# Optimal Urban Transport Design in a Quantitative Spatial Model*

Pol Cosentino†

Université Paris-Dauphine, PSL

May 19, 2026

## Abstract

Which metro design should a city adopt to maximise worker welfare? This paper develops a framework combining a quantitative urban model with a simulated annealing algorithm. The model features heterogeneous workers, multiple transport modes, and general equilibrium effects: as commuting costs fall, workers reallocate across residence and workplace and land prices adjust endogenously. The algorithm solves over one million spatial equilibria to sample (sub-)optimal metro designs. Applied to Paris in 1896, the results show that strategic metro stations are driven by location centrality, productivity, and amenities. Among the four competing 19th-century proposals, the adopted Fulgence Bienvenüe network (1896) achieves the highest welfare-efficiency score: by routing lines along Haussmann boulevards, it exploited cheap cut-and-cover construction to connect the city's most strategic locations at minimum cost.

**Key words**: Optimal Urban Transport Infrastructure, Quantitative Spatial Model
**JEL code**: O18, R12, R23

*I am particularly grateful to Clément Bosquet and Clément Gorin for their guidance and support. I benefited from conversations with Sara Bagagli, Pierre-Philippe Combes, Tilman Graff, Stephan Heblich, Camille Hémet, Gabriel Kreindler, and Gabriel Loumeau. I thank participants of AFET conference, and the UEA conference for their valuable comments and suggestions. Responsibility for results, opinions and errors lies with the author alone.

†Université Paris-Dauphine, Université PSL, LEDA, CNRS, IRD, 75016 PARIS, FRANCE
Email: pol.cosentino@dauphine.psl.eu. First version: September 2024.

# 1 Introduction

Urban public transport is one of the most consequential levers available to city planners. Over the past 25 years, more than 200 metro projects have been built or extended worldwide, each pursued with the goal of reducing commuting costs, easing congestion, and widening workers' access to jobs and amenities. Yet designing these systems optimally is challenging. The space of possible network configurations is vast, locations differ in productivity, amenities, and housing supply, and improvements in commuting costs trigger general equilibrium responses. As workers reallocate across residence and workplace, land prices adjust, and the effects of new infrastructure propagate across the entire city (Redding and Turner, 2015; Heblich et al., 2020; Tsivanidis, 2026).

In this paper, I develop a framework to identify the optimal design of urban transport infrastructure. The framework incorporates general equilibrium effects, which arise as workers reallocate toward different residence and workplace within the city in response to public policy interventions, thereby influencing the land market where workers and firms compete for housing. Additionally, it accommodates multiple transport modes, and accounts for neighborhood heterogeneity in terms of amenities, productivity, and housing supply. It also allows for production and residential agglomeration forces. Taking these elements into account is crucial to enhance the efficiency of urban transport systems.

In this framework, the planner's problem involves choosing a metro network $M$ based on an economic objective $W(M;\theta)$, that depends on structural parameters $\theta$. Additionally, the planner's preferences are influenced by factors beyond the economic objective, such as unobserved preferences for specific characteristics of the metro network. I model this as an idiosyncratic shock, denoted by $\iota_M$. Therefore, the planner faces a discrete choice problem within the space of all possible metro networks. While it is impossible to derive a unique global metro network due to the high-dimensional optimization problem, a simulated annealing algorithm is used to sample (sub)optimal metro networks.

First, I provide a numerical example of a city with random distributions of fundamentals and demonstrate how to rationalize the within-city spatial equilibrium using a canonical quantitative urban model (QUM) *à la* Ahlfeldt et al. (2015), with structural parameters calibrated to the literature. Subsequently, I apply the developed framework to compute the optimal metro design in this numerical example.

Second, I apply my methodology to address the following question: How should the

Parisian metro have been planned at the beginning of the 20th century? I combine several historical data and divide the Parisian space into a $500m \times 500m$ grid to create locations and potential metro stations. In this application, a calibrated quantitative urban model (QUM) with heterogeneous workers allows to recover location type-specific amenities and productivity through the rationalization of the within-city spatial equilibrium. In the baseline scenario, I define an economic objective $W(M;\theta)$ as the gains in rateable value in open city settings, following the Henry George Theorem (George, 1879).[1] Additionally, I estimate the construction cost of the metro $M$ based on fixed costs, the number of stations, and lines. In this Parisian application, I solve one million spatial equilibria and study the features that characterize 500 optimal metro designs.

First, the numerical example of a city with a data generating process and a canonical QUM without agglomeration forces provides generalizable results. I find that the strategic locations of stations are driven by amenities and productivity. In addition, the station's centrality in the network is also an important determinant in the planner's choice of metro network. This result goes in line with Borusyak and Hull (2023) who emphasize that central stations are much more likely to be chosen to increase market potential.

Second, I validate these findings in my historical application on Paris using the economic objective based on rateable values gains and without considering agglomeration forces. As in the numerical simulation, I find that the strategic locations of stations are driven by centrality, productivity and low-skilled amenities (since low-skilled workers constituted the main labor force at that time). Additionally, over the sampled metro networks, the average economic impact factor is approximately 7.89, aligning with the findings of Heblich et al. (2020). This means that the economic benefits are 7.89 times the metro construction costs. The average number of metro stations is around 51, leading to an average distance to a metro station of approximately 913 meters[2]. The average increases in total production and population within the city are 33.9% and 22.1%, respectively. Furthermore, the average decrease in total commuting time within the city

[1] In the classical approach to valuing public goods, landlords experience the welfare gains from new transport technology through changes in the value of land and buildings (George, 1879). This methodology has also been used in Heblich et al. (2020) to evaluate the metro in London.

[2] The total potential area consists of 367 grid cells. Each grid cell covers $0.25\,\text{km}^2$, resulting in a total area of $91.75\,\text{km}^2$. With 51 stations, each station covers approximately $1.799\,\text{km}^2$. Assuming a square area per station, the side length is $\sqrt{1.799\,\text{km}^2} \approx 1.341\,\text{km}$. The diagonal of this square is $1.341\,\text{km} \times \sqrt{2} \approx 1.897\,\text{km}$, so the average distance is half the diagonal: $1.897\,\text{km}/2 \approx 0.913\,\text{km}$.

is 20.4%, with significant disparities and greater improvements in locations closer to the city center.

Then, I conduct several comparative experiments to assess the role of structural parameters and metro construction cost parameters. First, I confirm the role of productivity in determining strategic metro stations by creating a counterfactual CBD. I demonstrate that the developed framework reacts to this experiment and predicts a different urban transport design which follows the productivity shock. This is reflected in a higher proportion of strategic metro stations in the new CBD. Second, I assess the role of metro construction fixed costs (such as technology acquisition) and show that these costs determine the size of the metro network, as they create incentives for planners to amortize these large investments. The higher the fixed cost, the larger the metro network. Lastly, introducing agglomeration forces to the quantitative urban model (QUM) results in larger people inflows from the outside economy, which in turn enables a larger optimal metro network, while strategic stations become less concentrated into specific locations. Hence, agglomeration forces both expand the size of the optimal metro design and reshape its spatial pattern.

Next, I consider a budget-constrained planner that tries to build a metro system in this historical context. I calibrate the total metro budget achieved at that time, and I compare a utilitarian scenario, where the planner aims to maximize the sum of utilities for both high-skilled and low-skilled workers, with an egalitarian scenario, where the planner seeks to maximize the minimum utility of both groups. I find that the utilitarian scenario predicts relatively more strategic metro stations in the west side of Paris compared to the egalitarian scenario. Hence, it tends to favor high-skilled workers more than low-skilled workers, as high-skilled amenities are relatively more prevalent in the west of Paris than low-skilled amenities.

Finally, I define an optimal urban transport design proximity index and use it to evaluate metro projects proposed during the 19th century. The results show that the adopted Fulgence Bienvenüe network (1896) achieves the highest proximity index, both in raw terms and once normalised by network length. This result stems directly from the introduction of Haussmann-boulevard cost heterogeneity: the Bienvenüe design was explicitly routed along the wide Haussmann boulevards, making extensive use of the cheaper cut-and-cover technology.This finding suggests that the adopted metro design was not only

politically convenient but also economically sound, and offers a novel interpretation of the decision-making process documented in Cottereau (2004).

My work relates to two strands of literature. First, I connect to recent research in quantitative spatial models by using and extending a canonical quantitative urban model (QUM) developed by Ahlfeldt et al. (2015). Since this seminal work, a large literature has emerged to study the welfare and distributional effects of transport infrastructure improvements (Heblich et al., 2020; Severen, 2021; Allen and Arkolakis, 2022; Brinkman and Lin, 2022; Bagagli, 2023; Weiwu, 2024; Tsivanidis, 2026; Cosentino, 2026), neighborhood effects (Redding and Sturm, 2024), and forward-looking migration decisions (Desmet et al., 2018; Caliendo et al., 2019; Balboni, 2025; Allen and Donaldson, 2020; Heblich et al., 2021; Kleinman et al., 2023; Takeda and Yamagishi, 2023). These models are sufficiently rich to capture the organization of economic activity within cities with low data requirements and facilitate counterfactual analyses that take into account general equilibrium effects to quantify the welfare effects of public policies.

Second, this paper aligns with the literature on the optimal design of transport networks. Recent works have studied the endogenous choice of transport infrastructure using a trade perspective (Balboni, 2025; Fajgelbaum and Schaal, 2020; Santamaria, 2020; Allen and Arkolakis, 2022). These studies provide frameworks for optimal investment in inter-city infrastructure. With the exception of Kreindler et al. (2025), who focus on the bus network in Jakarta, and Cosentino (2026), who compare circular and radial transit systems in historical settings, little has been done on the design effectiveness of urban transport infrastructure. In their paper, Kreindler et al. (2025) use a travel demand model to estimate public transport preferences (such as directness, speed, and waiting times) and provide a set of potential optimal networks under the assumption that commuting patterns are fixed. I contribute to this literature on optimal transport design by building on Kreindler et al. (2025) to provide a new framework that incorporates general equilibrium effects, as workers reallocate toward different residence and workplace following improvements in urban transport infrastructure. This framework allows for agglomeration forces and considers several transport modes. Additionally, it enables planners to set metro construction cost parameters and the minimum number of lines to be built.

The remainder of the paper is structured as follows. Section 2 presents a canonical

quantitative urban model with homogeneous workers residing in a closed city and provides a numerical example of a city. Section 3 introduces a simulated annealing algorithm to determine an optimal urban transport infrastructure design and discusses the findings of this numerical example. Section 4 introduces the application to Paris by presenting the data and discussing the historical context, outlining and quantifying the quantitative spatial model by solving the spatial equilibrium, and presenting the findings. Section 5 presents an index for comparing the different historical metro network projects with the results from the simulated annealing algorithm based on various scenarios. Finally, Section 6 concludes.

# 2 Canonical Quantitative Urban Model

## 2.1 Model set-up

### 2.1.1 Workers

I consider a mass of workers living and working within a city of $\mathbb{N}$ discrete locations embedded in an economy $\mathbb{M}$. Workers choose a pair of locations $\{n, i\}$, where $n$ is the location of residence and $i$ is the workplace. A worker $o$ choosing the pair of locations $\{n, i\}$ derives the following indirect utility,

$$U_{ni}(o) = \frac{B_n w_i}{d_{ni} P_n^{1-\beta} Q_n^{\beta}} \epsilon_{ni}(o), \tag{1}$$

where $B_n$ represents the amenities in residence $n$, $w_i$ is the wage in workplace $i$, and $d_{ni} = e^{\kappa t_{ni}}$ is the commuting cost, which takes an exponential functional form (following Ahlfeldt et al. (2015)). $P_n$ is the price of the final consumption good (taken as numéraire, $P_n = 1$ for all $n$), and $Q_n$ is the rent paid in location $n$. The worker $o$ derives an idiosyncratic shock $\epsilon_{ni}(o)$ following a Fréchet distribution. Finally, $\beta$ represents the share of income that the worker spends on housing.

Using standard properties of the Fréchet distribution (McFadden, 1974), the probability that a worker chooses to live in $n$ and work in $i$ is given by,

$$\lambda_{ni} = \frac{L_{ni}}{R_{\mathbb{N}}} = \frac{(B_n w_i)^{\epsilon} (d_{ni} Q_n^{\beta})^{-\epsilon}}{\sum_{k \in \mathbb{N}} \sum_{l \in \mathbb{N}} (B_k w_l)^{\epsilon} (d_{kl} Q_k^{\beta})^{-\epsilon}} = \frac{\Phi_{ni}}{\sum_{k \in \mathbb{N}} \sum_{l \in \mathbb{N}} \Phi_{kl}}, \tag{2}$$

where $L_{ni}$ is the number of commuters from $n$ to $i$ and $R_{\mathbb{N}}$ is the total number of workers in the city. As explained in Redding (2022), the probability of commuting between locations $n$ and $i$ depends on location characteristics, workplace characteristics, and the bilateral cost (numerator), as well as all other residence and workplace characteristics, and bilateral commuting costs (denominator).

By summing across workplaces, I obtain the probability that a worker lives in $n$,

$$\lambda_n^R = \frac{R_n}{R_{\mathbb{N}}} = \frac{\sum_{i\in\mathbb{N}} \Phi_{ni}}{\sum_{k\in\mathbb{N}} \sum_{l\in\mathbb{N}} \Phi_{kl}}, \tag{3}$$

and by summing across residence places, I obtain the probability that a worker works in $i$,

$$\lambda_i^L = \frac{L_i}{R_{\mathbb{N}}} = \frac{\sum_{n\in\mathbb{N}} \Phi_{ni}}{\sum_{k\in\mathbb{N}} \sum_{l\in\mathbb{N}} \Phi_{kl}}. \tag{4}$$

Finally, population mobility ensures that everyone derives the same expected utility in the city,

$$\overline{U} = \delta \left[ \sum_{k\in\mathbb{N}} \sum_{l\in\mathbb{N}} \Phi_{kl} \right]^{1/\epsilon}, \tag{5}$$

where $\delta \equiv \Gamma\left(\frac{\epsilon-1}{\epsilon}\right)$ and $\Gamma(.)$ is the gamma function.

### 2.1.2 Firms

The representative firm produces the final good with a Cobb-Douglas technology under constant returns to scale,

$$Y_i = A_i \left(\frac{L_i}{\alpha}\right)^{\alpha} \left(\frac{H_i^L}{1-\alpha}\right)^{1-\alpha}, \tag{6}$$

where $A_i$ represents productivity in workplace $i$, $L_i$ is the total labor supply in workplace $i$, and $H_i^L$ is the housing used commercially in workplace $i$. The share of labor in the production function is $\alpha$.

The first-order conditions of the firm's profit maximization with respect to labor and housing are,

$$w_i = \alpha \frac{Y_i}{L_i}, \tag{7}$$

$$Q_i = (1-\alpha) \frac{Y_i}{H_i^L}. \tag{8}$$

### 2.1.3 Housing

Following Combes et al. (2021), housing is produced by developers with a Cobb-Douglas technology combining land and machinery capital. Using profit maximization, the housing supply can be written as,

$$H_i^S = k_i \, Q_i^{\frac{(1-\mu)}{\mu}}, \tag{9}$$

where $k_i = (1-\mu)^{\frac{(1-\mu)}{\mu}} K_i$, and $\frac{(1-\mu)}{\mu}$ is the housing supply elasticity. Since land availability $K_i$ is assumed to be fixed, only rent variation drives the housing supply (see appendix B.1.1 for more details).

### 2.1.4 Markets clearing

**Commuter market clearing** The commuter market clearing condition implies a gravity equation, requiring that the number of workers employed at each workplace equals the total number of residents commuting there.

Using the commuting probability equation (2) and the residential probability equation (3), it is possible to write the probability that a worker commutes from $n$ to $i$, conditionally on living in location $n$ as,

$$\lambda_{ni|n}^R = \frac{\lambda_{ni}}{\lambda_n^R} = \frac{(w_i/d_{ni})^\epsilon}{\sum_{l \in N}(w_l/d_{nl})^\epsilon}. \tag{10}$$

Therefore, the commuter market clearing can be expressed as,

$$L_i = \sum_{n \in N} \frac{(w_i/d_{ni})^\epsilon}{\sum_{l \in N}(w_l/d_{nl})^\epsilon} R_n. \tag{11}$$

**Housing market clearing** The housing market clearing condition implies that the housing supply, $H_n^S$, equals the total housing demand by residents and firms at equilibrium,

$$H_n^S = H_n^R + H_n^L = \beta \, \nu_n \, \frac{R_n}{Q_n} + \frac{1-\alpha}{\alpha} \, \frac{w_n}{Q_n} \, L_n, \tag{12}$$

where $H_n^R$ and $H_n^L$ are the residential and commercial housing demand in $n$, respectively, and $\nu_n \equiv \sum_{i \in N} \lambda_{ni|n}^R w_i$ is the expected income from residents living in location $n$. The housing demand of firms in workplace $i$ is derived from the first-order conditions equations (7) and (8).

### 2.1.5 General equilibrium

In a closed-city setting with an integrated housing market, and given the structural parameters $\{\alpha, \beta, \kappa, \epsilon\}$, a spatial equilibrium is defined by a set of endogenous variables $\{R_n, L_i, Q_n, w_i, \overline{U}\}$ that rationalize the spatial distribution of fundamentals $\{B_n, A_i\}$, housing supply $\{H_n^S\}$, commuting costs $\{d_{ni}\}$, and a calibrated total population $\{R_{\mathbb{N}}\}$.

### 2.1.6 Model extension

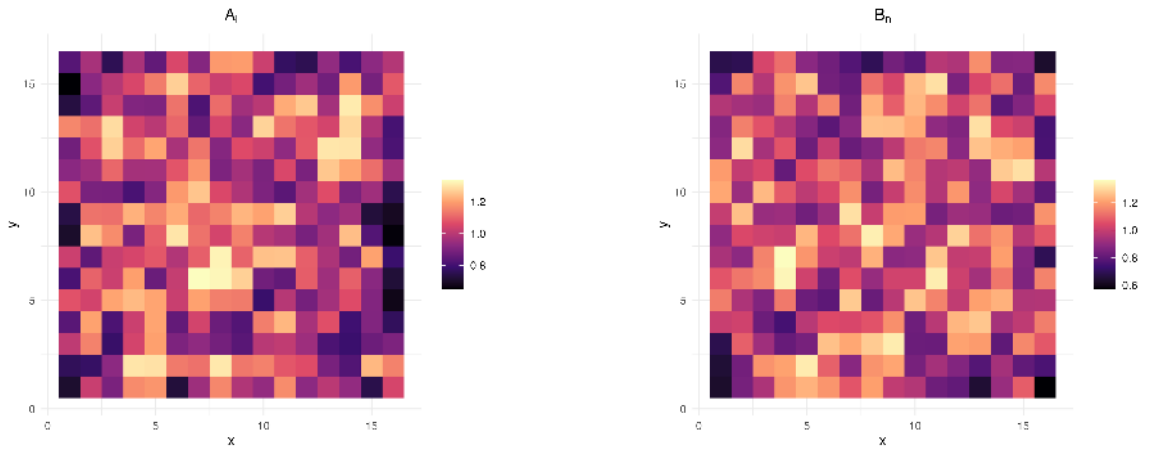
Several model extensions can be made from this canonical quantitative urban model. For example, some papers have considered different types of workers (Tsivanidis, 2026; Loumeau, 2024; Weiwu, 2024; Redding and Sturm, 2024; Champalaune and Cosentino, 2026), dynamics due to mobility friction (Warnes, 2021; Takeda and Yamagishi, 2023; Greaney et al., 2024), congestion (Herzog, 2024), trade of goods (Monte et al., 2018; Cosentino, 2026), consumption access (Miyauchi et al., 2025; Cosentino, 2026), school choices (Loumeau, 2023; Pietrabissa, 2023), and transport mode choices (Koster, 2024; Champalaune and Cosentino, 2026). These extensions of the canonical QUM à la Ahlfeldt et al. (2015) are specific to the context and the research question. Although tractable, they require additional data than the usual observation of residence employment, workplace employment, commuting time and housing.

