

Loumeau, Gabriel

Working Paper

Locating public facilities: Theory and micro evidence from Paris

KOF Working Papers, No. 452

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Loumeau, Gabriel (2019) : Locating public facilities: Theory and micro evidence from Paris, KOF Working Papers, No. 452, ETH Zurich, KOF Swiss Economic Institute, Zurich,
<https://doi.org/10.3929/ethz-b-000332703>

This Version is available at:

<https://hdl.handle.net/10419/204789>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Paper

Locating Public Facilities: Theory and Micro Evidence from Paris

Author(s):

Loumeau, Gabriel

Publication Date:

2019-03

Permanent Link:

<https://doi.org/10.3929/ethz-b-000332703> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

KOF Swiss Economic Institute

**Locating Public Facilities:
Theory and Micro Evidence from Paris**

Gabriel Loumeau

KOF Working Papers, No. 452, March 2019

KOF

ETH Zurich
KOF Swiss Economic Institute
LEE G 116
Leonhardstrasse 21
8092 Zurich, Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Locating Public Facilities: Theory and Micro Evidence from Paris

Gabriel Loumeau *
ETH Zürich

March 12, 2019

Abstract

This paper investigates the problem of the optimal location of public facilities. I develop a quantifiable model in which the central planner decides on a location strategy, which includes the geographical location and the capacity of public facilities, while anticipating how individuals and firms will react. The central planner's objective is to maximize aggregate welfare. I calibrate the model to fit the economic and geographic characteristics of the Paris metropolitan area at a $1\text{km} \times 1\text{km}$ geographic resolution and focus on secondary schools as an example of public facilities. The counterfactual analysis, which compares the optimal and observed location strategy between 2001 and 2015, suggests that adopting the optimal strategy in any year would have increased welfare growth by about 12%. Half of the effect is attributable to improvements in channels other than shorter commutes to the public facility, mostly via lower housing prices and shorter commutes to the workplace. The analysis also reveals that the observed location strategy disproportionately favored short commutes in central locations and led to a mis-allocation of residential and commercial activities between the center and the periphery.

Keywords: Public facility location, agglomeration, commuting, gravity, Big Cube-Small Cube.

JEL classification: H11; R53; R41.

* *Corresponding author:* ETH Zürich, Department of Management, Technology, and Economics, Leonhardstrasse 21, 8092 Zurich, Switzerland; E-mail: loumeau@kof.ethz.ch; +41 44 632 67 58.

I am grateful to seminar and conference participants in Barcelona, Aix-en-Provence, Munich, Tampere, Cologne and Paris. In particular, I thank Gabriel Ahlfeldt, Peter Egger, Marko Koethenburger, Etienne Lehmann, Nicole Loumeau, Mohammed Mardan, Yasusada Murata, Dávid Nagy, Ariell Reshef, Frédéric Robert-Nicoud, Laurent Simula, and Nora Strecker for helpful comments and suggestions; and Juliette Caillaudeau from the "Direction de l'Evaluation, de la Prospective et de la Performance" (DEPP) for great help in accessing the data.

1 Introduction

How to choose the location of new public facilities? Naturally, minimizing the distance between users and facilities is crucial for the facilities' accessibility and efficiency. In the case of health care services, [Buchmueller, Jacobson, and Wold \(2006\)](#) and [Nicholl, West, Goodacre, and Turner \(2007\)](#) show that proximity to a hospital strongly influences the chances of recovery after an accident. Looking at educational services, shorter travel distances imply shorter daily commutes and higher attendance rates ([Card, 1993](#); [Frenette, 2006](#)).¹ However, travel distance to a public facility is not the only factor to consider. Individuals and firms react to the location of facilities in deciding their own locations. Typical evidence of such an endogenous reaction to facility location is the capitalization of public goods provision into housing prices.² Hence, to answer the question of where to locate public facilities appropriately, one needs to analyze not only the pure distance-minimizing problem, but also how individuals and firms react to facility locations. Despite widespread policy interest in this question, an approach that can investigate the problem of locating public facilities in a comprehensive manner has not been proposed in the literature.

In this paper, I develop and apply a quantifiable spatial general equilibrium model of public service³ location to evaluate location decisions. In the model, the central planner decides on a location strategy, i.e., discrete location and capacity of public facilities, while observing the location's characteristics and anticipating how individuals and firms will react. The modeling approach is based on a framework à la [Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#) which applies tools developed in quantifiable models of international economics (see, among others, [Eaton and Kortum, 2002](#)) to an urban area setting with commuting. The novelty in the model is twofold. First, individuals decide on the commute between residential location, workplace location, and a public facility in order to maximize their utility. Utility is affected by the distance to public facilities, the

¹Longer distances to public services also induce several adverse indirect effects (see, among others, [Alvarez-Pedrerol, Rivas, López-Vicente, Suades-González, Donaire-Gonzalez, Cirach, de Castro, Esnaola, Basagana, Dadvand, Nieuwenhuijsen, and Sunyer, 2017](#) and [Liu, Ma, Liu, Han, Chuang, and Chuang, 2015](#)).

²Early literature includes [Tiebout \(1956\)](#), [Oates \(1969\)](#), and [Yinger \(1982\)](#), among others. See [Reback \(2005\)](#), [Cellini, Ferreira, and Rothstein \(2010\)](#), and [Fack and Grenet \(2010\)](#) for more recent estimates of the magnitude of such capitalization effects.

³In the remainder of the paper, I will refer to a "facility" as a particular establishment (e.g. a specific school) and to a "service" as the full set of facilities of a particular type (e.g. all schools).

distance to the workplace, wages, and residential prices, among other factors. Hence, the central planner’s location strategy, the individual residential and workplace locations, and the public facility used by each individual are simultaneously chosen. Second, the optimal location of public facilities is an outcome of the model, rather than an exogenous location fundamental.

I refer to an optimal location strategy as one that maximizes aggregate welfare. The methodological novelty in solving the optimization problem is that locations are set in an Euclidean space that permits the approximation of the welfare problem with a convex function which can then be used as the upper bound for the problem in a Branch-and-Bound approach. Precisely, the approach employed in this paper closely follows the Branch-and-Bound algorithm labeled *Big Cube - Small Cube* developed by [Schöbel and Scholz \(2010\)](#), which in turn follows early work by [Hansen, Peeters, Richard, and Thisse \(1985\)](#). One key advantage of this approach is that it does not require the welfare problem to be convex or even differentiable. I show that the approach converges to the *global* optimum.

The framework of the paper offers sufficient flexibility to take it to the data. I proceed in two steps. First, the model is calibrated to fit the characteristics of the Paris metropolitan area in terms of residential and commercial densities, as well as in terms of housing and labor prices. The choice of the application is directed by the availability of precise data on public facilities over time, population location, as well as on bilateral commutes between residential, public facility, and workplace locations. As the unit of analysis, I use a grid of $1\text{km} \times 1\text{km}$ cells, covering a square defined by 50km sides and taking the Paris city center as the centroid.⁴ Using precise geo-localized data, I attribute the total number of individuals, the total number of secondary-school-age pupils, the share of residential and commercial land usage, housing prices and local income to each cell.

Second, I run counterfactual exercises to evaluate the location choices of newly-built secondary schools made between 2001 and 2015. As an example of public services, I use French “Collèges”, i.e., lower secondary schools. As an example of public facilities, French lower secondary schools offer several advantages, which I detail in Section 6. Let me introduce already the most important ones. First, differences across public lower secondary schools are minimal. The same national degree is offered, teachers’

⁴The grid is created using the Geographic Information System ArcGIS.

salaries do not vary by location, teachers must pass the same qualifying exams, and state spending per pupil is equal across facilities. Second, due to the centralized financing of the secondary school system, there are (almost) no spatial differences in the amount that individuals pay for schooling. Third, the location of private secondary schools has little effect on the location strategy of public secondary schools. This is due to, both, the location strategy adopted by private secondary schools, and their generally non-secular nature. Fourth, I can precisely identify the pupils in the public service, since education is compulsory in France until the age of 16.⁵ Lastly, both matrices of the universe of bilateral commutes from residential locations to public schools and to workplaces are available for the region.

The counterfactual exercise reveals that adopting the welfare-maximizing strategy in a given year leads to a 12,07% increase in welfare growth. Only half of the welfare gains are attributable to reduced commutes to public facilities, the other half arises via general equilibrium channels (mostly via lower housing prices and shorter commutes to the workplace). Furthermore, this exercise also shows that welfare gains from adopting the optimal strategy only arise when more than two new facilities are located. The mean difference in welfare growth, conditional on three facilities being located, is then 18% and up to 33% in 2005. The more new facilities are located, the easier it is to gear the general equilibrium towards higher welfare levels.

Using the difference between the observed and optimal strategies at the cell level as a measure of inefficiencies, I identify three general patterns. First, the observed strategy disproportionately favors lower commutes in central places (i.e. within 15km of the city center). Second, the observed strategy leads peripheral places (i.e. places located between 15 and 25km from the center) to be over-specialized in commercial activity and central places to be over-specialized in residential activity. Third, places with high amenity or high productivity fundamentals are both inefficiently specialized in residential activity compared to low amenity places.

The paper proceeds as follows. In Section 2, I briefly discuss the related literature. I present a case study of how a facility location is chosen in Section 3. I describe the model and its main assumptions in Section 4 and solve for the equilibrium and optimal solution in Section 5. In Section 6, I introduce the data. I estimate the model structurally, calibrate it, and assess its performance in Section 7, before presenting the counterfactual

⁵In France, the average pupil attends lower secondary education between the ages of 11 and 15.

analysis in Section 8. I predict the optimal future location strategy and the optimal school district boundaries in Section 9. Section 10 concludes.

2 Relation to the literature

This paper relates to the literature in various ways. In its general intention, this paper is part of a recent strand of the literature that has approached economic geography questions from an optimal design perspective. Examples of such work are [Fajgelbaum and Schaal \(2017\)](#), who study the optimal design of a transport network, [Albouy, Behrens, Robert-Nicoud, and Seegert \(2016\)](#), who investigate optimal city size, and [Allen, Arkolakis, and Li \(2015\)](#), who look at the optimal structure of a city. In its approach, this paper builds upon recent developments in urban economics (see [Ahlfeldt et al., 2015](#)), which applies quantitative spatial general equilibrium models in rich geographic contexts.

The specific question of the endogenous location of discrete public facilities has received relatively little attention in economic research.⁶ [Fujita \(1986\)](#) and [Thisse and Wildasin \(1992\)](#) were the first to analyze the endogenous location problem in the context of an urban area. In these models, the location decisions of firms and households are a function of the location of the public facility.⁷ However, compared to this paper, these analyses remain theoretical. In fact, most of the work on the issue at hand comes from operations research and the question has been to solve various forms of the classical Fermat-Weber location problem.⁸ This body of literature often takes the location of individuals (or costumers) as well as land prices as given; hence, ignoring the endogeneity of individuals' and firms' location decisions. [Thisse and Zoller \(1983\)](#) provide a comprehensive review of the literature on public facility location at the intersection of economics and operations research.

More generally, abstracting from the question of the location of public facilities at discrete points, a large literature has analyzed how local amenities and public goods

⁶The case of transport infrastructure constitutes an exception: As a key determinant of both trade and commuting costs, the relationship between transport infrastructure and the spatial distribution of economic activity has received significantly more attention (see, among others, [Baum-Snow, 2007](#); [Redding, Sturm, and Wolf, 2011](#); [Fajgelbaum and Schaal, 2017](#); [Donaldson, 2018](#)).

⁷More recently, [Berliant, Peng, and Wang \(2006\)](#) and [Belletini and Kempf \(2013\)](#) have also modeled the location problem by accounting for the endogeneity of individual locations.

⁸For a review of the facility location problem literature in Operations Research, see [ReVelle and Eiselt \(2005\)](#) and [Farahani and Hekmatfar \(2009\)](#).

determine individuals' and firms' residential locations (see, among others, [Schuler, 1974](#); [Turnbull, 1989](#); [Bénabou, 1993, 1996](#); [Brueckner, Thisse, and Zenou, 1999](#)). Recently, [Gagné, Koster, Moizeau, and Thisse \(2017\)](#) focus on within-city sorting of income-heterogeneous individuals with exogenously given amenities. Naturally, this paper also relates to the literature on the capitalization effects of public good provision. The capitalization of a better tax-benefit linkage into housing prices has been described in early literature ([Tiebout, 1956](#); [Oates, 1969](#); [Hamilton, 1976](#); [Yinger, 1982](#)). More recently, [Reback \(2005\)](#), [Cellini et al. \(2010\)](#) and [Fack and Grenet \(2010\)](#) have focused on estimating the magnitude of the capitalization effect into property prices. [Epple, Gordon, and Sieg \(2010\)](#) propose a multi-jurisdiction public finance model with location amenities.

With secondary schools as the subject of the application, this paper also relates to the literature studying education systems in a general equilibrium context. [Epple and Romano \(1998\)](#) study the provision of education with competition between public and private schools. [Epple, Romano, and Sieg \(2006\)](#) and [Epple, Romano, Sarpça, and Sieg \(2017\)](#) focus on higher education in the U.S. context. This literature has generally focused more on the interaction between tax-financed public provision and tuition-financed private provision and how these affect the quality and level of provision, rather than on the question of the location of facilities.

Finally, this paper also contributes to the general discussion about urban structure formation by focusing on a relatively understudied type of agglomeration force, namely the provision of public services.⁹ In addition to the more standard agglomeration forces studied in the literature on New Economic Geography following [Krugman \(1991\)](#), this paper emphasizes the importance of public services in explaining the formation of core-periphery structures. In a way, the reasoning presented in this paper follows the line of classical economic geography theories, such as the Central Place Theory developed by [Christaller \(1933\)](#), which focuses on the provision of services to surrounding areas in order to explain the existence of cities.

⁹See [Fujita, Krugman, and Venables \(2001\)](#) for an overview of the theoretical literature on urban structures and [Overman, Redding, and Venables \(2001\)](#) for an overview of the empirical literature on the same topic.

3 Case study: Locating a secondary school in Ile-de-France

In this section, I provide anecdotal evidence of how a facility location is chosen. Understanding the actual decision process leading to a particular location choice helps the development of a realistic and relevant framework. Precisely, I detail the process that led to the opening of a new secondary school in the municipality of Montreuil in September 2018.¹⁰

Who took the decision? The decision *to open* a new school in Montreuil was taken by the Département of Seine-Saint-Denis (NUTS3 region).¹¹ The decision *on the location* of the newly-built secondary school in Montreuil was taken jointly by a large number of administrative bodies. On top of the municipality and the Département (NUTS3 region), the central government, the Région (NUTS2 region), the cooperation of municipalities in which Montreuil is a member, the public transport authority, the department of national and regional parks, as well as several local associations were all involved in deciding where to locate the new school.

Where was it located? The school was ultimately built on unoccupied land bordering a municipal park owned by the Département.¹² The first striking observation when investigating the chosen location is that five existing public schools and one private schools are within a radius of 1km from the new location. The second observation is that the new school location is next to two development zones (Zone d'Aménagement Concerté (ZAC), in French) created in 2011. These zones are areas in which new housing space was built; and hence, where the population density increased.

How was the location justified? The project proposal¹³ argues that, without a new school, the local demand for school seats would get close to full capacity in 2022.¹⁴ It was argued that building a new school would lead to a total capacity that is significantly

¹⁰Montreuil is a municipality on the eastern border of Paris intra-muros; hence, within the area studied in the application.

¹¹The “Plan Ambition Collège 2015-2020” was passed on November 27th, 2014. [Link to project website](#) (in French).

