LOCATION-ALLOCATION OF PUBLIC SERVICES – CITIZEN ACCESS, TRANSPARENCY AND MEASUREMENT. A METHOD AND EVIDENCE FROM BRAZIL AND SWEDEN.[1]

AUGUST 2016

ANDERS FREDRIKSSON

CORS — Center for Organization Studies, FEA-USP, Universidade de São Paulo, Av. Prof. Luciano Gualberto, 908, São Paulo CEP 05508-900, SP, Brazil

anders.fredriksson@usp.br

Abstract

Access to public services is an important determinant of economic opportunities and well-being. We argue that location-allocation analysis can be used not only to determine the optimal spatial allocation of public services, but also as a tool to increase transparency in the allocation decisions. We illustrate by analyzing Citizen Service Centers in Brazil and Sweden. We first show how the optimal spatial allocation of such centers can be obtained with data that is largely publicly available, which allows for transparency. Based on the suggested optimum, we then introduce different public services misallocation metrics, beyond what the literature has previously discussed. We show how such misallocation metrics can be related to socioeconomic outcomes and political- and other variables, thus suggesting a method that can become a useful tool in analyses of public sector performance.

Keywords: Public services, Location-allocation, Citizen Service Center, Spatial access, Transparency, Misallocation

[1] I thank Matheus Albergaria, Erlend Berg, Love Ekenberg, Wouter Gelade, Catherine Guirkinger, Thiago Paixão, Imran Rasul, Sylvia Saes, Rohini Somanathan and Frédéric Wautelet for helpful discussions; participants at the CRED workshop (UNamur), the 2014 CSAE/Oxford development economics workshop and the 2015 Nordic Conference in Development Economics; Eric Thorén and Anette Petersson at the Swedish Tax Agency and Ilídio Machado, Cristina Onaga and Carlos Torres at Poupatempo; the Swedish Transport Administration and the Agency for Growth Policy Analysis for road data, Nordregio for the Swedish population density map, the Consortium des Équipements de Calcul Intensif (CÉCI) - Fonds de la Recherche Scientifique de Belgique (Grant No. 2.5020.11), where computational resources for some robustness tests were used, and two anonymous referees and the editor for much helpful suggestions. A salary grant from the Swedish Research Council is acknowledged. Any errors are the sole responsibility of the author.

1. Introduction

The goal of this paper is to develop a method that can increase transparency in the spatial allocation of public services. *Method* refers both to the application of location-allocation models in the analysis of the spatial distribution of public services, and the metrics needed to quantify access to such services. *Transparency* refers to that the location-allocation analysis should be reproducible, that the metrics of spatial access be objective and quantifiable, and the data taken as input should be largely publicly available. Preferably, the location-allocation analysis and metrics should also be intuitive and easy to comprehend.

The method can give citizens a tool for an objective analysis of the spatial allocation of public services (e.g. a quantification of statements such as "all the health clinics are in the capital region"). The end goal of improving spatial access is to affect socioeconomic- and related outcomes, such as increasing the fraction of citizens that have the personal documents needed to exercise basic citizen's rights, or improving schooling outcomes or health indicators. Importantly, the same method would be available to government planners, in order to facilitate location-allocation decisions. It can also be used to better understand the reasons for spatial "misallocation" of public services, and for a quantification of the welfare effects of such a mismatch between actual allocations and citizens' needs.

The study is inherently interdisciplinary, drawing on several fields in (primarily) Social Science, and makes three contributions. First, it outlines how a transparent analysis of the spatial distribution of public services can be conducted, with data that is largely publicly available. Second, it discusses how access to public services can be compared across countries, regions, municipalities, etc., including the degree of misallocation, a precisely defined concept, which is introduced below. Two examples of public services – those of Citizen Service Centers in São Paulo, Brazil (called Poupatempo, meaning "Savetime") and in Sweden (called Servicekontor, meaning "Service offices") are analyzed throughout the paper, and illustrate the transparency argument, and the access and misallocation comparisons, respectively. Third, the Brazilian and Swedish cases are of interest in themselves.

The analysis in the paper concerns, and is restricted to, spatial access, as measured by travel distances or travel times, two terms used interchangeably.[2] Section 2 serves as a broader

[2] One classification of different components of access to public services is the five A's in Penchansky and Thomas (1981): Availability, Accessibility, Accommodation, Affordability and Acceptability. Accessibility (or

motivation, defines concepts and reviews some of the related literature, positions the study, i.e. for which public services and settings it is deemed as most realistic, and discusses its limitations. Section 3 states the formal location-allocation problem, followed by the analysis of Poupatempo in section 4. The location-allocation of Swedish Service offices is discussed in section 5 (with some details in the appendix). Section 5 also defines the misallocation metric. Section 6 discusses public services access comparisons and section 7 discusses the results.

2. Spatial allocation of public services

Spatial access to public services affects socioeconomic outcomes, through several channels. First, large distances may imply that services are not used, which in turn affects outcomes. This is in line with the findings of Corbacho et al. (2012A, 2012B), where a larger distance to a civil registry implies a lower probability of registering births, which, later in life, has an impact on the years of schooling (in a study from the Dominican Republic). Kondylis and Manacorda (2012) find that distance to school affects school attendance (in Tanzania), Currie and Reagan (2003) that distance to a health facility affects the probability of having a checkup, and Buchmueller et al. (2005) that an increased distance to a hospital increases the mortality count from emergent conditions such as heart attacks (both in the US). Guagliardo (2004) and Burkey et al. (2012) cite additional studies, in which the spatial access to health services affects health outcomes. Second, conditional on using a public service, say going to school, outcomes may be affected, in this case learning, due to fatigue from traveling, etc. A third effect of distance is on the travel time itself, which can be analyzed from an opportunity cost of time perspective. In Fredriksson (2016A), we conduct such an analysis of the time costs of using Poupatempo.

Awareness of the socioeconomic impacts of spatial access has led researchers, at least since Hakimi (1964), to develop *location-allocation models* that aim at determining the optimal spatial distribution of public services. Reviews of the extensive location-allocation literature can be found in e.g. Revelle et al. (1970), Beaumont (1981), Rushton (1984), Marianov and Serra (2002, 2004), Revelle and Eiselt (2005) and Revelle et al. (2008). The literature broadly

---

Spatial access) concerns the physical location of the public services and the users, distances and travel times/costs, etc. By excluding e.g. "Accommodation", a concept which would include opening hours of facilities and the organization of how citizens are received at the location of the public service delivery, issues such as teacher and doctor absence, closed facilities, etc., are excluded. These are well-known problems of service delivery in many developing countries (World Bank, 2003; Banerjee and Duflo, 2007). Interpreting "spatial access" in its strictly physical/geographical meaning also excludes the analysis of access to public services and space in socio-cultural terms, as is done in e.g. sociology and ethnography (e.g. Claval, 1984).

specifies problems as one of three types: i. p-median - "where to place p facilities in order to minimize citizen average travel distance (or travel time)", which is the specification that will be used in this paper, as motivated below, ii. p-center – "where to place p facilities such that the citizen furthest away is as close as possible" or iii. coverage problems, which aim at "covering" the population and in which the number of facilities is endogenous. Each category contains subdivisions and alternative specifications. [3]

The analysis here suggested is primarily intended for rural (or non-urban) areas, for relatively homogenous public services. This has several implications, among which is that travel distance as a measure of access, inherent in the p-median specification, can be rationalized. In a literature that is broadly related to location-allocation analysis, Weibull (1976) discusses different metrics of accessibility. These metrics are, on the "supply side", typically a sum of components of the type (Service level in area A/Distance to area A), where an inhabitant or a population at a certain location has access to services A, B, C, D etc., at different distances (i.e. $S_A/d_A + S_B/d_B + ...$). Such a sum is likely to have more components in an urban area, where a city inhabitant can use e.g. bus/metro/car or walk to different services in different directions. In a rural area, a resident will have fewer de facto options. Examples of homogenous public services can be "getting a planned vaccination", "registering a birth" or "getting an ID". The São Paulo Citizen Service Centers and similar programs in other states and countries have as a specific objective to provide a similar set of public services to all citizens, with an impersonal treatment. As a result, there is less of a rationale to include public service units other than the closest, in a metric of spatial accessibility. In a data collection project focused on renewal of driver's license, in the interior of São Paulo state, 94% of all interviewees that used Poupatempo had visited the closest unit (Fredriksson, 2016A).[4] These data support a simple distance/accessibility specification.