## 2.2 Quantitative illustration

### 2.2.1 City structure

To illustrate an urban spatial equilibrium from the canonical quantitative urban model developed in section 2, I present a numerical example based on a data-generating process (DGP) and calibrated parameters. The city consists of locations arranged on a $16 \times 16$ grid, with each point spaced 500 meters apart. I abstract from agglomeration forces and randomly assign residential and production fundamentals ($B_n$ and $A_i$), as shown in Figure 1. Housing supply is assumed to be uniform across locations, with $H_n = 100$. Commuting times, $t_{ni}$, are based on Euclidean distances and a walking speed of 5 km/h.

Figure 1
Numerical example: Spatial distribution of fundamentals




(a) Productivity (b) Amenities

Notes: This figure presents the distribution of fundamentals in each location following a data generating process, with Panel (a) depicting productivity and Panel (b) depicting amenities. The color intensity in each panel corresponds to the magnitude of the respective variable, with the color bar indicating the scale.

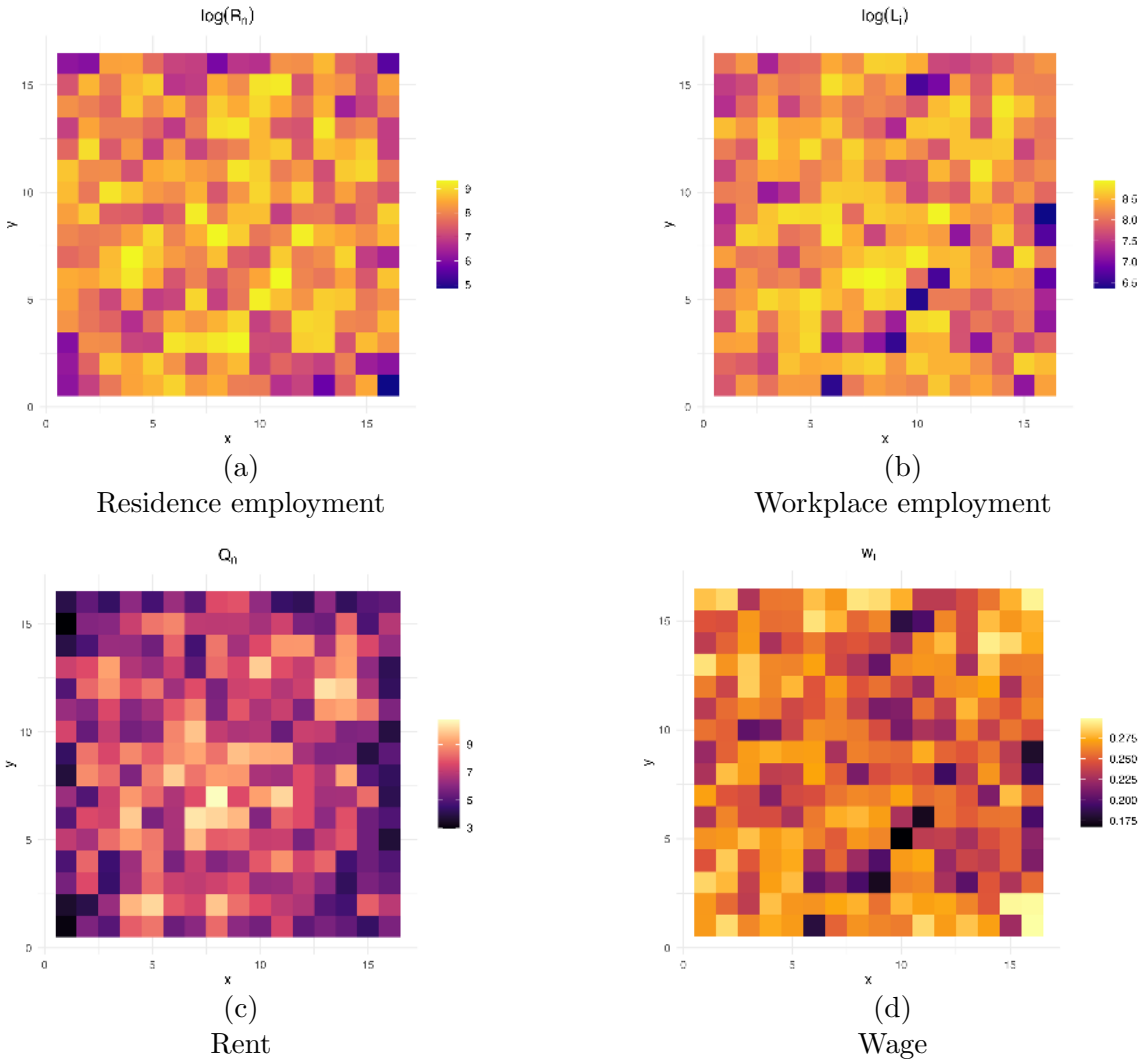
### 2.2.2 Parametrization

I calibrate the structural parameters using central values from the literature. First, the Fréchet parameter $\epsilon$ is set to 5, within the typical range of 2 to 7 reported by Severen (2021) and Ahlfeldt et al. (2015). Second, the commuting cost elasticity $\kappa$ is calibrated to 0.01, consistent with the estimate in Ahlfeldt et al. (2015). Finally, I set $\alpha$ and $\beta$ to 0.7 and 0.25, respectively, aligning with commonly used values in the literature.

### 2.2.3 Initial spatial equilibrium

Figure 2 illustrates the initial spatial equilibrium by plotting the distribution of residence employment ($R_n$), workplace employment ($L_i$), rent ($Q_n$), and wage ($w_i$) in each location.

Figure 2
Numerical example: Initial spatial equilibrium




(a) Residence employment

(b) Workplace employment

(c) Rent

(d) Wage

Notes: This figure presents the initial spatial equilibrium resulting from a canonical QUM with the numerical example. Panel (a) shows the logarithm of residence employment $\log(R_n)$, illustrating the distribution of residential employment across different locations. Panel (b) depicts the logarithm of workplace employment $\log(L_i)$, showing the distribution of workplace employment. Panel (c) represents rent $Q_n$, indicating the spatial variation in rent levels. Panel (d) displays wages $w_i$, reflecting the spatial distribution of wage levels. The color intensity in each panel corresponds to the magnitude of the respective variable, with the color bar indicating the scale.

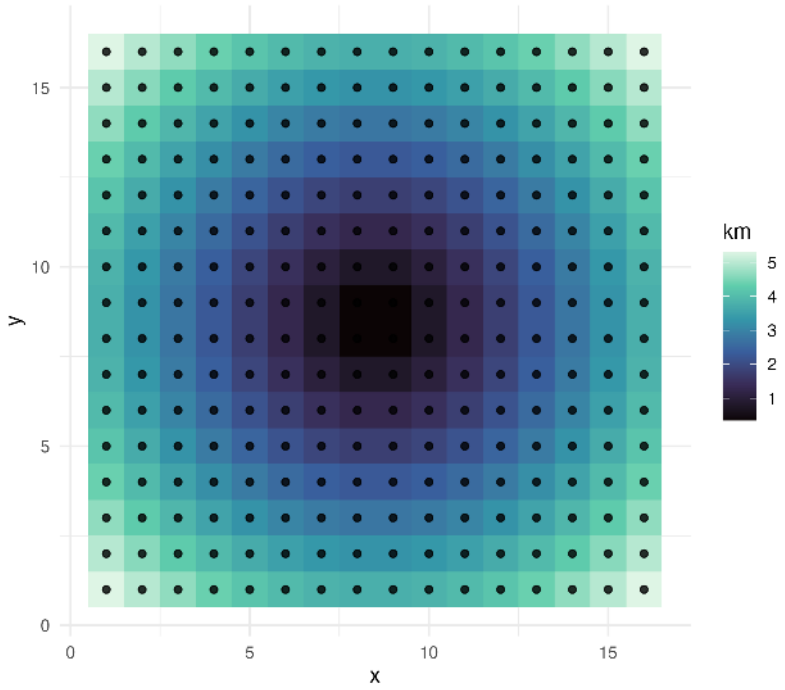
# 3 Optimal Urban Transport Design

## 3.1 Optimization Environment

The planner chooses the metro network $M$ consisting of $L$ metro lines and $S$ selected metro stations that maximizes an economic objective $W(M;\theta)$. The economic objective can take various forms (such as expected utility, total population, total rateable values, total agglomeration forces, etc.) and is context-specific.

The planner can face different constraints in the maximization process, such as network constraints (connecting the metro network to the existing network, a limit on the

Figure 3
Numerical example: Potential metro stations




Notes: This figure shows the potential metro stations of the numerical example. Each black dot represents a potential metro station. The background color gradient is based on the distance to the city center, with darker colors indicating closer proximity to the center.

total number of lines, etc.) or construction constraints (metro stations cannot be implemented in some areas due to monuments, land availability, etc.).

The economic objective $W(M;\theta)$ results from a spatial equilibrium computed through a quantitative spatial model (previously developed in section 2), which is a function of the metro network $M$ and model parameters ($\theta \equiv \{\alpha, \beta, \epsilon, \mu, \kappa, A_i, B_n, H_n, R_{\mathbb{N}}\}$) capturing structural parameters and the spatial distribution of fundamentals.

Finally, the planner needs to divide the space into grid cells to create potential metro stations $\mathcal{S}$ to choose a metro network $M$. The more granular the space of potential stations, the better it captures the complexity of the transport design. However, a large set of potential metro stations also implies a large solution space.

In the numerical example, I consider a planner implementing 3 metro lines without constraints on the number of stations. The planner can build a potential metro station in each location, as shown in Figure 3. Finally, the planner maximizes the expected utility of workers in this city, $\overline{U}$, given by equation (5). However, the optimization problem is discrete and high-dimensional, as highlighted by Kreindler et al. (2025), and it does not appear to have an analytical solution.

### 3.1.1 Planner problem

Considering the framework where the planner needs to choose a metro network $M$ among $\mathcal{M}$, she solves,

$$\max_{M \in \mathcal{M}} W(M;\theta) + \iota_M, \tag{13}$$

where $W(M;\theta)$ is the economic objective (here $\overline{U}$) given the metro network $M$ and model parameters $\theta$, and $\iota_M$ represents an idiosyncratic shock for a network $M$ following a Gumbel distribution with a shape parameter $\sigma$. For example, the idiosyncratic shock captures a specific planner's preference for a particular design. Hence, the probability that the network $M$ is chosen by the planner is,

$$\pi(M;\theta) = \frac{\exp(\sigma\, W(M;\theta))}{\sum_{M' \in \mathcal{M}} \exp(\sigma\, W(M';\theta))}. \tag{14}$$

### 3.1.2 Optimization process

While the probability distribution $\pi$ cannot be computed due to the high dimensionality of $\mathcal{M}$, the probability ratio of two different metro networks can be computed as follows,

$$\frac{\pi(M';\theta)}{\pi(M;\theta)} = \frac{\exp(\sigma\, W(M';\theta))}{\exp(\sigma\, W(M;\theta))}. \tag{15}$$

Therefore, I use a simulated annealing (SA) algorithm to explore the space of metro networks as in Kreindler et al. (2025). This algorithm is a step-by-step process. It starts with an initial metro network (randomly draw) $M_k$ with $L$ metro lines and $S$ metro stations, which gives an initial economic objective $W(M_k;\theta)$. In addition, there is an "inverse temperature" parameter $\sigma$ that increases over the $K$ steps.[3] This parameter is crucial and ensures that the algorithm explores the space of networks widely.

At step $k$, a perturbation function $\Psi$ creates a candidate network $M'$ given the current network $M_k$. The acceptance probability of the candidate network is defined as,

$$Pr(M_{k+1} = M' | M_k) = \min\left(1, \frac{\exp(\sigma\, W(M';\theta))}{\exp(\sigma\, W(M_k;\theta))}\right). \tag{16}$$

Hence, the candidate network $M'$ is always accepted if it has a greater economic objective ($W' > W_k$), and otherwise $M'$ is accepted with a probability that increases with $W'$ and

[3] For more on the "inverse temperature", I refer the reader to Kreindler et al. (2025).

decreases with $\sigma$.

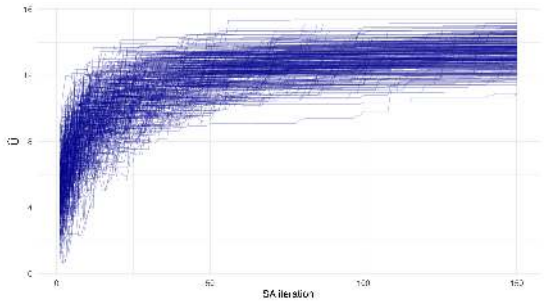
By doing so, this Markov chain samples from the planner's distribution $\pi$, starting with low values of $\sigma$ to more often accept $M'$ and widely explore the space of metro networks, and ending with high values of $\sigma$ to put more weight on the economic objective $W'$.

The perturbation function $\Psi$ modifies the current metro network $M$ based on probabilistic scenarios. For simplicity, I consider only the possibility to remove the current network and create a new one in this numerical example. After this perturbation, which gives a candidate network $M'$, the matrix of commuting cost is updated by computing bilateral commuting time between locations, which in turn allows the computation of a candidate economic objective $W'(M';\theta)$ through the quantitative urban model.

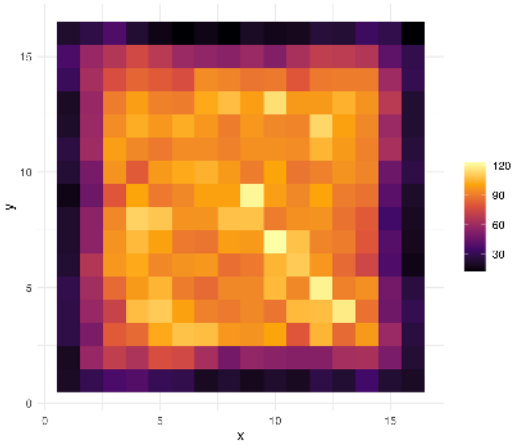
## 3.2 Simulation results with a numerical example

Using this numerical example as an experiment, I sample 500 metro networks through independent runs with 200 iterations and $\sigma_K = 1{,}000$. Panel (a) of Figure 4 shows the convergence of each run over the iterations, with most stabilizing around 13% welfare gains. Over the 500 sampled metro networks, it may be more efficient for a metro line to pass through certain locations than others. In order to test whether certain metro stations are more strategic than others, I sum up the frequency of appearance of the stations across all runs. Panel (b) of Figure 4 displays all potential metro stations, with their size being proportional to the frequency of their appearance across all runs, providing an indication of the systematic importance of certain stations. We observe that stations in the city center are sampled more frequently than those on the periphery. This pattern is a mechanical result, as central stations inherently possess greater market potential than others (Borusyak and Hull, 2023). Controlling for the distance to the city center to capture centrality, as well as longitude and latitude, Panels (c) and (d) illustrate the correlation between location fundamentals and the frequency of metro stations. The results indicate that amenities and productivity are positively correlated with the frequency of metro stations. Therefore, these factors appear to be crucial in the design of optimal urban transport infrastructure.

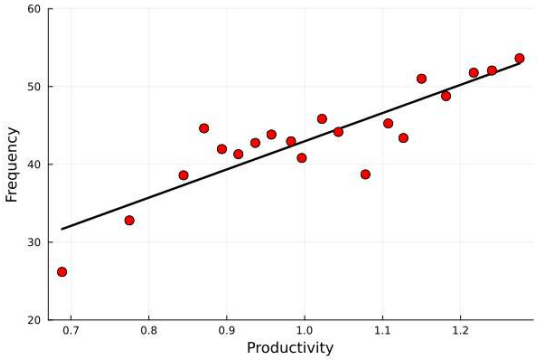
Figure 4
Numerical example: Simulated annealing algorithm results



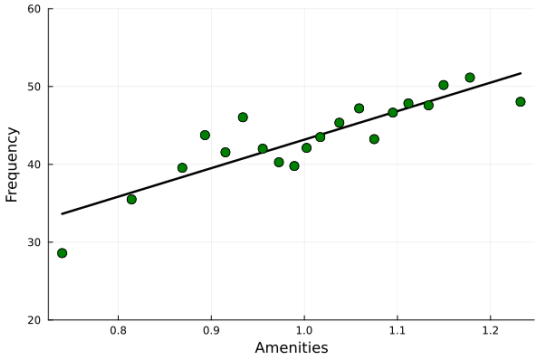

(a)
SA runs




(b)
Metro stations frequency




(c)
Productivity and frequency




(d)
Amenities and frequency

Notes: This figure presents the results of the simulated annealing algorithm in a numerical example. Panel (a) shows the results of SA runs over 200 iterations, with each line representing a single run. Panel (b) displays all potential metro stations, with their size being proportional to the frequency of their appearance across all 500 runs. Panel (c) plots a binned correlation between metro station frequency and their associated productivity, illustrating how productivity influences station frequency. Panel (d) plots a binned correlation between metro station frequency and their associated amenities, showing the relationship between amenities and station frequency.

# 4 An application to historical Paris

## 4.1 Context and data

In this subsection, I first present the data that enables quantitative analysis. Next, I detail the historical context during which the Parisian metro was constructed.

### 4.1.1 Data

**Geography** I use a grid of $500\,\text{m} \times 500\,\text{m}$, made up of 450 spatial units. This granularity matters as it allows capturing more complexity in location characteristics, and therefore in urban transport design. Each spatial unit varies in terms of amenities, productivity, centrality, housing supply, and transport network. This variation allows for a better understanding of the different factors at play in explaining urban transport design.

**Residence and workplace employments** First, I use data from a companion paper Cosentino (2026) to obtain residential information across Parisian neighborhoods in 1896. Overall, the Parisian population reached $2,536,834$ in 1896, but this aggregate figure masks significant spatial disparities. Specifically, the city center experienced a decrease in population, while most residents moved to the periphery throughout the 19th century, as documented by Cosentino (2026). I combine this residential information with occupational data provided by the 1891 census, which allows for the differentiation between high-skilled and low-skilled residents by neighborhood. Then, I proceed to distribute the population by type across the grid cells using a newly digitized urban footprint map shown in Figure 5. This map, consisting of around 4,600 buildings and areas, enables me to distinguish between permanent structures (such as monuments and scenic gardens) and non-permanent footprints (such as commercial and residential buildings). I supplement the footprint information with building height data from the national buildings database (BDNB), which provides the year of construction and height of each building not yet demolished. This allows me to compute the building volume within each grid cell and distribute the population of each neighborhood accordingly. Panel (a) and Panel (b) of Figure 6 show the spatial distribution of high-skilled and low-skilled residents, respectively. While both types of workers seem to concentrate in the city center, there is spatial sorting of high-skilled workers toward the west and low-skilled workers

toward the east, aligning with the pattern described by Heblich et al. (2021) (see Figure A2 for an illustration by showing the ratio of the high-skilled residence employment with respect to the low-skilled residence employment).

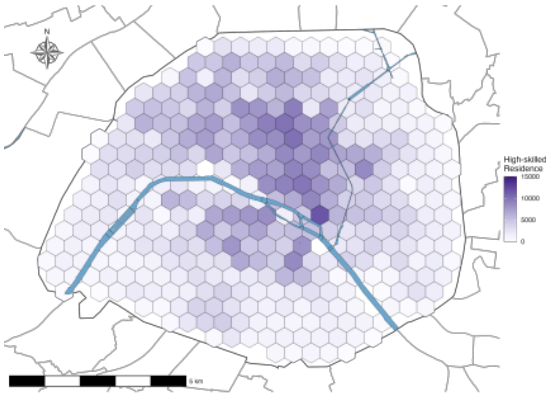
Figure 5
Parisian buildings in 1894


Notes: This figure illustrates the distribution of 4,600 buildings and areas in Paris in 1894, highlighting the urban footprint of the period. The map differentiates between permanent areas (such as monuments and gardens), depicted in dark red, and non-permanent areas (such as commercial and residential buildings), shown in light gray.

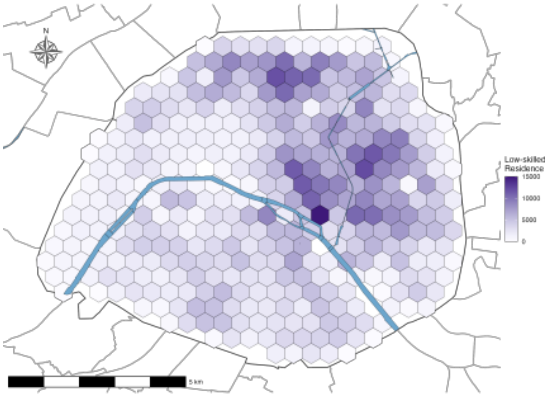
Second, I use city directories of Paris from 1896 provided by *ANR SoDuCo* to obtain a proxy for neighborhood workplace employment, as in Cosentino (2026). These city directories provide detailed information on firm locations and their respective industries.