¹²Figure 14 displays the Local Urban Plan (Plan Local d'Urbanisme (PLU), in French) which was active when the new school location was selected. Author's own illustration based on the project proposal and the PLU from September 13, 2012.

¹³[Link to project proposal](#) (in French).

¹⁴Figure 15 display the evolution of offer and demand of school seats with and without a new school.

higher than the number of students (by about 23% in 2018 and by about 18% in 2022). Note, however, that full capacity would not have been reached by 2022, even without a new school.

Why is the justification not enough? In looking at the project proposal, two pitfalls become apparent: (i) lack of accounting for effects on economic variables, such as housing prices and commutes and (ii) lack of comparison to other locations in order to justify the choice of this particular location in that municipality. Such pitfalls are not specific to the French context. The same shortcomings can be observed in the otherwise very informative “Oregon School Siting Handbook” (Kunec, 2005) or in more standard research in urban planning (Moussa, Mostafa, and Elwafa, 2017). In both practical and more academic literature in urban planning such effects are often ignored.¹⁵

4 Theoretical framework

Consider a metropolitan area that occupies a closed and bounded subset S of a two-dimensional Euclidean space. A location is a point in S . There are two sets of locations: $I \in S$ city blocks and $K \in S$ public facility locations. Each block i has an effective supply of floor space L_i which can be used for residential or commercial purposes. The fraction of floor space used commercially or residentially are denoted θ_i and $1 - \theta_i$, respectively. The distance between two locations is given by the Euclidean norm. To study changes in expected utility, the metropolitan area is populated by an exogenous measure of H individuals and assumed to be closed.

4.1 Individuals

The utility of individual o living in place i , working in place j , and using a facility in site k is linear in an aggregate consumption index C_{ijko} , such that: $U_{ijko} = C_{ijko}$. This consumption index depends on consumption of the single final good (c_{ijko}), consumption of residential floor space (l_{ijko}), the utility from residential amenities in i (B_i), the disutility from commuting from residential place i to public facility location k ($d_{ik} \geq 1$), the disutility from commuting from residential location i to workplace j ($d_{ij} \geq 1$), and an idiosyncratic shock that is specific to the individual and varies with the individual’s

¹⁵Vincent (2006) constitute a notable exception by arguing in favor of a higher integration of school location decisions into general city planning in the U.S. context.

location choice (z_{ijko}) . The disutility from commuting from place i to place k is modeled as an iceberg cost $d_{ik} = e^{\nu\rho_{ik}} \in [1, \infty)$, which increases with the Euclidean distance ρ_{ik} between places i and k . Similarly, the disutility of commuting from residential location i to workplace j is defined as $d_{ij} = e^{\eta\rho_{ij}} \in [1, \infty)$, where ρ_{ij} is the distance between i and j . The aggregate consumption index is assumed to take the following Cobb-Douglas form:

$$C_{ijko} = z_{ijko} \frac{B_i}{d_{ik}d_{ij}} \left(\frac{c_{ijko}}{\beta} \right)^\beta \left(\frac{l_{ijko}}{1-\beta} \right)^{1-\beta}, \quad 0 < \beta < 1. \quad (1)$$

The idiosyncratic shock (z_{ijko}) describes the heterogeneity in utility that individuals derive from living in i , working in j , and using the public service in k . For each individual, this idiosyncratic component is drawn from an independent Fréchet distribution following [McFadden \(1974\)](#) and [Eaton and Kortum \(2002\)](#):

$$F(z_{ijko}) = e^{-z_{ijko}^{-\epsilon}}, \quad (2)$$

where $\epsilon > 1$ is the shape parameter that controls the dispersion of idiosyncratic utility.

After observing their realization of the idiosyncratic utility, individuals choose their place of residence, their workplace, and their public facility to maximize their utility. Assuming the price of the final good to be the numeraire, the individual demands for the final good (c_{ijko}) and residential floor space (l_{ijko}) are as follows:

$$c_{ijko} = \beta w_j \quad (3)$$

and

$$l_{ijko} = (1-\beta) \frac{w_j}{Q_i}, \quad (4)$$

where w_j is the wage received by individuals working in j and Q_i is the housing price in location i . Here, I make the standard assumption that rent is accrued by absentee landlords and is therefore not spent within the city. Substituting (3) and (4) into (1), I obtain the following indirect utility function:

$$U_{ijko} = \frac{B_i z_{ijko} w_j Q_i^{\beta-1}}{d_{ik} d_{ij}}. \quad (5)$$

4.2 Individuals' location choices

I look at the combined location choices of individuals across the metropolitan area. Following [Ahlfeldt et al. \(2015\)](#), I derive the probability that individuals choose a particular

combination of residence, workplace and public service location:^{16 17}

$$\begin{aligned}\pi_{ijk} &= Pr[u_{ijk} \geq \max\{u_{ijk}\}; \forall i, j, k] \\ &= \frac{(d_{ik}d_{ij}Q_i^{1-\beta})^{-\epsilon}(B_iw_j)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_rw_s)^\epsilon} \equiv \frac{\Phi_{ijk}}{\Phi}.\end{aligned}\quad (6)$$

Equation (6) shows that individuals sort across all combinations of residence and public service locations depending on their idiosyncratic preferences and the characteristics of these locations.

I can derive the probability that an individual decides to live in i among all possible locations in the metropolitan area by summing Φ_{ijk} across all workplaces and public service locations.

$$\pi_{Ri} = \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it}d_{is}Q_i^{1-\beta})^{-\epsilon}(B_iw_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_rw_s)^\epsilon} \equiv \frac{\sum_{s=1}^I \sum_{t=1}^K \Phi_{ist}}{\Phi} \quad (7)$$

The same can be done for the probability that an individual decides to use the public service in k out of all possible public service location choices:

$$\pi_{Sk} = \frac{\sum_{r=1}^I \sum_{s=1}^I (d_{rk}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_rw_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_rw_s)^\epsilon} \equiv \frac{\sum_{i=1}^I \sum_{j=1}^I \Phi_{rsk}}{\Phi}. \quad (8)$$

Similarly, the probability of working in j can be expressed as:

$$\pi_{Mj} = \frac{\sum_{r=1}^I \sum_{t=1}^K (d_{rt}d_{rj}Q_r^{1-\beta})^{-\epsilon}(B_rw_j)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_rw_s)^\epsilon} \equiv \frac{\sum_{j=1}^I \sum_{k=1}^K \Phi_{rjt}}{\Phi}. \quad (9)$$

Note that the number of people residing in i , working in j , and using the facility in k , H_{ijk} ; residing in i , H_{Ri} ; working in j , H_{Mj} ; and using the service in k , H_{Sk} ; can be obtained by multiplying the above probabilities by the total number of individuals, H , or the total number of users, \tilde{H} .¹⁸

4.3 Production

I assume a single final good is traded without costs within the metropolitan area. Production occurs under perfect competition and constant returns to scale. For simplicity, I assume the following Cobb-Douglas production function:

$$y_j = A_j \left(H_{Mj} \right)^\alpha \left(L_{Mj} \right)^{1-\alpha}, \quad (10)$$

¹⁶For full derivation, see Section C of the Appendix.

¹⁷In what follows, I will refer to residential locations using indices i and r , to workplace locations using indices j and s , and to facility locations by k and t .

¹⁸In the case of schools, the total number of users is equal to the number of pupils.

where y_j is the output of the final good in workplace j , A_j is the final goods productivity, H_{Mj} is the total workplace employment, and L_{Mj} is the land used.

Firms choose their place of production, their level of employment, and their level of floor space consumption. From the first order conditions for profit maximization, the number of workers working in j is:

$$H_{Mj} = \left(\frac{\alpha A_j}{w_j} \right)^{\frac{1}{1-\alpha}} L_{Mj}. \quad (11)$$

Therefore, employment in place j is increasing in productivity (A_j), decreasing in wages (w_j), and increasing in commercial land use (L_{Mj}).

The equilibrium commercial floor price is determined by the zero profit condition along with profit maximization:

$$q_j = (1 - \alpha) \left(\frac{\alpha}{w_j} \right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}}. \quad (12)$$

Given the Cobb-Douglas form of the production function, the price index p_j in location j is:

$$p_j = A_j \left[\left(\frac{\alpha}{1-\alpha} \right)^{1-\alpha} + \left(\frac{\alpha}{1-\alpha} \right)^{-\alpha} \right] w_j^\alpha q_j^{1-\alpha}. \quad (13)$$

Assuming that individuals can consume from every location within the city without costs, the consumption price index P can be expressed as:

$$P = \left(\sum_{j=1}^I p_j^{-\epsilon} \right)^{-1/\epsilon}. \quad (14)$$

Note that all price variables are consequently expressed in Consumption Equivalent Units (CEU) or, in other words, relative to the price of the single final consumption good such that: $Q = \frac{\mathbf{Q}}{P}$, $q = \frac{\mathbf{q}}{P}$, $w = \frac{\mathbf{w}}{P}$, where bold variables refer to observed prices.

4.4 Land market

The fraction of floor space used for commercial purposes is $\theta_i \in (0, 1)$. At the equilibrium, no location is fully specialized.¹⁹ Therefore, in equilibrium, $q_i = Q_i$ in all locations.

¹⁹I assume no full specialization and an exogenous θ for two reasons. First, in many real world metropolitan areas, slow moving zoning policies are in place to restrict the transition of land from residential to commercial use (or from commercial to residential). Therefore, taking θ as given appears a straightforward way to account for such policies without explicitly modeling them in detail. Second, taking θ as given speeds up the solution for the model. This is crucial when I later optimize welfare using a Branch-and-Bound approach. With this assumption, the optimization algorithm converges to

In the model, land prices are determined by the individuals' and firms' location choices. As is often done in the urban literature, I consider that floor space L is provided by a competitive construction sector. I follow [Epple et al. \(2010\)](#) and [Ahlfeldt et al. \(2015\)](#) and assume that the production function takes the Cobb-Douglas form: $L_i = M_i^\mu N_i^{1-\mu}$, where M is capital and N is land. The corresponding dual cost function for floor space is then $Q_i = \mu^{-\mu}(1-\mu)^{-(1-\mu)}P^\mu R_i^{1-\mu}$, where Q_i is the price for floor space, P is the common price for capital, and R_i is the price for land. Since the price for capital is the same across all locations, the relationships between the quantities and prices for geographical land and floor space can be expressed as:

$$L_i = \phi_i N_i^{1-\mu}, \quad (15)$$

$$Q_i = \chi R_i^{1-\mu}, \quad (16)$$

where $\phi_i = M_i^\mu$ determines the density of development and χ is a constant.

The residential land market clearing implies that the demand for residential and commercial floor space equals the supply of floor space in each location: L_i . The residential land market clearing condition can then be written as:

$$\mathbb{E}[l_i]H_{Ri} = (1-\beta)\frac{w_j}{Q_i}H_{Ri} = (1-\theta_i)L_i. \quad (17)$$

The commercial land market clearing imposes:

$$\left(\frac{(1-\alpha)A_j}{q_j}\right)^{1/\alpha} H_{Mj} = \theta_j L_i. \quad (18)$$

The clearing condition for total demand for space can then be expressed as:

$$(1-\theta_i)L_i + \theta_i L_i = L_i = \phi_i N_i^{1-\mu}. \quad (19)$$

4.5 Central planner

The central planner decides on the location and capacity of public facilities, under consideration of location characteristics and anticipating how individuals and firms will react. To model the fact that facilities are often built on publicly owned, available land, I assume that the central planner does not consume land. Note also that, in order to properly investigate the location side of the problem, I deliberately abstract from the the global optimal *within* a reasonable time. However, a version of the model with endogenous share of land uses is available upon request. The uniqueness and existence properties of the equilibrium also hold in that case.

financing side by considering the number of facilities as exogenous. This is done in order to avoid entering lengthy discussions on the appropriate level of public spending, as well as discussions on the appropriate capacity of a public facility.²⁰

For brevity, I refer to any planner’s combined choice as a *location strategy*. The planner’s objective is to maximize the aggregate welfare associated with providing the public service. Formally, the planner decides on a location strategy by choosing the set of geographic coordinates $\mathbf{k} = [\tilde{k}_1, \dots, \tilde{k}_K]$ of the facilities.²¹ Hence, the planner’s maximization problem is:²²

$$\max_{\tilde{k}_1, \dots, \tilde{k}_K} \mathbb{E}[u] = \Gamma \left(\frac{\epsilon - 1}{\epsilon} \right) \left[\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon \right]^{(1/\epsilon)}. \quad (20)$$

$\mathbb{E}[u]$ represents the expected utility level in the metropolitan area. Γ is the Gamma function.

4.6 Discussion of assumptions

To complete the description of the model, a number of dimensions deserves further attention. First, the model assumes that all individuals are affected by distance to the workplace and distance to the public facility. Clearly, not everyone works or uses a given public service. Hence, individuals in the model can be thought of as members of a family with working parents and children. All family members value the dis-utility of commuting incurred by other family members. Second, the model deliberately ignores administrative borders as the location decision is, in the vast majority of real cases, taken by many different entities (see Section 3). Hence, ignoring administrative borders is therefore a straightforward way to model this issue in a simplified manner. Finally, distances between locations are given by the Euclidean norm, which ignores existing transport networks. This implies that the benefit of transport networks for a given location will feed into residential amenities, final goods productivity, and density of development. Locations close to transport nodes will be more attractive along these dimensions.

²⁰See [Afsa \(2014\)](#) for an empirical analysis of the various effects of a facility’s capacity using the French secondary education facilities.

²¹Note that, while k indexes a facility, \tilde{k}_k refers to its coordinates.

²²For full derivation of the expected utility in the metropolitan area, see Section D of the Appendix.

5 Equilibrium and optimality

I solve for the equilibrium and optimality of the model in two steps. First, I derive the optimal behavior of individuals and firms conditional on a planner’s location strategy. Second, I solve for the optimal location strategy anticipating the behavior of individuals and firms.

5.1 Individuals’ behavior conditional on a location strategy

Conditional on a location strategy, the general equilibrium of the model can be referenced by the following five vectors $\{\boldsymbol{\pi}_R, \boldsymbol{\pi}_S, \boldsymbol{\pi}_M, \boldsymbol{Q}, \boldsymbol{w}\}$ and a scalar $\{P\}$. Given this equilibrium vector and scalar, all other endogenous variables of the model can be determined. This equilibrium vector solves the following five of equations: population share in residential location (7), population share in public facility location (8), population share at workplace (9), the price index (14), firm profit maximization (12), and residential land market clearing (17).

Proposition 1 *Assuming strictly positive, finite, and exogenous characteristics ($H \in (0, \infty)$, $\tilde{H} \in (0, \infty)$, $A_i \in (0, \infty)$, $B_i \in (0, \infty)$, $N_i \in (0, \infty)$, $\rho_{ik} \in (1, \infty) \times (1, \infty)$, $\rho_{ij} \in (1, \infty) \times (1, \infty)$), there exist unique general equilibrium vectors $\{\boldsymbol{\pi}_R, \boldsymbol{\pi}_M, \boldsymbol{\pi}_S, \boldsymbol{Q}, \boldsymbol{w}\}$ and scalar $\{P\}$.*

PROOF: See the proof of Proposition 1 in Appendix E.

Q.E.D.

5.2 Deriving the optimal location strategy

Solving the maximization problem of the central planner is complex for at least two reasons. First, the dimensionality of the problem makes any standard search-type approach infeasible. Second, the central planner’s problem does not have an explicitly convex representation. In this paper, I propose a branch-and-bound algorithm inspired by the “Big Cube - Small Cube (BCSC)” method by Schöbel and Scholz (2010).²³ This algorithm possesses several key properties which makes it particularly attractive in this setting.