[3] Public services could, in principle, be placed anywhere (in the continuous two-dimensional plane), but the discussion here restricts the choice set of locations to the population points themselves (cities, municipalities, etc.), which are connected with roads (i.e. a discrete location-allocation problem on a network). A much related and overlapping literature is the analysis of private sector facility siting (e.g. warehouse location).

[4] Out of 267 individuals that renewed their license at Poupatempo, no one of 185 individuals with a Poupatempo in their home municipality renewed at a Poupatempo elsewhere, although, in principle, there are no restrictions to do so. Out of 82 individuals not living in a Poupatempo municipality, but who still used Poupatempo, 65 renewed at the closest unit. 17 renewed elsewhere (typically at the second closest Poupatempo). In addition, hypothesizing that personal contacts may change people's behavior, we observe that citizens that use Poupatempo self-report having less personal contacts at the entity, than do those who use the traditional government bureaucracy (which Poupatempo is intended to replace).

The p-median problem maximizes efficiency, as measured by aggregate travel distance. Under the assumption that the cost per public service unit is constant, solving the p-median problem is equivalent to asking how a fixed government budget can be spent as efficiently as possible. Minimizing average distance is an intuitive starting point, at least when the (negative) effect of some individuals being far from the service can be assumed to be approximately linear in distance. This assumption seems reasonable for the above cited examples of public services (vaccination, birth certificate, ID), but not for e.g. ambulance placement, where a maximum travel time and stochastic demand are important components. Another justification of the p-median is that it is possible to define a set of transparent distance metrics (the average travel distance in a region, the aggregate reduction in travel distance for the inhabitants of a municipality if the optimum is implemented, etc.) that are directly related to the p-median objective (minimize average distance). If we accept the objective function, we can thus translate analyses of the distance metrics into welfare statements.

With a Rawlsian-type objective function, the p-center specification would lead to some public service units with little usage and in remote locations, as such locations implicitly have a larger weight in determining placement. In the case of the allocation of Citizen Service Centers in Brazil and Sweden, for which we have inquired about detailed implementation criteria (Fredriksson and Saes, 2016 and Fredriksson, 2016B), a p-center specification would be somewhat counter to explicitly stated lower bounds on municipality population for implementation of service units. In addition, a one-to-one correspondence between the objective function and different metrics is not as straightforward to define for the p-center problem, which involves an equity-efficiency tradeoff and in which a unique solution is not guaranteed (Beaumont, 1981). The coverage specifications are typically used for emergency services, which are thus excluded from the analysis.

The constant budget assumption is an attractive feature in order to get a transparent p-median analysis, in which actual and optimal allocations can be compared. Two other assumptions, implicit in the p-median formulation, is that travel distances/times to the public service affects all individuals' utility to the same extent (normally expressed as that "tastes and incomes" do not matter) and that all individuals have the same weight in a social welfare function. With a focus on homogenous public services with non-stochastic demand, where citizens visit the location of service delivery, analyzed in a non-urban/rural context and where distance/travel time is used as the measure of spatial access, we next discuss data.

2.1 Data needs, data availability and transparency

The basic data needed for the p-median analysis is population and a road distance/travel time matrix with closest distances or fastest travel times between all population points. Disaggregate population data is typically available from statistical authorities. An approximate distance matrix can be constructed based on the latitude and longitude of population points, and the Euclidean norm. Slightly more accurate will be the use of great-circle distances, i.e. correcting for the fact that the distances are calculated on a sphere, and not in the plane. More accurate still, and used in the present paper, are real travel distances. One advantage of the latter is that geographical conditions other than distance are implicitly considered. That is, if a certain city is easier to reach, it will be reflected in the distances/travel times, and hence impact the calculations.

Several points should be made with respect to data. First, the less aggregate is the distance data, the more accurate will be the analysis. In principle, we would like a matrix with the travel distance/time between each individual and each location that is a candidate for placement of the public service, but in practice aggregation of the population data is necessary. This may lead to three types of errors that could affect the solution: i. loss of information when individuals are assigned to a point (center, centroid) for which we have distance data, rather than the area (e.g. a municipality) where they actually live, ii. the distance to the closest public service for individuals that live in a municipality (thus represented by a population point) where the public service is implemented will be zero, which underestimates their true distance (this is a special case of i.) and iii. it may be that aggregated areas/zones/regions, represented by a (center or centroid) point for which we have distance data, have subareas for which the closest other point with road distance data is not the same as for its center/centroid (Marianov and Serra, 2004). In the paper, we conduct a robustness test of the allocation of Servicekontor in Sweden, with two different road distance databases and hence level of aggregation of the population data, to assess to what extent these considerations impact the optimal solution, in the largely rural Swedish setting.

The second point, which is key to the transparency argument in the paper, and discussed at some length, is that the technological development makes spatial data, more specifically road- and travel time data, increasingly available and useable. Ever more data is collected, stored and made available for use, and, in addition, there is a public good argument for making such data available (think of web applications aiming at diminishing traffic

congestion, with positive externalities). In order to detail this point, consider first the growing number of spatial data sources. These include the digitization of traditional maps, aerial photography, satellite images, data collected by vehicles and mobile phones with GPS equipment, etc. Such data collection is done by public entities (e.g. national transport authorities), private firms (e.g. Google) and non-governmental/voluntary organizations (e.g. OpenStreetMap), where the latter project has an explicit objective to make the data freely available. Although firms using OpenStreetMap data are allowed to charge for their products (e.g. routing services), free availability of the raw data assures competition in such services. The development is likely to remedy a problem in quantitative evaluations of spatial public services, which has been the lack of data (Gibbons et al., 2014). In the academic sector, the number of studies using e.g. geocoded zip codes or actual road distances/travel times is increasing, and several authors either compare how different distance measures perform (e.g. Wallace et al., 2014), or criticize the use of the Euclidean norm (e.g. Brew, 2015). In a recent book, Arsanjani et al. (2015) report 224 research articles based on OpenStreetMap data, from an August 2014 search. We specifically reviewed the last five years of papers in the European Journal of Operational Research (EJOR) and Socio-Economic Planning Sciences (SEPS), based on a keyword search. Out of 38 reviewed papers that had conducted any real-world location-allocation or distance analysis, and from which it was possible to discern the type of data used, 27 used Euclidean distances (or other norms not needing map data), whereas 11 used real distance/travel time data. Four of these used data from specialized mapping/routing firms, three used Google Maps, two used their own spatial data, one used Microsoft MapPoint, and one used OpenStreetMap data.[5] In parallel, there is a development of tools embedded in statistical software packages, which automate the retrieval of spatial/distance data. Such scripts exist, to retrieve data from Google Maps, for STATA (Ozimek and Miles, 2011) and for R (Kahle and Wickham, 2013); from MapQuest, for STATA (Voorheis, 2015); and from different OpenStreetMap developers, for STATA (Huber and Rust, 2016 and Zeigermann, 2016). The data retrieval typically operates through the provider's Application Programming Interface (API). Although there are user limitations for both Google Maps and MapQuest (currently 2500 queries/day and 15000 queries/month without charges, respectively), and potential limitations as to how the data can be used and/or presented, these

[5] For EJOR, searching for "location-allocation" resulted in 53 papers published since May 2011. For SEPS, which publishes fewer papers, we searched for both of the words "location" and "allocation", which resulted in 48 papers. Out of this total of 101 papers, 47 used distance measures in location-allocation analyses (or the like), and in 38 it was possible to discern what type of distance data was used.

tools are now in use by the research community.[6] In a rapid development, it seems likely that i). distance matrices retrieved using queries on open data sources, or on commercial data made available for research purposes, will become ever more available and of increasing quality, and ii.) if the purpose is to use such distance matrices for location-allocation analyses, the data retrieval and processing can be done without using traditional Geographical Information Systems (GIS) software. The latter point serves to broaden the group of potential users of distance data, to those not having expert software/GIS knowledge. For the below analysis of Poupatempo, distance data was retrieved using the MapQuest API, based on OpenStreetMap, similar to Buchmueller et al. (2005).

The third point with respect to distance data is that the travel time matrix can be adapted to accommodate different travel times for different modes of transport, if we know the ratio of individuals in each municipality that use e.g. their own vehicle vs. public transport, and have the matrix for each mode of transport. The above considerations about the social welfare function in the p-median specification would still be true, but it would relax another implicit assumption, i.e. that each individual travels at the same speed.