To compute the number of low-skilled workers in each grid cell, I apply weights based on the worker intensity by industry of the 1872 industrial survey to transition from individual firm data to aggregate counts. For high-skilled workers, I assume that each firm is managed by an owner, who is considered a high-skilled worker. Panel (c) and Panel (d) of Figure 6 show the spatial distribution of high-skilled and low-skilled workers, respectively. Both are concentrated in the city center, where most economic activities took place during this period.

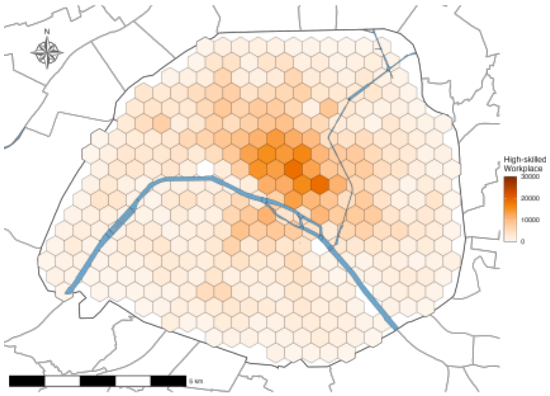
Figure 6
Spatial distribution of residents and workers by skill in 1896



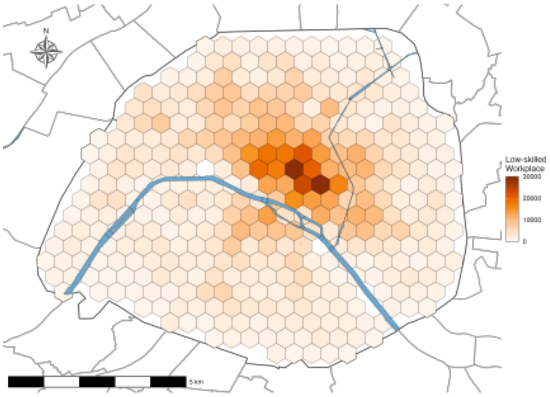

(a)
Residence employment high-skilled




(b)
Residence employment low-skilled




(c)
Workplace employment high-skilled




(d)
Workplace employment low-skilled

Notes: This figure illustrates the spatial distribution of residence and workplace employment by skill level in 1896. Panel (a) shows the residence employment for high-skilled workers, while Panel (b) depicts the residence employment for low-skilled workers. Residence employment at the grid cell level is estimated through the footprint and building heights. Panel (c) presents the workplace employment for high-skilled workers, and Panel (d) shows the workplace employment for low-skilled workers. Workplace employment at the grid cell level is estimated through city directories. The color intensity in each panel represents the density of employment, with darker shades indicating higher concentrations.

**Rents** Since I lack accurate and detailed data on rents, I use Paris fiscal matrices provided by the tax authorities to obtain the property taxes (*contribution foncière*) by neighborhoods. These data are crucial for recovering location-specific productivity and amenities in the structural analysis. I use the same previous building information to assign an average housing price per grid cell.

**Transport network** I combine several data sources in order to accurately construct the historical Parisian transport network. In 1896, workers mainly relied on three commuting modes: the *Petite Ceinture* (PC) railroad, omnibus, and tramways.

To begin with the *Petite Ceinture*, I rely on the *Association Sauvegarde Petite Ceinture* and OpenStreetMap to obtain detailed information on train stations and railroad locations. This circular line of 32 km was the metro precursor and offered the first railroad passenger service within the Paris intra-muros. Its average speed was 20 km/h, which is similar to the metro.

Finally, I construct omnibus and tramway networks by using the *Tableau des lignes de tramways* from the *Bibliothèque nationale de France*. These archival data provide information on the starting and ending stations of omnibuses and tramways, as well as connections between lines. Overall, the network consisted of 35 omnibuses and 19 tramways (see Figure A1). According to Passalacqua (2012), the average speed of omnibuses and tramways was 9 km/h.

Therefore, this transportation information provides me with a set of stations connected within a network $\Xi = \left[\xi^{\text{OT}}, \xi^{\text{PC}}\right]$, where subscripts OT and PC indicate the omnibuses/tramways, and the PC railroad, respectively. I assume a vector of travel time $\Omega = \left[\omega^{\text{OT}}, \omega^{\text{WA}}, \omega^{\text{PC}}\right]$, where subscripts WA indicate walking. I calibrate the walking speed to 5 km/h and consider a 5-minute waiting time $\gamma^C$ if there is a connection between different modes.

Overall, this rich transport network allows the computation of the commuting time $t_{ni} = t_{ni}(\Omega, \Xi, \gamma^C)$ between all locations considering different transport modes, and by using a least cost path method based on a Dijkstra algorithm. It aims to: (i) rationalize the initial spatial equilibrium in Paris, (ii) obtain a fixed network where an optimal metro will be added.

### 4.1.2 Historical context: metro planning

According to Cottereau (2004), the first legal initiative to create the Parisian metro came from a deliberation of the Seine General Council on November 10, 1871. Several goals were considered, such as unifying the Seine department through a railroad network, connecting the city center of Paris, and making interconnections between major train stations.[4] This was followed by a series of metro projects with different designs and objectives, but some of these metro projects shared a common philosophy. The London metro, built in 1863, had the effect of separating work from home by creating internal migration between suburbs and the city center, as shown by Heblich et al. (2020). In contrast to London, the Parisian metro was designed to connect residences to workplaces, rather than pushing them into the suburbs.

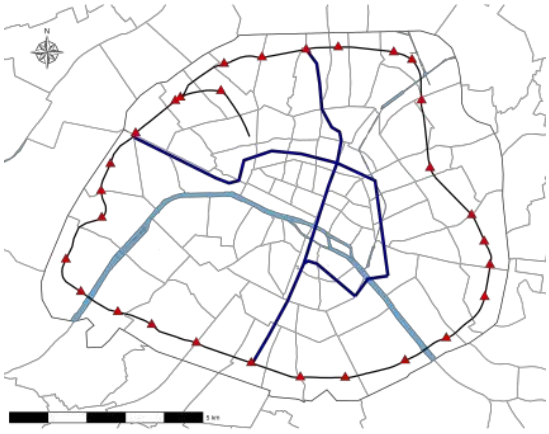
In 1872, Léon Say proposed a metro network that would use existing railways by adding two radial metro lines within Paris to complete the PC railroad (see Panel (a) of Figure 7). In 1876, another project by Alphand-Huet suggested diverting as much traffic as possible to the old network (see Panel (b) of Figure 7). In 1883, a metro project that did not rely on the existing transport network was proposed (see Panel (c) of Figure 7). Finally, the final metro project by Fulgence Bienvenüe was adopted in 1896 (see Panel (d) of Figure 7). This metro was designed to be self-sufficient and to guarantee access to all residents through a wide range of connections. In addition, it had to maintain the residential-workplace ties within Paris itself.

## 4.2 QUM extension

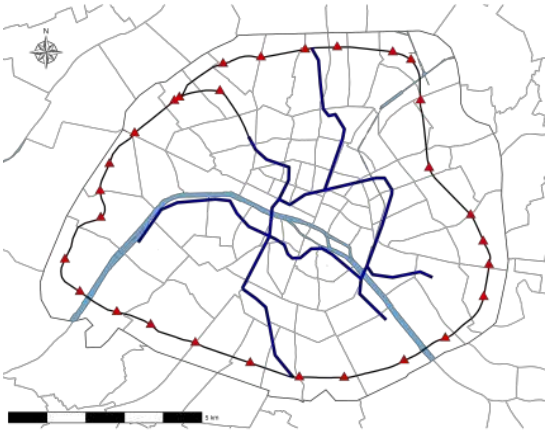
In this subsection, I adapt the canonical urban model developed in section 2 to match the historical Parisian setting and the empirical facts. I modify several components presented below: (i) worker heterogeneity with high-skilled and low-skilled workers, (ii) an open city setting, and (iii) worker substitution by considering a CES nested within the firm production function, with machinery capital.

[4] I do not include the major train stations in my analysis, as this would increase the computation time for commuting times between neighborhoods, and these stations are mainly intended to carry passengers from the suburbs.

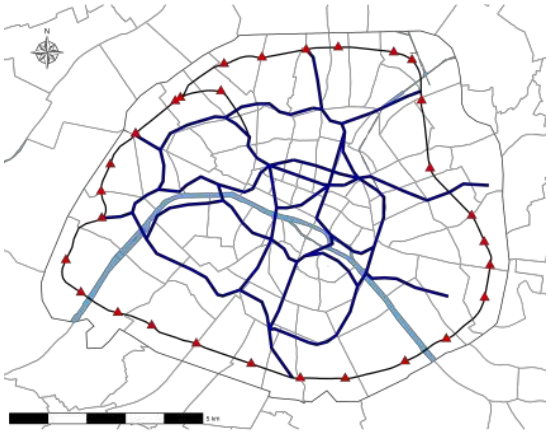
Figure 7
Metro Network Projects



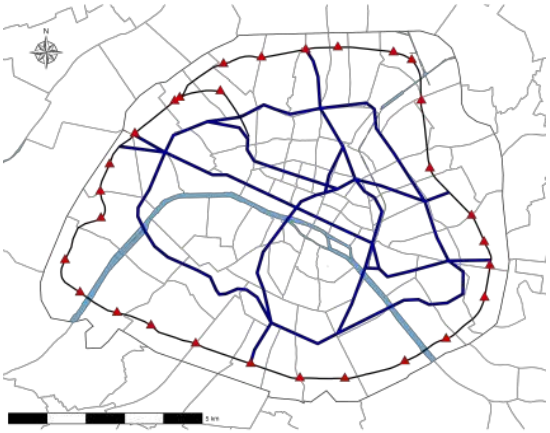
(a)
Léon Say Metro Network (1872)



(b)
Alphand-Huet Metro Network (1876)



(c)
Council Metro Network (1883)



(d)
F. Bienvenüe Metro Network (1896)

Notes: This figure illustrates various historical proposals for metro networks in Paris. Panel (a) depicts the Léon Say Metro Network proposed in 1872. Panel (b) presents the Alphand-Huet Metro Network from 1876. Panel (c) shows the Council Metro Network proposed in 1883. Panel (d) illustrates the F. Bienvenüe Metro Network from 1896, which was ultimately implemented.

### 4.2.1 Workers heterogeneity

As shown in subsection 4.1, the occupational sorting of that time in Paris highlights the need to consider both high- and low-skilled workers. This addition allows to rationalize factors influencing spatial sorting at equilibrium. I consider a mass of workers of type $g$ living and working within a city embedded in an economy. A worker $o$ of type $g$ choosing the residence-workplace pair $\{n, i\}$ derives the following indirect utility,

$$U_{ni,g}(o) = \frac{B_{n,g} w_{i,g}}{d_{ni} P_n^{1-\beta_g} Q_n^{\beta_g}} \epsilon_{ni,g}(o), \tag{17}$$

where only amenities ($B_{n,g}$), the wage ($w_{i,g}$), the housing share ($\beta_g$), and the idiosyncratic shock ($\epsilon_{ni,g}(o)$) differ from equation (1), and are type-specific. Hence, the same properties of the Fréchet distribution can be applied to recover the type-specific commuting probability $\lambda_{ni,g}$, the type-specific residential probability $\lambda^R_{n,g}$, the type-specific workplace probability $\lambda^L_{i,g}$, and the type-specific conditional commuting probability $\lambda^R_{ni|n,g}$ (see appendix B.2.1 for more details).[5]

### 4.2.2 Open city setting

I consider an open city setting where workers from the outside economy are free to enter the city. This setup reflects the historical context in which the population of Paris doubled during the 19th century, as documented by Cosentino (2026). Within this open city framework, the expected utility of a worker of type $g$ residing and working in Paris can be expressed as,

$$\mathbb{U}_g \left( \frac{R_{\mathbb{N}_g}}{R_{\mathbb{M}_g}} \right)^{1/\phi} = \overline{U}_g, \tag{18}$$

with $\overline{U}_g \equiv \delta_g \left[ \sum_{k \in \mathbb{N}} \sum_{l \in \mathbb{N}} \Phi_{kl,g} \right]^{1/\epsilon_g}$ the type-specific expected utility that a worker derives from living and working in Paris (which is a type-specific version of equation (5)), $\phi$ the migration elasticity, $\mathbb{U}_g$ the type-specific expected utility in the wider economy, and $R_{\mathbb{N}_g}/R_{\mathbb{M}_g}$ the share of workers of type $g$ living in the economy choosing a pair of locations within Paris. To satisfy the population mobility condition, a relatively higher expected utility within the city compared to the economy implies a greater proportion of workers

[5] The extension that allows locations to have different consumption prices by relaxing the assumption of free trade is not implemented in order to maintain tractability and limit computational complexity.

opting for Paris, as expressed by the ratio,

$$\frac{R_{\mathbb{N}_g}}{R_{\mathbb{M}_g}} = \left(\frac{\overline{U}_g}{\mathbb{U}_g}\right)^{\phi}. \tag{19}$$

### 4.2.3 Firms

Additionally, I consider an extension of the firm production function by incorporating a Cobb-Douglas framework that includes machinery capital and a CES for labor supply. These two features reflect the extensive use of machinery in Parisian production during 1896 and accommodate the diverse skill levels of the workforce at that time. Therefore, the representative firm produces the final good with Cobb-Douglas technology under constant returns to scale,

$$Y_i = A_i \left(\frac{L_i}{\alpha_L}\right)^{\alpha_L} \left(\frac{H_i^L}{\alpha_F}\right)^{\alpha_F} \left(\frac{M_i}{\alpha_M}\right)^{\alpha_M}, \tag{20}$$

where $A_i$ represents the productivity in workplace $i$, $L_i$ is the total labor supply in workplace $i$, $H_i^L$ is the housing used commercially in workplace $i$, and $M_i$ is the machinery capital used in workplace $i$. The total labor supply is expressed as $L_i = \left(\sum_g a_{i,g} L_{i,g}^{\rho}\right)^{1/\rho}$ and takes the form of a constant elasticity of substitution (CES) function, where $a_{i,g}$ is the group-specific skill intensity and $\rho$ is the parameter governing substitution. The shares of labor, machinery, and housing dedicated to the production function are $\alpha_L$, $\alpha_M$, and $\alpha_F$, respectively.

The first-order conditions of the firm's profit with respect to type-specific labor, housing, and machinery are given by,

$$w_{i,g} = \alpha_L^{(1-\alpha_L)}\, a_{i,g}\, L_{i,g}^{\rho-1}\, A_i\, L_i^{\alpha_L-\rho} \left(\frac{H_i^L}{\alpha_F}\right)^{\alpha_F} \left(\frac{M_i}{\alpha_M}\right)^{\alpha_M}, \tag{21}$$

$$Q_i = \alpha_F^{(1-\alpha_F)}\, A_i \left(\frac{L_i}{\alpha_L}\right)^{\alpha_L} \left(H_i^L\right)^{\alpha_F-1} \left(\frac{M_i}{\alpha_M}\right)^{\alpha_M}, \tag{22}$$

$$r_i = \alpha_M^{(1-\alpha_M)}\, A_i \left(\frac{L_i}{\alpha_L}\right)^{\alpha_L} \left(\frac{H_i^L}{\alpha_F}\right)^{\alpha_F} (M_i)^{\alpha_M-1}. \tag{23}$$

## 4.3 Model inversion

In contrast to the numerical example presented above, the distribution of housing supply and fundamentals (amenities by skill, as well as productivity) are not known. Therefore, the model needs to be inverted step by step in order to rationalize the spatial equilibrium with observed residence, workplace, and rent data.

### 4.3.1 Wages

Using the type-specific conditional commuting probability combined with the labor market clearing condition,

$$L_{i,g} = \sum_{n \in \mathbb{N}} \frac{(w_{i,g}/d_{ni})^{\epsilon_g}}{\sum_{l \in \mathbb{N}} (w_{l,g}/d_{nl})^{\epsilon_g}} R_{n,g}, \tag{24}$$

there exists a unique vector of wages that solves the labor market, given a calibration of $\epsilon_g$, the observed vectors of residence employment, workplace employment, and the commuting cost parametrization.

### 4.3.2 Skill intensity

Using type-specific wages, the first-order condition of the firm equation (21), and the condition $a_{i,L} + a_{i,H} = 1$, the intensity in low-skilled workers $a_{i,L}$ can be recovered,

$$\frac{1 - a_{i,L}}{a_{i,L}} = \frac{w_{i,H}}{w_{i,L}} \left( \frac{L_{i,L}}{L_{i,H}} \right)^{\rho - 1}, \tag{25}$$

where $w_{i,H}$ is the high-skilled wage in workplace $i$, $w_{i,L}$ is the low-skilled wage in workplace $i$, $L_{i,H}$ is the total high-skilled labor employed in workplace $i$, and $L_{i,L}$ is the total low-skilled labor employed in workplace $i$.

### 4.3.3 Productivity

From the zero profit condition due to free entry, and the first-order conditions of the firm's profit maximization, it is possible to express the location productivity as,

$$A_i = W_i^{\alpha_L} Q_i^{\alpha_F} r_i^{\alpha_M}, \tag{26}$$

which is a function of aggregate wages $W_i \equiv \left( \sum_g a_{i,g}^{\frac{1}{1-\rho}} w_{i,g}^{\frac{\rho}{\rho-1}} \right)^{\frac{\rho-1}{\rho}}$ (see appendix B.2.2 for more details), the price of machinery capital $r_i$, and rents $Q_i$. I make the assumption that machinery capital is freely traded in the city and normalize the price of the machinery capital $r_i = 1$ for all workplaces $i$. Intuitively, higher costs (wages, rents) imply a higher productivity in order for the zero profit condition to hold.

### 4.3.4 Amenities

Combining the residential probability equation (3) and expected utility equation (18), it is possible to derive type-specific amenities as,

$$B_{n,g} = \frac{(\lambda_{n,g}^R)^{1/\epsilon_g} Q_n^{\beta_g}}{\left( \sum_{i \in \mathbb{N}} (w_{i,g}/d_{ni})^{\epsilon_g} \right)^{1/\epsilon_g}}, \tag{27}$$

where amenities are considered as residuals explaining residential choices in this commuting model. For a given residential market access (denominator), and a given rent, a higher share of workers living in location $n$ implies higher amenities.

### 4.3.5 Rateable values

Finally, the housing market clears and rateable values ($\mathbb{Q}_n$) are determined by the total housing demand by residents and firms, and rents,

$$\mathbb{Q}_n = Q_n(H_n^R + H_n^L) = \sum_g \beta_g \nu_{n,g} R_{n,g} + \frac{\alpha_F}{\alpha_L} w_{n,g} \frac{L_n^{\rho} L_{n,g}^{1-\rho}}{a_{n,g}}, \tag{28}$$

where $H_n^R$ and $H_n^L$ are the residential and commercial housing demand in $n$, respectively, and $\nu_{n,g} \equiv \sum_{i \in \mathbb{N}} \lambda_{ni|n,g}^R w_{i,g}$ is the type-specific expected income from residents living in location $n$. The housing demand of firms in workplace $i$ is derived from the first-order conditions equations (21) and (22).

### 4.3.6 Equilibrium

**Spatial Equilibrium Definition:** *Given the structural parameters* $\{\beta_g, \alpha_M, \alpha_F, \alpha_L, \epsilon_g, \kappa, \rho\}$, *and a commuting cost parametrization, a spatial equilibrium is defined with observed vectors* $\{L_{i,g}, R_{n,g}, d_{ni}, Q_n\}$, *and unobserved vectors* $\{w_{i,g}, a_{i,g}, A_i, B_{n,g}, \mathbb{Q}_n\}$ *such that: i)*

*commuting market clears (11); ii) skill-intensity (25) and zero profit condition (26) are satisfied; iii) population mobility is satisfied (27); and iv) housing market clears (28).*

## 4.4 Model Quantification and Spatial Equilibrium

### 4.4.1 Parameters calibration

**Internally calibrated** I internally calibrate the commuting cost elasticity ($\kappa$) and the housing supply elasticity ($1 - \mu/\mu$). Using bilateral flow data from the PC railroad prior to the metro's implementation, Cosentino (2026) estimates a commuting cost elasticity of 0.0204. While this value is relatively high compared to the literature, it is robust to using alternative specifications. Then, Cosentino (2026) also estimates a housing supply elasticity of 1.76 with rateable values and housing units data.