²³This approach generalizes the “Big Square-Small Square” approach by Hansen et al. (1985) to the multi-dimensional case.

First, the algorithm converges to the *global* optimum of the problem. Second, it does not require the analyzed problem to be convex, or even differentiable.

The Big Cube - Small Cube (BCSC) approach The Big Cube - Small Cube is a branch-and-bound algorithm which converges by successively eliminating sub-spaces from the overall space considered until a small solution space around the global optimum remains. The removal of a sub-space from the possible solution space is done by comparing, an upper bound for the problem if the solution lies within the sub-space, to, the best solution obtained at that point. If the upper bound is lower than the best solution, the sub-space does not contain the optimum. The upper bound is obtained by approximating the central planner's problem with a convex function $\bar{f}(\mathbf{k})$ with $\mathbf{k} = [\tilde{k}_1, \dots, \tilde{k}_K]$ which fulfills the following condition: $\bar{f}(\mathbf{k}) \geq f(\mathbf{k})$, where $f(\mathbf{k})$ is the central planners problem. The name of the algorithm comes from the fact that the sub-spaces considered are polyhedrons (i.e. cubes in the two facilities case).

More precisely, the BCSC approach proceeds by splitting the search space into equal polyhedrons and computing the value of $f(\mathbf{k})$ at the center of each of them. The best solution is stored and an upper bound is then calculated for each polyhedron. All polyhedrons with an upper bound *lower* than the highest obtained value can be discarded. All other polyhedrons are divided into smaller polyhedrons, and added to the list of polyhedrons to be searched.

For simplicity, I present the version of the algorithm with two new facilities to locate ($K = 2$) in detail. Remember that in the setting at hand, the central planner faces a *maximization* problem. Denote by $f(\tilde{k}_1, \tilde{k}_2)$, the function that the central planner maximizes when locating the additional facilities. Consider the feasible area ψ where the new facilities can be located:

$$\psi = [\underline{x}^1, \bar{x}^1] \times [\underline{y}^1, \bar{y}^1] \times [\underline{x}^2, \bar{x}^2] \times [\underline{y}^2, \bar{y}^2] \subset \mathbb{R}^{2K}, \quad (21)$$

where x and y refer to coordinates. $\underline{\cdot}$ and $\bar{\cdot}$ denote the lower and the upper bound of the feasible area, respectively. At first, $\psi = S \times S$. The method proceeds as follows:

Big Cube-Small Cube algorithm

Start an iteration count, $iter = 1$.

1. Create a list Ψ of feasible areas. At first, Ψ only contains ψ .

2. Select the first entry of Ψ .
3. Evaluate $f(c(\psi))$ at the midpoint of ψ , $c(\psi)$:

$$c(\psi) = \left[\frac{1}{2}(\underline{x}^1 + \bar{x}^1), \frac{1}{2}(\underline{y}^1 + \bar{y}^1) \right] \times \left[\frac{1}{2}(\underline{x}^2 + \bar{x}^2), \frac{1}{2}(\underline{y}^2 + \bar{y}^2) \right], \quad (22)$$

and take $f(c(\psi))$ as a possible solution ($V(\psi)$).

4. Calculate an upper bound ($UB(\psi)$) for ψ (The derivation of an appropriate upper bound is detailed below).
 - (a) If $UB(\psi) \geq V_{max}$, split ψ into 8 ($4K$) equally sized smaller feasible areas. Add the smaller areas to Ψ and delete ψ from Ψ .
 - (b) If $UB(\psi) < V_{max}$, delete ψ from Ψ .
5. Termination rule: if $iter < \bar{iter}$, return to Step 2 and set $iter := iter + 1$. If $iter = \bar{iter}$, take V_{max} and the corresponding $c(\psi)$ as the optimal solution. \bar{iter} denotes the arbitrarily given maximum number of allowed iterations.

Deriving upper bounds using convex approximation For each $\psi \in \Psi$, a suitable upper bound, $\bar{f}(\mathbf{k})$, of the central planner's objective function, $f(\mathbf{k})$, must always be greater than $f(\mathbf{k})$ for a given vector of facility locations \mathbf{k} and must be convex. Formally, $\bar{f}(\mathbf{k})$ must meet the following properties:

$$\bar{f}(\mathbf{k}) \geq f(\mathbf{k}), \quad \text{and} \quad \bar{f}''(\mathbf{k}) \geq 0. \quad (23)$$

For a given number of facility locations K and without loss of generality, I can reformulate the central planner's maximization problem (20) using the zero profit condition as:

$$\begin{aligned} \max_{k_1, \dots, k_K} f(\mathbf{k}) &= \Gamma \left(\frac{\epsilon - 1}{\epsilon} \right)^\epsilon \sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon \\ &= \sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K \mathbb{W} \times d_{rt}^{-\epsilon}, \end{aligned} \quad (24)$$

where:

$$\mathbb{W} = \Gamma \left(\frac{\epsilon - 1}{\epsilon} \right)^\epsilon (d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r \alpha (1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_s^{\frac{1}{\alpha}} Q_s^{\frac{\alpha-1}{\alpha}})^\epsilon. \quad (25)$$

Next, I define $\bar{\mathbb{W}}$ such that $\bar{\mathbb{W}} = \mathbb{W}(\min\{\mathbf{Q}\})$, where $\min\{\mathbf{Q}\}$ is the observed minimum of \mathbf{Q} . Hence, $\bar{\mathbb{W}}$ is independent of any facility location strategy. Intuitively, this is equivalent to shutting down the equilibrium effects through housing prices. Let me then define:

$$\bar{f}(\mathbf{c}) = \sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K \bar{\mathbb{W}} \times d_{rt}^{-\epsilon}. \quad (26)$$

\bar{f} is a convex function and, under Assumption 1, $\bar{f}(\mathbf{k}) \geq f(\mathbf{k}), \forall \mathbf{k}$.

Assumption 1 *Modeled housing prices are greater than or equal to the minimum observed housing price for all places. Formally, $Q_i \geq \min\{\mathbf{Q}\}, \forall i$.*

Proposition 2 follows knowing that to find the maximum of a convex function over a polyhedron, it is sufficient to investigate the extreme points of the polyhedron.

Proposition 2 *Let $\psi \subset \mathbb{R}^{2K}$ be a polyhedron and let V_1, \dots, V_{4K} , denote the $2^{2K} = 4^K$ vertices of ψ . Then, given Assumption 1,*

$$UB(\psi) = \max\left\{\bar{f}(V_1), \dots, \bar{f}(V_{4K})\right\} \quad (27)$$

is an upper bound of the central planner's problem (20).

PROOF: See the proof of Proposition 2 in Appendix F.

Q.E.D.

Simulated annealing The convergence speed of the BCSC relies heavily on the quality of the upper bound. After eliminating potentially large cubes in the first iterations, the BCSC approach progresses in relatively small steps due to the exponentially increasing number of new cubes that need to be evaluated. In practice, it is then useful to apply a heuristic approach once the obtained solution is close to the global optimum. To that end, I employ a standard simulated annealing technique. This heuristic method is widely used to search for global optima in high-dimensional problems (Fajgelbaum and Schaal, 2017).²⁴

²⁴The steps of the method are detailed in Appendix B.

5.3 Solving the distance-minimizing problem

Solving the pure distance-minimizing problem, i.e., the central planner’s problem without considering the reaction of individuals and firms to facility locations, is useful for at least two reasons. First and foremost, as seen in Section 3, minimizing the distance between individuals and public facilities is on clearly-stated goal of locating a new facility. Therefore, comparing the difference between the distance-minimizing strategy and the observed location strategy to the difference between the optimal strategy to the observed strategy allows me to relate the general equilibrium gains under the optimal strategy to the more standard gains in commuting. Second, even though in practice, a central planner is rarely facing a problem in which a large number of new facilities have to be located,²⁵ such an eventuality is possible and deserving of attention. Solving such high-dimensional problems with the Big Cube - Small Cube approach is impracticable. Hence, in the case of a high number of new facility locations, I substitute the BCSC with a gradient-based approach. Specifically, I solve the distance-minimizing problem using a version of the Weiszfeld algorithm following [Weiszfeld \(1937\)](#), modified to fit the multi-facility problem.²⁶

6 Data description

This section describes the data used in the application, motivates the use of French secondary schools in the application over other public services, and provides a short overview of the French secondary education system.

Data sources The core data employed in the application combines three main datasets: geo-localized individual data, geo-localized public service data, and commuting data (from 2010). First, information about the location and capacity of the French secondary school system is provided by the French Ministry of Education. This data provides the exact coordinates of each establishment, as well as the number of students registered in each school. Second, geo-localized individual data is given by the 2010-gridded population data, which registers the number of persons residing in $1\text{km} \times 1\text{km}$ squares covering the whole country. This data also includes the population by age group

²⁵In this paper, I arbitrarily define “large” as more than 10 facilities. Note that, in our application, a maximum of 7 new facilities have to be located in a given year within the Paris region.

²⁶See Appendix G for a detailed presentation of the method.

in each grid. Third, data about the average commuting time to school is provided by the National Institute for Statistics and Economic Studies (INSEE).²⁷

From any given residential location, the probability of commutes to school, π_{ik} , and to workplace, π_{ij} , are independently recorded at the municipal level. To retrieve the probability to commute from residential location i to workplace j and facility k , π_{ijk} , I assume that conditional on a residential location, the choice of the public facility and the workplace are independent. This allows the derivation of π_{ijk} as follows:

$$\pi_{ijk} = \pi_{ijk|i} \times \pi_i = \pi_{ij|i} \times \pi_{ik|i} \times \pi_i = \frac{\pi_{ij} \times \pi_{ik}}{\pi_i}. \quad (28)$$

I create the units of analysis by designing a grid of $1\text{km} \times 1\text{km}$ cells over the total area of $50\text{km} \times 50\text{km}$ with Paris' Hotel de Ville as the centroid.²⁸ Out of the total 2'500 cells, 2'331 cells of this grid contain built-up areas. As all data described above are geo-localized, it is then straightforward to match the data onto the chosen grid. Lastly, I compute Euclidean distances between each cell.

French secondary school system To empirically analyze the location of public services, it is useful to focus on each public service separately, apart from all others. Considering the location of all public services jointly, without drawing distinctions between them might lead to bad spatial coverage for each particular service. In the application, I focus on the location of "Collèges", i.e., French lower secondary education facilities. In France, secondary education is organized in two stages: the lower secondary education level called "Collège" for pupils aged between 11 and 15, and the higher secondary education level called "Lycée" for pupils aged between 15 and 18.²⁹

Compared to other public services and other national settings, studying the location of French secondary schools offers several unique advantages:

1. *Minimal difference across facilities:* Differences across secondary schools are min-

²⁷Additionally, I obtain wage data and other demographic information at the municipal level from the INSEE. Workplace wages are computed based on workplace and individual characteristics. The website "Meilleursagents.com" provides municipal housing price data. These data are based on the notary database "BIEN" and their own transaction records. Land use data is provided by the regional urban planning institute.

²⁸I chose a square over the more traditional circle to facilitate the definition of the bounds enclosing the candidate locations in the simulated annealing approach. The bounds can be more easily defined by a square in terms of latitude and longitude.

²⁹For simplicity, I will refer to the French "Collège" as secondary school from this point forward.

imal. The same national degree is offered to pupils, teachers earn the same salary, teachers must pass the same qualifying exams, and average spending per pupil by the State is equal across facilities.

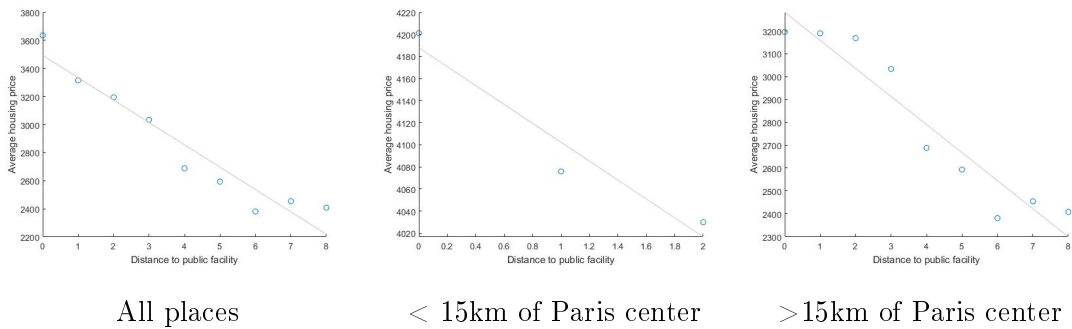
2. *No spatial differences in financing:* Due to the centralized financing of the secondary school system, there are (almost) no spatial differences in the amount households pay for the schooling of one student. Hence, individuals do not locate strategically to pay more or less for education.
3. *Perfect identification of users:* As education is compulsory in France until the age of 16, I am able to easily and precisely identify the users of the public service. According to UNESCO statistics, the enrollment rate in lower secondary education in France was 99,79% in 2010. I, therefore, consider that all individuals between the ages 11 and 15 are users of secondary schools. Also, as it is compulsory, the central planner must provide a seat to every pupil aged between 11 and 15.
4. *Commuting data along two dimensions is available:* The full matrix containing the universe of commutes from residential locations to public schools and from residential locations to workplaces is available in 2010. Hence, the commuting elasticities can be precisely estimated in a manner that is consistent with the model.

However, the use of lower secondary schools raises several concerns which I will try to answer in the remainder of this section. First, when looking at public education, one important consideration to keep in mind is that the private sector offers a competing service. In France, 17% of all pupils attend a private establishment. For simplicity and clarity of the approach, our model assumes that no private education is offered and that only pupils going to a public establishment need to be offered a seat. This assumption should not harm the generality of the framework for two reasons. First, the strategy of private schools in France is to locate very close to public schools and to draw away and enroll the good students from the respective public schools. Hence, the key moment targeted by the framework – the spatial coverage – should not be significantly impacted by the existence of private schools. Second, 97% of all private schools are religious schools while public schools obey strict secular rules. Hence, the switch from public to private (or vice-versa) might not be as straightforward as in other countries.

Second, even though the school financing and the educational attainment of teachers is the same across all lower secondary schools, the assumption that schools are homogeneous (i.e. that quality is the same across all schools) may be difficult to verify empirically. Differences in peer effects, for example, could be observed across schools. However, this model includes difference in wages and, hence, allows for spatial income sorting. School heterogeneity would be problematic if it is not correlated with income. Non-income-correlated heterogeneity is unlikely given the characteristics of the French secondary education system discussed above. One possible alternative public service with similarly low heterogeneity would be primary schools. However, as such facilities are significantly cheaper to build, maintain, and operate, many municipalities operate their own. Therefore, the mechanism targeted by our model would not be best illustrated using primary schools.

Third, the literature on the capitalization of public goods into housing prices highlights the importance of school quality to determine the magnitude of the effect, (see, among others, [Gibbons and Machin, 2003](#); [Fack and Grenet, 2010](#)). For tractability, I posit that distance matters for all schools; therefore, I abstract from that literature by ignoring the heterogeneity with respect to quality. This abstraction appears relatively minor when looking at the distribution of housing prices as a function of distance to secondary schools by kilometers (see Figure 1). Overall, a strong negative correlation is observed, both at the center and in the periphery of Paris. Hence, distance appears to matter, at least on average. Naturally, the correlations in Figure 1 might be due to other channels, such as firms locations. Evaluating the location of public facilities, accounting for all these channels, is the precise goal of this paper.