2.2 Level of analysis

The location-allocation model assumes that the decision of where to locate public services is taken at a higher level than the level at which the service is implemented (hence an allocation problem). This is one reason we also include the analysis of Citizen Service Centers in Sweden. A national authority (Skatteverket – the Tax Agency) decides upon locations in municipalities, which generates a distribution of public services coverage at the intermediate administrative level (i.e. counties, which is also a political-administrative level). It is therefore of interest to construct indices and compare access, not only at the municipality level, but also at the county level. The Brazilian analysis does not have such intermediate administrative regions.

In addition to comparing access itself, we also suggest a precisely quantified misallocation metric, which can be defined at different levels of aggregation, as the difference between the actual distance to the public service, and the distance if the (p-median) optimal solution were

[6] See e.g. Gibson et al. (2013) for a study of the effect of distance on voting, and Shen and Hsia (2015), for a study of ambulance routing, both using the Ozimek and Miles (2011) STATA command based on Google Maps. Buchmueller et al. (2005) is an early study using MapQuest data. Some authors argue that the Ozimek and Miles (2011) script is obsolete, but as the underlying functionality, i.e. the Google Maps API, is not, updated scripts are likely to appear.

implemented. This metric can be compared across e.g. municipalities or more aggregate regions (such as the Swedish counties), and can also be regressed on potential explanatory factors of misallocation, such as political factors. More generally, any potential determinant of misallocation can be correlated with the misallocation metric. Such an analysis thus has the potential to connect the outcome of the purely spatial p-median analysis with (quantifiable) explanatory factors from political science, sociology and other fields of Social Science. These aspects are further discussed in sections 6 and 7.[7]

The Brazilian analysis thus illustrates the transparency argument, whereas the Swedish case serves primarily to discuss different metrics that can be used for e.g. regional comparisons of access to public services (in addition to the above discussed robustness analysis).

3. The p-median problem – specification and solution method

The problem is where to locate p public service facilities in order to minimize the total travel distance, where demand is originated from a set I of population points, by choosing from a set J of candidate locations. Let n be the number of elements in J, with $n > p$. Let $i \in I$, $j \in J$, and let h_i be the demand from population point i and d_{ij} the distance from population point i to candidate location j. If there are no capacity constraints for the public service units (which will imply that each population point will be served by its closest unit), the problem can be formulated as choosing two sets of indicator variables: $x_j \in \{0,1\} \forall j$, which indicates if there will be a service unit at location j, and $y_{ij} \in \{0,1\} \forall i, \forall j$, indicating which entity j serves demand from i. We thus solve

$$\min_{x_j,\, y_{ij}} \sum_{i \in I} \sum_{j \in J} h_i d_{ij}\, y_{ij} \text{ s.t. } \sum_{j \in J} x_j = p,\ \sum_{j \in J} y_{ij} = 1\ \forall i \text{ and } y_{ij} \leq x_j \forall i, \forall j. \qquad (1)$$

The three constraints assure that p units will be "opened", that each demand point is served by one unit and that a demand point cannot be served by an "unopened" unit, respectively. If there are pre-existing "fixed" units which citizens can use (as in the Poupatempo analysis), then let the candidate set J consist of two subsets, including/excluding these fixed units, respectively, i.e. $J = J_F \cup J_{-F}$. Let the number of fixed units (i.e. the number of elements in J_F) be f. Define $p' = p + f$ and $n' = n + f$, such that n', p' replace n, p in the above

[7] The nature of the allocation problem, as a choice between (e.g.) municipalities, makes it conceptually connected to the literature on fiscal federalism (e.g. Oates, 1997), which is one of several reasons why political factors are inherent to the analysis.

formulation. Then the only other change is that the first constraint is replaced by $\sum_{j \in J_F} x_j = f$ and $\sum_{j \in J_{-F}} x_j = p$.

With respect to the above discussion, two points should be made, First, d_{ij} can be Euclidean- or great-circle distances, actual travel distances on the road network (e.g. Dijkstra-distances), travel times on the same network, or any other general "cost" parameter. Each entry d_{ij} could, in principle, also be a weighted sum of travel times for different modes of transport, where the weights would be the fraction of individuals using each mode (e.g. from surveys of transport habits). Second, h_i can be the entire population, or subgroups thereof, depending on the public service studied.

The problem in (1) was solved, for the below Brazilian and Swedish cases, with the IBM C-PLEX software, which uses a Branch and Bound algorithm. The Swedish problem involves 10^{71} possible combinations (with n=257, p=90), and is solved in around two minutes (with an Intel i5 processor, 1.6GHz base speed, 8GB RAM). This problem is of the same magnitude that Smith et al. (2015) argue is the largest problem that can be solved exactly with a software such as the C-PLEX (n=500, p=50), but smaller than the current upper limit for getting exact solutions with any method (n=1000, p=50). This is promising, as the method suggested can be used on real world problem sizes.

4. Citizen Service Centers in São Paulo, Brazil

Poupatempo ("Savetime") is a government service "one stop shop" that co-locates many different state authorities. It was first established in metropolitan São Paulo in 1997, in order to address several concerns with a malfunctioning government bureaucracy, citizens' dependence on bureaucracy intermediaries, etc. At these centers, citizens can typically resolve a number of errands, involving different government authorities, at the same physical location, and, ideally, complete errands in one visit/one day. A citizen can use any Poupatempo unit. Examples of agencies located at Poupatempo are DETRAN (the Department of Transit), IIRGD (the Institute for civic identification), SERT (the Secretary of Labor and Labor Relations), public utility companies, the consumer complaints bureau, a post office, a public bank, etc. With 43 million inhabitants and 43 million Poupatempo visits in 2015, i.e. more than one visit per adult per year, this state government program is an important and widely used public service (Fredriksson, 2016A, and references therein,

discuss the Poupatempo reform and similar programs in other countries). Figures 1A and 1B show maps of São Paulo state and the Poupatempo expansion, described next.

We analyze the spatial optimality of the 2008-2011 Poupatempo expansion into the interior and coastal areas of the state, with 16 new units, which increased the number of units in this area from 5 to 21.[8] This is the period when the service extends well beyond metropolitan São Paulo and the largest urban centers. The interior has 606 municipalities, the entire population of which constitutes the demand in our p-median specification. Based on information from Poupatempo about the expansion decision, we consider 42 municipalities to have been candidates, in 2007, to get a unit implemented. These candidate locations are the most populous cities in the interior of the state, excluding those that already had a Poupatempo unit (and a few municipalities in the adjacency to the pre-existing units). Poupatempo also stated that the densely populated Bandeirantes/Anhangüera highway region (running in a north-northwest direction from the metropolitan area) should be underweighted, in order for the units to be spread out across the state. Such a spreading out is inherent to the p-median analysis, and the discussions with Poupatempo inspired the original problem formulation. In an alternative specification, we consider 52 municipalities to have been candidates in 2007, based on a lower population threshold. Poupatempo had originally stated that only cities with more than 100.000 inhabitants had been candidates to get a unit, but, as one municipality with fewer inhabitants, Caraguatatuba, had a unit implemented, we decided to also use a lower population threshold (i.e. an exogenous threshold of 80.000, rather than 94.000, which is adapted to just include Caraguatatuba in the candidate list). In Fredriksson and Saes (2016), we further discuss the Poupatempo implementation criteria.

We use real distance (fastest) road data from MapQuest, the data of which is based on OpenStreetMap. The distance matrix contains all distances between the 606 municipalities and the candidate cities + the pre-existing units. The distance data is between the city centers (originally the location of the municipality's first church, which almost always corresponds with today's center). We use the 2007 (i.e. pre-reform) population data from SEADE (the São Paulo state data entity), and the aggregation of the population data is thus at the individual municipality level. With four pre-existing metropolitan area Poupatempo units that could potentially be closest to citizens in the interior/coastal areas, and with five pre-existing units in the interior, we thus solve the p-median problem with $p = 16, f = 4 + 5 = 9$ which gives

[8] In the paper we refer to the "current allocation", as that of 2013, i.e. after the expansion. No units were implemented in the interior in 2012-2013. A new build-out, not analyzed here, was undertaken in 2014-2015.

$p' = 25; n = 42, n' = 51, J = \{j_1, .. j_{51}\}$ and $J_F = \{j_1, .. j_9\}$ (and thus $J_{-F} = \{j_{10}, .. j_{51}\}$); and $I = \{i_1, .. i_{606}\}$. In the alternative specification, we have $n = 52, n' = 61$. The solution to these two "n-p" specifications, called "42-16" and 52-16", respectively, are displayed in figures 1C and 1D.

