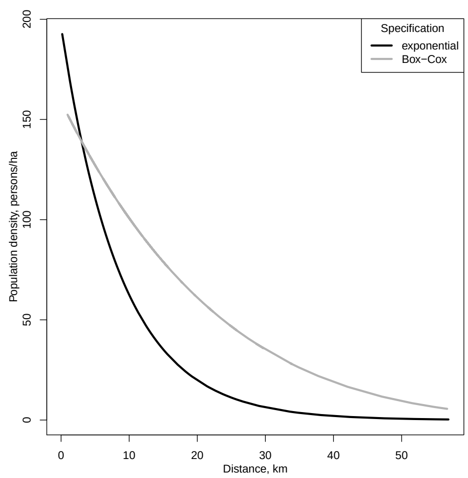
Figure 8: Population and employment density and density gradients



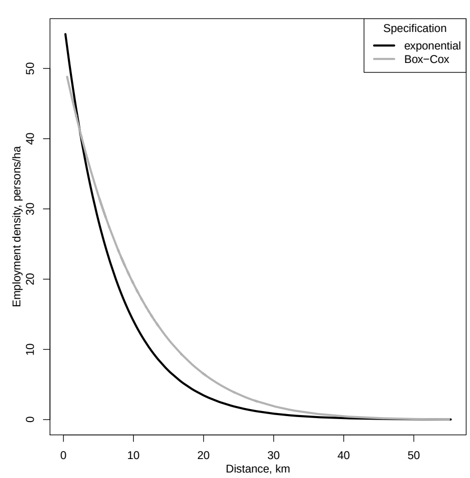
(a) population by subdistricts



(b) employment by subdistricts



(c) population density function



(d) employment density function