Figure 1: Housing prices and distance to the facility



Secondary schools in Paris' region between 2001 and 2015 I focus on public secondary school openings and closures between 2001 and 2015³⁰ in the Paris' region. Table 1 summarizes the number of openings and closures by year. There are between 0 and 7 school openings and between 0 and 2 closures per year. The overall stock increased from 660 in 2000 to 698 in 2015 in the Paris region.

Table 1: Openings and closures of facilities in Paris region (2001-2015, 50km square)

Year	Stock	Open	Close	Year	Stock	Open	Close	Year	Stock	Open	Close
2001	660	1	0	2006	679	5	2	2011	688	0	2*
2002	665	5	0	2007	682	3	0	2012	689	1	0
2003	670	5	0	2008	686	5	1*	2013	689	0	0
2004	671	1	0	2009	687	1	0	2014	696	7	0
2005	676	5	0	2010	690	3	0	2015	698	2	0

*: Facility closures for renovation or asbestos removal.

Moreover, a first glance at the observed distribution of secondary schools provides clear and interesting patterns. Figure 2 illustrates the average school capacity as a function of distance to the city center in 2010. I observe that the average capacity increases from 475 to 650 pupils with increasing distance to the center (with the variance in capacity also increasing). Given the high population density in the center, one might have expected a flatter or even negative relationship. Figure 3 displays the average number of schools as a function of distance to the city center. A bell shaped curve is observed with a maximum around 10km.

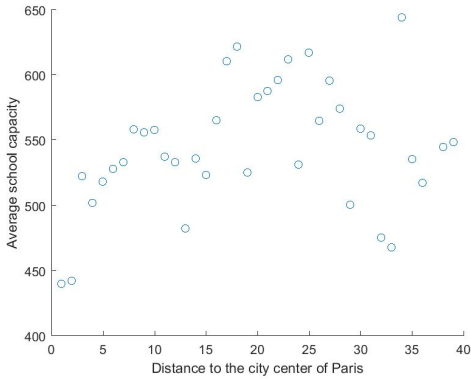


Figure 2: Secondary school seats

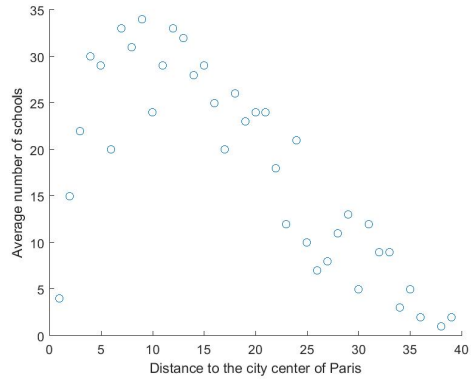


Figure 3: Number of schools

³⁰The choice of the period is driven by data availability.

7 Structural Analysis

In this section, I estimate the commuting parameters and calibrate the model for residential amenities, final good productivity, and density of development.

7.1 Estimation of commuting parameters

From the commuting shares in (6), the model predicts a semi-log gravity equation for commuting flows between the place of residence i , the workplace j , and the public facility location k :

$$\ln \pi_{ijk} = -\nu \rho_{ik} - \xi \tau_{ij} + \omega_i + \zeta_j + v_k + \epsilon_{ijk}, \quad (29)$$

where ω_i are residence-fixed effects capturing residence characteristics, ζ_j refers to workplace characteristics, and v_k are public-facility-fixed effects. The denominator of (6) is a constant. The parameters ν and η are the semi-elasticities of commuting flows with respect to the Euclidean distance. They are defined as $\nu = \kappa\epsilon$ and $\xi = \eta\epsilon$, where κ is the commuting cost parameter for journeys to the public facilities, η is the commuting cost parameter for journeys to the workplace, and ϵ is the heterogeneity parameter from the Fréchet-distributed shock on individuals' utility.

To empirically estimate the semi-elasticities of commuting flows, I use municipal data from the 232 municipalities and the 20 Parisian districts (covered by the 50km \times 50km grid). In total, 773'606 tri-dimensional links are observed within the grid. As bilateral distances are measured at the cell level, I compute the Euclidean distance between two municipalities by taking the average distance between all cells in a municipality to all cells in another.

Table 2 displays the estimation of (29) using these data. In Column (I), I estimate a linear fixed effects model. I obtain a semi-elasticity of commuting to school of -0.14 and a semi-elasticity of commuting to work of -0.11, both significant at the 1-percent level. Columns (II) and (III) report the estimation results of a Pseudo Poisson Maximum Likelihood (PPML) model and of a Negative Binomial (NB) model, respectively. I obtain a semi-elasticity of commuting to school of -0.88 (-0.88) for the PPML (NB) and of -0.21 (-0.21) for the PPML (NB) for commutes to work.³¹ I use these estimates in the

³¹Note that the estimated semi-elasticity of commutes to work is larger than what Ahlfeldt et al. (2015) obtain for Berlin (i.e. -0.07). Two factors help explain this difference. First, the higher density of Paris (about 5 times higher) can lead to greater difficulty in moving from the residency location to the

application. The semi-elasticity of commute to school is significantly larger than the semi-elasticity of commuting to work. This reflects the fact that parents generally prefer that their kids not commute long distances to their school, as pupils (aged 11 to 15) might commute on their own.

Table 2: Gravity estimations of commutes

	(I)	(II)	(III)
Great circle distance ($-\kappa\epsilon$) to school	-0.14*** (0.002)	-0.88*** (0.001)	-0.88*** (0.001)
Great circle distance ($-\eta\epsilon$) to work	-0.11*** (0.002)	-0.21*** (0.001)	-0.21*** (0.001)
Estimation	OLS	PPML	NB
Fixed effects	Yes	Yes	Yes
Observations	773,606	773,606	773,606
R^2	0.62	-	-

Notes: Estimates based on bilateral commuting probability in 2010. PPML stands for Pseudo Poisson Maximum Likelihood, and NB for Negative Binomial. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

7.2 Calibration of location characteristics

Using the structure of the model, I show that there is a one-to-one mapping between the observed and the unobserved location characteristics. I can then derive the unobserved location characteristics, namely, residential amenities, final goods productivity, and the density of development.

Proposition 3 *Given known values for the parameters $\{\alpha, \beta, \kappa, \mu, \eta\}$ and the observed data $\{Q, H_M, H_R, K, \rho_{ij}, \rho_{ik}\}$, there exist unique vectors of the unobserved location characteristics $\{B, A, \phi\}$ that close the model for a given location strategy.*

PROOF: See the proofs of Proposition 3 in Appendix H.

Q.E.D.

workplace. Second, the zone considered in this paper includes a number of rural areas where individuals often work in the municipality in which they reside.

Residential amenities B_i as a function of distance to the city center are presented in Figure 4. Similarly, productivity A_j and density of development ϕ_i as functions of distance to the center are presented in Figures 5 and 6, respectively. Figure 4 illustrates that residential amenities are relatively constant across all locations. The variance of residential amenities increases with distance from the center as some locations may compensate a larger distance to the center with higher local amenities. Interpreting the density of development is less straightforward – available floor space is a function of available land space which decreases with distance to the city center.

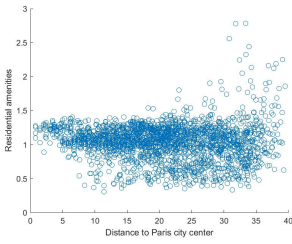


Figure 4: Residential amenities (B_i)

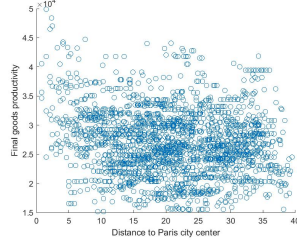


Figure 5: Final goods productivity (A_j)

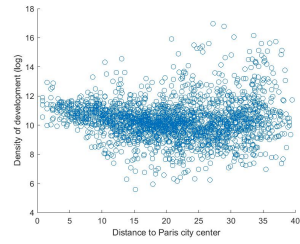


Figure 6: Density of development (ϕ_i)

7.3 Calibration of remaining parameters

Apart from the semi-elasticities of commuting flows estimated above, I choose the Fréchet shape parameter ϵ to minimize the squared difference between the variances across places of log-adjusted wages from the model and the data following [Ahlfeldt et al. \(2015\)](#). This leads to $\epsilon = 6.83$. Using this value of the shape parameter ϵ and the estimated commuting probability parameters ν and η , I can retrieve the commuting cost parameters $\kappa = \frac{0.88}{6.83} = 0.13$ and $\eta = \frac{0.21}{6.83} = 0.03$. I set the values of the remaining parameters in the model using standard sources in the literature. Table 3 presents the values and sources of these parameters.

Table 3: Calibration overview

PARAMETERS COMMON TO ALL LOCATIONS		
1. Preferences		
$\beta = 0.75$	Consumption share in utility	Davis and Ortalo-Magne (2011)
$\epsilon = 6.83$	Shape parameter	Own estimation
2. Commuting		
$\kappa = 0.13$	Elasticity of commuting to public facility	Own estimation
$\nu = 0.03$	Elasticity of commuting to workplace	Own estimation
3. Technology		
$\alpha = 0.8$	Labor share in production	Valentinyi and Herrendorf (2008)
4. Land market		
$\mu = 0.8$	Capital share in construction	Combes, Duranton, and Gobillon (2016)
LOCATION-SPECIFIC VARIABLES		
1. Land Endowments		
K_i	Residential land mass in location i	Own derivation (using ArcGIS)
ϕ_i	Density of development in i	Own estimation
2. Amenities		
B_i	Residential amenities	Own estimation
3. Productivity		
A_i	Final good productivity	Own estimation

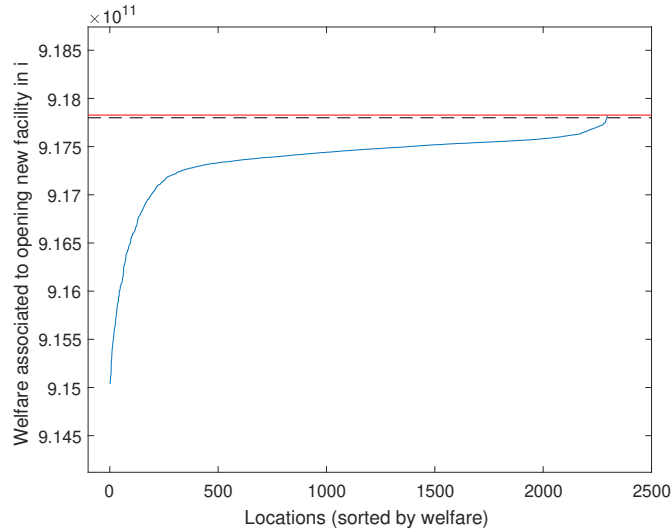
7.4 Performance of the framework

To assess the performance of the proposed framework, I exploit the special case of opening *one additional facility*. This particular scenario offers the advantage that locating a single facility is a “simpler” problem than locating several at the same time, as there are no interaction effects across new facilities. Taking the centroid of the 2’331 cells with built-up area as location candidates, it is feasible to use a “pseudo”-exhaustive search approach to determine the welfare-maximizing location of one new facility. There are only 2’331 possibilities of where to locate one facility, while locating two (three) facilities already requires searching through $2’331^2 = 5’433’561$ ($2’331^3 = 1.267 \times 10^{10}$) possible location combinations. The outcome of this exhaustive search can serve as a benchmark to evaluate the performance of the framework.

Figure 7 displays the outcome of the exhaustive search for one new facility location. All candidate locations are ranked (horizontal axis) according to the aggregate welfare (vertical axis) generated by opening a facility in each location. Aggregate welfare on the vertical axis is measured in private consumption equivalent units. The blue curve displays

the (sorted) welfare level associated with building a new facility at the centroid of one of the 2'331 cells. The dashed, black line represents the aggregate welfare reached after the last iteration of the BCSC approach, while the red line indicates the final aggregate welfare derived by the framework. The framework reaches a welfare level slightly higher than the brute-search approach. This is to be expected as the brute-search ignores the space outside the centroids of the grid cells, where the optimal location is likely to lie.

Figure 7: Performance of the framework



8 Evaluating location strategies

Using the framework described above, I now aim to evaluate the location decisions of secondary schools in the Paris region in the period 2001-2015. In a first step, I measure the difference between the optimal and the observed location strategies for each year between 2001 and 2015, given the existing location strategy of each year (Section 8.1). In a second step, I identify key drivers of this difference using the structure of the model.

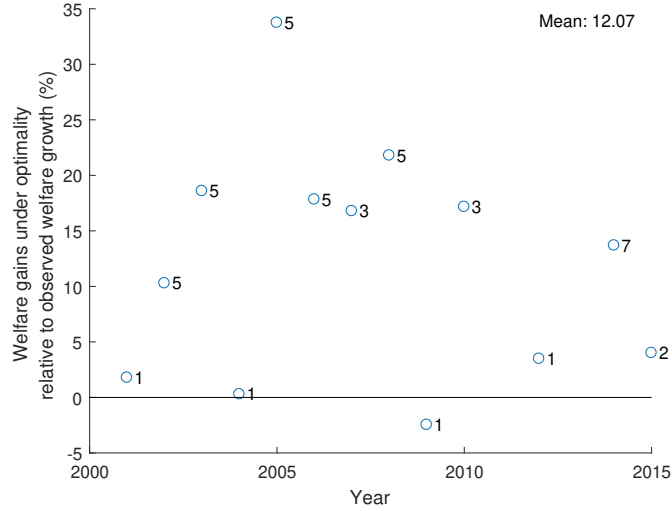
8.1 Measuring the difference across strategies

Given the already existing secondary schools, I quantitatively compare the optimal and observed location strategy in each year between 2001 and 2015. Importantly, the analysis is conducted separately for each year to avoid carrying early differences throughout the sample period. Hence, the optimal strategy in a given year is always computed based on the previous year's observed strategy.

Figure 8 reports the differences between the strategies. The horizontal axis displays the years, while the vertical axis reports the welfare gains (in percent) under the optimal strategy relative to the observed welfare growth in that year. Precisely, this welfare gains are computed as: $100 \times \frac{W_t^{opt} - W_t^{obs}}{W_t^{obs} - W_{t-1}^{obs}}$, where t is a year index. W_t^{opt} is the welfare level under the optimal strategy in year t . W_t^{obs} is the welfare level under the observed strategy in year t . Additionally, Figure 8 also reports the number of secondary schools that opened in that year next to each observation marker.³²

On average, welfare growth is 12.07% larger under the optimal strategy, ranging from -2.48% (in 2009) and 33.74% (in 2005).³³ For comparison, the average welfare growth is *only* 6% larger when locating public facilities to minimize the commutes to the facility. Furthermore, while the relative gains under the optimal strategy are relatively small when one or two facilities are built (on average, $\approx 1\%$), these relative gains are much larger when more than two facilities are built (on average, $\approx 18\%$).