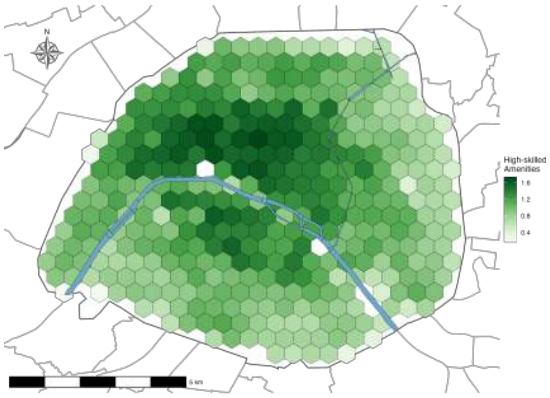
**Externally calibrated** The remaining parameters are calibrated using values from the existing literature. The housing expenditure share ($\beta_g$) is set to 0.3 for low-skilled and 0.2 for high-skilled workers, reflecting the empirical observation that higher-income households spend a smaller proportion of their income on housing (Tsivanidis, 2026; Redding and Sturm, 2024; Weiwu, 2024). The Fréchet dispersion parameter ($\epsilon_g$) is calibrated to 6 for low-skilled and 4 for high-skilled workers, following estimates in Redding and Sturm (2024), matching the fact that the poor are more sensitive to location characteristics than the rich. The production function parameters are drawn from Heblich et al. (2020), who study London during a comparable historical period: the shares are set to 0.2 for housing ($\alpha_F$), 0.2 for machinery capital ($\alpha_M$), and 0.6 for labor ($\alpha_L$). The elasticity of substitution between skill groups ($\rho$) is set to 0.3, consistent with the estimates provided by Card (2009), which are largely used in the literature (e.g., Tsivanidis, 2026; Herzog, 2024; Redding and Sturm, 2024; Weiwu, 2024). Finally, the migration elasticity $\phi$ is calibrated to 3 (Monte et al., 2018; Bryan and Morten, 2019; Bilal, 2023).

### 4.4.2 Results from spatial equilibrium

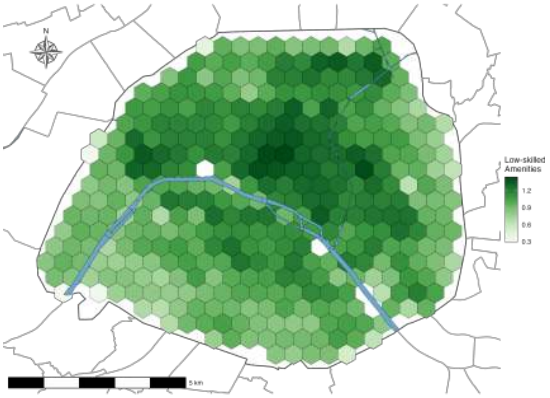
**Fundamentals** Figure 8 shows the results of the spatial equilibrium, with spatial distributions of high-skilled amenities in Panel (a), low-skilled amenities in Panel (b), and productivity in Panel (c). Overall, these fundamentals recovered by the model are higher in the middle and west of Paris. For amenities, the pattern is consistent with the east-side

story described in Heblich et al. (2021), where high-skilled workers value the west side of Paris relatively more than low-skilled workers, and sort themselves there.

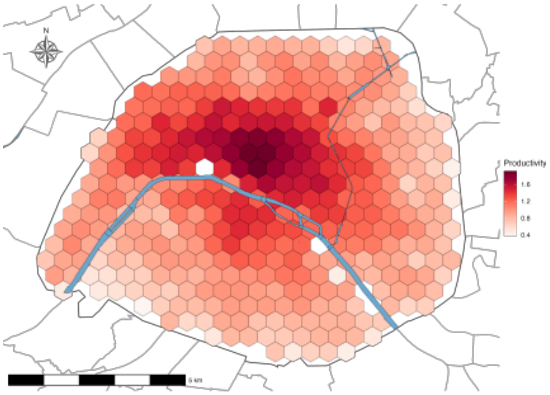
Figure 8
Fundamentals distribution




(a)
High-skilled amenities




(b)
Low-skilled amenities

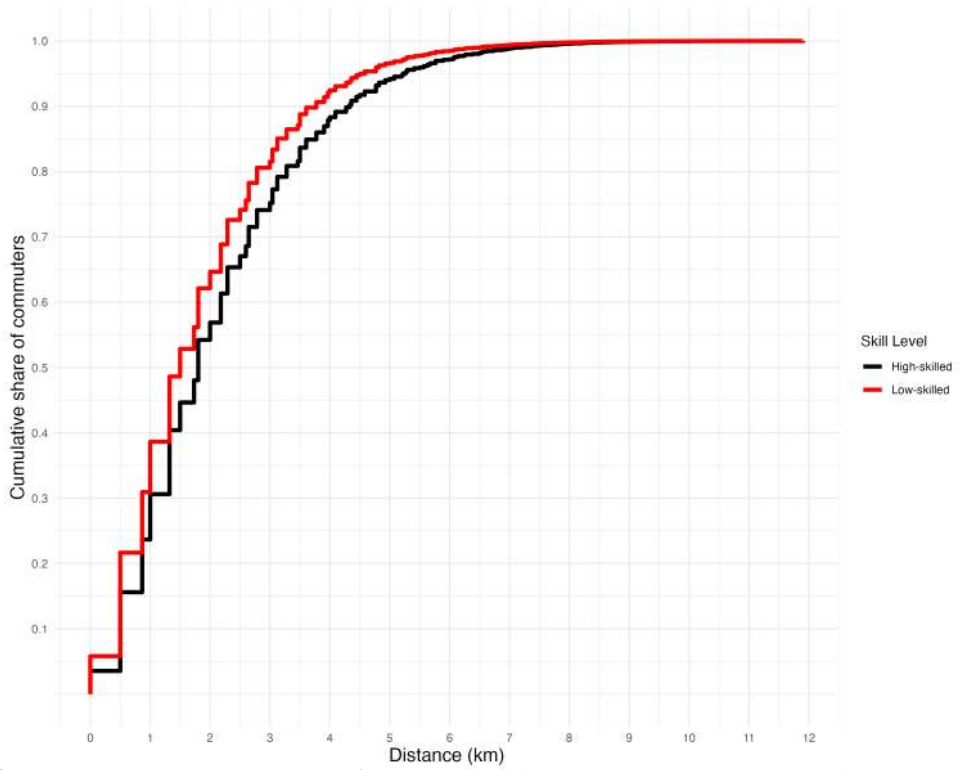



(c)
Productivity

Notes: This figure illustrates the distribution of amenities and productivity fundamentals, resulting from a heterogeneous agents QUM and rationalizing within-city spatial equilibrium in 1896. Panel (a) depicts the distribution of high-skilled amenities. Panel (b) shows the distribution of low-skilled amenities. Panel (c) presents the distribution of productivity across the city. Darker shades indicate areas with higher productivity/amenities.

**Commuting** Regarding commuting patterns predicted by the quantitative urban model, Figure 9 shows the cumulative distribution of commuting with respect to distance in the initial equilibrium by skill level. By sorting each residence-workplace pair based on Euclidean distance and plotting the cumulative share of commuters by skill level, it aims to show how many workers commute less than each distance. Based on the initial equilibrium, high-skilled workers commute more and further than low-skilled workers. More than 65% of Parisian low-skilled workers commute less than 2 km, compared to 57% for high-skilled workers.

Figure 9
Cumulative commuting distance distributions




Notes: This figure shows the cumulative distribution of commuters by skill level in the initial equilibrium. Each residence-workplace pair is sorted based on Euclidean distance. The black line represents high-skilled workers, and the red line represents low-skilled workers. The x-axis indicates the distance in kilometers, and the y-axis shows the cumulative share of commuters.

## 4.5 Optimal Metro Design

### 4.5.1 Optimization Environment

As before, the planner chooses the metro network $M$ consisting of $L_M$ metro lines and $S_M$ selected metro stations that maximizes an economic objective $W(M;\theta)$.

In this application to Paris, I consider several economic criteria $W$, and I constrain the starting and ending metro stations of each line to be connected with the PC railroad.[6] The baseline $W$ is defined following the Henry George Theorem (George, 1879), implying that public infrastructure gains are valued through the land market.

Hence, the economic impact factor $W$ is defined as the ratio between the net present value of changes in rateable values considering an infinite lifetime, and metro construction costs,

$$W(M;\theta) \equiv \frac{\sum_{n\in\mathbb{N}} \Delta\mathbb{Q}_n(M;\theta)}{r} \times \frac{1}{\text{Cost}(M)}, \tag{29}$$

where $\sum_{n\in\mathbb{N}} \Delta\mathbb{Q}_n(M;\theta)$ is the total change in rateable values occurring in the city fol-

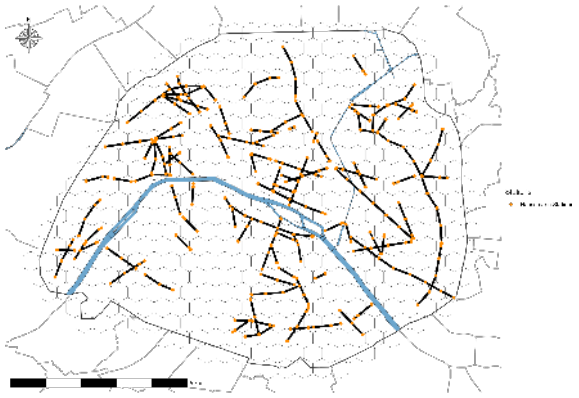
[6]This provides a natural way to start with existing transport infrastructure and, importantly, decreases the optimization space.

lowing the implementation of the metro $M$, $r$ is a discount rate calibrated to 5%, and Cost($M$) is the cost of infrastructure $M$.[7]

**Construction technology: cut-and-cover vs. deep-bore tunneling** The cost of building a metro line in late 19th-century Paris was not uniform across the city. Two distinct construction technologies were available, each with very different cost implications. The first, *cut-and-cover* (or open-trench), consists of digging a trench from the surface, laying the tunnel structure, and backfilling. This method is fast and cheap but requires a wide, unobstructed street overhead. The second, *deep-bore tunneling*, involves drilling under the existing urban fabric using a tunnel boring machine; it avoids surface disruption but is substantially more expensive. A critical feature of 19th-century Paris is that Baron Haussmann's urban renewal programme (1853-1870) had carved wide, straight boulevards through the old medieval street network. These Haussmann boulevards made cut-and-cover construction directly feasible in dense Paris, dramatically lowering the cost of tunneling under them relative to digging elsewhere. Figure 10 illustrates the cut-and-cover construction method (Panel (a)) and shows the set of potential metro stations and segments eligible for the cheaper cut-and-cover technology since they are lying along Haussmann boulevards (Panel (b)).

Figure 10
Cut-and-cover construction method and Haussmann boulevard stations

(a)
Cut-and-cover construction method

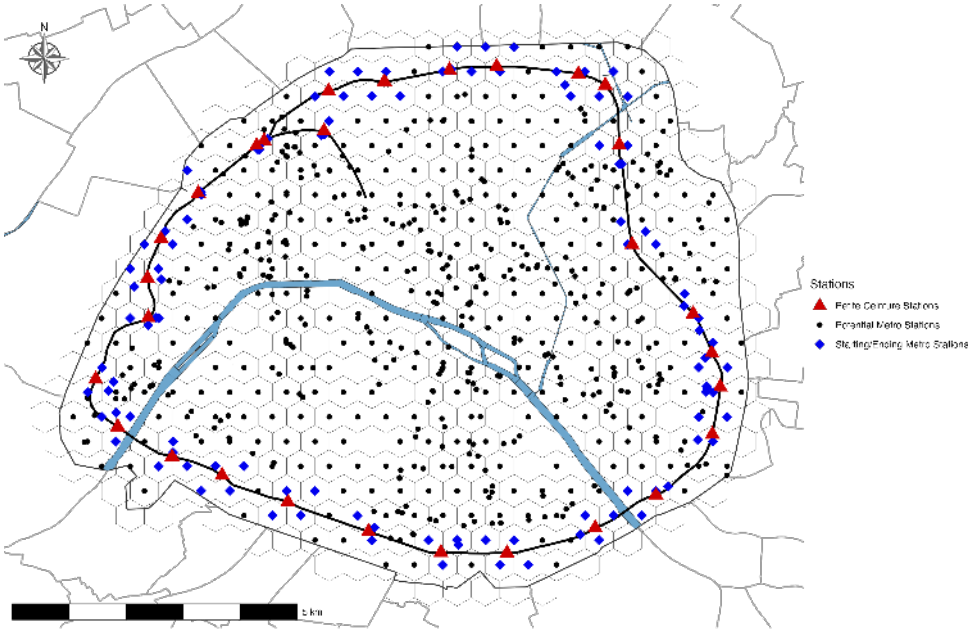


(b)
Haussmann stations and segments

Notes: Panel (a) illustrates the cut-and-cover (open-trench) construction method. A trench is dug from the surface, the tunnel structure is built inside, and the street is restored. The method requires a wide street overhead and is substantially cheaper than deep-bore tunneling. Panel (b) shows all potential metro stations (dots) and segments (lines) eligible for cut-and-cover construction, i.e. those located along Haussmann boulevards. Stations and segments outside these boulevards require the more expensive deep-bore tunneling technology.

[7] Rateable values equal the sum of prices times quantities for residential floor space and commercial floor space. I calibrate the baseline rateable value by using the per head revenue of the 1872 industrial survey.

Figure 11
Potential metro stations




Notes: This figure shows all potential metro stations. Red triangles represent *Petite Ceinture* stations, blue diamonds indicate starting/ending metro stations, black dots denote potential metro stations, and the black line represents the *Petite Ceinture* railroad.

The construction cost function therefore assigns a technology type: cut-and-cover or deep-bore, depending on whether each inter-station link runs along a Haussmann boulevard. I make the assumption that $\text{Cost}(M)$ is a linear function of the total number of unique stations and the total length of the metro,

$$\text{Cost}(M) = \underbrace{\vartheta_0}_{\text{fixed}} + \sum_{s \in S_M} \underbrace{\vartheta_1 Q_n}_{\text{per station}} + \sum_{c \in C_M} \underbrace{\vartheta^{\text{cc}}}_{\text{cut-and-cover}} L_c + \sum_{t \in T_M} \underbrace{\vartheta^{\text{tun}}}_{\text{tunneling}} L_t + \sum_{x \in X_M} \underbrace{\delta^{\text{Seine}}}_{\text{Seine crossing}}, \tag{30}$$

where $\vartheta_0$ is a fixed cost (representing, for example, technology acquisition), $\vartheta_1$ is the average cost per unique metro station, $\vartheta^{\text{cc}}$ is the per-kilometer cost of cut-and-cover construction, and $\vartheta^{\text{tun}}$ is the per-kilometer cost of deep-bore tunneling. $C_M$ denotes the set of segments eligible for cut-and-cover (along Haussmann boulevards) and $T_M$ the set of segments requiring tunneling (elsewhere). $X_M$ denotes segments crossing the Seine, which incur an additional fixed surcharge $\delta^{\text{Seine}}$ per crossing. $Q_n$ is the average rent in location $n$ (with a geometric mean equal to 1), which acts as a local shifter allowing for heterogeneity in construction costs across space and captures the fact that implementing a metro station in a rich location is more costly (representing, for example, land acquisition

or building destruction) than in a poor location.

To estimate the construction cost parameters $\vartheta = \{\vartheta_0, \vartheta_1, \vartheta^{\text{cc}}, \vartheta^{\text{tun}}, \delta^{\text{Seine}}\}$, I rely on historical cost data from Heblich et al. (2020) and used for the London Metropolitan area reported: £330,000 per mile for cut-and-cover and £500,000 per mile for deep-bore tunneling, as of 1921. Converting at a rate of £1 = 25 francs, netting out the fixed-cost and the station-cost, and adjusting for mileage (1 mile $\approx$ 1.609 km) yields $\vartheta^{\text{cc}} \approx$ 2,820,000 francs per km and $\vartheta^{\text{tun}} \approx$ 4,272,000 francs per km. The fixed cost $\vartheta_0$ = 46,136,926 francs and per-station cost $\vartheta_1$ = 512,632.5 francs are derived from the total Fulgence Bienvenüe metro budget, with 15% allocated to fixed acquisition costs and 30% to stations; the Seine-crossing surcharge $\delta^{\text{Seine}}$ is calibrated to bridge the gap between the cut-and-cover and deep-bore per-km rates (see Appendix C.1 for full details).

To choose a metro network $M$, I divide the Parisian space into grid cells, each 500 meters apart.[8] From the information in Figure 5, I remove the metro stations that lie within a permanent area (for example, a monument). I then augment this set by adding all stations located along Haussmann boulevards, yielding a total of 711 potential metro stations $\mathcal{S}$, as illustrated in Figure 11. Then, I select all stations that are within 500 meters of the PC railroad to construct the sample of potential starting/ending stations, resulting in 103 potential metro stations connected to the PC. Again, the optimization problem is discrete and high-dimensional, without having an analytical solution. Consequently, I use the same simulated annealing algorithm previously developed in section 3 with the same optimization process. I adapt my perturbation function $\Psi$ to my context and consider 3 probabilistic scenarios to create a candidate metro network $M'$:

(i) Deviation: randomly selects an existing line, maintains its departure and arrival stations, and makes it pass through a randomly selected point within the grid.

(ii) Add: creates a new line by randomly picking starting/ending metro stations, or randomly selects an existing line, deletes it, and a new line is created by randomly picking starting/ending metro stations.

(iii) Remove: randomly selects an existing line and deletes it.

After this perturbation, which gives a candidate network $M'$, the matrix of commuting cost is updated by computing bilateral commuting time between locations, taking into

[8] This corresponds to the average length between stations of the actual Parisian metro network.

account fixed transport modes (omnibus, tramways, and the PC railroad), which in turn allows the computation of a candidate economic impact factor $W'(M';\theta)$ through the quantitative urban model.

### 4.5.2 Need for speed

Since the number of steps $K$ needs to be high in order to widely explore the space of metro networks, each iteration must be optimized to reduce computation time. Therefore, I combine R software to handle the perturbation function creating a candidate network $M'$ and an associated commuting costs matrix, and Julia software to efficiently solve the new spatial equilibrium and welfare using exact-hat algebra (Dekle et al., 2007). To speed up the process, I use the *JuliaCall* package to create a permanent Julia session in R, while parallelizing the computation.