Figure 9: The crossroads based city center region

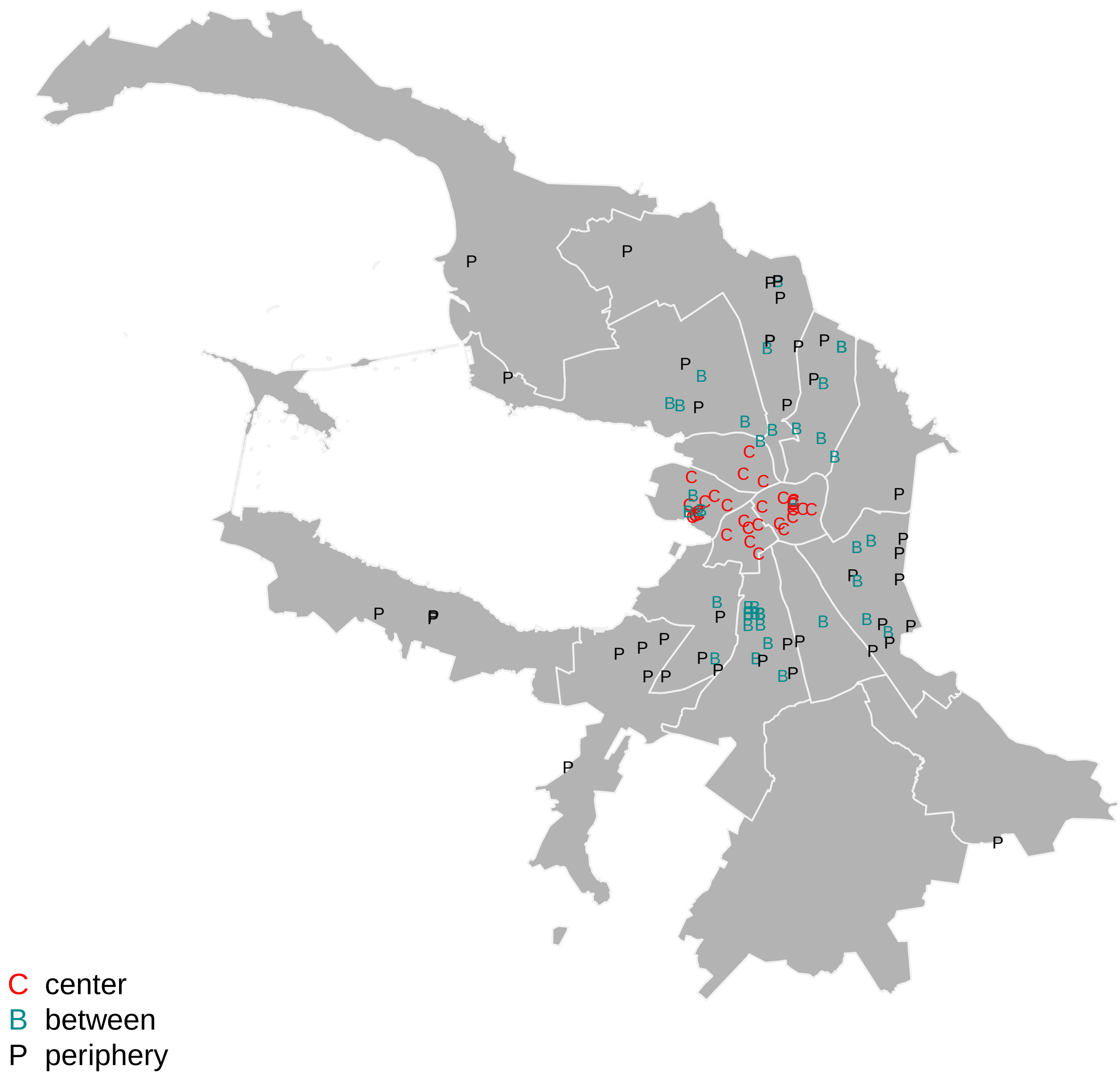


Figure 10: The averaged manually drawn city center region

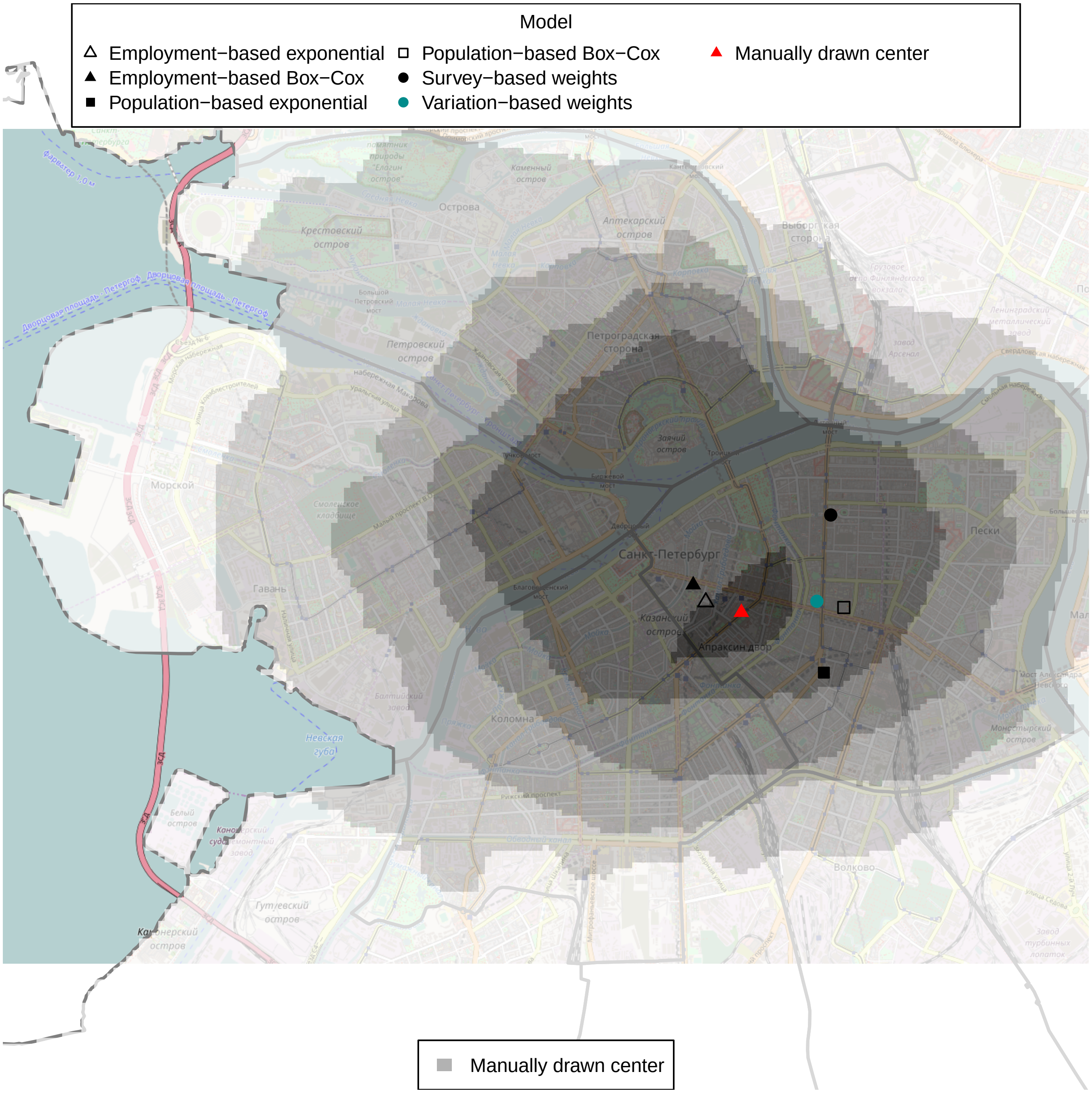


Figure 11: Spatially smoothed distribution of Instagram photographs