Figure 8: Observed and optimal location strategies (2001-2015)



There are three main take-aways from this analysis. First, welfare could be improved

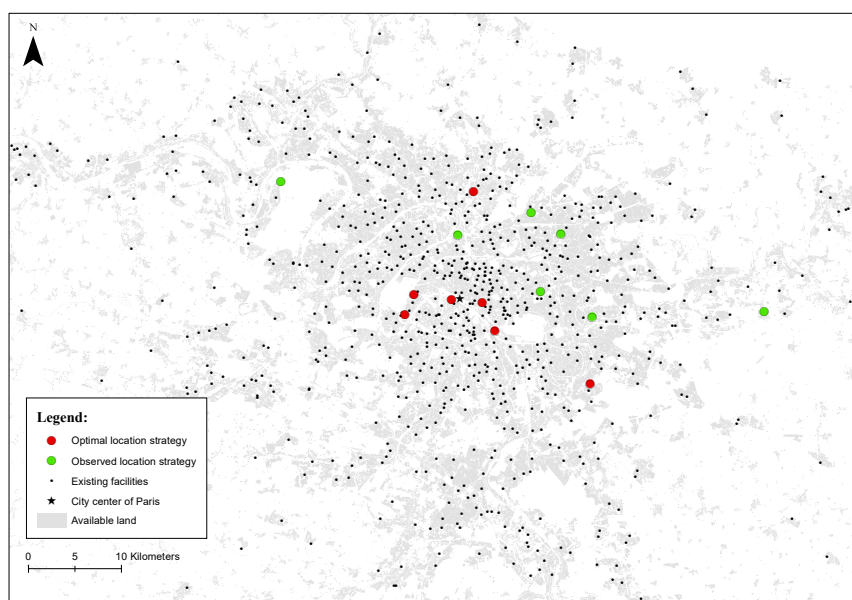
³²Table 6 in the Appendix displays the data behind Figure 8. Column (I) indicates the number of facilities that were built in each year. Columns (II) and (III) report each year's welfare level associated with the observed and the optimal strategies, respectively. Column (IV) relates gains in welfare under the optimal strategy to the gains from the observed strategy.

³³The proposed framework performed better than the observed strategy in all years except in 2009, when welfare growth would have been 2.48% lower under the optimal strategy. The framework of the model is not able to reach the optimal point in that year due to a strong gradient of the welfare function in its neighborhood. Note that a brute-search approach using the framework in this paper would easily detect such cases.

by adopting the optimal location strategy. Second, these gains mostly occur when more than two facilities are to be built. When one or two facilities are built, the observed strategy is almost as good (and in one case better) than the proposed optimal strategy. Hence, the framework proposed appears to be a good predictor of the observed world when only few facilities are built. Third, half of the welfare gains can be attributed to a reduction in commuting to the facility, the rest is derived from general equilibrium effects on individuals. This finding strongly highlights the importance of studying public facility location in a general equilibrium framework.

To illustrate these findings in a more intuitive way, I map the observed and optimal location strategies in 2014 (Figure 9).³⁴ The observed location strategy (green circles) appears largely different from the optimal location strategy (red circles). While the optimal strategy locates five facilities near the center of Paris where population density is high and two in the periphery (at about 15-20km), the observed location strategy saw facilities built in the North-East of Paris, five are within 10 to 20km and two are more than 25km from the center.

Figure 9: Map of the observed and optimal location strategies (2014)



³⁴2014 saw the highest number of secondary school openings than any other year in the sample period (see Table 6).

8.2 Cell-level analysis of the differences across strategies

In this section, I aim to explain the differences between the optimal and observed strategies revealed in the previous subsection using the structure of the model. To start, I take the difference between the *observed* and *optimal* strategies of residential density ($\overline{H_{Ri}}$), workplace density ($\overline{H_{Mj}}$), commutes to public facility ($\overline{\sum_k H_{ik}d_{ik}}$), commutes to workplace ($\overline{\sum_j H_{ij}d_{ij}}$), housing prices ($\overline{Q_i}$), and wages at the residence location ($\overline{w_i}$) at the 1km \times 1km cell level.

$$\begin{aligned}\overline{H_{Ri}} &= H_{Ri}^{\text{Obs}} - H_{Ri}^{\text{Opt}}, & \overline{\sum_k H_{ik}d_{ik}} &= \left(\sum_k H_{ik}d_{ik}\right)^{\text{Obs}} - \left(\sum_k H_{ik}d_{ik}\right)^{\text{Opt}}, \\ \overline{H_{Mj}} &= H_{Mj}^{\text{Obs}} - H_{Mj}^{\text{Opt}}, & \overline{\sum_j H_{ij}d_{ij}} &= \left(\sum_j H_{ij}d_{ij}\right)^{\text{Obs}} - \left(\sum_j H_{ij}d_{ij}\right)^{\text{Opt}}, \\ \overline{Q_i} &= (H_{Ri}Q_i)^{\text{Obs}} - (H_{Ri}Q_i)^{\text{Opt}}, & \overline{w_i} &= (H_{Ri}w_i)^{\text{Obs}} - (H_{Ri}w_i)^{\text{Opt}}\end{aligned}\quad (30)$$

A positive value of $\overline{H_{Ri}}$ implies that the cell is a more attractive residential location under the observed than under the optimal strategy. Conversely, a negative value implies that the cell is more attractive under the optimal than the observed strategy. The same reasoning applies to the attractiveness of commercial locations measured ($\overline{H_{Mj}}$), housing prices ($\overline{Q_i}$), and wages at the residence location ($\overline{w_i}$). Similarly, a positive value of $\overline{\sum_k H_{ik}d_{ik}}$ implies that for a given location the sum of commutes to public facilities is larger under the observed than under the optimal strategy. As they describe the distance to optimality, these simple differences can be interpreted as a measure of inefficiency.

Table 4 reports simple descriptive statistics based on cell-level differences. On average, opting for the optimal strategy leads to large improvements in commutes (to work and public facilities) and housing prices. In contrast, residential and commercial densities are slightly higher under the optimal strategy. Similarly, wages decrease on average under the optimal strategy; however, this decrease is very small compared to the decrease in housing prices.

However, the gains and losses vary significantly by cell as shown by the range and the standard deviation of the differences. Hence, it appears useful to further investigate how these gains and losses are distributed geographically. Analyzing the distribution of gains and losses also enables me to identify zones where inefficiencies are large. More than a general measure of distance to optimality, policy makers would benefit from an identification of such zones, as well as the identification of the present types of inefficiencies.

Table 4: Descriptive statistics of the difference between the strategies (2001-2015)

	Residential density	Distance to facility	Workplace density	Distance to workplace	Housing Prices	Wage at residence
Average	-0.03	33.15	-0.38	33.02	39.43	0.26
Min	-42.32	-439.68	-194.33	-378.60	-137.07	-195.71
Max	84.10	110.23	37.77	90.93	145.69	39.00
St. dev.	4.73	47.11	4.72	48.13	17.18	4.94

Questions such as: Does the current location strategy favor central or peripheral areas? Does the strategy shape land usage in an inefficient manner? If so, where does it do so? can only be answered by looking at the geographic heterogeneity in the differences across the observed and the optimal strategy.

Inefficiencies and distance to the city center Analyzing the heterogeneity of the inefficiencies as a function of distance to the city center reveals interesting patterns, as shown in Figure 10. Consider a set of binary variables equal to unity for places within 2km-rings around the Paris city center (13 binary indicators in total). Figure 10 plots the point estimates (and the 95% confidence intervals) resulting from regressing the differences in residential density and sum of commutes to facility on these binary variables. The excluded category corresponds to all places further away than 24km from the Paris city center (10'712 out of 29'861 places).

Figure 10a plots the parameter estimates for residential density. Places at the outer periphery (16 to 24km) appear to be the most disadvantaged.³⁵ This is true compared to central places, but also compared to more distant areas (or simply put, compared to rural areas). The observed strategy increases residential density in central and rural places at the expense of intermediary locations. The picture is reversed when looking at workplace density (Figure 10b). Here the observed strategy favors intermediary locations, over rural areas and over more central places. Hence, combining these two pictures, the optimal strategy argues in favor of a more mixed land use – increasing workplace density in central and rural locations and residential density at intermediary locations.

Figure 10 also reveals that individuals in central places (within 15km of the Paris city center) commute much shorter distances, both to the workplace and the public facility. Figures 10c and 10d suggest that individuals within 5km of the city center

³⁵Due to the high concentration of administrative buildings and museums in the first ring, the negative estimate obtained for the first rings (0-4km) should be interpreted with caution.

commute in total about 200km less³⁶ than individuals located at more than 14km from the center under the observed strategy compared to the optimal one. Less striking but potentially more surprising is the fact that (almost) no significant difference in commuting is observed between the two strategies after 14km. Hence, potential welfare gains in these locations are not obtained through shorter commutes but through general equilibrium effects.

Finally, Figures 10e-10f highlight the difference between the two strategies in terms of housing prices and wages (at the residential location). Housing prices are higher under the observed strategy and decrease with distance to the city center. Wages are relatively higher under the optimal strategy in places within 24km of the city center compared to places further away. These findings are particularly interesting if placed in the context of the very high housing prices in a metropolitan area such as Paris. Better facility locations that account for such effects might help reduce housing prices and increase wages in central places.

Analyzing these findings in the context of the urban-rural divide provides an interesting policy context. In many countries, representatives of rural areas often point to the increasing centralization of the public services³⁷ as the source of rural decline, which the central government often counters by reversing the causality, i.e., centralization reforms do not cause rural decline, it is rural decline that renders centralization reforms necessary. The findings in this section quite literally defend the middle ground between these two positions. Representatives of rural areas are correct in arguing that central places are over-provided with public facilities (in this case, secondary schools). However, the central government is correct in defending that building facilities solely in rural areas does not maximize aggregate welfare. The findings show that priority should be given to a better allocation of activity in intermediary places which can be achieved by locating public facilities in a strategic manner.³⁸

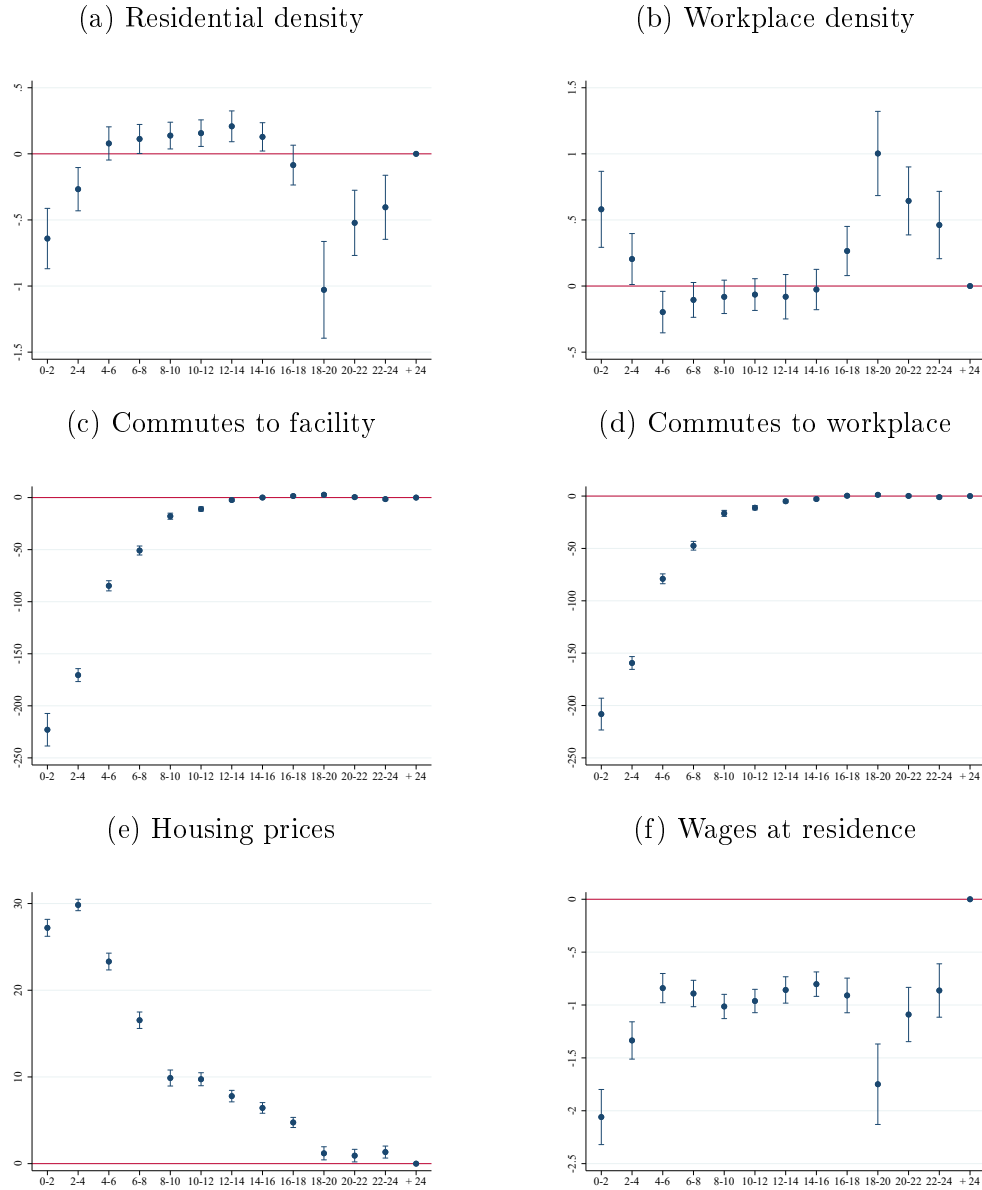
Inefficiencies and location fundamentals I now present a similar heterogeneity analysis, but focus on the heterogeneity of the inefficiencies with respect to the distribution of fundamentals. Figure 11 displays the expected value of the difference across

³⁶Remember that this is a measure combining distance and the number of commuters.

³⁷Local border reforms are a typical example of such centralization reforms as their aim is to regroup local services to enjoy economies of scale (see, among others Reingewertz, 2012; Tavares, 2018).

³⁸I explore the question of the central planner's future optimal strategy in Section 9.1.

Figure 10: Cell-level inefficiencies as a function of distance to city center



the strategies by decile of residential amenities. Difference in residential and workplace density by residential fundamental are reported in graphs 11a-b and by commercial fundamentals in graphs 11c-d.³⁹ The bottom decile is the omitted category in all graphs of Figure 11.

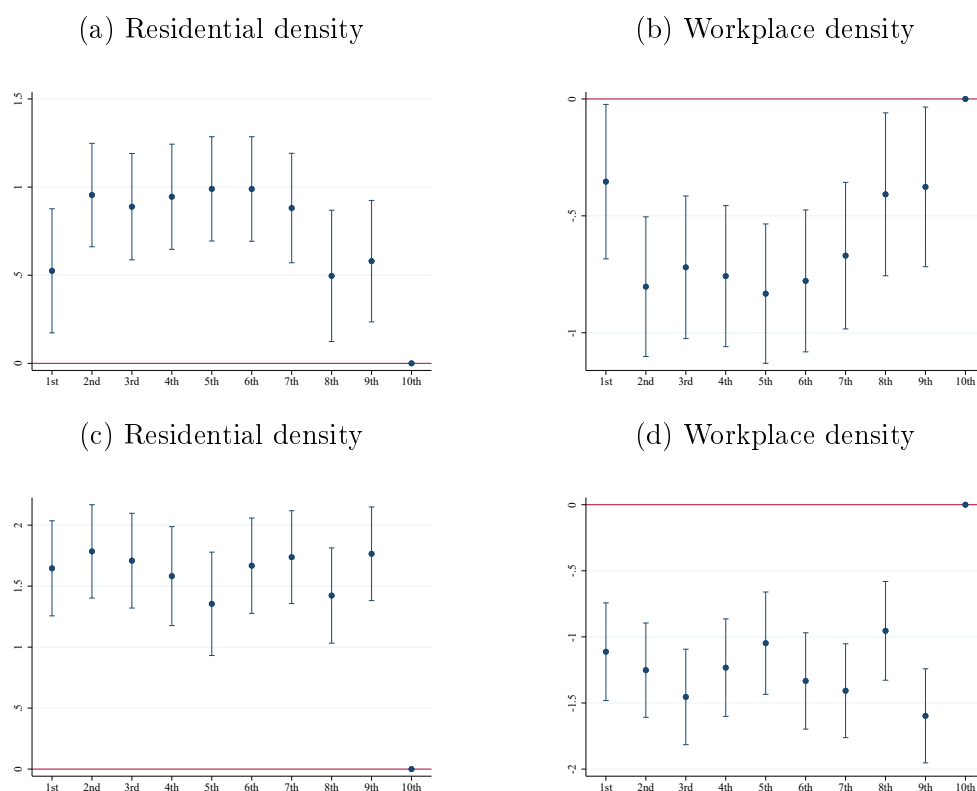
Figure 11a-11b show that places with a high level of residential amenities also have a higher residential density and a lower workplace density under the observed strategy compared to low amenity places. Hence, relative to low amenity areas, cells between

³⁹Note that the correlation coefficient between the two types of fundamentals is small (i.e. 0.098).

the 2nd and 7th decile are particularly over-specialized in residential use. Welfare maximization would encourage higher workplace density in such places. These findings reveal that planning authorities have been more successful (or simply more willing) to transform high amenity places into residential areas than lower amenity places. This relatively unsurprising result contrasts with the findings when considering the same question in terms of the distribution of final good productivity.