Figure 1A. São Paulo state and its municipality division. The area in black is metropolitan São Paulo (which has 39 municipalities), excluded from the study. The remaining area, i.e. the interior and the coast, has 606 municipalities. The state's area is 248.000 km^2.

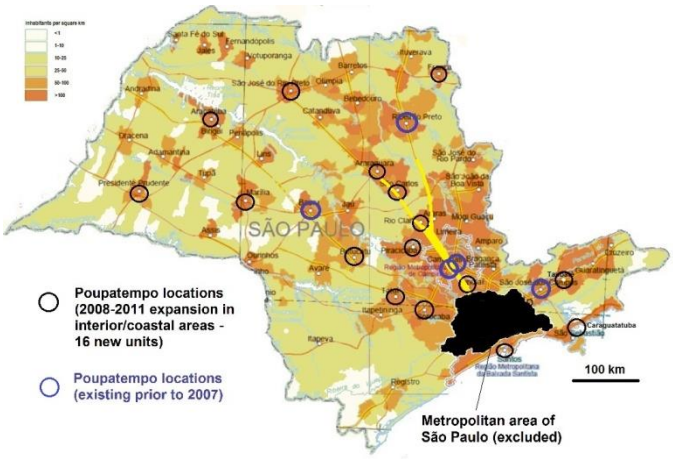

Figure 1B. Population density and municipalities with Poupatempo. The five units in blue existed prior to 2007. The 16 units in black are those of the 2008-2011 expansion. The pre-existing Poupatempo units in (the outskirts of) metropolitan São Paulo, which inhabitants in the interior may potentially have as their closest unit, are not shown, but are considered when solving the p-median problem. The Bandeirantes-Anhangüera corridor is shown in yellow.

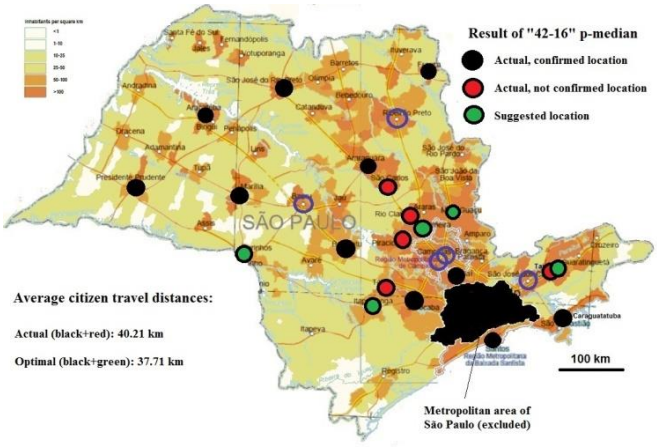

Figure 1C. Solution to the "42-16" p-median problem. The municipalities in black are those with Poupatempo units, part of the 2008-2011 expansion, that were confirmed by the analysis. The five units in red are those actual units that were not confirmed by the p-median analysis, and the five units in green are those units instead suggested.

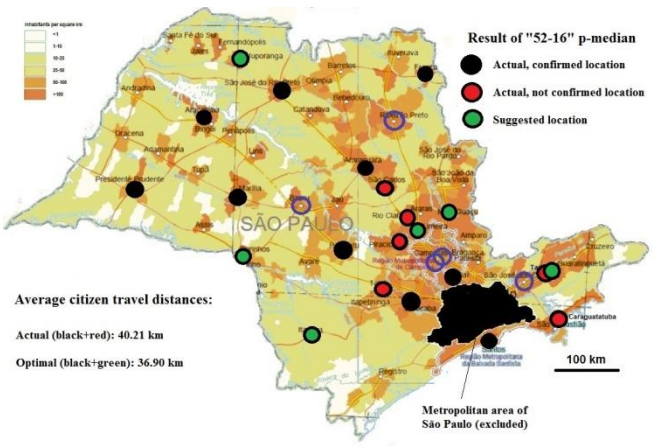

Figure 1D. Solution to the "52-16" p-median problem. The municipalities in black are those with Poupatempo units, part of the 2008-2011 expansion, that were confirmed by the analysis. The six units in red are those actual units that were not confirmed by the p-median analysis, and the six units in green are those units instead suggested.

With the objective of minimizing average citizen distance, the units in São Paulo are, notwithstanding the above criteria, somewhat too concentrated along the Bandeirantes/Anhangüera highways (see figures 1B-1D). The optimal allocations reflect a larger dispersion of units, as seen by comparing the red and green units in figures 1C-1D. Five out of 16 units have a different location in the 42-16 optimum than in the actual allocation, and the average distance to the closest Poupatempo unit decreases from 40.21 km to 37.71 km. The 52-16 specification sees an additional non-optimal unit, with a distance reduction from 40.21 km to 36.90 km. These differences in average travel distance between the actual allocation and the optimum are quite modest (6.2% and 8.2%, respectively). Certain areas, however, such as the south-west, will have much smaller distances in the 42-16 optimum than in the actual allocation. This is even more so in the 52-16 optimum, where somewhat smaller municipalities were allowed to be candidates, which affected the solution.

Municipality population and distance data from an OpenStreetMap provider were used to get these results, and the calculations can be replicated for many other public services and locations. The distances are real travel distances, thus taking into account natural obstacles such as mountains, rivers, etc. One use of the results would be to convert (the actual-optimal differences in) travel distances/times into opportunity cost of time, in a social welfare calculation. Before these results are discussed further, however, we first analyze Citizen Service Centers in the Swedish context, and introduce different distance metrics.

5. Citizen Service Centers in Sweden

Servicekontor ("Service offices") are single-window Citizen Service Centers that handle errands related to the Tax-, Social Insurance- and Pensions Agencies. The offices were implemented following a reorganization of the Tax Agency, starting in 2007, and the Social Insurance Agency has since closed down its separate offices. Servicekontor were placed where there was previously a tax- or insurance office, but new offices opened as well. The Pensions Agency has never had separate offices. A citizen can visit any Servicekontor unit. Although there is a general move towards e-government, a previous attempt by the Social Insurance Agency to shut down its channel of in situ attendance met with strong resistance and was abandoned, and the last years have seen an increase in the number of errands at the Citizen Service Centers. Reading government reports on the topic, there seems to be a consensus that the Servicekontor initiative will remain in place for the foreseeable future, and the number of offices has remained the same since 2013. In a country of 10 million

inhabitants, the offices have around 4.5 million visits per year. There are 103 offices, 90 of which are outside of Stockholm County, Göteborg, Malmö and the island of Gotland. We analyze the optimality of the spatial distribution of these 90 offices. Figure 2 shows the Swedish county division, population density, and the Servicekontor locations.

The p-median problem of where to place 90 Servicekontor, choosing from 257 candidate locations, the number of municipalities in the study area, taking the entire population as the demand variable, was solved in Fredriksson (2016B). The paper, in Swedish, is concerned with suggesting variables, other than the number of visitors, than can be used by government authorities when expanding/reducing the number of public service offices. We are instead interested in comparing access to public services in different regions, and in constructing metrics for such comparisons. Table 1, columns 1-7, reproduces the p-median solution. In this table, the counties are in north-south order. The p-median problem was solved using population data for 1937 localities, from Statistics Sweden (Statistiska Centralbyrån), including the principal town in each of the 257 municipalities, and an average of 6.5 other built-up areas per municipality. In this data, the 15% of the population that does not live in built-up areas, was aggregated to each municipality's most populous town. The distance matrix, provided by the Swedish Agency for Growth Policy Analysis (Tillväxtanalys) contains 1937×1937 entries, i.e. the shortest route between each pair of locations. The distances are between the centroids of the population points.

The average distance to a Servicekontor is 9.44 km, or 10.74 km if Stockholm County, which was not analyzed, is excluded. In the optimum, in which 11 units would be "re-allocated" (the names of which are displayed in table 1 in Fredriksson, 2016B), the average distances are instead 9.04 km and 10.23 km, representing decreases with 4.2% and 4.7%, respectively. The robustness analysis, in which we use a less detailed distance database, has the same optimal locations, except for three municipalities (i.e. 87 out of 90 optimal locations are the same). This analysis is reported in the appendix, together with some more details of the problem specification. The aggregation problem discussed in section 2.1 thus does not impact allocations much for this specific setting, and we proceed with the here reported solution.