### 4.5.3 Simulation Results

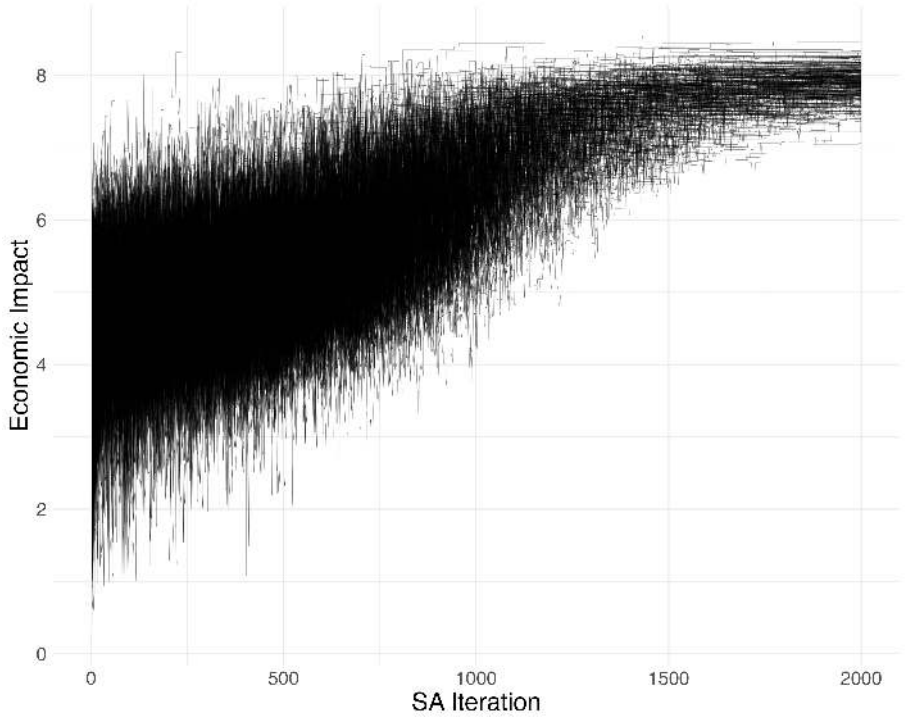
Using the vector of structural parameters $\theta$, and calibrated metro costs ($\vartheta = \{\vartheta_0, \vartheta_1, \vartheta^{\text{cc}}, \vartheta^{\text{tun}}, \delta^{\text{Seine}}\}$, I set the number of steps $K$ to 2,000 and consider 500 independent simulated annealing runs to sample from $\pi$. I consider a minimum number of metro lines of 2, and a maximum of 10. I set the inverse temperature parameter $\sigma_K = 30$, and gradually increase it from 1.

**Convergence and examples** Figure 12 shows the convergence of all runs. Overall, they depict a consistent pattern with a progressive increase towards similar values, with the distribution of final welfare values concentrated between 7 and 9. Finally, Figure 13 shows sampled metro networks with similar economic impact factors but different designs.

**Size of metro network and economic impact factor** The simulations provide 500 selected metro networks sampled from $\pi$ varying in the number of metro stations $S$, residence and workplace employments across locations, total population, commuting patterns, as well as economic impact factor $W$.

Panel (a) of Figure 14 shows the distribution of the economic impact factor from sampled metro networks. They are centered around 7.89, with a minimum of 7.06 and a maximum of 8.46. This range is comparable in magnitude to the findings of Heblich

Figure 12
Simulated annealing algorithm convergence



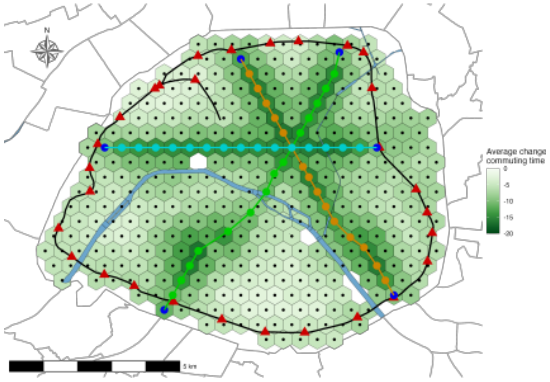

Notes: This figure shows the convergence of the simulated annealing algorithm over 2,000 iterations, with each line representing a different run.

et al. (2020) for the London metro. Panel (b) of Figure 14 shows the distribution of the total number of metro stations from sampled metro networks. While there is a peak at 51 metro stations, some metro networks are extremely different, with a minimum of 30 stations and a maximum of 82 stations. Panel (c) of Figure 14 displays the correlation between the number of metro stations $S$ and economic impact $W$ for the sampled metro networks.[9] Since $W$ is a function of the metro construction cost, adding other metro stations does not necessarily lead to a higher economic impact factor. However, there is still heterogeneity between metro networks with the same number of stations, as shown by the R-squared of 0.007. It suggests that urban transport infrastructure design matters since metro networks display heterogeneity in economic impact factor for a given number of stations.

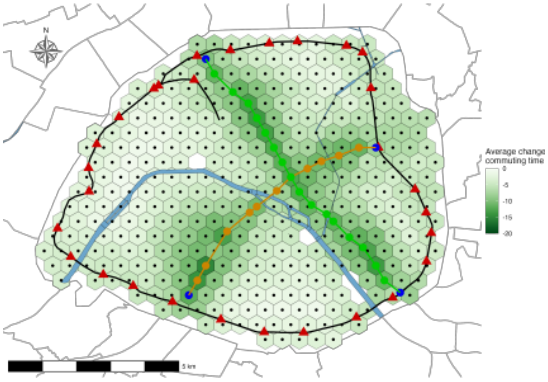
**City, location, and commuting characteristics** Over the 500 sampled metro networks, the total production and population increase by, on average, 33.9% and 22.1%, respectively. Figure 15 shows the relative increase in population by skill. Both have similar inflows within the city with medians around 1.22 and 1.24 for high-skilled and

[9]Figure A3 shows the relationship for all the iterations and not only the final sampled metro networks.

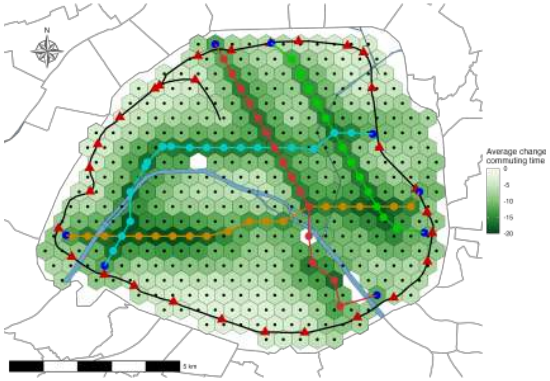
Figure 13
Optimal urban transport example



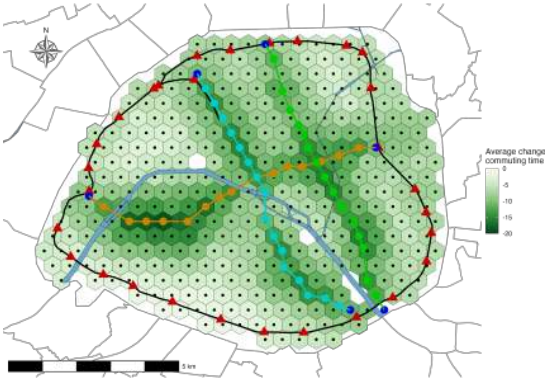
(a)
Impact factor: 8.1, $S = 51$ stations



(b)
Impact factor: 8.1, $S = 31$ stations



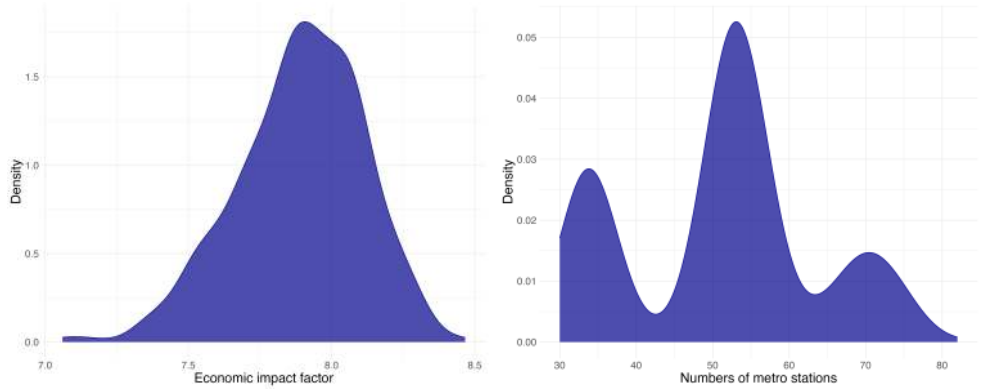
(c)
Impact factor: 7.4, $S = 74$ stations



(d)
Impact factor: 7.4, $S = 53$ stations

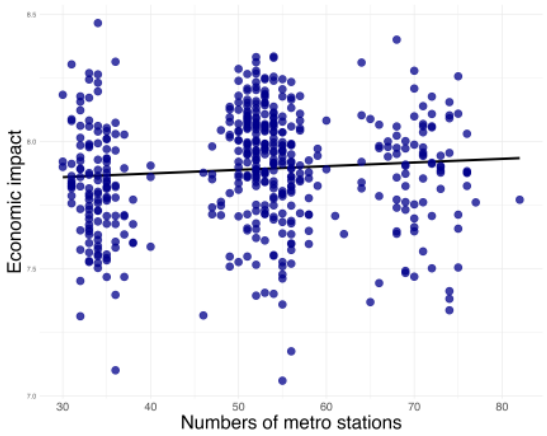
Notes: Each subfigure represents an example of an optimal urban transport, showing the number of train stations ($S$) and the associated impact factor. The economic impact factor is defined as the ratio between the net present value of change in rateable values (considering an infinite lifetime and a 5% discount rate) and the metro construction costs. Darker shades indicate areas with a higher decrease in average commuting time with respect to the initial. Colored dots indicate metro stations, with a single color for each line. Red triangles represent *Petite Ceinture* stations, and the black line represents the *Petite Ceinture* railroad.

Figure 14
Size and economic impact of sampled metro networks




(a) Distribution of economic impact

(b) Metro size

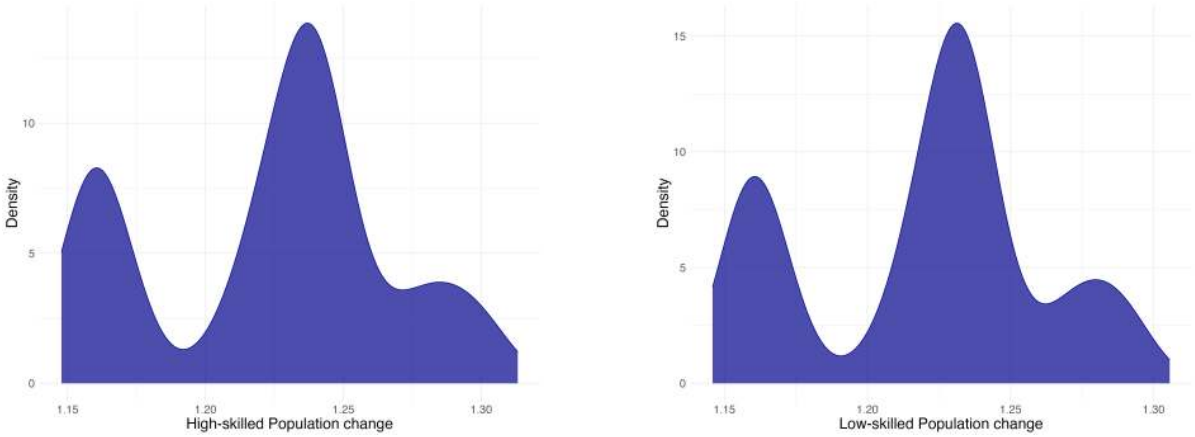



(c) Metro size and economic impact

Notes: Panel (a) depicts the distribution of the economic impact factor from the SA runs. Panel (b) shows the distribution of the total number of metro stations in the SA runs. Panel (c) illustrates the relationship between the number of metro stations and the economic impact, with each blue dot representing a SA run. This figure shows results from 500 independent runs of the SA algorithm with 2,000 iterations. The economic impact factor is defined as the ratio between the net present value of change in rateable values (considering an infinite lifetime and a 5% discount rate) and the metro construction costs.

low-skilled workers, respectively.

Figure 15
Population distribution by skill level

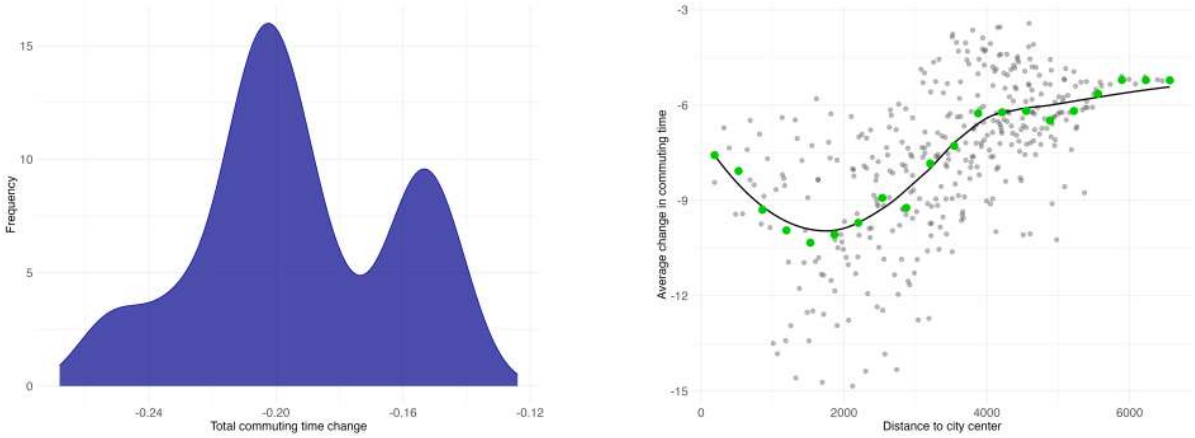



(a) High-skilled (b) Low-skilled

Notes: This figure shows the relative changes in population distribution by skill level with respect to the initial. Panel (a) depicts the distribution of population change for high-skilled workers. Panel (b) shows the distribution of population change for low-skilled workers.

Regarding savings in commuting time, Panel (a) of Figure 16 shows the distribution of change in total commuting time over the 500 sampled metro networks. The average total decrease in commuting time within the city is 20.4%. However, this aggregate value hides large heterogeneity. Panel (b) of Figure 16 shows the change in average commuting time by location with respect to distance from the city center. While all locations experience a decrease in average commuting time, those that are closer to the city center experience a relatively higher decrease in commuting time.

Figure 16
Commuting time change



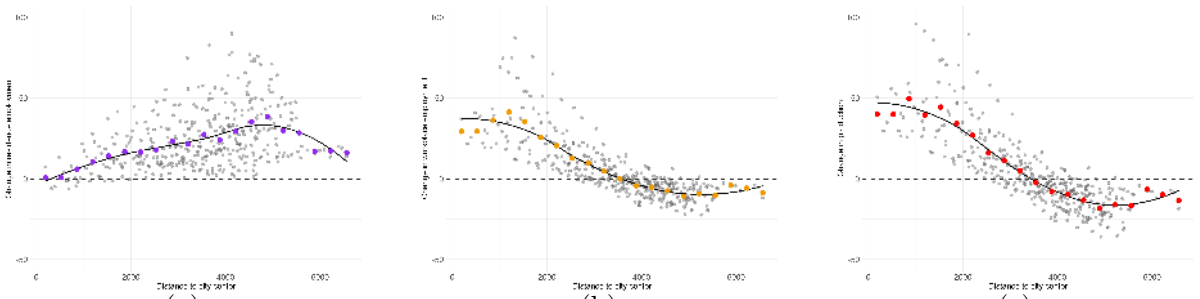

(a) Total commuting time change (b) Location commuting time change

Notes: This figure illustrates changes in commuting time with respect to the initial. Panel (a) shows the relative changes in total commuting time. Panel (b) depicts the average change in percentage points of each location with respect to distance to the city center.

Figure 17 shows the average change of economic activities for each location with re-

spect to distance from the city center over the 500 sampled metro networks (see Figure A4 for maps in appendix). While Panel (a) displays that locations close to the periphery experience higher growth in residence employment, Panels (b) and (c) show that locations close to the city center experience higher growth in workplace employment and production, respectively. This aligns with Heblich et al. (2020), who show that implementing a railroad connecting the city center to the suburbs creates a separation between workplace and residence.

Figure 17
Change in economic activities with respect to city center

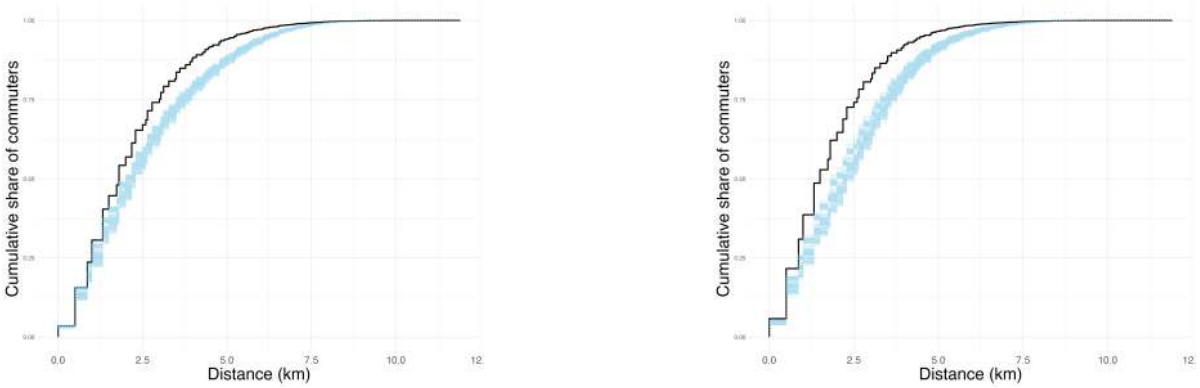


(a) Residence employment (b) Workplace employment (c) Production

Notes: This figure shows changes in percentage points of location economic activities with respect to distance from the city center. Panel (a) depicts the change in residence employment. Panel (b) shows the change in workplace employment. Panel (c) illustrates the change in production.

Figure 18 examines the change in commuting patterns by plotting the cumulative distribution of commuting by skill level over distance (as in Figure 9) following the implementation of the 500 sampled metro networks. Overall, these infrastructure changes create similar disconnections between workplace and residence, with a decline in the share of commuters working in their place of residence and a rise in long-distance commuting for both skill groups.

Figure 18
Commuting decisions by level of workers




(a) High-skilled workers

(b) Low-skilled workers

Notes: This figure shows the cumulative distribution of commuters by skill level. Each blue line represents one cumulative distribution of commuters resulting from a SA run, while the black line represents the initial cumulative distribution of commuters. Panel (a) depicts the distribution for high-skilled workers, and Panel (b) shows the distribution for low-skilled workers.

**Strategic locations** Previous Figure 13 shows that different metro designs can achieve similar economic criteria. However, some stations may be more strategic and appear more frequently than others across the 500 sampled metro networks. To explore this pattern, Figure 19 shows the frequency of appearance of each metro station across the 500 sampled metro networks. Metro stations within the city center of Paris appear to be strategic locations as they depict a high level of frequency. Consequently, simulations of the simulated annealing algorithm converge toward this design where the lines tend to pass through the city center, regardless of its initial design and the starting/ending metro stations. However, this result could be mechanical, due to the centrality of these stations in the network.

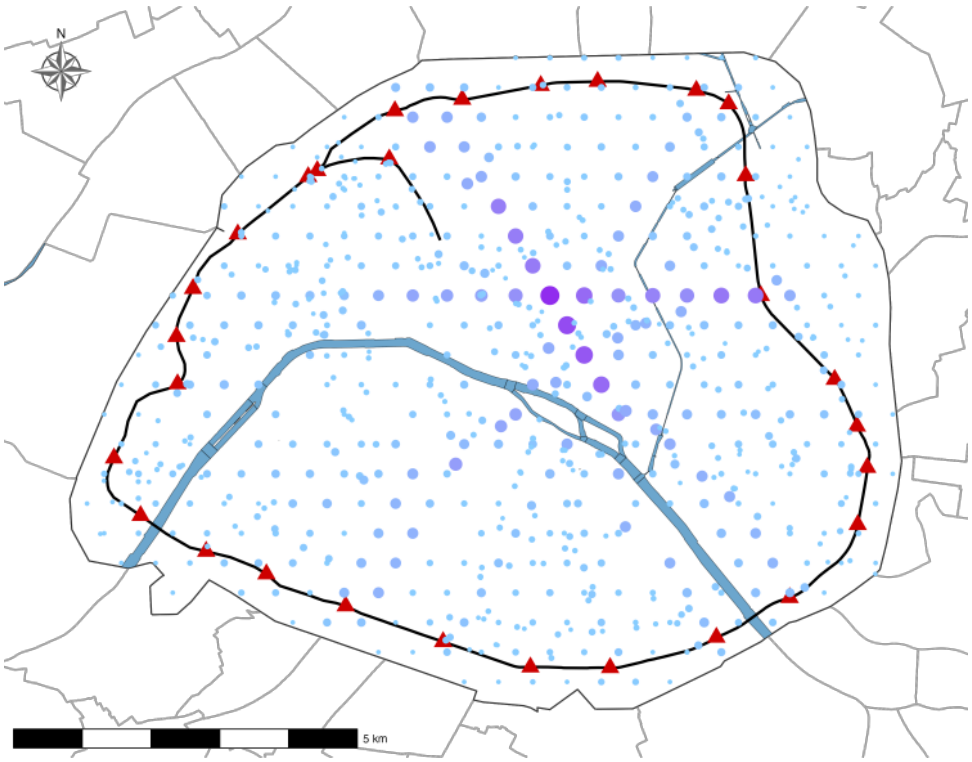
To further investigate which factors explain strategic metro stations, I implement the following OLS regression,

$$\mathrm{F}_m = \beta_0 + \mathbf{X}_m\boldsymbol{\beta} + \varepsilon_m, \tag{31}$$

where $\mathrm{F}_m$ is the frequency of a metro station $m$, $\mathbf{X}_m$ is a vector of location characteristics associated with metro station $m$ including: $A_m$ the (log) productivity (see Panel (c) of Figure 8), $B_{m,g}$ the (log) type-specific amenities (see Panel (a) and Panel (b) of Figure 8), $\mathrm{Costs}_m$ the (log) construction costs, and the (log) distance to the city center (defined as Notre-Dame, which is at the center of Paris) of metro station $m$.