Compared to places with low productivity, residential density is higher and workplace density is lower in highly productive places under the observed strategy. One might have expected the opposite relationship. Hence, it appears that places with high productivity are also over-specialized in residential as opposed to commercial usage. This analysis reveals that welfare would be increased by public policies encouraging commercial activity in highly productive areas.

Figure 11: Cell-level inefficiencies as a function of residential amenity



9 Going further: Future location strategy and school districts under optimality

So far, the framework has been used to evaluate past location strategies. However, the framework also offers sufficient flexibility to shed light on other important aspects of facility locations. This section highlights two of these aspects. First, the calibrated model can be used to predict the optimal location strategy to adopt by 2025 (i.e. 10 years from the last observed year) given the number of new facilities to be located (Section 9.1). Second, to function efficiently, many public facilities are associated with a district (or catchment area). A district is an area around a facility k such that all users residing in this area will use the same facility k . Hence, in the case of secondary schools around Paris, I derive the optimal school districting policies, accounting for the effects on commutes to the workplace, housing prices, and wages, in addition to the effects on the commutes to the facility. This holistic accounting is rarely done when designing districts in the real world (Section 9.2).

9.1 Future location strategy

I use the previously-derived framework to determine the optimal location strategy that should be adopted in the future. Specifically, given the existing facilities in 2015 (i.e. facilities built by December 2015), I derive the optimal strategy that should be adopted, anticipating the 2025 population count.⁴⁰ I analyze the arbitrarily chosen scenarios under which 1, 10, 20, or 30 new secondary schools are opened.

Table 5 summarizes the results. Welfare is maximized when 30 new facilities are opened. Under that scenario, welfare would increase by about 4.5% between 2015 and 2025. Note that this increase has two components. First, the location strategy is chosen to maximize welfare (about 1%). Second, the population increase in these 10 years would also lead to an increase in the overall welfare level (about 3.5%).

Figure 12 illustrates the optimal location strategy when 30 new facilities are built between 2015-2025.⁴¹ *A priori*, no clear geographical pattern is observed. New facilities are located both close to the center and further away. However, using cell-level exogenous

⁴⁰The 2025 population count is predicted by linear extrapolation of the population growth between 2010 and 2015.

⁴¹30 new facilities is reasonable given the number of facilities that were opened in an average 10- year window between 2001 and 2015.

Table 5: Future location strategies

Current location strategy	Optimal strategy ($\times 10^{11}$)	Difference relative to total (%)	Optimal strategy ($\times 10^{11}$)	Difference relative to total (%)	Optimal strategy ($\times 10^{11}$)	Difference relative to total (%)	Optimal strategy ($\times 10^{11}$)	Difference relative to total (%)
K = 698	K + 1		K + 10		K + 20		K + 30	
9.36	9.69	3.48	9.72	3.83	9.75	4.12	9.78	4.51

characteristics, Section 8.2 showed that the optimal location strategy between 2001 and 2015 led to an improved distribution of commutes and activity across the metropolitan area.

Figure 12: Optimal future location strategy (2025, + 30)



9.2 Optimal school districts

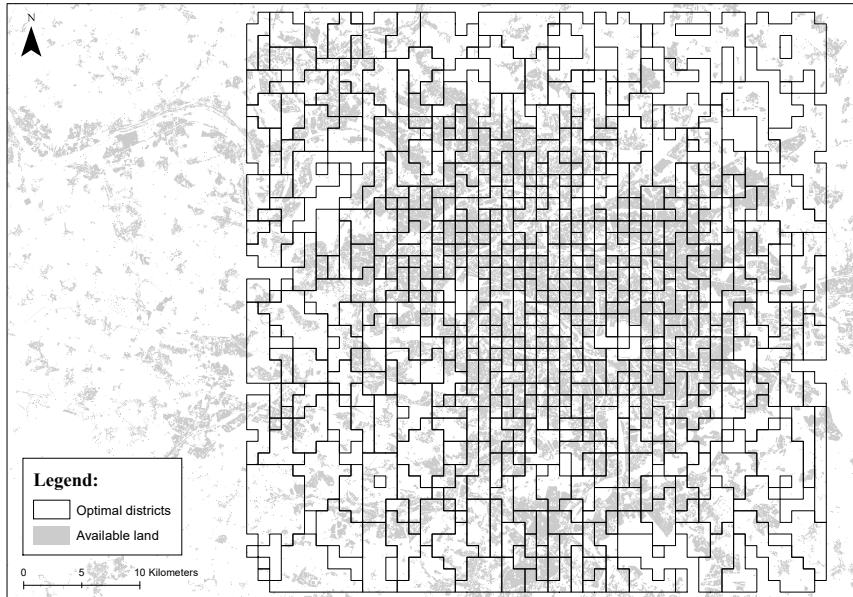
Until now, this paper has focused on locating facilities. To do so, an intermediary step was to probabilistically determine which individual uses which facility. This step already constitutes an interesting result. Under optimality, the assignment defines a district that maximizes welfare, accounting for all the presented general equilibrium effects.⁴² Hence, even without new facilities to locate, the framework can be used to inform planning

⁴²It is important to remember that social goals such as achieving equality of chances by reducing school differences in peer effects, which has motivated some districting policies in the real world, are not considered in this paper.

authorities (i.e. how to best design school districts).

Formally, as assignment of individuals to a given facility is done probabilistically, it follows that, for each residential location, there is one facility for which the assignment probability is highest.⁴³ Hence, the district of a given facility k can be defined as the set of residential locations $i \in S$ for which $\pi_{ik|i}$ is the highest. Formally, the district of facility k is the set: $\{i \in S \mid \pi_{ik|i} > \pi_{ik'|i}, \forall k' \neq k\}$. Figure 13 displays the district boundaries predicted by the model given the observed location strategy in 2010. As expected, districts are small in the city center and increase in size with distance to the Paris city center.

Figure 13: Optimal districts



10 Conclusion

Overall, this paper highlights the need to approach facility location strategies as active policies that will shape the economic geography of an area. Ignoring the endogenous reactions of individuals and firms by simply building where residential density has increased leads to significant inefficiencies, such as the mis-allocation of residential and commercial activity and/or inefficient commuting patterns.

In this paper, I develop a flexible structural and quantifiable framework to analyze

⁴³For simplicity, I ignore the case of where $\pi_{ik|i} = \pi_{ik'|i}, \forall k' \neq k$ as this is only theoretical and never occurs empirically.

the location of public facilities. To determine the optimal location strategy, the model combines a quantitative economic geography approach in a rich urban setting with an optimal location problem, in which the central planner anticipates how individuals and firms will react. Using stochastic shocks to commuting decisions that yield a gravity equation for commuter flows, I determine the global optimal location strategy using a novel hybrid optimization approach.

In the application, I match detailed economic and demographic variables to a grid of $1\text{km} \times 1\text{km}$ cells covering an area of $50\text{km} \times 50\text{km}$ around the Paris city center centroid. Using this grid in combination with exact location data on all public secondary schools active in the Paris region between 2001 and 2015, I calibrate the model and compare the observed location strategy to the optimal strategy predicted by the model. The welfare growth under the optimal location strategy is 12.07% larger than under the observed strategy between 2001 and 2015. The analysis also reveals that gains through general equilibrium channels such as lower housing prices, higher wages, or lower commutes to the workplace are as large as gains from reduced commutes to the facility. Finally, focusing on the heterogeneity of the difference between the two strategies, there are three main take-aways from the analysis. First, the observed strategy disproportionately favors lower commutes in central places. Second, the observed strategy made intermediary areas inefficiently less attractive for residential usage and inefficiently more attractive for commercial usage compared to more rural and more central places. Third, places with high amenity or high productivity fundamentals are over-specialized in residential usage under the observed strategy.

The proposed framework is sufficiently flexible such that it could also be used to analyze the location of other types of public services. Natural candidates for such further analyses are education facilities other than secondary schools, health care facilities, and police and fire stations. It follows that the framework could be used to inform planning authorities in deciding where to locate a wide range of public services.

References

- AFSA, C. (2014): “Une question de taille,” *Education & Formations*, 85.
- AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): “The Economics of Density: Evidence From the Berlin Wall,” *Econometrica*, 83, 2127–2189.
- ALBOUY, D., K. BEHRENS, F. ROBERT-NICOUD, AND N. SEEGER (2016): “The Optimal Distribution of Population across Cities,” NBER Working Papers 22823, National Bureau of Economic Research, Inc.
- ALLEN, T., C. ARKOLAKIS, AND X. LI (2015): “Optimal City Structure,” Mimeograph, Yale University.
- ALVAREZ-PEDREROL, RIVAS, LÓPEZ-VICENTE, SUADES-GONZÁLEZ, DONAIRE-GONZALEZ, CIRACH, DE CASTRO, ESNAOLA, BASAGANA, DADVAND, NIEUWENHUIJSEN, AND SUNYER (2017): “Impact of commuting exposure to traffic-related air pollution on cognitive development in children walking to school,” *Environmental Pollution*, 231, 837–844.
- BAUM-SNOW, N. (2007): “Did Highways Cause Suburbanization?” *The Quarterly Journal of Economics*, 122, 775–805.
- BELLETTINI, G. AND H. KEMPF (2013): “Why not in your backyard? On the location and size of a public facility,” *Regional Science and Urban Economics*, 43, 22–30.
- BÉNABOU, R. (1993): “Workings of a City: Location, Education, and Production,” *The Quarterly Journal of Economics*, 108, 619–652.
- (1996): “Equity and Efficiency in Human Capital Investment: The Local Connection,” *The Review of Economic Studies*, 63, 237–264.
- BERLIANT, M., S.-K. PENG, AND P. WANG (2006): “Welfare analysis of the number and locations of local public facilities,” *Regional Science and Urban Economics*, 36, 207–226.
- 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.
- BUCHMUELLER, T., M. JACOBSON, AND C. WOLD (2006): “How far to the hospital?: The effect of hospital closures on access to care,” *Journal of Health Economics*, 25, 740–761.
- CARD, D. (1993): “Using Geographic Variation in College Proximity to Estimate the Return to Schooling,” NBER Working Papers 4483, National Bureau of Economic Research, Inc.
- CELLINI, S. R., F. FERREIRA, AND J. ROTHSTEIN (2010): “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design,” *The Quarterly Journal of Economics*, 125, 215–261.

- CHRISTALLER, W. (1933): *Die zentralen Orte in Süddeutschland: Eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*, University Microfilms.
- COMBES, P.-P., G. DURANTON, AND L. GOBILLON (2016): “The Production Function for Housing: Evidence from France,” PSE Working Papers halshs-01400852, HAL.
- DAVIS, M. A. AND F. ORTALO-MAGNE (2011): “Household Expenditures, Wages, Rents,” *Review of Economic Dynamics*, 14, 248–261.
- DONALDSON, D. (2018): “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” *American Economic Review*, 108, 899–934.
- EATON, J. AND S. KORTUM (2002): “Technology, Geography, and Trade,” *Econometrica*, 70, 1741–1779.
- EPPLE, D., B. GORDON, AND H. SIEG (2010): “A New Approach to Estimating the Production Function for Housing,” *American Economic Review*, 100, 905–24.
- EPPLE, D., R. ROMANO, S. SARPÇA, AND H. SIEG (2017): “A general equilibrium analysis of state and private colleges and access to higher education in the U.S.” *Journal of Public Economics*, 155, 164 – 178.
- EPPLE, D., R. ROMANO, AND H. SIEG (2006): “Admission, Tuition, and Financial Aid Policies in the Market for Higher Education,” *Econometrica*, 74, 885–928.
- EPPLE, D. AND R. E. ROMANO (1998): “Competition between Private and Public Schools, Vouchers, and Peer-Group Effects,” *American Economic Review*, 88, 33–62.
- FACK, G. AND J. GRENET (2010): “When do better schools raise housing prices? Evidence from Paris public and private schools,” *Journal of Public Economics*, 94, 59–77.
- FAJGELBAUM, P. D. AND E. SCHAAL (2017): “Optimal Transport Networks in Spatial Equilibrium,” Working Paper 23200, National Bureau of Economic Research.
- FARAHANI, R. AND M. HEKMATFAR (2009): *Facility Location: Concepts, Models, Algorithms and Case Studies*, Contributions to Management Science, Physica-Verlag HD.
- FRENETTE, M. (2006): “Too Far to Go On? Distance to School and University Participation,” *Education Economics*, 14, 31–58.
- FUJITA, M. (1986): “Optimal location of public facilities: Area dominance approach,” *Regional Science and Urban Economics*, 16, 241 – 268.

- FUJITA, M., P. KRUGMAN, AND A. VENABLES (2001): *The Spatial Economy: Cities, Regions, and International Trade*, vol. 1, The MIT Press, 1 ed.
- GAIGNÉ, C., H. R. KOSTER, F. MOIZEAU, AND J.-F. THISSE (2017): “Amenities and the Social Structure of Cities,” CEPR Discussion Papers 11958, C.E.P.R. Discussion Papers.
- GIBBONS, S. AND S. MACHIN (2003): “Valuing English primary schools,” *Journal of Urban Economics*, 53, 197 – 219.
- HAMILTON, B. W. (1976): “Capitalization of Intra-jurisdictional Differences in Local Tax Prices,” *American Economic Review*, 66, 743–53.
- HANSEN, P., D. PEETERS, D. RICHARD, AND J.-F. THISSE (1985): “The Minisum and Minimax Location Problems Revisited,” *Operations Research*, 33, 1251–1265.
- KRUGMAN, P. (1991): “Increasing Returns and Economic Geography,” *Journal of Political Economy*, 99, 483–499.
- KUNEC, M. (2005): *Planning for Schools and Liveable Conditions: The Oregon School Siting Handbook*, Oregon Transportation and Growth Management Program.
- LIU, W.-T., C.-M. MA, I.-J. LIU, B.-C. HAN, H.-C. CHUANG, AND K.-J. CHUANG (2015): “Effects of commuting mode on air pollution exposure and cardiovascular health among young adults in Taipei, Taiwan,” *International Journal of Hygiene and Environmental Health*, 218, 319 – 323.
- MCFADDEN, D. (1974): “The measurement of urban travel demand,” *Journal of Public Economics*, 3, 303–328.
- MOUSSA, M., Y. MOSTAFA, AND A. A. ELWABA (2017): “School Site Selection Process,” *Procedia Environmental Sciences*, 37, 282 – 293, green Urbanism (GU).
- NICHOLL, J., J. WEST, S. GOODACRE, AND J. TURNER (2007): “The relationship between distance to hospital and patient mortality in emergencies: an observational study,” *Emergency Medicine Journal*, 24, 665–668.
- OATES, W. (1969): “The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis,” *Journal of Political Economy*, 77, 957–71.
- OVERMAN, H., S. REDDING, AND A. VENABLES (2001): “The Economic Geography of Trade Production and Income: A Survey of Empirics,” CEPR Discussion Papers 2978, C.E.P.R. Discussion Papers.