In order to start analyzing regional differences in the spatial access to Servicekontor, it is interesting to first note, from table 1, that there are large differences between different counties. The county with the largest average distance, Jämtland (23.75 km), has more than four times the average of Västmanland (5.48 km, from column 4). Such uneven access is well

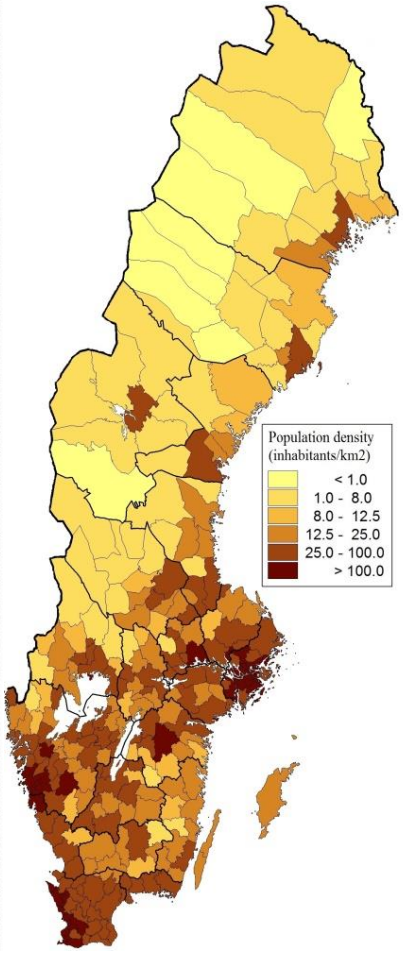

Figure 2A. Sweden and its counties. (grey striped=excluded, grey=office locations not analyzed). **2B.** Population density. **2C.** Servicekontor.

1	2	3	4	5	6	7	8	9
		Actual allocation		Suggested optimal allocation			Actual-optimal distance differences	
County	**Population**	**# offices**	**Average distance (km)**	**# offices**	**Δ # offices**	**Average distance (km)**	**Per inhabitant (km)**	**Aggregate (Mkm)**
Norrbotten	248609	6	22.10	7	+1	16.33	5.77	1.43
Västerbotten	259286	4	15.67	4		15.67	0.00	0.00
Jämtland	126691	3	23.75	3		22.17	1.58	0.20
Västernorrland	242625	4	12.11	5	+1	8.77	3.34	0.81
Gävleborg	276200	6	7.23	6		7.23	0.00	0.00
Dalarna	277047	5	15.91	6	+1	12.67	3.24	0.90
Uppsala län	336159	3	10.04	4	+1	7.17	2.87	0.96
Värmland	273265	4	16.58	3	-1	18.31	-1.73	-0.47
Västmanland	253560	4	5.48	3	-1	7.57	-2.09	-0.53
Stockholms län	2053921							
Örebro län	280230	3	11.07	3		10.94	0.13	0.04
Södermanland	270049	4	7.92	4		7.92	0.00	0.00
Östergötland	429642	3	9.92	3		9.66	0.26	0.11
Jönköpings län	336866	6	8.48	6		8.37	0.11	0.04
Västra Götaland (excl. Göteborg)	1576025	12	9.78	13	+1	8.55	1.23	1.95
Kronoberg	183940	3	12.10	2	-1	15.37	-3.27	-0.60
Kalmar län	233536	3	19.84	4	+1	15.04	4.80	1.12
Halland	303392	4	7.43	4		7.42	0.02	0.00
Blekinge	153163	2	12.20	2		12.20	0.00	0.00
Skåne (excl. Malmö)	1243393	11	6.83	8	-3	8.60	-1.77	-2.20
Number of offices		**90**		**90**				
Average distance			**9.44 km**			**9.04 km**		
(excluding Stockholm county)			**10.74 km**			**10.23 km**		

Table 1. Allocation of Servicekontor in Sweden, per county, listed from north to south (see figure 2A). Column 2 shows county population. Columns 3 and 4 show the number of offices, and the average travel distance to the closest unit, in the actual allocation. Column 5 shows the number of units in the p-median optimum, column 6 the change from the actual amount of units, and column 7 the average distance in the optimum. Column 8 is the difference in average distance to the closest unit, between the actual and optimal allocations, and column 9 the county level misallocation metric, i.e. the distance difference times population.

documented, from a Swedish/European regional perspective (Dahlgren, 2008).[9] The average distance in different counties/regions does not indicate, however, if allocations are fair, in terms of the p-median objective. As the p-median specification considers both population and distance, a metric that compares counties must contain both components. Such a *misallocation metric* is calculated in column 9, which is *the difference in distance between the actual and optimal allocations*, as calculated in column 8, *multiplied by population*, from column 2. This metric is reproduced in figure 3, in the upper panel, where the counties have been ranked from the smallest value (most advantaged), to the largest value (most disadvantaged).[10]

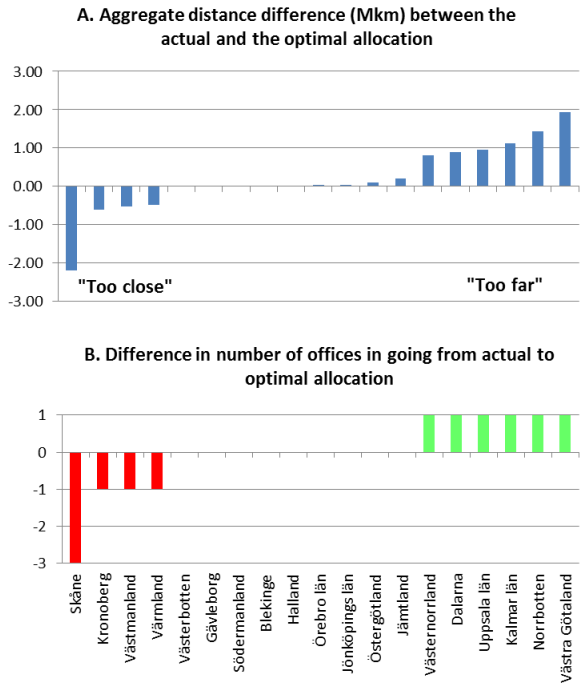

Figure 3. Upper panel: County level misallocation metric, i.e. county population times the difference between the average distance to the closest public service unit, in the actual and optimal allocations. A negative value indicates that citizens, on average, are "too close", compared to if the optimum had been implemented, and a positive value indicates that citizens are "too far". Note that the sum of the bars is positive (see text). Lower panel: Suggested changes in the number of Servicekontor to reach the optimum, per county (red=reduction, green=addition). The values close to zero (instead of exactly zero) in the upper panel, are due to "border" effects. Citizens living close to county with a change will be affected if there is a change in which unit is their closest.

[9] Although the paper is not concerned with equity, the disparity in access can also be characterized by Lorenz curves, as suggested by Burkey et al. (2012). As many individuals have zero travel distance in our model, but not in reality (those living in a population point with a Servicekontor, which is the majority of the population), the inequality will be exaggerated (it is thus an effect of the second type of aggregation error, discussed in the introduction). Burkey et al. (2012) also note that distance distributions are typically more unequal than travel time distributions. Our measure is not directly comparable with other studies or between, say, São Paulo and Sweden. We can study the change, however, between the actual allocation and the optimum. For Sweden, 75% of the population does (only) 15% of the traveling in the actual allocation, and 16% in the optimum. There is thus almost no change in equity, as measured by such a Gini coefficient, in implementing the p-median optimum.

[10] I thank Rohini Somanathan for discussions regarding the definition of the disaggregate misallocation metric.

The graph is accompanied, in the lower panel, with the changes suggested in the number of units, between the actual and optimal allocations (i.e. column 6 of table 1). This county level misallocation metric shows that neither of the above mentioned counties (Jämtland and Västmanland) are the most disadvantaged/advantaged. This is because of their (small) population. It is instead the ranking in figure 3 that shows the relative disadvantage/advantage between counties, if the optimum is to be reached. Skåne is the most advantaged county (far left in figure 3) and Västra Götaland is the most disadvantaged (far right). A comparison of the panels in figure 3 shows that there is a very high correlation between the misallocation metric and the suggested changes in the number of offices. The misallocation metric can be defined on various levels of aggregation, which we discuss further below. Also note that the sum of the metric is always positive, whenever the actual and optimal allocations differ. We next suggest a method, based on the above analysis, for how a comparison of access to a public services can be made.