Table 1 shows the results. Each column reports a separate regression. The first four columns investigate the effect of productivity, low-amenities, high-amenities, and dis-

Figure 19
Metro stations frequency in sampled metro networks




Notes: This figure illustrates the frequency of metro stations occurring in 500 simulated annealing runs. The size of each station is proportional to its frequency, with larger dots indicating higher frequency, while the varying shades of purple dots indicate different levels of station frequency. Red triangles represent *Petite Ceinture* stations.

tance to the city center on the metro station frequency separately, holding construction costs constant. These estimates suggest a positive and significant effect of productivity and type-specific amenities on the frequency of sampled metro stations. By contrast, the coefficient associated with the distance to the city center is negative and significant, suggesting that potential stations that are closer to the city center are more represented in the sampled metro stations. This result could be partly mechanical since radial lines necessarily cross the city center. However, productivity and amenities are highly geographically correlated, and both are higher in the city center (see Figure 8).

Controlling for the five independent variables in column (5) decreases the magnitude of the effects associated with productivity and low-skilled amenities variables. However, the coefficient associated with high-skilled amenities becomes negative. Since the economic objective is based on gains in rateable values, connecting locations with higher low-skilled amenities is more strategic than connecting those with higher high-skilled amenities. This is primarily because these locations attract low-skilled workers, representing the city's main workforce with 54% of the Parisian population being low-skilled. Additionally, low-skilled workers tend to spend a higher proportion of their income on housing. Therefore,

implementing a metro network within these areas increases low-skilled workers' market access and total income, which in turn increases payments to landlords. Moreover, estimates suggest that, on average, after accounting for the location centrality, construction costs, and the level of amenities, a 1% increase in productivity leads to a 1.3 increase in frequency within the sampled metro networks compared to other stations. Overall, results from this historical application confirm the findings of the numerical example and highlight the importance of targeting metro stations that connect locations with high productivity and amenities. (These results are robust to using standardized $\beta$ coefficients, which allow direct comparison of effect magnitudes. See Table A3 in Appendix.)

Table 1
Determinants of metro stations frequency

| | Metro Station Frequency | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Log. Productivity | 157.4*** | | | | 131.0*** |
| | (30.9) | | | | (32.6) |
| Log. Costs | -12.2 | 24.7*** | 27.4*** | 9.0 | -18.4** |
| | (9.7) | (8.4) | (9.2) | (7.7) | (9.3) |
| Log. Low-skilled Amenities | | 89.5*** | | | 91.7** |
| | | (26.3) | | | (41.8) |
| Log. High-skilled Amenities | | | 41.9*** | | -85.2*** |
| | | | (11.9) | | (32.1) |
| Log. distance to CC | | | | -80.3*** | -64.9*** |
| | | | | (20.0) | (19.0) |
| Observations | 334 | 334 | 334 | 334 | 334 |
| $R^2$ | 0.14 | 0.11 | 0.10 | 0.17 | 0.20 |

Notes: Each column reports estimates from a separate regression. Regressions use potential metro stations derived from 500mx500m grid cells. Dependent variable is the frequency of metro stations within 500 SA runs. Heteroskedasticity-robust standard errors are in parentheses and ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

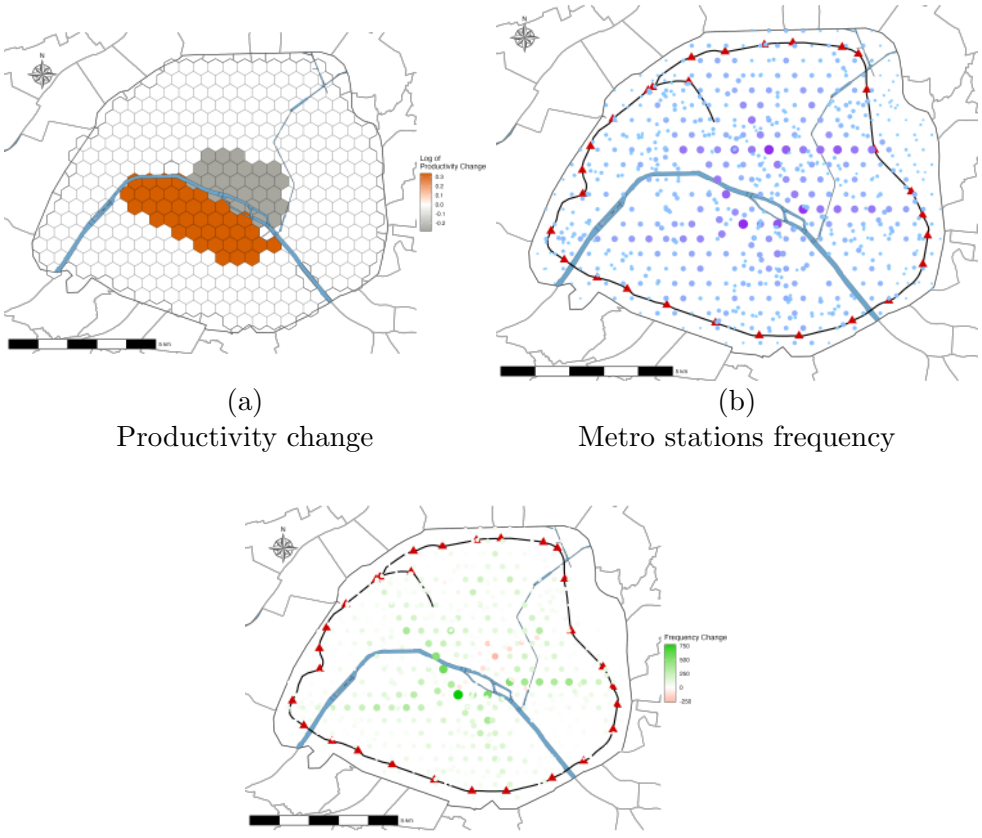
### 4.5.4 Comparative static

Next, I run comparative statics exercises on parameters and show how the characteristics of sampled metro networks change. These experiments allow to refine the results obtained previously and investigate what other factors determine convergence towards specific metro networks.

**An alternative CBD** First, I conduct a comparative static experiment in order to check whether productivity is the main factor explaining metro station frequency with the sampled metro networks. To achieve this, I simulate a productivity shock in locations

near the city center, with positive shock for those located on the southern bank of the Seine in Paris and negative shock for those located on the northern bank as illustrated in Panel (a) of Figure 20. Then, I sample again 500 metro networks using the simulated annealing algorithm and considering the same parameters. Panel (b) of Figure 20 shows the counterfactual frequency of metro stations within the sampled metro networks, and another pattern emerges from it. Compared to previous simulations in Figure 19, metro stations in the South seem to appear more often in the SA runs. Panel (c) of Figure 20 shows the change between baseline and counterfactual metro station frequency, and it depicts an increase around areas that experienced a productivity shock. Hence, this experiment suggests that productivity is one of the main drivers of strategic location, confirming previous results in Table 1.

Figure 20
Comparative static: new city center



(a)
Productivity change

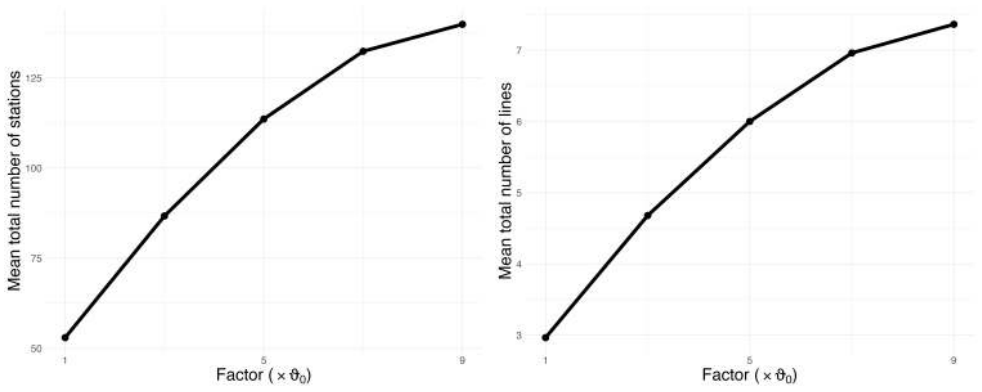
(b)
Metro stations frequency

(c)
Strategic stations change

Notes: Panel (a) shows the productivity shock with an alternative central business district (CBD) in orange, and the old CBD in grey. Panel (b) illustrates the frequency of metro stations occurring in 500 simulated annealing runs following this new CBD. Panel (c) shows the change in frequency of metro stations between the baseline and the new CBD scenario. Green indicates a positive change in frequency, while red indicates a negative change. The size of the dots is proportional to the absolute value of the frequency change.

**Metro construction fixed cost** Second, I conduct a comparative static experiment investigating the role of the metro construction fixed-cost $\vartheta_0$. The metro cost functional form equation (30), by assuming fixed and linear costs, should provide an incentive for planners to build larger metro networks in order to smooth large fixed-investments. To explore this, I implement simulations and investigate the size of the metro network while gradually increasing the fixed-cost by a factor. For a given $\vartheta_0$, I consider 100 runs of the simulated annealing algorithm with 1,000 iterations each in order to sample metro networks from the distribution $\pi$. Then, I compute the average number of stations and lines of the sampled metros. Figure 21 shows the results, where the x-axis represents the multiplication factor of $\vartheta_0$, the y-axis represents the average number of metro stations in Panel (a), and the average number of metro lines in Panel (b). Overall, this experiment shows that increasing the fixed-cost $\vartheta_0$ leads to sampled metro networks being larger in terms of lines and stations.

Figure 21
Comparative static: metro construction fixed-cost




(a) Mean number of total stations (b) Mean number of total lines

Notes: Panel (a) shows the average total number of stations over the SA runs as a function of the cost factor. Panel (b) illustrates the average total number of metro lines over the SA runs as a function of the cost factor.

**Agglomeration forces** In this last comparative static experiment, I investigate the role of agglomeration forces. To do so, I allow agglomeration forces by letting type-specific amenities and productivity be functions of two components: (i) exogenous type-specific residential ($b_{n,g}$) and production ($a_n$) fundamentals, and (ii) own location residence or workplace density, as in Heblich et al. (2020). Hence, both productivity and amenities

can be written as,

$$B_{n,g} = b_{n,g} \left( \frac{R_n}{K_n} \right)^{\eta^R}$$
$$A_n = a_n \left( \frac{L_n}{K_n} \right)^{\eta^L}$$

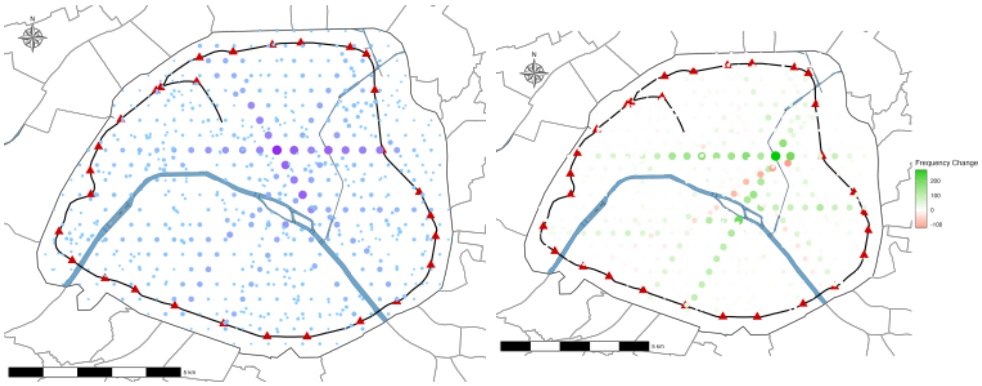
where endogenous residential externalities depend on the location's own residence employment density weighted by a constant elasticity ($\eta^R$) parameter, and endogenous production externalities depend on the location's own workplace employment density weighted by a constant elasticity ($\eta^L$) parameter. I calibrate $\eta^L$ to 0.07, and $\eta^R$ to 0.16, which are standard values estimated in the urban literature (Ahlfeldt et al., 2015; Heblich et al., 2020). Then, I repeat the baseline simulation procedure considering agglomeration forces, resulting in 500 sampled metro networks.

Across the 500 sampled metro networks, total production and population rise by 132.2% and 66.8% on average, respectively, substantially larger than in the baseline scenario without agglomeration forces. Underlying these gains, total productivity, total low-skilled amenities, and total high-skilled amenities increase on average by 1.4%, 6.3%, and 5.9%, respectively. Due to the large population inflow, the additional rateable value gains make it possible to implement a larger metro network: the average number of metro stations rises to 65.8 (compared to 51 in the baseline), while the average number of lines increases from 2.9 to 3.7.

Panel (a) of Figure 22 displays the spatial distribution of metro station frequency within the sampled metro networks. A similar pattern emerges from it, compared to Figure 19, with higher strategic stations in the city center. Panel (b) of Figure 22 displays the change in metro station frequency between the baseline and this comparative static experiment. Overall, there is a general increase in frequency, consistent with the larger network size: most locations display positive changes. However, a more nuanced pattern also emerges: some stations in the city center, particularly in the center near the Seine, experience a decrease in frequency, losing relative strategic importance compared to the baseline. At the same time, a horizontal east-west corridor of stations above the center becomes more prominent, suggesting that the introduction of agglomeration forces shifts part of the optimal design away from isolated central nodes and towards a more structured line pattern. Panel (c) of Figure 22 also explores this pattern by displaying

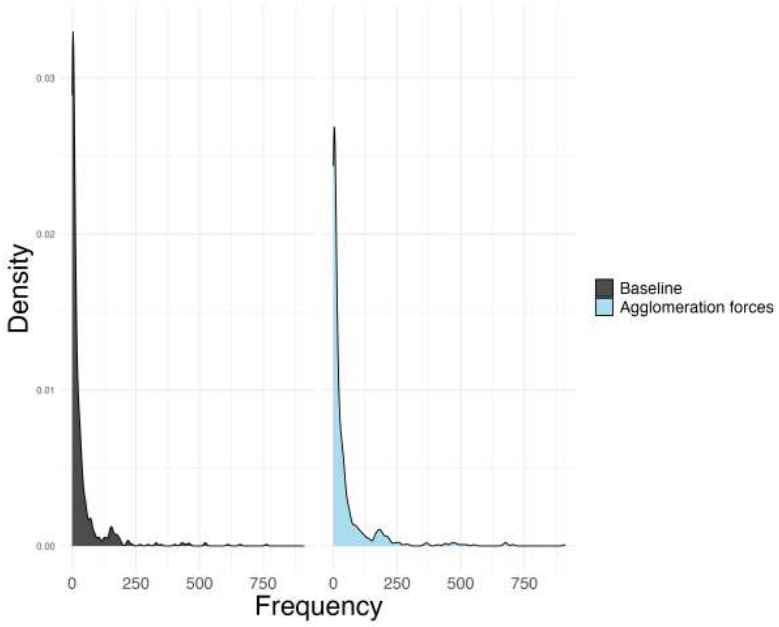
the distribution of metro station frequency, indicating the concentration of strategic stations. As a result, introducing agglomeration forces seems to decrease the concentration of strategic metro stations in certain locations and the number of metro stations with zero appearances in the simulations. Overall, this comparative static experiment shows that introducing agglomeration forces leads the simulated annealing algorithm to converge towards an optimal metro pattern that goes through the city center, with overall more metro stations, and a different design.

Figure 22
Comparative static: agglomeration forces



(a)
Metro stations frequency

(b)
Strategic stations change




(c)
Distribution of metro stations frequency

Notes: Panel (a) shows the frequency of metro stations occurring in the simulated annealing runs considering agglomeration forces. The size of each station is proportional to its frequency, with larger dots indicating higher frequency, while the varying shades of purple dots indicate different levels of station frequency. Panel (b) illustrates the change in frequency of metro stations between the baseline and the agglomeration forces scenario. Green indicates a positive change in frequency, while red indicates a negative change. The size of the dots is proportional to the absolute value of the frequency change. Panel (c) depicts the distribution of frequency in both the baseline and agglomeration forces scenarios.

## 4.6 Egalitarian versus Utilitarian

In this subsection, I explore whether the maximization of utility by a planner, distinguishing between an egalitarian and a utilitarian approach, leads to different metro designs. Previous results used an economic impact factor based on rateable value gains and metro construction costs as the economic objective $W$. However, a planner can also build a metro network within a given budget $\mathcal{B}$, aiming to maximize an economic objective $W$.

A utilitarian planner seeks to maximize the sum of the utilities of the two agents, treating all individuals as having equal weight within the city. Conversely, an egalitarian planner aims to maximize the minimum utility between the two agents, following a Rawlsian welfare approach that focuses on enhancing the well-being of the worst-off workers within the city.

To investigate whether different planner objectives result in different metro designs, I define the following utilitarian planner maximization problem,

$$\begin{aligned} \max \quad & \widehat{\overline{U}}_H(M;\theta) + \widehat{\overline{U}}_L(M;\theta) \\ \text{subject to} \quad & \mathcal{B} \geq \text{Cost}(M), \end{aligned} \tag{32}$$

and the following egalitarian planner maximization problem,

$$\begin{aligned} \max \quad & \min\left(\widehat{\overline{U}}_H(M;\theta), \widehat{\overline{U}}_L(M;\theta)\right) \\ \text{subject to} \quad & \mathcal{B} \geq \text{Cost}(M). \end{aligned} \tag{33}$$

In both equations, $\widehat{\overline{U}}_H$ and $\widehat{\overline{U}}_L$ represent the indirect utility change in percentage points for high-skilled and low-skilled workers, respectively. $\mathcal{B}$ denotes the total budget for metro construction, and $\text{Cost}(M)$ represents the construction costs of metro $M$.