- REBACK, R. (2005): “House prices and the provision of local public services: capitalization under school choice programs,” *Journal of Urban Economics*, 57, 275–301.
- REDDING, S. J., D. M. STURM, AND N. WOLF (2011): “History and Industry Location: Evidence from German Airports,” *The Review of Economics and Statistics*, 93, 814–831.
- REINGEWERTZ, Y. (2012): “Do municipal amalgamations work? Evidence from municipalities in Israel,” *Journal of Urban Economics*, 72, 240 – 251.
- REVELLE, C. AND H. EISELT (2005): “Location analysis: A synthesis and survey,” *European Journal of Operational Research*, 165, 1 – 19.
- SCHÖBEL, A. AND D. SCHOLZ (2010): “The big cube small cube solution method for multidimensional facility location problems,” *Computers & OR*, 37, 115–122.
- SCHULER, R. E. (1974): “The Interaction between Local Government and Urban Residential Location,” *The American Economic Review*, 64, 682–696.
- TAVARES, A. F. (2018): “Municipal amalgamations and their effects: a literature review,” *Miscellanea Geographica*, 22, 5–15.
- THISSE, J. AND H. ZOLLER (1983): *Locational Analysis of Public Facilities*, Studies in Regional Science and Urban Economics, North-Holland Publishing Company.
- THISSE, J.-F. AND D. E. WILDASIN (1992): “Public facility location and urban spatial structure : Equilibrium and welfare analysis,” *Journal of Public Economics*, 48, 83–118.
- TIEBOUT, C. M. (1956): “A Pure Theory of Local Expenditures,” *Journal of Political Economy*, 64.
- TURNBULL, G. K. (1989): “Household behavior in a monocentric urban area with a public sector,” *Journal of Urban Economics*, 25, 103 – 115.
- VALENTINYI, A. AND B. HERRENDORF (2008): “Measuring factor income shares at the sectoral level,” *Review of Economic Dynamics*, 11, 820 – 835.
- VINCENT, J. M. (2006): “Public Schools as Public Infrastructure: Roles for Planning Researchers,” *Journal of Planning Education and Research*, 25, 433–437.
- WEISZFELD, E. (1937): “Sur le point pour lequel la somme des distances de n points donnés est minimum,” *Tôhoku Mathematical Journal*, 43, 355–386.
- YINGER, J. (1982): “Capitalization and the Theory of Local Public Finance,” *Journal of Political Economy*, 90, 917–43.

A Supplemental material, Section 3

Figure 14: Montreuil's local urban plan (2012)

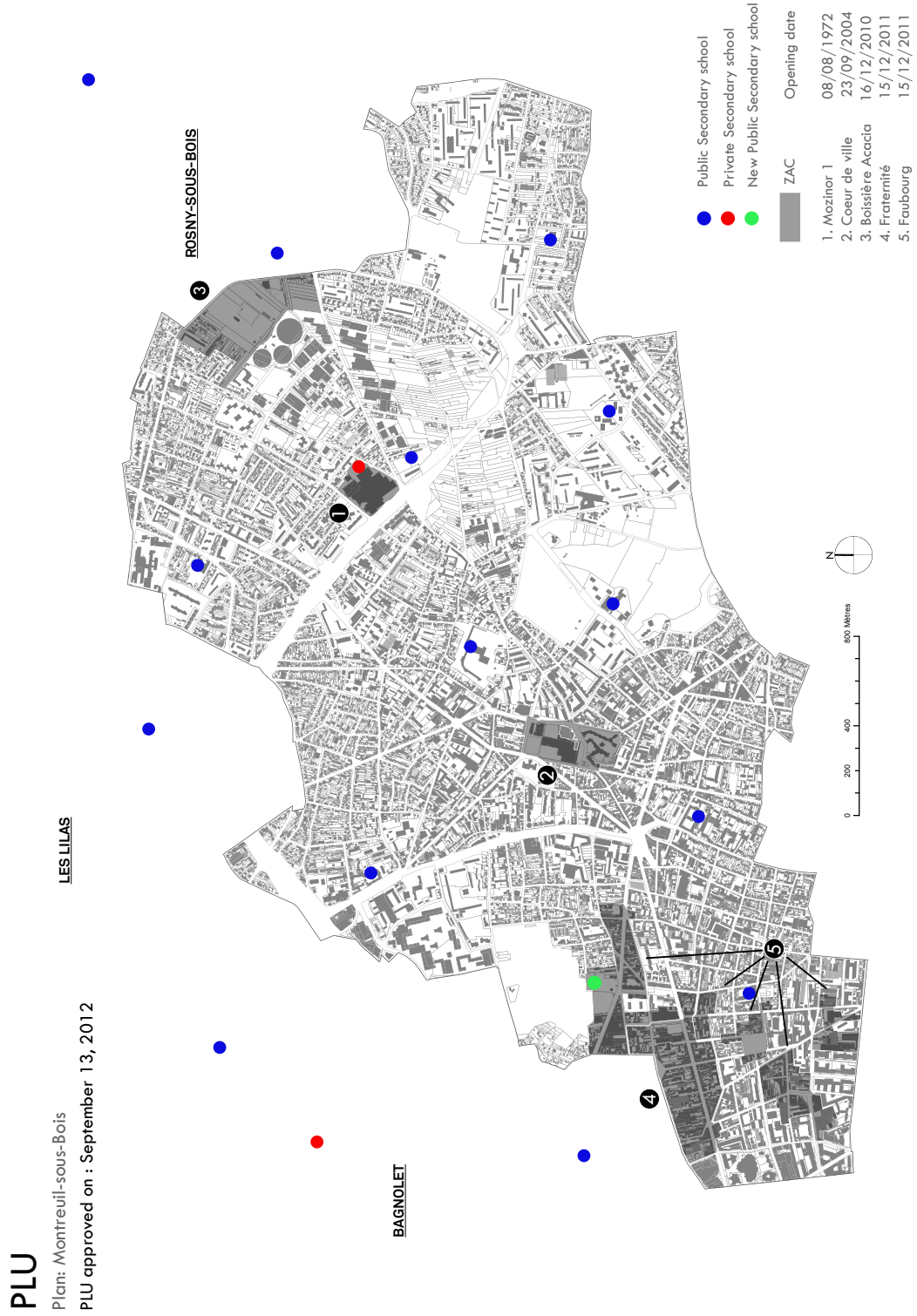
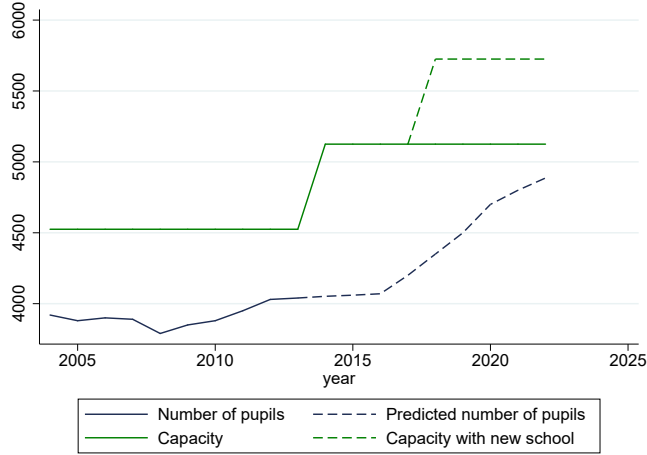


Figure 15: Capacity and demand for secondary schools in Montreuil



B Supplemental material, Section 5

The simulated annealing approach functions as follows:

Simulated annealing algorithm

Define $a < 1$ as the convergence criteria that controls the speed of convergence. Start an iteration count, $iter := 1$.

1. Set the initial locations \mathbf{k}^l to the \mathbf{k} locations obtained by the Big Cube - Small Cube method. Set the initial temperature of the system T to some arbitrary threshold (which should not be too small at first).
 2. Obtain a new candidate strategy by perturbing \mathbf{k}^{l+1} . Compute the corresponding $f(\mathbf{k}^{l+1})$.
 3. Accept the new candidate strategy if the corresponding $f(\mathbf{k}^{l+1})$ is such that $f(\mathbf{k}^{l+1}) > f(\mathbf{k}^l)$ or, with probability $\frac{1}{1+\exp(\frac{f(\mathbf{k}^{l+1})-f(\mathbf{k}^l)}{T})}$, accept a new strategy \mathbf{k}^{l+1} for which $f(\mathbf{k}^{l+1}) < f(\mathbf{k}^l)$.
 4. Stop when the average $(f(\mathbf{k}^{l+1}) - f(\mathbf{k}^l))$ in the last 1'000 iterations is less than some specified small threshold. Otherwise, let $iter := iter + 1$ and $T := aT$ and return to Step 2.
-

C Theory Appendix: Deriving individual location choices

In this section of the Appendix, I detail the analytical derivation of individual location choices. For the sake of completeness, I will sometimes repeat what has already been presented in the main part of the paper. As the relationship between the aggregate consumption index (1) and the idiosyncratic component of utility is monotonic, the distribution of the utility of an individual living in i , working in j , and using public services in k is also Fréchet-distributed:

$$G_{ijk} = Pr[U \leq u] = F\left(\frac{ud_{ik}d_{ij}Q_i^{1-\beta}}{w_j B_i}\right), \quad (31)$$

$$G_{ijk} = e^{-\Phi_{ijk}u^{-\epsilon}}, \quad \Phi_{ijk} = (d_{ik}d_{ij}Q_i^{1-\beta})^{-\epsilon}(B_i w_j)^\epsilon.$$

I first derive the probability that individuals choose a particular combination of residence and public service location.

$$\begin{aligned} \pi_{ijk} &= Pr[u_{ijk} \geq \max\{u_{rst}\}; \forall r, s, t] \\ &= \int_0^\infty \prod_{t \neq k} G_{ijt}(u) \left[\prod_{s \neq j} \prod_k G_{ist}(u) \right] \left[\prod_{r \neq i} \prod_j \prod_k G_{rst}(u) \right] g_{ijk}(u) du \\ &= \int_0^\infty \prod_r \prod_s \prod_t \epsilon \Phi_{ijk} u^{-(\epsilon+1)} e^{-\Phi_{rst}u^{-\epsilon}} du \\ &= \int_0^\infty \epsilon \Phi_{ijk} u^{-(\epsilon+1)} e^{-\Phi u^{-\epsilon}} du \end{aligned} \quad (32)$$

Noting that

$$\frac{d}{du} \left[-\frac{1}{\Phi} e^{-\Phi u^{-\epsilon}} \right] = \epsilon u^{-(\epsilon+1)} e^{-\Phi u^{-\epsilon}}, \quad (33)$$

I obtain the probability that a individual resides in i and use public service in k :

$$\pi_{ijk} = \frac{(d_{ik}d_{ij}Q_i^{1-\beta})^{-\epsilon}(B_i w_j)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon}(B_r w_s)^\epsilon} \equiv \frac{\Phi_{ijk}}{\Phi}. \quad (34)$$

Equation (34) shows that individuals will sort across all combinations of residence and public service locations depending on their idiosyncratic preferences and the characteristics of these locations. To ensure tractability of the general equilibrium, and because I do not observe the necessary individual characteristics in the data, I abstract from other dimensions of heterogeneity among individuals.

I can derive the probability that an individual decides to live in i among all possible locations in the metropolitan area by summing Φ_{is} across all public service locations.

$$\pi_{Ri} = \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} (B_i w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \equiv \frac{\sum_{s=1}^I \sum_{t=1}^K \Phi_{ist}}{\Phi} \quad (35)$$

The same can be done for the probability that an individual decides to use the public service in k among all possible public service locations:

$$\pi_{Sk} = \frac{\sum_{r=1}^I \sum_{s=1}^I (d_{rk} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \equiv \frac{\sum_{i=1}^I \sum_{j=1}^I \Phi_{rsk}}{\Phi}. \quad (36)$$

Probability of working in j is then:

$$\pi_{Mj} = \frac{\sum_{r=1}^I \sum_{t=1}^K (d_{rt} d_{rj} Q_r^{1-\beta})^{-\epsilon} (B_r w_j)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \equiv \frac{\sum_{j=1}^I \sum_{k=1}^K \Phi_{rjt}}{\Phi}. \quad (37)$$

D Theory appendix: Expected utility

Here, I detail the derivation of the expected utility in the metropolitan area. The derivation follows the derivation in Section C.

Among all possible combinations of place of residence and public service, individuals choose the combination that offers the highest utility. Since the maximum of a sequence of Fréchet-distributed random variables is also Fréchet-distributed, the distribution of utilities across all combinations is:

$$1 - G(u) = 1 - \prod_{r=1}^S \prod_{s=1}^S \prod_{t=1}^S e^{-\Phi_{rst} u^{-\epsilon}}, \quad (38)$$

where the left-hand side is the probability that an individual has a utility lower than u and the right-hand side is one minus the probability that an individual has a utility level lower than u for all possible pairs of blocks of residence and public service. This leads to:

$$G(u) = e^{-\Phi u^{-\epsilon}}, \quad \Phi = \sum_{r=1}^S \sum_{s=1}^S \sum_{t=1}^S \Phi_{rst}. \quad (39)$$

Given that utility is Fréchet-distributed, I can derive the expected utility of moving to the metropolitan area:

$$\mathbb{E}[u] = \int_0^\infty \epsilon \Phi u^{-\epsilon} e^{-\Phi u^{-\epsilon}} du. \quad (40)$$

Setting the following variable changes,

$$y = \Phi u^{-\epsilon}, \quad dy = -\epsilon \Phi u^{-(\epsilon+1)} du, \quad (41)$$

I can then write the expected utility of moving to the city as:

$$\mathbb{E}[u] = \int_0^\infty \Phi^{1/\epsilon} y^{-1/\epsilon} e^{-y} dy. \quad (42)$$

This is equivalent to:

$$\mathbb{E}[u] = \gamma \Phi^{1/\epsilon}, \quad \gamma = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right), \quad (43)$$

where $\Gamma(\cdot)$ is the Gamma function, \mathbb{E} is the expectation operator, and the expectation is taken over by the distribution of idiosyncratic utility.

$$\mathbb{E}[u] = \gamma \Phi^{1/\epsilon} = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right) \left[\sum_{r=1}^S \sum_{s=1}^S \sum_{t=1}^S (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon \right]^{(1/\epsilon)} \quad (44)$$

E Theory Appendix: Existence and uniqueness of equilibrium given a location strategy

The existence and uniqueness proofs with commutes along two dimensions closely follow the proof for commutes along one dimension by [Ahlfeldt et al. \(2015\)](#). In addition to commutes along a second dimension, the proof differs from the [Ahlfeldt et al. \(2015\)](#) model in that it assumes given shares of land devoted to housing and commercial uses.