6. Comparing and analyzing spatial access to public services

First consider a hypothetical comparison of the access to Citizen Service Centers in different Brazilian states, thus repeating the above São Paulo analysis, which was based on data that is largely publicly available, obtained without involvement of the entity being evaluated. More than 20 states (out of 27) have implemented Citizen Service Centers, with similar objectives. The comparison would thus be restricted to one public service and a specific institutional setting, in order to increase comparability. The centers are state level programs however, meaning that there are no federal political decisions involved. A first comparison of access is to calculate the *actual average distance to the service*, per state. Although unequal access across states is potentially driven, primarily, by differences in geography and population density, a transparent presentation is still useful in order to analyze reasons for differences in access. We can also solve the p-median, for each state's current amount of service centers, to obtain the *optimal average distance*, per state. Such a comparison, of hospital access in four US states, is conducted by Burkey et al. (2012). The percentage reduction between the actual and optimal distances can also be calculated, as do said authors, Athreya and Somanathan (2008), for post offices in south India, and as was done above for São Paulo. Importantly, such an analysis takes as given each state's absolute level of public spending, as well as the relative priorities between different public services. We can still make *quantifiable* welfare- and efficiency comparisons between states, however. By *multiplying the state population with the actual-optimal distance difference*, we get the *aggregate* "unnecessary" travel, per

state, due to the non-implementation of the optimum. This is, as discussed above, the p-median measure of the aggregate welfare difference between the actual and optimal allocations, and is straightforward to calculate. Brazilian (or U.S.) states, or Nordic countries, etc., could thus be ranked in terms of this aggregate misallocation measure (perhaps, most illustratively, presented along with the average "unnecessary" travel distance per citizen, and the percentage measure of misallocation).

Next consider a public service for which there is an intermediate administrative/political level, between the entity/authority/government responsible for the service, and the level at which it is implemented. An example is the above analyzed Swedish Service offices, where the national Tax Agency decides upon the implementation, and the public service itself is implemented in municipalities. Although a technical/administrative agency, rather than a political body, the analysis could serve to study whether certain counties are more successful than others in attracting public services. Ultimately we want to know the impact on citizens of the location decisions. Whereas the national level misallocation metric measures the aggregate inefficiency, the regional level, or *disaggregate*, misallocation metric quantifies, in citizen travel distance, how the aggregate misallocation is distributed. In some regions the public service will on average be closer than it "should be", and in other regions services will be "too far" (and on average there is a loss). The construction of such a metric was exemplified in the above Swedish analysis. Again, we can make welfare comparisons, now between different regions (column 9 in table 1, upper panel of figure 3). Such a comparison is, indeed, more straightforward to interpret, than the above suggested comparison of (e.g.) Brazilian states. In the Brazilian case each state has its own budget and public services priorities, whereas the Swedish allocation decision is made by a national entity. This serves to illustrate that the political-administrative level at which allocation decisions are taken, will have an impact on how spatial access comparisons should be made and interpreted.

The measures here presented can also be used to study reasons for misallocation. If we think of Service offices, one could certainly imagine that the Tax Agency consults counties about their needs for the public service, or the county political leadership trying to attract services to the region. Differently from Poupatempo, which is handled by a secretary under the state governor, Swedish Servicekontor are handled by a non-political agency. One could still imagine, however, that national level politicians (e.g. members of parliament, representing regions), collaborate with the county leadership, to try to influence allocations. The county level misallocation metric can be used to study such potential effects.

Similar to the above, it is straightforward to define the (also *disaggregate*) misallocation metric at the municipality level, i.e. municipality population times the municipality's distance difference to the closest actual and the closest optimal unit. A topic that we analyze further in a parallel project (Fredriksson and Saes, 2016), figure 4 illustrates some correlates of the municipality level misallocation metric in the state of São Paulo. There is a correlation between being "too close" and having a higher GDP/capita (panels A & B), a lower level of illiteracy (A & C), and a higher probability of having a mayor from PSDB, the political party that also holds the São Paulo state governor position since 1994, and is responsible for the public program (panels A & D). The opposite holds for the municipalities that are "too far". The correlations displayed are weak, and are precisely correlations, not causal links. One purpose of the paper, however, is to provide a method to display such regularities. Take illiteracy as an example. Reforms of the public bureaucracy may have as an objective to provide public services to traditionally disadvantaged socioeconomic groups. Figure 4 does not lend support to such a "pro-poor" profile of the Poupatempo expansion. Figure 4 was constructed with publicly available data, allowing for a transparent analysis of potential determinants of the misallocation metric, which, in addition, has a precise welfare interpretation. In a regression of the metric on e.g. illiteracy and other factors, and if causality can be established, a significant coefficient on illiteracy would have an interpretation as follows: "a 1% higher degree of illiteracy in a municipality leads to x kms of additional travel distance to the public service" (the topic is further discussed in Fredriksson and Saes, 2016).

Although not used in this paper, we experimented with exhaustive iteration as a solution method for smaller location-allocation problems, as this method can allow for robustness tests of some of the suggested uses of the misallocation metric.[11]

[11] We found that the largest feasible problem for exhaustive iteration on a standard desktop computer was a problem involving 10^{10} possible combinations (e.g. n=37, p=16, as in a previous version of the Poupatempo problem). With exhaustive iteration, one can choose a distance cutoff and get all "good enough" solutions (e.g. the top 1 million solutions, or every allocation that has an average citizen distance within 50 meter from the optimum). These near-optima can be used to check if regressions aiming to explain potential misallocation give different results depending on if, say, the second, third, or 1000th ranked solution is used, instead of the optimum. Methods aiming to explain misallocation should probably be robust to such changes.