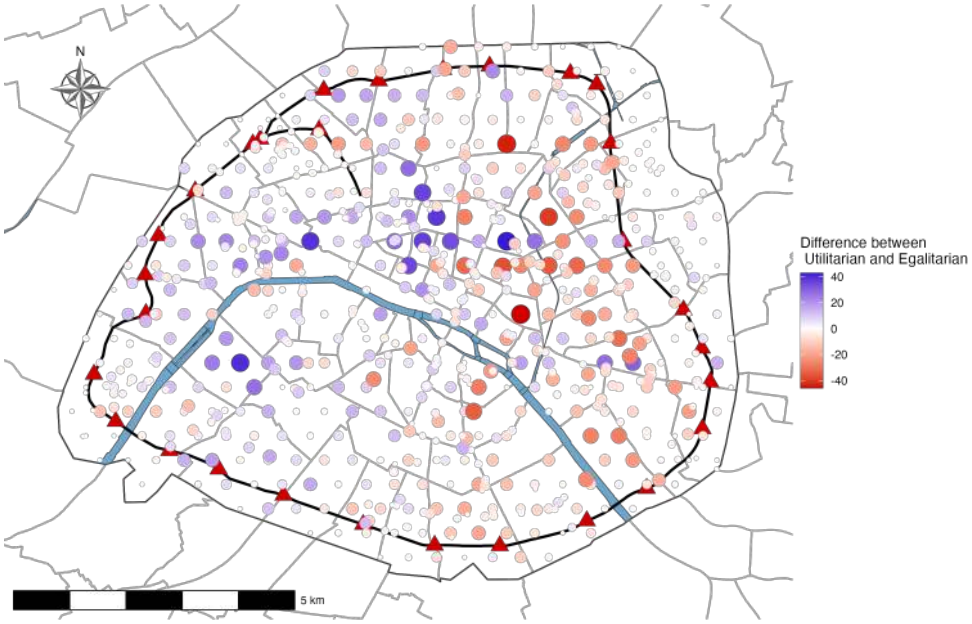
Again, I sampled 500 metro networks using the simulated annealing algorithm for both a utilitarian and an egalitarian planner problem. Figure 23 shows the difference in metro station frequency between the utilitarian scenario and the egalitarian scenario. Blue dots indicate metro stations that appear more often in the utilitarian scenario than in the egalitarian scenario, while red dots indicate the opposite. This visualization shows a two-sided pattern: metro stations in the west appear more frequently in the utilitarian scenario, whereas those in the east appear more frequently in the egalitarian scenario. As a consequence, the utilitarian scenario tends to implement more often a metro within

locations that have high-level amenities for high-skilled workers, while the egalitarian scenario tends to implement more often a metro within locations that have high-level amenities for low-skilled workers. This overall pattern can be explained by the fact that, in an open city setting, an inflow of workers from the outside economy creates pressure on the housing market, increasing housing prices. Since low-skilled workers spend a relatively higher share of their income on housing, they suffer relatively more from an increase in housing prices. Hence, within the scenario of a utilitarian planner maximizing the sum of utilities, the change in expected utility for high-skilled workers will be relatively larger than that for low-skilled workers. As a consequence, this will drive the metro network design towards locations with a high level of amenities for high-skilled workers.

Given that the egalitarian scenario maximizes the minimum utility between the two types of workers, the metro design converges on locations where the amenities of low-skilled workers are large, as their expected utility increases relatively less than that of high-skilled workers.

Note that I find similar results when considering a general maximization problem where the planner assigns a political representation weight $z \in [0, 1]$ to low-skilled workers ($\max : \ (1-z) \times \widehat{\overline{U}}_H(M;\theta) + z \times \widehat{\overline{U}}_L(M;\theta)$). Specifically, when comparing a weight $z = 1$ (targeting low-skilled workers) with a weight $z = 0$ (targeting high-skilled workers), the difference in frequency displays the same pattern as in Figure 23 (see Figure A5 and the formalization in Appendix C.5).

Figure 23
Difference between Utilitarian and Egalitarian sampled metro networks




Notes: This figure shows the change in frequency of metro stations occurring in the simulated annealing runs between the Utilitarian and the Egalitarian scenarios. Blue indicates a positive change in frequency, while red indicates a negative change. The size of the dots is proportional to the absolute value of the frequency change.

# 5 Comparison with metro network projects

Once the planner has sampled metros from $\mathcal{M}$, she can use this information to compute proximity indexes for her metro projects,

$$\text{Index}_m = \sum_s \left( \frac{wgt_s}{\text{Dist}_{sm}} \right), \tag{34}$$

with $\text{Dist}_{sm}$ the distance between lines of the metro project $m$, and the potential metro stations $s$. Moreover, $wgt_s$ is the station weight based on the frequency, which is specific to the simulated annealing algorithm parameters (represented in Figure 19 for the baseline results). Hence, this index aims to capture whether a metro project is more or less far from strategic metro locations resulting from runs of the SA algorithm. Since distance appears in the denominator, a higher index will translate to a greater proximity of the metro project to strategic metro stations.

Table 2 shows the indexes for the metro projects presented in section 4.1, and based on the different scenarios with general equilibrium and without agglomeration forces:

maximization of economic impact factor (baseline results) in column (1), maximization of utility with a utilitarian planner in column (2), maximization of utility with an egalitarian planner in column (3). Column (4) shows the total length for each metro project, and the last 3 columns show the indexes normalized by the length.

Regarding the Fulgence Bienvenüe metro project, which was selected and implemented by public authorities, it exhibits the highest index in the baseline scenario that maximizes the economic impact factor, both in raw terms and once normalised by network length. This result reflects the fact that the Bienvenüe network was systematically routed along Haussmann boulevards, allowing extensive use of the cheaper cut-and-cover construction technology and thereby lowering the denominator of the welfare-to-cost ratio. Properly accounting for this spatial cost heterogeneity thus rehabilitates the adopted design as the most welfare-efficient of the four historical proposals. Among the three remaining projects, the sensitivity to the planner's objective varies considerably. The Léon Say index is virtually unchanged across scenarios: general and utilitarian are identical (5.64), while the egalitarian scenario falls by only 3.4% (5.45), reflecting a design broadly insensitive to the skill composition. By contrast, Alphand-Huet and the Council 1883 network both score markedly higher under utility-maximisation objectives relative to the general baseline: Alphand-Huet by +6.2% under the utilitarian scenario and +7.7% under the egalitarian, and the Council 1883 by +12.9% and +15.1%, respectively. Both projects also score marginally higher under the egalitarian scenario than the utilitarian one (+1.4% for Alphand-Huet and +2.0% for the Council 1883), consistent with routes passing through locations with relatively higher amenities for low-skilled workers. Finally, while the Council 1883 holds the highest raw index in all scenarios, this largely reflects its greater length (63 km): once normalised, it ranks last in the general scenario (0.24, vs. 0.27 for Léon Say and 0.26 for Alphand-Huet), though it recovers to match Alphand-Huet under the egalitarian scenario (both 0.28).

# 6 Conclusion

This paper develops a framework to identify the optimal design of urban transport infrastructure. The framework combines a quantitative urban model featuring heterogeneous workers and multiple transport modes with a simulated annealing algorithm that

Table 2
Metro projects

| Metro project | General | Uti | Ega | Length (km) | Normalized Index | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1)/(4) | (2)/(4) | (3)/(4) |
| Léon Say | 5.64 | 5.64 | 5.45 | 21.22 | 0.27 | 0.27 | 0.26 |
| Alphand-Huet | 7.39 | 7.85 | 7.96 | 27.94 | 0.26 | 0.28 | 0.28 |
| Council 1883 | 15.32 | 17.29 | 17.64 | 63.01 | 0.24 | 0.27 | 0.28 |
| Fulgence Bienvenüe | 140.87 | 103.06 | 98.24 | 57.90 | 2.43 | 1.78 | 1.70 |

efficiently explores a high-dimensional discrete network space. It incorporates general equilibrium effects, workers reallocating across residence and workplace in response to commuting cost changes, with land prices adjusting endogenously, and allows for production and residential agglomeration forces. Applied to historical Paris and to a canonical numerical example, results show that strategic metro stations are predominantly located within areas of high centrality, high productivity, and high amenities.

A potential application of this framework is the use of strategic station locations as instrumental variables, given that they are driven by welfare considerations. This approach allows to differentiate parts of the network designed to maximize well-being from those influenced by other non-random factors, thereby offering valuable insights for urban planning and policy-making.

# References


Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189.

Allen, T. and Arkolakis, C. (2022). The welfare effects of transportation infrastructure improvements. *The Review of Economic Studies*, 89(6):2911–2957.

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research.

Bagagli, S. (2023). The (express)way to segregation: Evidence from chicago. Technical report, Department of Economics, Harvard University. Job Market Paper.

Balboni, C. (2025). In harm's way? infrastructure investments and the persistence of coastal cities. *American Economic Review*, 115(1):77–116.

Bilal, A. (2023). The geography of unemployment. *The Quarterly Journal of Economics*, 138(3):1507–1576.

Borusyak, K. and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

Brinkman, J. and Lin, J. (2022). Freeway revolts! the quality of life effects of highways. *The Review of Economics and Statistics*, pages 1–45.

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica*, 87(3):741–835.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Champalaune, P. and Cosentino, P. (2026). Commuting, air quality and welfare. *Available at SSRN 6134726*.

Combes, P.-P., Duranton, G., and Gobillon, L. (2021). The production function for housing: Evidence from france. *Journal of Political Economy*, 129(10):2766–2816.

Cosentino, P. (2026). Trade, commuting and city structure.

Cottereau, A. (2004). Les batailles pour la création du métro: un choix de mode de vie, un succès pour la démocratie locale. *Revue d'histoire du XIXe siècle. Société d'histoire de la révolution de 1848 et des révolutions du XIXe siècle*, (29):89–151.

Dekle, R., Eaton, J., and Kortum, S. (2007). Unbalanced trade. *American Economic Review*, 97(2):351–355.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The geography of development. *Journal of Political Economy*, 126(3):903–983.

Fajgelbaum, P. D. and Schaal, E. (2020). Optimal transport networks in spatial equilibrium. *Econometrica*, 88(4):1411–1452.

George, H. (1879). Poverty and progress. *An Inquiry into the Cause of*.

Greaney, B., Parkhomenko, A., and Van Nieuwerburgh, S. (2024). Dynamic urban economics. *Available at SSRN 4897077*.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*, 135(4):2059–2133.

Heblich, S., Trew, A., and Zylberberg, Y. (2021). East-side story: Historical pollution and persistent neighborhood sorting. *Journal of Political Economy*, 129(5):1508–1552.

Herzog, I. (2024). The city-wide effects of tolling downtown drivers: Evidence from london's congestion charge. *Journal of Urban Economics*, 144:103714.

Kleinman, B., Liu, E., and Redding, S. J. (2023). Dynamic spatial general equilibrium. *Econometrica*, 91(2):385–424.

Koster, H. R. (2024). The welfare effects of greenbelt policy: Evidence from england. *The Economic Journal*, 134(657):363–401.

Kreindler, G., Gaduh, A., Graff, T., Hanna, R., and Olken, B. A. (2025). Optimal public transportation networks: Evidence from the world's largest bus rapid transit system in jakarta. *conditionally accepted, American Economic Review*.

Loumeau, G. (2023). Locating public facilities: Theory and micro evidence from paris. *Journal of Urban Economics*, 135:103544.

Loumeau, G. (2024). Accommodating the rise in urbanisation: Are new towns a good solution? *The Economic Journal*, 134(662):2530–2557.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328.

Miyauchi, Y., Nakajima, K., and Redding, S. J. (2025). The economics of spatial mobility: Theory and evidence using smartphone data. *The Quarterly Journal of Economics*, 140(4):2507–2570.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–3890.

Passalacqua, A. (2012). La vitesse ferroviaire comme point de mire: le monde des transports parisiens et le rail. *Revue d'histoire des chemins de fer*, (42-43):155–173.

Pietrabissa, G. (2023). School access and city structure. *Unpublished manuscript*.

Redding, S. J. (2022). Trade and geography. *Handbook of International Economics*, 5:147–217.

Redding, S. J. and Sturm, D. M. (2024). Neighborhood effects: Evidence from wartime destruction in london. Working Paper.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. *Handbook of regional and urban economics*, 5:1339–1398.

Santamaria, M. (2020). Reshaping infrastructure: Evidence from the division of germany.

Severen, C. (2021). Commuting, labor, and housing market effects of mass transportation: Welfare and identification. *Review of Economics and Statistics*, pages 1–99.

Takeda, K. and Yamagishi, A. (2023). History versus expectations in the spatial economy: Lessons from hiroshima. *Available at SSRN 4339170*.

Tsivanidis, N. (2026). Evaluating the impact of urban transit infrastructure: Evidence from bogotá's transmilenio. *American Economic Review*, 116(2):418–463.

Warnes, P. E. (2021). Transport infrastructure improvements and spatial sorting: Evidence from buenos aires. *Unpublished manuscript*.

Weiwu, L. (2024). Unequal access: Racial segregation and the distributional impacts of interstate highways in cities. Technical report, MIT. Job Market Paper.

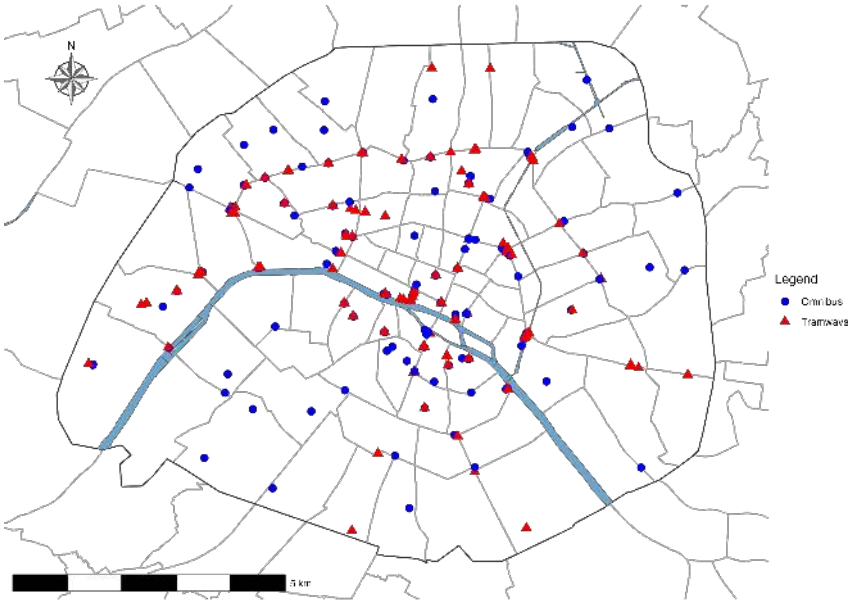
# Online Appendix

Pol Cosentino

# A Data

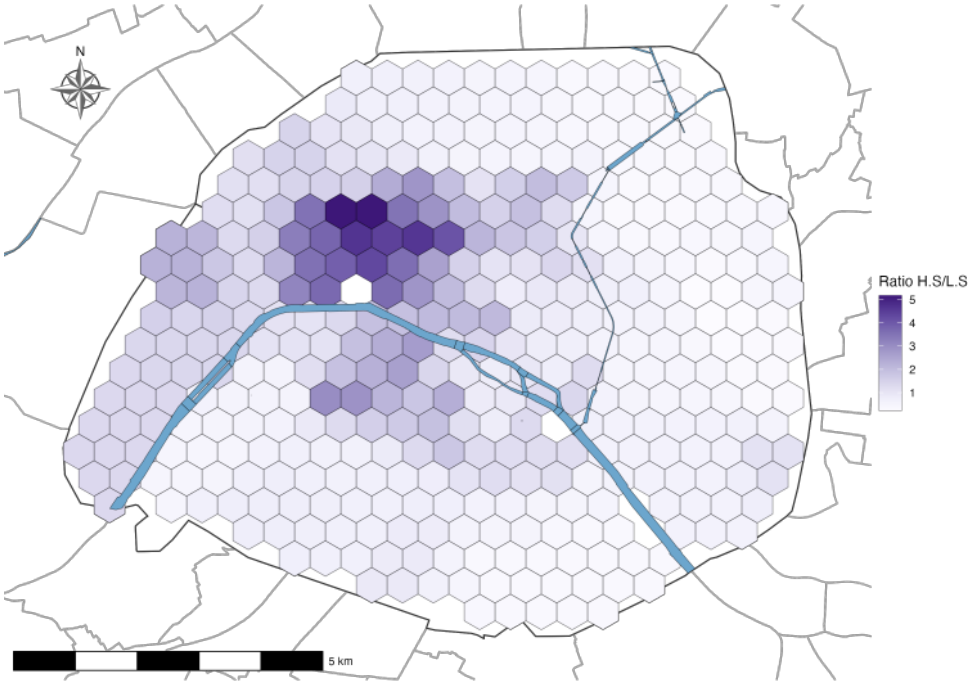
## A.1 Map of omnibus and tramways

Figure A1
Omnibus and tramways in 1896



## A.2 Map of residential sorting by skill

Figure A2
Spatial sorting of residents by skill in 1896




Notes: This figure illustrates the spatial residential sorting by skill level in 1896 by showing the ratio of the high-skilled residence employment with respect to the low-skilled residence employment. The darker shades indicate relatively higher concentrations of high-skilled residents.

# B Quantitative urban model

## B.1 Canonical

### B.1.1 Housing supply

The supply of housing is determined by a competitive construction sector with a Cobb-Douglas technology following Combes et al. (2021),

$$H_i = K_i^{\mu} \Lambda_i^{1-\mu}, \tag{B.1}$$

where $H_i$ is the housing supplied in location $i$, $K_i$ is the land in location $i$, and $\Lambda_i$ is the capital in location $i$. Finally, $\mu$ is the share of land within the construction technology, and the capital is assumed to be supplied without any cost in the wider economy at a price $\mathbb{P} = 1$. Therefore, the maximization of the firm's profit in the construction sector

gives,

$$(1-\mu)^{1/\mu} Q_i^{1/\mu} K_i = \Lambda_i. \tag{B.2}$$

Plugging equation (B.2) into equation (B.1) gives,

$$H_i = K_i (1-\mu)^{\frac{1-\mu}{\mu}} Q_i^{\frac{1-\mu}{\mu}}, \tag{B.3}$$

which is equal to equation (9) in the main paper.

### B.1.2 Iterative algorithm to solve a spatial equilibrium

Table A1
Algorithm for solving spatial equilibrium

| Line | Code | Description/Comment |
|---|---|---|
| 0 | $\{L_i^0, R_n^0, w_i^0, Q_n^0\}$ | Initial guess |
| 1 | `while er > tol` | |
| 2 | $w_i = A_i^{\frac{1}{\alpha}} Q_i^{\frac{\alpha-1}{\alpha}}$ | Compute location wages |
| 3 | $U_{ni} = \frac{B_n w_i}{d_{ni} Q_n^{\beta}}$ | Compute pair locations utility |
| 4 | $\Phi_{ni} = \left(U_{ni}\right)^{\epsilon}$ | —— |
| 5 | $\Phi = \sum_{k \in N} \sum_{l \in N} \Phi_{kl}$ | —— |
| 6 | $\lambda_{ni} = \frac{\Phi_{ni}}{\Phi}$ | Compute commuting flows probability |
| 7 | $R_n = \sum_{i \in N} \lambda_{ni} \; R_{\mathbb{N}}$ | Compute residential choices |
| 8 | $L_i = \sum_{n \in N} \lambda_{ni} \; R_{\mathbb{N}}$ | Compute working choices |
| 9 | $\lambda^R_{ni\|n} = \frac{\lambda_{ni}}{\sum_{i \in N} \lambda_{ni}}$ | Compute conditional commuting probability |
| 10 | $H_n^R = \beta \; \frac{\sum_{i \in N} \lambda^R_{ni\|n} w_i R_n}{Q_n}$ | Compute residential floorspace demand |
| 11 | $H_n^F = \frac{1-\alpha}{\alpha} \; \frac{w_n L_n}{Q_n}$ | Compute commercial floorspace demand |
| 12 | $H_n = H_n^R + H_n^F$ | Compute total floorspace demand |
| 13 | $Q_n^1 = Q_n \left(\frac{H_n}{H_n^S}\right)$ | Update rent with floorspace market clearing condition |
| 14 | $\text{er} = \max\left(\left\|(Q_n^1 - Q_n)\right\|\right)$ | Compute error based on previous and updated rents |
| 15 | $Q_n = 0.1 \times Q_n^1 + 0.9 \times Q_n$ | Update rent |
| 16 | `end` | |

## B.2 Extension

### B.2.1 Residential-workplace choices

Following the type-specific indirect utility equation (17), the type-specific commuting probability can be expressed as,

$$\lambda_{ni,g} = \frac{L_{ni,g}}{R_{\mathbb{N}_g}} = \frac{(B_{n,g}w_{i,g})^{\epsilon_g}(d_{ni}Q_n^{\beta_g})^{-\epsilon_g}}{\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}(B_{k,g}w_{l,g})^{\epsilon_g}(d_{kl}Q_l^{\beta_g})^{-\epsilon_g}} = \frac{\Phi_{ni,g}}{\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}\Phi_{kl,g}}, \tag{B.4}$$

where only amenities ($B_{n,g}$), the wage ($w_{i,g}$), the share of income spend on housing ($\beta_g$), and the Fréchet parameter ($\epsilon_g$) differ from equation (2) since they are type-specific.