Conditional on a location strategy, the equilibrium of the model is referenced by the vectors $\{\boldsymbol{\pi}_R, \boldsymbol{\pi}_S, \boldsymbol{\pi}_M, \mathbf{Q}, \mathbf{w}\}$ and scalar $\{P\}$. The following expressions define the elements of the equilibrium.

$$\pi_{Ri} = \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} (B_i w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \quad (45)$$

$$\pi_{Mj} = \frac{\sum_{r=1}^I \sum_{t=1}^K (d_{rt} d_{rj} Q_r^{1-\beta})^{-\epsilon} (B_r w_j)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \quad (46)$$

$$\pi_{Sk} = \frac{\sum_{r=1}^I \sum_{s=1}^I (d_{rk} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \quad (47)$$

$$q_j = (1 - \alpha) \left(\frac{\alpha}{w_j} \right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}} \quad (48)$$

$$P = \left(\sum_{j=1}^I p_j^{-\epsilon} \right)^{-1/\epsilon} \quad (49)$$

$$(1 - \theta_i)\phi_i N_i^{1-\mu} = \frac{(1 - \beta)w_j}{Q_i} H_{Ri} \quad (50)$$

Assuming strictly positive, finite, and exogenous characteristics ($H \in (0, \infty)$, $\tilde{H} \in (0, \infty)$, $A_i \in (0, \infty)$, $B_i \in (0, \infty)$, $N_i \in (0, \infty)$, $\rho_{ik} \in (1, \infty) \times (1, \infty)$, $\rho_{ij} \in (1, \infty) \times (1, \infty)$), there exist unique general equilibrium vectors $\{\boldsymbol{\pi}_R, \boldsymbol{\pi}_M, \boldsymbol{\pi}_S, \boldsymbol{Q}, \boldsymbol{w}\}$ and scalar $\{P\}$.

Proof:

The zero-profit condition implies:

$$w_j = \alpha(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_j^{\frac{1}{\alpha}} Q_j^{\frac{\alpha-1}{\alpha}}. \quad (51)$$

Let me rewrite (45) as follows:

$$\pi_{Ri} = \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it}d_{is}Q_i^{1-\beta})^{-\epsilon} (B_i\alpha(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_s^{\frac{1}{\alpha}} Q_s^{\frac{\alpha-1}{\alpha}})^{\epsilon}}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt}d_{rs}Q_r^{1-\beta})^{-\epsilon} (B_r\alpha(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_s^{\frac{1}{\alpha}} Q_s^{\frac{\alpha-1}{\alpha}})^{\epsilon}}. \quad (52)$$

Assuming that the utility of the chosen units is such that $(\frac{\gamma}{\mathbb{E}[u]})^{\epsilon} = 1$, it follows that the residential land market clearing provides the following system:

$$D_i(Q) = \frac{\alpha(1 - \beta)(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_i^{\frac{1}{\alpha}}}{Q_i^{\frac{1}{\alpha}}} \times \sum_{s=1}^I \sum_{t=1}^K \frac{(B_i\alpha(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_s^{\frac{1}{\alpha}})^{\epsilon}}{(d_{it}d_{is}Q_i^{1-\beta} Q_s^{\frac{1-\alpha}{\alpha}})^{\epsilon}} = (1 - \theta_i)L_i. \quad (53)$$

This land market condition provides a system of I equations for I unknown residential floor space prices, Q_i , which have the following properties:

1. $\lim_{Q_i \rightarrow 0} D_i(Q) = \infty > L_i$
2. $\lim_{Q_i \rightarrow \infty} D_i(Q) = 0 < L_i$
3. $\frac{dD_i(Q)}{dQ_i} < 0$
4. $\frac{dD_i(Q)}{dQ_j} < 0$
5. $\left| \frac{dD_i(Q)}{dQ_i} \right| > \left| \frac{dD_i(Q)}{dQ_j} \right|$

Hence, there exists a unique vector \mathbf{Q} which clears the land market system.

Given housing prices, I can then obtain a vector \mathbf{w} of wages using the zero-profit condition. I can then derive the vectors $\boldsymbol{\pi}_R$, $\boldsymbol{\pi}_M$, and $\boldsymbol{\pi}_S$. Lastly, I can derive the price index P .

Q.E.D.

F Theory Appendix: Proof of global convergence of the BCSC approach

For each $\psi \in \Psi$, I am interested in a function $\bar{f}(\mathbf{k})$ which (i) must always be greater than $f(\mathbf{k})$ for any given vector of facility locations \mathbf{k} and (ii) must be convex. Formally, $\bar{f}(\mathbf{k})$ must meet the following properties:

$$\bar{f}(\mathbf{k}) \geq f(\mathbf{k}), \quad \text{and} \quad \bar{f}''(\mathbf{k}) \geq 0. \quad (55)$$

Let me define $\bar{f}(\mathbf{k})$ as:

$$\bar{f}(\mathbf{k}) = \sum_{t=1}^K \sum_{r=1}^I \sum_{s=1}^I \bar{\mathbb{W}} \times d_{rt}^{-\epsilon}, \quad (56)$$

where

$$\bar{\mathbb{W}} = \mathbb{W}(\min\{\mathbf{Q}\}). \quad (57)$$

Lemma 1 Given Assumption 1, it follows that: $\bar{f}(\mathbf{k}) \geq f(\mathbf{k}), \forall \mathbf{k}$.

Lemma 2 Given that $\bar{\mathbb{W}}$ is *not* a function of d_{ik} , the convexity of \bar{f} for a given k can be established using the second derivative of $\bar{f}_k(\mathbf{k})$ as follows:

$$\bar{f}_k''(\mathbf{k}) = \sum_{r=1}^I \sum_{s=1}^I \epsilon^2 \kappa^2 \bar{\mathbb{W}} \left(\frac{x_r - \tilde{k}_k}{\|x_r - \tilde{k}_k\|} \right)^2 e^{(-\epsilon \kappa \|x_r - \tilde{k}_k\|)} > 0, \quad \forall \tilde{k}_k \neq x_r. \quad (58)$$

Given that the sum of convex functions is also convex, we have: $\bar{f}''(\mathbf{k}) > 0$.

Q.E.D.

G Theory Appendix: The pure distance-minimizing problem

The distance-minimizing problem can be formulated as follows:

$$\min_{\tilde{k}_1, \dots, \tilde{k}_K} = \sum_K \sum_I H_{Ri} d_{ik}. \quad (59)$$

The optimization of the center location can be done using a standard gradient-optimization approach. Note that (59) is a separable function of the centers, which allows us to decompose the multi-facility problem into multiple single-facility problems as:

$$f(\tilde{k}_1, \dots, \tilde{k}_K) = \sum_{k=1}^K f_k(\mathbf{k}), \quad \text{where} \quad f_k(\mathbf{k}) = \sum_{i=1}^I H_{Ri} e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)}, \quad k \in K. \quad (60)$$

Given that (60) is a convex function of \tilde{k}_k ,⁴⁴ $\nabla f_k(\mathbf{k}) = 0$ is necessary and sufficient for obtaining the minimum. The gradient of $f_k(\mathbf{k})$ is given by:

$$\nabla f_k(\mathbf{k}) = \sum_{i=1}^I \frac{H_{Ri} \epsilon \kappa e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)}}{\|x_i - \tilde{k}_k\|} (x_i - \tilde{k}_k). \quad (61)$$

Setting $\nabla f_k(\mathbf{k})$ equal to 0, I obtain the optimal centers:

$$\tilde{k}_k^* = \sum_{i=1}^I \delta_k(x_i) x_i, \quad (62)$$

as convex combinations of the data points, with weights $\delta_k(x_i)$ given by:

$$\delta_k(x_i) = \frac{e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)} H_{Ri} / \|x_i - \tilde{k}_k\|}{\sum_{i=1}^I e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)} H_{Ri} / \|x_i - \tilde{k}_k\|}. \quad (63)$$

From (62) and (63), I get K mappings of $T_k : \mathbf{k} \rightarrow T_k(\mathbf{k})$:

$$T_k(\mathbf{k}) = \sum_{i=1}^I \left[\frac{H_{Ri} e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)} / \|x_i - \tilde{k}_k\|}{\sum_{i=1}^I H_{Ri} e^{(-\epsilon\kappa\|x_i - \tilde{k}_k\|)} / \|x_i - \tilde{k}_k\|} \right] x_i, \quad \forall \tilde{k}_k \neq x_i. \quad (64)$$

The algorithm converges to the global solution of the distance-minimizing problem. Empirically, I stop the iterative process when the sum of the distances between the centers of the last two iterations is less than an arbitrarily given criterion. In the application,

⁴⁴See convexity proof in Appendix F.

the iterations stop when the last move made by each center is less than 0.001 decimal degree (approx. 70m in Paris' metropolitan area). Formally,

$$\sum_{k=1}^K \mathbb{1}[d(\tilde{k}_k^+, \tilde{k}_k) < 0.001] = 0, \quad (65)$$

where $d(\tilde{k}_k^+, \tilde{k}_k)$ is the distance between the last two iterations of a center's location.

H Calibration Appendix: Calibration of location amenities, productivity, and density of development

H.1 Residential amenities

The share of individuals residing in i as been defined as follows:

$$\pi_{Ri} = \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} (B_i w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon}. \quad (66)$$

I can rewrite this condition as a system of I equations for the I unknown residential amenities B_i as follows:

$$D_i(B) = \pi_{Ri} - \frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} (B_i w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} = 0. \quad (67)$$

Lemma 3: Given the parameters $\{\epsilon, \kappa, \zeta, \beta\}$, and observables $\{\mathbf{Q}, \boldsymbol{\rho}, \boldsymbol{\rho}, \mathbf{w}, \boldsymbol{\pi}_R\}$, the system in (67) exhibits the following properties:

Property 1: $D(B)$ is continuous.

Property 2: $D(B)$ is homogeneous of degree zero.

Property 3: $\sum_{i=1}^I D_i(B) = 0$.

Property 4: $D(B)$ exhibits gross substitution:

$$\frac{\partial D_i(B)}{\partial B_j} > 0, \quad \forall i, j, i \neq j \quad (68)$$

$$\frac{\partial D_i(B)}{\partial B_i} < 0, \quad \forall i \quad (69)$$

PROOF: Properties 1 and 2 of Lemma 3 follow directly from an inspection of (67). Property 3 is satisfied by noting:

$$\begin{aligned}\sum_{i=1}^I D_i(B) &= \sum_{i=1}^I \pi_{Ri} - \sum_{i=1}^I \left[\frac{\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} (B_i w_s)^\epsilon}{\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} (B_r w_s)^\epsilon} \right] \\ &= 1 - 1 \\ &= 0.\end{aligned}\tag{70}$$

Property 4 can be established by noting:

$$\frac{\partial D_i(B)}{\partial B_j} = \frac{\epsilon B_i^{2\epsilon-1} (\sum_{s=1}^I \sum_{t=1}^K (d_{it} d_{is} Q_i^{1-\beta})^{-\epsilon} w_s^\epsilon)^2}{[\sum_{r=1}^I \sum_{s=1}^I \sum_{t=1}^K B_r^\epsilon (d_{rt} d_{rs} Q_r^{1-\beta})^{-\epsilon} w_s^\epsilon]^2} > 0.\tag{71}$$

Using property 2, which implies $\nabla D_i(B)B = 0$, it follows that:

$$\frac{\partial D_i(B)}{\partial B_i} < 0, \quad \forall i.\tag{72}$$

Thus, gross substitution is established.

Q.E.D.

Lemma 4: Given the parameters $\{\epsilon, \kappa, \zeta, \beta\}$ and observables $\{Q_i, \rho_{ik}, \rho_{ij}, w_j, \pi_{Ri}\}$, there exists a unique vector B which solves (67).

PROOF: I proceed in two steps. First, I show that there exists at most one (normalized) vector B which solves (67). Second, I show a vector B that solves (67) exists.

Gross substitution requires that $D(B) = D(B')$ cannot occur if B and B' non-collinear vectors. By homogeneity of degree zero, we can assume that $B' \geq B$ and $B_i = B'_i$ for some i . Now suppose that we lower (or keep constant) B' in all locations except in i one at a time. By gross substitution, B_i will increase in at least one step. Hence, $D(B) > D(B')$ which is a contradiction.

By homogeneity of degree zero, the search for an equilibrium amenity vector can be restricted to the unit simplex $\Delta = \{\sum_{i=1}^I B_i = 1\}$. Define on Δ the function $D^+(\cdot)$ by $D_i^+(B) = \max\{D_i(B), 0\}$. $D^+(\cdot)$ is continuous. Denote $\alpha(B) = \sum_{i=1}^I [B_i + D_i^+(B)]$ with $\alpha(B) \geq 1, \forall B$. Then define the function $f(\cdot)$ from the closed convex set Δ into itself as:

$$f(B) = [1/\alpha(B)][B + D^+(B)].\tag{73}$$

By Brouwer's Fixed Point Theorem, there exist a $B^* \in \Delta$ such that $B^* = f(B^*)$. Since $\sum_{i=1}^I D_i(B) = 0$, it follows that at the fixed point for amenity, $B^* = f(B^*)$ and $D_i(B) = 0$ for all i .

Q.E.D.

Homogeneity of degree zero implies that the equilibrium amenity vector is unique up to a normalization. I impose the normalization that the geometric mean amenity is equal to 1, i.e., $\left[\prod_{i=1}^I B_i \right]^{1/I} = 1$.

H.2 Final goods productivity

Given the parameters $\{\alpha\}$ and the observed data $\{\mathbf{Q}, \mathbf{w}, \mathbf{K}\}$, I can show, using the zero-profit condition that there is a unique mapping of the final goods productivity $\tilde{\mathbf{A}}$:

$$\tilde{A}_j = (1 - \alpha)^{\alpha-1} \left(\frac{\alpha}{w_j} \right)^{-\alpha} Q_j^{1-\alpha}. \quad (74)$$

H.3 Density of development

Similarly, given the parameters $\{\beta, \mu\}$ and the observed data $\{\mathbf{Q}, \mathbf{H}_R, \mathbf{w}, \mathbf{K}, \boldsymbol{\theta}\}$, I can show, using (17), that there is a unique mapping of the density of development $\tilde{\boldsymbol{\phi}}$.

$$\tilde{\phi}_i = \frac{\alpha(1 - \beta)(1 - \alpha)^{\frac{1-\alpha}{\alpha}} A_i^{\frac{1}{\alpha}}}{(1 - \theta_i) Q_i^{\frac{1}{\alpha}} K_i^{1-\mu}} H_{Ri}. \quad (75)$$

I Empirical Appendix: Additional graphs and tables

Table 6: Differences between optimal and observed location strategies

Year	(I)	(II)	(III)	(IV)
	Number of new facilities	Observed strategy (CEU, $\times 10^{11}$)	Optimal strategy (CEU, $\times 10^{11}$)	Welfare diff. relative to growth (%)
2001	1	8.78	8.79	-1.77*
2002	5	8.84	8.84	10.28
2003	5	8.89	8.90	18.58
2004	1	8.93	8.93	0.28
2005	5	8.98	9.00	33.74
2006	5	9.02	9.03	17.82
2007	3	9.06	9.07	16.78
2008	5	9.10	9.11	21.78
2009	1	9.14	9.14	-2.46
2010	3	9.18	9.18	17.15
2012	1	9.25	9.25	3.48
2014	7	9.32	9.33	13.69
2015	2	9.36	9.36	4.00
Average	-	-	-	12.065

Note: Welfare differences relative to growth, for the optimal strategy, is: $100 \times \frac{W_t^{opt} - W_t^{obs}}{W_t^{obs} - W_{t-1}^{obs}}$. CEU stands for Consumption Equivalent Units.

*: The negative sign comes from the fact that the observed strategy decreased welfare between 2000 and 2001.