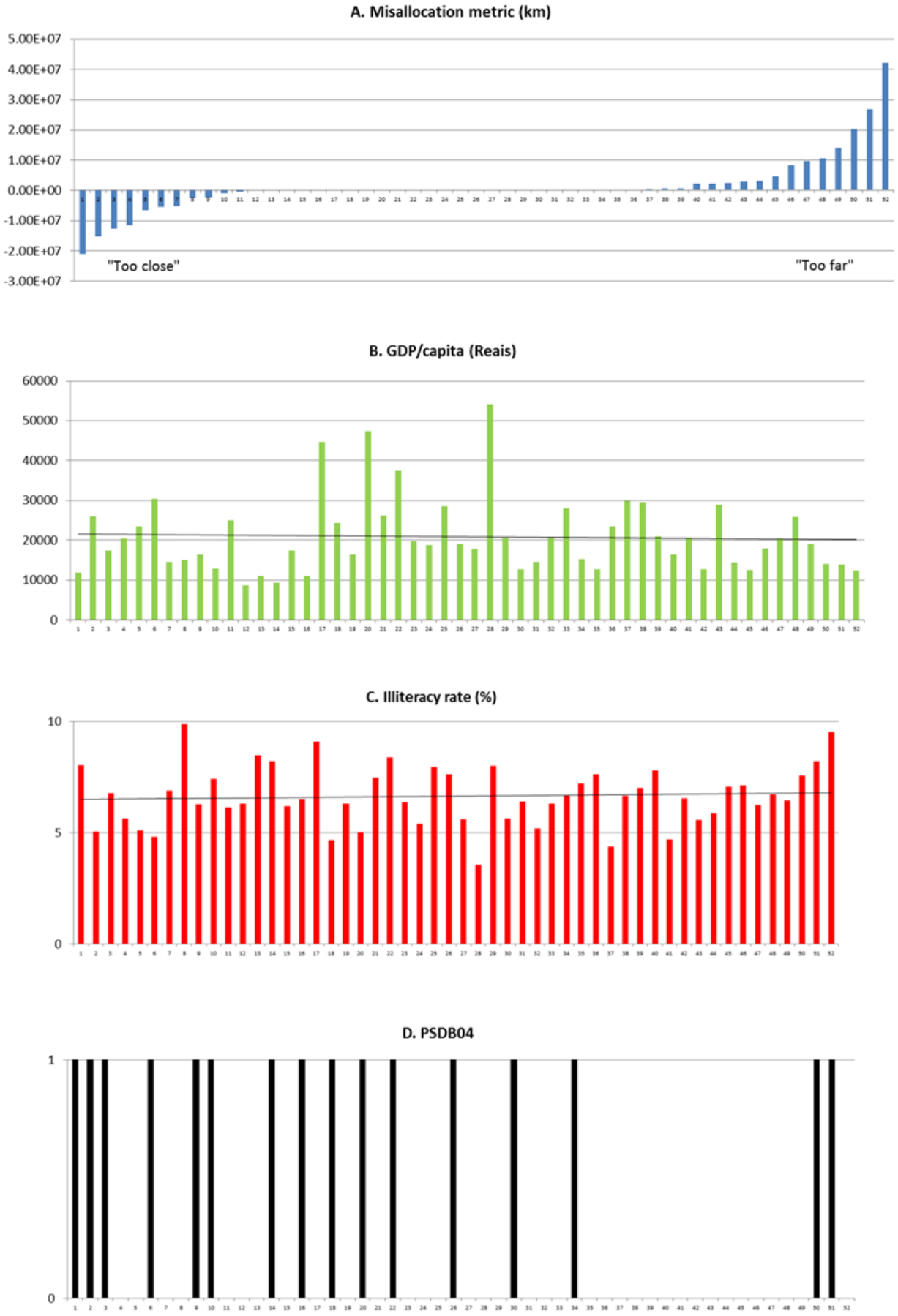

Figure 4. (A) Municipality level misallocation metric for the specification with 52 candidate locations, ranked in order of the metric, (B) GDP/capita in 2009 Reais (from SEADE), (C) adult illiteracy rate in 2000 (from SEADE), (D) dummy for whether PSDB held the mayor position in 2004-2008 (from TSE, the Superior Election Tribunal).

7. Discussion

The paper develops a method than can increase transparency in the spatial allocation of public services. It should be relatively straightforward to use the method and data sources here suggested, to construct e.g. country comparisons of access to certain public services, or of misallocation. One important characteristic of the suggested distance metric, be it the actual distance to the public service or the misallocation metric, is that it is quantitative, rather than ordinal/categorical. This is different from e.g. the Transparency International Corruption Perceptions Index or the World Bank Doing Business Index (Transparency International, 2016; World Bank, 2016). These and similar survey based measures are inherently difficult to interpret, as an increase in such indices does not have a quantitative meaning. The quantitative nature of the metric instead allows for its use in regression analysis, etc., with a precise quantitative interpretation. Although we have used travel distance in the paper, the same argument applies to travel times, and such data can also be retrieved through the methods discussed in section 2.

Consider a comparison between the access to public services in different countries. The discussion in section 6 takes each country's geography and population size as given, as well as the government budget and priorities, and then analyzes the spatial efficiency and the aggregate possible reduction of citizen travel distance for, say, an allocation of a number of Citizen Service Centers. Of interest in itself is to compare e.g. average travel distances across countries, for such services. One example of how efficiency in spending can be compared across countries would be to group countries with similar "average distance to public services", and then compare within these groups. As an example, if Sweden and Norway would have similar average distances to e.g. Citizen Service Centers or primary health care, it is likely that Norway, with more than 80% of the surface covered by mountains, spends more per inhabitant. Accepting such budget differences, we could still calculate the efficiency of the spending, by comparing the percentage degree of misallocation for the two countries.

The paper stresses transparency. To further discuss this point, the point of departure is that the allocation of public services occurs within the realms of political economy. A large literature is dedicated to modeling the political process and the incentives of politicians and voters, providing insight on e.g. redistribution and public goods allocations (Persson and Tabellini, 2000). Such models may explain why we observe policies that benefit certain groups, e.g. farmers, import-protected automobile industries, etc. Other authors study the

impact of bureaucrats' incentives, corruption, grassroots collective action, ethnic divisions, access to media and citizen monitoring, respectively, on the allocation of public goods and services.[12] A part of the literature is specifically dedicated to studying if the effects of increased information on part of the citizenry can affect outcomes. There are both successes (Björkman and Svensson, 2009; Reinikka and Svensson, 2011) and no success (Banerjee et al., 2008) in changing outcomes. If objective, quantifiable and transparent measures of the allocation of public services are available, it can be one component in providing citizens with arguments, most probably voiced through the political process, to change outcomes when access is insufficient or "unfair". The above referenced paper by Burkey et al. (2012) is an example of such information, and the reader/citizen may use it to inquire why hospital access, as compared to the optimum, is different in Tennessee than in the Carolinas and in Virginia, and also if it "should" be changed. Information such as that contained in the above figure 3 or in table 3 in Burkey et al. (2012), could become ever more available, with the method here outlined. One specific development that should become increasingly possible is to provide a web-based citizen monitoring platform of public services allocation. This could involve both the presentation of accessibility "indices" themselves, as well as a simple (limited-options) user application (i.e. "analyze the location of health care facilities in your region"). The method which should be applied is ultimately a trade-off between realistic allocation criteria, the possibility to scale up the analysis for replication in different regions/countries, and data availability. One advantage of relying on a simple location-allocation formulation is that data requirements are lower, hence allowing for a transparent analysis to be conducted.

Another use of the distance metrics is to study reasons for misallocation. Here the *disaggregate* nature of the region/county/municipality metrics is crucial. The "center", e.g. the São Paulo state government or the Swedish Tax Agency, determine a distribution of which municipalities should get services. Were some municipalities, in the case of São Paulo, politically favored? Do Swedish counties try to influence the nation-wide distribution of the allocation of Servicekontor, and other public services? A first step to analyze these questions is to have an objective measure of the allocation, and the disaggregate misallocation metric can be one such measure.

Some modifications of the location-allocation specification can be made, without changing the interpretation of the suggested metrics. These include different weights for different

[12] See e.g. Rasul and Rogger (2016), di Tella and Schargrodsky (2003), Platteau and Gaspart (2003), Alesina et al. (1999), Strömberg (2004) and Björkman and Svensson (2009), for studies in these six respective areas.

socioeconomic groups in the demand variable, different modes of transport (if data is available, which is a potential issue) and capacity constraints, which may result in that some citizens potentially will not use their closest unit. If the specification is changed to incorporate equity concerns, however, other aspects than a pure distance measure must be considered for comparisons of spatial access.

Another caveat to the interpretation of the spatial analysis is causality, and it is instructive to contrast the location-allocation assumptions to the literature on local public goods, originating with Tiebout (1956). Although this literature is largely aspatial, it predicts the level of public goods in a region/area/constituency based on different utility functions between regions and U-shaped cost curves for the establishment of public services. It assumes that citizens will move to locations with a level of public goods and taxes matching their preferences. The Tiebout model, or variants thereof, fair relatively well in explaining suburban/urban residential choices (Oates, 2006). This points to a fundamental contrast with location-allocation analysis, in which the causality goes from population distribution to public services location. A discussion of reverse causality should therefore be included when interpreting metrics of spatial access, even if the restriction to analyzing public services in non-urban areas likely diminishes the strength of the "Tiebout channel". The concerns can also be mitigated by using variables that pre-date the implementation of public services, when analyzing access, such as the use of the 2007 population figures in the above Poupatempo analysis, or to include pre-reform population growth as a potentially important determinant, as in Fredriksson and Saes, 2016. Gibbons et al. (2014) argue that causality is understudied in evaluations of spatial public policies, and also discuss the impact evaluation literature. In future work, we aim to further develop some of these methods.

References

- Alesina, A., Baqir, R., Easterly, W., 1999. Public Goods and Ethnic Divisions. The Quarterly Journal of Economics, 114(4): 1243-1284.
- Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (editors). 2015. OpenStreetMap in GI-Science: Experiences, Research, and Applications. Springer International Publishing.
- Athreya, S., Somanathan, R., 2008. Quantifying spatial misallocation in centrally provided public goods. Economics Letters, 98(2): 201-206.
- Banerjee, Abhijit V., and Esther Duflo. 2007. The Economic Lives of the Poor. Journal of Economic Perspectives, 21(1): 141-168.
- Banerjee, A., Banerji, R., Duflo, E., Glennerster, R., Khemani, S., 2008. Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education In India, NBER Working paper No. 14311.
- Beaumont, J., 1981. Location-allocation problems in a plane: A review of some models. Socio-Economic Planning Sciences, 15(5): 217-229.
- Björkman, M., Svensson, J., 2009. Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. The Quarterly Journal of Economics, 124(2): 735-769.
- Brew, J., 2015. (Mis)measurement in the study of food environment: we need better methods to solve the puzzle. Journal of Epidemiology and Community Health, 69(8): 817-818.
- Buchmueller, T., Jacobson, M., Wold, C., 2005. How far to the hospital? The effect of hospital closures on access to care. Journal of Health Economics, 25: 740-761.
- Burkey, M., Bhadury, J., Eiselt, H., 2012. A location-based comparison of health care services in four U.S. states with efficiency and equity. Socio-Economic Planning Sciences, 46: 157-163.
- Claval, P., 1984. The concept of social space and the nature of Social Geography. New Zealand Geographer, 40(2): 105-109.
- Corbacho, A., Rivas, R., 2012A. Travelling the Distance: A GPS Based study of the access to birth registration services in LA and the Caribbean. IDB Working Paper Series 307.
- Corbacho, A., Brito, S., Rivas, R., 2012B. Birth Registration and the Impact on Educational Attainment. IDB Working Paper Series 345.
- Currie, J., Reagan, P., 2003. Distance to hospital and children's use of preventive care: Is being closer better, and for whom? Economic Inquiry, 41(3): 378-391.
- Dahlgren, A., 2008. Geographical Accessibility Analysis – Methods and Applications. Licentiate thesis, Lund University.
- Di Tella, R., Schargrodsky, E., 2003. The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires. Journal of Law and Economics. 46(1): 269–92
- Fredriksson, A., 2016A. Citizen Service Centers in Brazil – Evidence from the Poupatempo reform. Manuscript.
- Fredriksson, A., 2016B. Servicekontorens rumsliga fördelning: Kan Operationsanalys bidra till en effektiv lokalisering av offentliga tjänster? Ekonomisk Debatt, 44(3): 49-61. http://nationalekonomi.se/sites/default/files/NEFfiler/44-3-Fredriksson.pdf
- Fredriksson, A., Saes, S., 2016. Determinants of Institutional Reform, Access to Public Services and Spatial Misallocation – Evidence from Citizen Service Centers in Brazil. Manuscript.
- Gibbons, S., Nathan, M., Overman, H., 2014. Evaluating spatial policies. Spatial Economics Research Centre Policy Paper 12.

- Gibson, J., Kim, B., Stillman, S., Boe-Gibson, G, 2013. Time to vote? Public Choice, 156: 517-536.
- Guagliardo, M., 2004. Spatial accessibility of primary care: concepts, methods and challenges. International Journal of Health Geographics, 3(3).
- Hakimi, S., 1964. Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. Operations Research, 12(3): 450-459.
- Huber, S., Rust, C., 2016. osrmtime: calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM). The Stata Journal, forthcoming.
- Kahle, D., Wickham, H., 2013. ggmap: Spatial Visualization with ggplot2. The R Journal 5(1): 144-161.
- Kondylis, F., Manacorda, M., 2012. School Proximity and Child Labor: Evidence from Rural Tanzania. Journal of Human Resources, 47(1):32-63.
- Marianov V., Serra D., 2002. Location Problems in the Public Sector. In Facility Location: Applications and Theory, Zvi Drezner and Horst Hamacker (Eds), Springer-Verlag, 119-150.
- Marianov V., Serra D., 2004. New Trends in Public Facility Location Modeling. Universitat Pompeu Fabra Economics and Business Working Paper No. 755.
- Oates, W., 1997. An essay on fiscal federalism. Journal of Economic Literature, 37(3): 1120-1149.
- Oates, W., 2006. The many faces of the Tiebout model. In: Fischel (Ed.), The Tiebout Model at Fifty. Lincoln Institute of Public Policy, pp. 21-45.
- Ozimek, A., Miles, D., 2011. Stata utilities for geocoding and generating travel time and travel distance information. The Stata Journal, 11(1):106-119.
- Penchansky, R., Thomas, J., 1981. The Concept of Access: Definition and Relationship to Customer Satisfaction. Medical Care, 19(2):127-140.
- Persson, T., Tabellini, G., 2000. Political Economics: Explaining Economic Policy. Cambridge : MIT Press
- Platteau, J-P., Gaspart, F., 2003. The Risk of Resource Misappropriation in Community-Driven Development. World Development, 31(10):1687-1703
- Rasul, I., Rogger, D., 2016. Management of bureaucrats and public service delivery: Evidence from the Nigerian civil service. Economic Journal (forthcoming)
- Reinikka, R., Svensson, J., 2011. The power of information in public services: Evidence from education in Uganda, Journal of Public Economics, 95(7): 956-966.
- ReVelle, C., Marks, D., Liebman, J., 1970. An analysis of private and public sector location models. Management Science, 16(11): 692-707.
- ReVelle, C., Eiselt, H., 2005. Location Analysis: A synthesis and survey. European Journal of Operational Research, 165: 1-19.
- ReVelle, C., Eiselt, H., Daskin., M., 2008. A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research, 184: 817-848.
- Rushton, G., 1984. Use of Location-Allocation Models for Improving the Geographical Accessibility of Rural Services in Developing Countries. International Regional Science Review, 9(3): 217-240.
- Shen, Y., Hsia, R., 2015. Ambulance Diversion Associated With Reduced Access To Cardiac Technology And Increased One-Year Mortality. Health Affairs, 34(8): 1273-1280.

- Strömberg, 2004. Radio's Impact on Public Spending, Quarterly Journal of Economics, 119(1):189-221
- Tiebout, C., 1956. A pure theory of local expenditures. The Journal of Political Economy, 64(5): 416-424.
- Transparency International, 2016. Corruption Perceptions Index. 15 Aug 2016 http://www.transparency.org/
- Voorheis, J., 2015. mqtime: A Stata tool for calculating travel time and distance using MapQuest web services. The Stata Journal, 15(3): 845-853.
- Wallace, D., Kahn, J., Angus, D., Martin-Gill, C., Callaway, C., Rea, T., Chhatwal, J., Kurland, K., Seymour, C., 2014. Accuracy of Prehospital Transport Time estimation. Academic Emergency Medicine, 21(1): 9-16.
- Weibull, J., 1976. An axiomatic approach to the measurement of accessibility. Regional Science and Urban Economics, 6: 357-379.
- World Bank, World Development Report 2004, 2003. Making Service Work for Poor People. Washington, DC: World Bank and Oxford University Press.
- World Bank, 2016. Doing business, ease of doing business. 15 Aug 2016 http://www.doingbusiness.org/
- Zeigermann, L., 2016. OPENCAGEGEO: Stata module for forward and reverse geocoding. Manuscript.

Appendix. Location-allocation of Swedish Servicekontor, specification and robustness

There is a total of 103 Servicekontor in Sweden. For Stockholm County, which has seven offices, three Servicekontor in the downtown area of the Stockholm municipality were assumed to be located in the same location (thus effectively considered as one unit in the analysis), and the remaining four units in their actual location (two of these units had to be "moved" slightly, as the exact locations did not have road data). The four units in the municipality of Göteborg, Sweden's second most populous city, were considered to be located in the same downtown location (thus effectively considered as one unit in the analysis). These 11 units (or 6 "merged" units) were thus all treated as "fixed", as was the unit in the municipality of Malmö (the third most populous). With one unit on the island of Gotland (which is also a county and a municipality), an area altogether excluded from the study, which reduces the total number of Servicekontor to 102, we thus analyzed the location of the remaining 90 units. Sweden has 290 municipalities (289 excluding Gotland), 26 of which are in the Stockholm county, which we excluded from the candidate list. We furthermore excluded, as candidates, Göteborg, and three municipalities in the same urban area. Malmö was also excluded, together with one other (adjacent) municipality. The p-median problem was thus solved with n=257 (289-26-4-2) and p=90 (102-7-4-1), n'=264 and p'=97 (according to the above reasoning), $J=\{j_1,...,j_{264}\}$, $J_F=\{j_1,...,j_7\}$, $J_{-F}=\{j_8,...,j_{264}\}$ and $I=\{i_1,...,i_{1937}\}$.

In order to study if the solution to the largely rural Swedish p-median problem is sensitive to the level of aggregation of the population data, we conducted an alternative analysis, based on a 2002 manually compiled (fastest) road distance database from the Swedish Transport Administration. Excluding Stockholm County and Gotland altogether, and excluding also a few small candidate locations due to limitations in the latter road data, we thus solved i) the p-median with n=246, p=90 and 404 population points, based on the 2002 road distance database, and ii) the p-median with n=246, p=90 and 1803 population points, based on the more detailed distance database, described in the main text. Comparing the two solutions, 87 out of 90 optimal locations are the same. The 2002 distance database is the fastest route between the city centers (in this sense similar to the São Paulo distance database). In addition to the difference in size of the areas for which the population is assigned to a single point, we thus also have a centroid-center aggregation difference. For the specific case studied, however, the aggregation problem discussed in section 2.1 does not seem to play a decisive role in determining allocations.