By summing across workplaces, I get the probability that a worker of type $g$ lives in $n$,

$$\lambda^R_{n,g} = \frac{R_{n,g}}{R_{\mathbb{N}_g}} = \frac{\sum_{i\in\mathbb{N}}\Phi_{ni,g}}{\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}\Phi_{kl,g}}, \tag{B.5}$$

and by summing across residence places, I get the probability that a worker of type $g$ works in $i$,

$$\lambda^L_{i,g} = \frac{L_{i,g}}{R_{\mathbb{N}_g}} = \frac{\sum_{n\in\mathbb{N}}\Phi_{ni,g}}{\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}\Phi_{kl,g}}. \tag{B.6}$$

Using the types-specific commuting probability, and the type-specific residential probability, the type-specific conditional commuting probability can be expressed as,

$$\lambda^R_{ni|n,g} = \frac{\lambda_{ni,g}}{\lambda^R_{n,g}} = \frac{(w_{i,g}/d_{ni})^{\epsilon_g}}{\sum_{l\in N}(w_{l,g}/d_{nl})^{\epsilon_g}}. \tag{B.7}$$

where only wage ($w_{i,g}$) differs from equation (10).

Finally, population mobility for a given type $g$ ensures that everyone derives the same expected utility in the city,

$$\overline{U}_g = \delta_g\Big[\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}\Phi_{kl,g}\Big]^{1/\epsilon_g}, \tag{B.8}$$

with $\delta_g \equiv \Gamma\Big(\frac{(\epsilon_g-1)}{\epsilon_g}\Big)$, $\Gamma(.)$ the gamma function.

### B.2.2 Wage cost index

Using the firm profit $\pi = Y_i - W_i L_i - r_i M_i - Q_i H_i^L$, the total firm labor payments $W_i L_i = \sum_g w_{i,g} L_{i,g}$, the FOC of firm with respect to the type-specific labor supply equation (21), and the FOC of firm with respect to the total labor supply $W_i = \alpha_L^{(1-\alpha_L)} A_i L_i^{\alpha-1} \left(\frac{H_i^L}{\alpha_F}\right)^{\alpha_F} \left(\frac{M_i}{\alpha_M}\right)^{\alpha_M}$, it is possible to express the aggregate wages as,

$$W_i = \left( \sum_g w_{i,g}^{\frac{\rho}{\rho-1}} a_{i,g}^{\frac{1}{1-\rho}} \right)^{\frac{\rho-1}{\rho}}. \tag{B.9}$$

### B.2.3 Housing and Machinery demands

The respective ratios of the first-order conditions (FOCs) from the housing equation (22) and the machinery equation (23) to the FOC of the labor supply equation (21) are given by,

$$M_i = \frac{\alpha_M}{\alpha_L} \frac{L_{i,g}^{1-\rho} L_i^{\rho}}{a_{i,g}} w_{i,g}, \tag{B.10}$$

$$H_i^L = \frac{\alpha_F}{\alpha_L} \frac{L_{i,g}^{1-\rho} L_i^{\rho}}{a_{i,g}} \frac{w_{i,g}}{Q_i}. \tag{B.11}$$

### B.2.4 Exact-hat algebra

Using "hat" expression with $\widehat{x} = \frac{x'}{x}$, every equation can be expressed in terms of change relative to the baseline equilibrium.

The relative change in type-specific amenities can be expressed as,

$$\widehat{B}_{n,g} = \widehat{R}_n^{\eta^R}, \tag{B.12}$$

which includes only changes in residential employment ($\widehat{R}_n$) since land is fixed.

Similarly, the relative change in productivity can be expressed as,

$$\widehat{A}_i = \widehat{L}_i^{\eta^L}, \tag{B.13}$$

which includes only changes in workplace employment ($\widehat{L}_i$) since land is fixed.

The relative change in location choices can be expressed as,

$$\widehat{\lambda}_{ni,g} = \frac{\left(\widehat{B}_{n,g}\widehat{w}_{i,g}\right)^{\epsilon_g}\left(\widehat{d}_{ni}\widehat{Q}_n^{\beta_g}\right)^{-\epsilon_g}}{\sum_k \sum_l \lambda_{kl,g}\left(\widehat{B}_{k,g}\widehat{w}_{l,g}\right)^{\epsilon_g}\left(\widehat{d}_{kl}\widehat{Q}_k^{\beta_g}\right)^{-\epsilon_g}}, \tag{B.14}$$

which is function of changes in living cost ($\widehat{Q}_n$), changes in commuting costs ($\widehat{d}_{ni}$), changes in type-specific amenities ($\widehat{B}_{n,g}$), and changes in type-specific wages ($\widehat{w}_{i,g}$).

The relative change in type-specifc Parisian total population can be expressed as,

$$\left(\widehat{R}_{\mathbb{N}_g}\right)^{1/\phi} = \Big[\sum_k \sum_l \lambda_{kl,g}\left(\widehat{B}_{k,g}\widehat{w}_{l,g}\right)^{\epsilon_g}\left(\widehat{d}_{kl}\widehat{Q}_k^{\beta_g}\right)^{-\epsilon_g}\Big]^{1/\epsilon_g}, \tag{B.15}$$

which is function of changes in living cost ($\widehat{Q}_n$), changes in commuting costs ($\widehat{d}_{ni}$), changes in type-specific amenities ($\widehat{B}_{n,g}$), and changes in type-specific wages ($\widehat{w}_{i,g}$).

The relative change in commercial housing demand can be expressed as,

$$\widehat{H}_n^L = \widehat{L}_n^{\rho}\ \widehat{L}_{n,g}^{1-\rho}\ \frac{\widehat{w}_{n,g}}{\widehat{Q}_n}, \tag{B.16}$$

and the relative change in machinery demand can be expressed as,

$$\widehat{M}_n = \widehat{L}_n^{\rho}\ \widehat{L}_{n,g}^{1-\rho}\ \widehat{w}_{n,g}. \tag{B.17}$$

The change in type-specific residential housing demand can be expressed as,

$$\widehat{H}_{n,g}^R H_{n,g}^R = \beta_g \left[\sum_{i\in\mathbb{N}} \left(\frac{\lambda_{ni|n,g}^R \left(\frac{\widehat{w}_{i,g}}{\widehat{d}_{ni}}\right)^{\epsilon_g}}{\sum_{l\in\mathbb{N}} \lambda_{nl|n,g}^R \left(\frac{\widehat{w}_{i,g}}{\widehat{d}_{nl}}\right)^{\epsilon_g}}\right)\right] \frac{\widehat{\lambda}_{n,g}^R \lambda_{n,g}^R \widehat{R}_{\mathbb{N}_g} R_{\mathbb{N}_g}}{\widehat{Q}_n Q_n}, \tag{B.18}$$

and the change in commercial housing demand can be expressed as,

$$\widehat{H}_n^L H_n^L = \frac{\alpha_F}{\alpha_L}\ \widehat{L}_n^{\rho} L_n^{\rho}\ \widehat{L}_{n,g}^{1-\rho} L_{n,g}^{1-\rho}\ \frac{\widehat{w}_{n,g} w_{n,g}}{\widehat{Q}_n Q_n}. \tag{B.19}$$

The relative change in total production can be expressed as,

$$\widehat{Y}_i = \widehat{A}_i\ \left(\widehat{L}_i\right)^{\alpha_L}\ \left(\widehat{H}_i^L\right)^{\alpha_F}\ \left(\widehat{M}_i\right)^{\alpha_M}. \tag{B.20}$$

The relative change in type-specific wages can be expressed as,

$$\widehat{w}_{i,g} = \widehat{A}_i^{1/\alpha_L} \, \widehat{L}_{i,g}^{\rho-1} \, \widehat{L}_i^{1-\rho} \, \widehat{Q}_i^{-\alpha_F/\alpha_L}. \tag{B.21}$$

Finally, the change in rateable values can be expressed as,

$$\widehat{\mathbb{Q}}_n \mathbb{Q}_n = \widehat{Q}_n^{1+\mu} \mathbb{Q}_n. \tag{B.22}$$

## B.3 Calibration

Table A2
Calibration of structural parameters for heterogeneous QUM

| Parameter | Description | Method | Value |
|---|---|---|---|
| | **Calibrated** | | |
| $\alpha_F$ | Share of housing in production function | Heblich et al. (2020) | 0.2 |
| $\alpha_M$ | Share of machinery in production function | Heblich et al. (2020) | 0.2 |
| $\alpha_L$ | Share of labor in production function | CRS ($\alpha_L + \alpha_M + \alpha_F = 1$) | 0.6 |
| $\rho$ | Elasticity of substitution between skill groups | Card (2009) | 0.3 |
| $\beta_L$ | Share of housing consumption for low-skilled workers | Weiwu (2024) | 0.3 |
| $\beta_H$ | Share of housing consumption for high-skilled workers | Weiwu (2024) | 0.2 |
| $\epsilon_L$ | Fréchet parameter for low-skilled workers | Redding and Sturm (2024) | 6 |
| $\epsilon_H$ | Fréchet parameter for high-skilled workers | Redding and Sturm (2024) | 4 |
| $d_{ni} = e^{\kappa \, t_{ni}}$ | Commuting cost elasticity | Cosentino (2026) | $\kappa = 0.0204$ |
| $(1-\mu)/\mu$ | Housing supply elasticity | Cosentino (2026) | 1.76 |

# C Simulated annealing algorithm

## C.1 Calibration of construction cost

As detailed in the main paper, the metro construction costs are defined based on historical cost data from Heblich et al. (2020) for London. The costs are £330,000 per mile for cut-and-cover and £500,000 per mile for deep-bore tunneling, as of 1921. Using an exchange rate of £1 = 25 francs and a conversion of 1 mile ≈ 1.609 km, the gross per-km rates are:

$$\frac{330{,}000 \times 25}{1.609} \approx 5{,}128{,}000 \text{ francs/km (cut-and-cover gross)},$$
$$\frac{500{,}000 \times 25}{1.609} \approx 7{,}768{,}000 \text{ francs/km (tunneling gross)}.$$

These gross rates are consistent with Cottereau (2004), who reports an all-in cost of approximately 5,000,000 francs per km for the Parisian metro without expropriation, consistent with the Bienvenüe network being predominantly cut-and-cover.. However, these gross rates cover the full cost per kilometre of metro infrastructure, including fixed technology costs, station construction, and tunnel segments. Since fixed costs ($\vartheta_0$) and station costs ($\vartheta_1$) are accounted for separately in equation (30), the segment-level rates $\vartheta^{\text{cc}}$ and $\vartheta^{\text{tun}}$ reflect only the remaining 55% of the gross per-km cost, after netting out the 15% fixed-cost share and the 30% station-cost share:

$$\vartheta^{\text{cc}} = 0.55 \times \frac{330{,}000 \times 25}{1.609} \approx 2{,}820{,}000 \text{ francs/km},$$
$$\vartheta^{\text{tun}} = 0.55 \times \frac{500{,}000 \times 25}{1.609} \approx 4{,}272{,}000 \text{ francs/km}.$$

I use some assumptions to derive the remaining cost parameters. The total length of the Fulgence Bienvenüe metro project (60 km) is used as a benchmark, yielding a total budget of approximately 307,579,500 francs under a blended rate. Within this total budget, 15% is allocated to technology acquisition and other fixed costs:

$$\vartheta_0 = 307{,}579{,}500 \times 0.15 = 46{,}136{,}926 \text{ francs}.$$

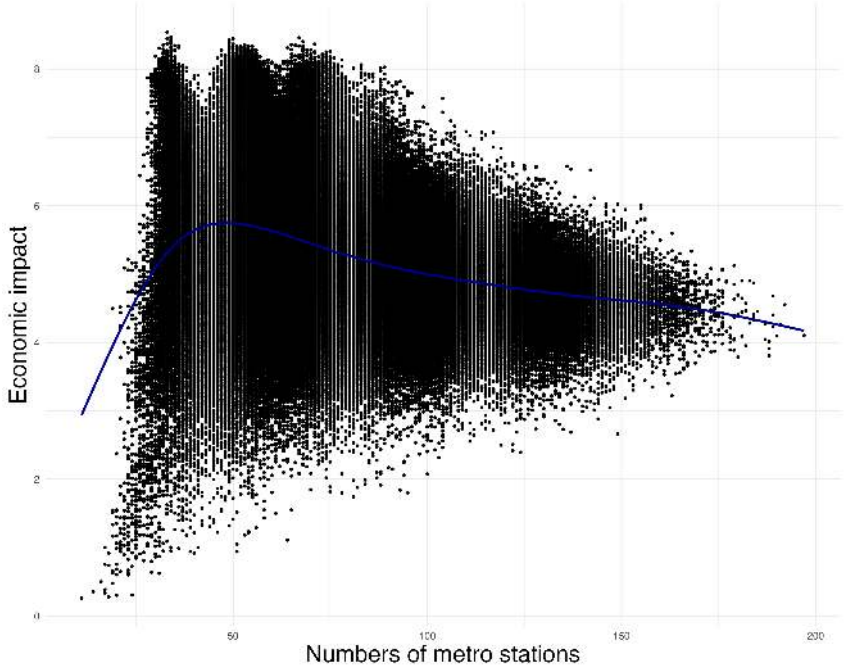
Furthermore, I assume that 30% of the total construction costs per km is dedicated to stations. Given that, on average, three stations are built per km (since the average inter-station distance is 500 m), the cost per station is:

$$\vartheta_1 = 5{,}126{,}325 \times \frac{0.3}{3} = 512{,}632.5 \text{ francs}.$$

The Seine-crossing surcharge $\delta^{\text{Seine}}$ is calibrated to capture the additional cost of using a bored tube under the river, and is set equal to the difference between the deep-bore and cut-and-cover per-km rates over the typical length of a Seine-crossing segment.

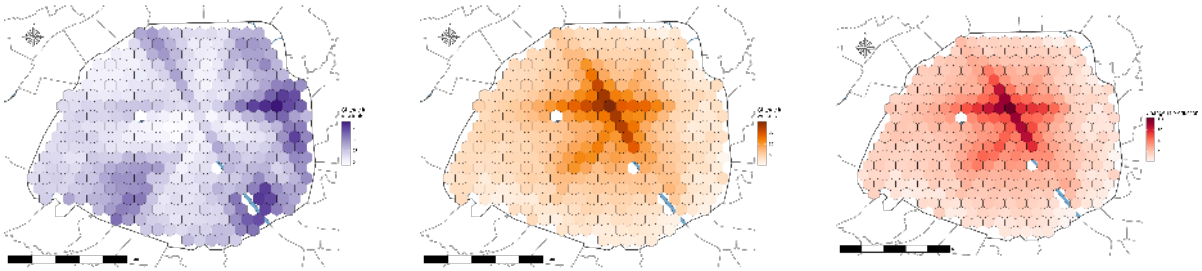
## C.2 Metro stations and economic impact - whole iterations

Figure A3
Metro size and economic impact for all iterations




## C.3 Baseline results - change in economic activities

Figure A4
Change in economic activities



(a) Residence employment (b) Workplace employment (c) Production

Notes: This figure shows changes in percentage points of location economic activities with respect to the initial. Panel (a) depicts the change in residence employment. Panel (b) shows the change in workplace employment. Panel (c) illustrates the change in production. Darker shades indicate areas with higher positive changes with respect to the baseline.

## C.4 Standardized coefficients

To assess the relative importance of each location characteristic in explaining metro station frequency, Table A3 replicates the regressions of Table 1 using standardized (z-scored) variables. Specifically, prior to estimation, each variable, both the dependent

variable and all regressors, is transformed to have mean zero and standard deviation one. The resulting coefficients are standardized $\beta$ coefficients: a coefficient of $\beta$ indicates that a one-standard-deviation increase in the regressor is associated with a $\beta$ standard-deviation change in metro station frequency. This normalization removes the influence of different units and scales, allowing direct comparison of effect magnitudes across productivity, amenities, centrality, and construction costs.

Table A3
Determinants of metro stations frequency (standardized coefficients)

| | Metro Station Frequency (std.) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Log. Productivity | 0.60*** | | | | 0.50*** |
| | (0.12) | | | | (0.12) |
| Log. Costs | -0.12 | 0.24*** | 0.27*** | 0.09 | -0.18** |
| | (0.10) | (0.08) | (0.09) | (0.08) | (0.09) |
| Log. Low-skilled Amenities | | 0.24*** | | | 0.25** |
| | | (0.07) | | | (0.11) |
| Log. High-skilled Amenities | | | 0.18*** | | -0.37*** |
| | | | (0.05) | | (0.14) |
| Log. distance to CC | | | | -0.48*** | -0.39*** |
| | | | | (0.12) | (0.11) |
| Observations | 334 | 334 | 334 | 334 | 334 |
| $R^2$ | 0.14 | 0.11 | 0.10 | 0.17 | 0.20 |

Notes: Each column reports estimates from a separate regression using z-scored variables (mean 0, standard deviation 1). Coefficients are standardized $\beta$s: a one-standard-deviation increase in the regressor is associated with a $\beta$ standard-deviation change in station frequency. Regressions use potential metro stations derived from 500m×500m grid cells. Dependent variable is the frequency of metro stations within 500 SA runs. Heteroskedasticity-robust standard errors are in parentheses and ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

## C.5 Targeting specific population

This appendix formalizes the general planner maximization problem introduced in Section 4.6, where the planner assigns an explicit weight $z \in [0, 1]$ to the utility gains of low-skilled workers.

**General planner problem.** Let $z \in [0, 1]$ denote the *political representation weight* of low-skilled workers. The planner solves,

$$\begin{aligned} \max \quad & (1 - z) \times \widehat{\overline{U}}_H(M; \theta) + z \times \widehat{\overline{U}}_L(M; \theta) \\ \text{subject to} \quad & \mathcal{B} \geq \text{Cost}(M), \end{aligned} \tag{C.1}$$

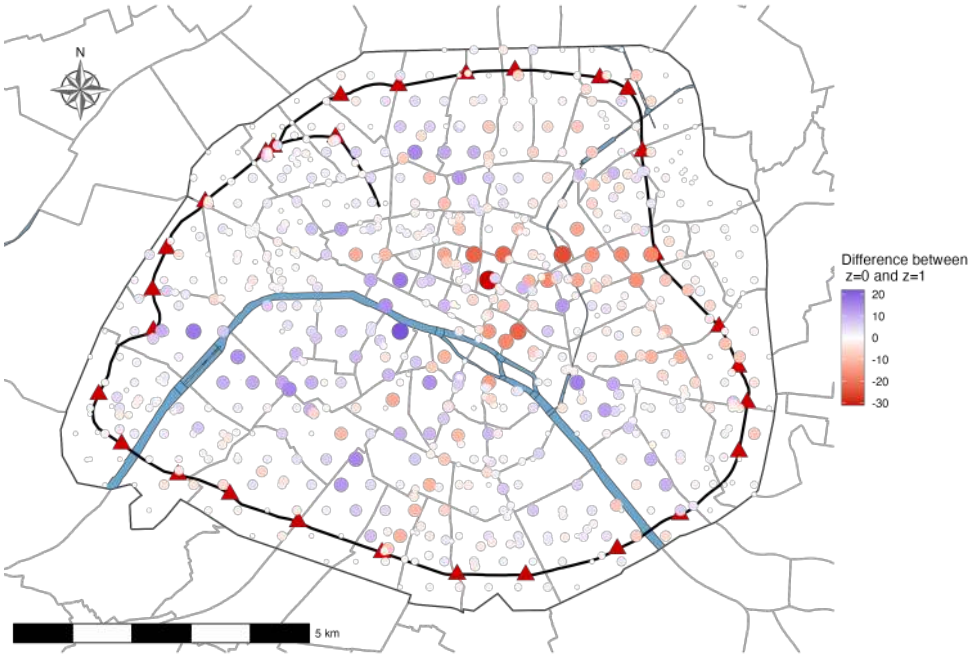
where $\widehat{\overline{U}}_H$ and $\widehat{\overline{U}}_L$ represent the indirect utility change in percentage points for high-skilled and low-skilled workers, respectively, $\mathcal{B}$ denotes the total budget for metro construction, and $\text{Cost}(M)$ represents the construction costs of metro $M$.

**Relationship to the utilitarian and egalitarian cases.** The two objectives defined in Section 4.6 are special cases of equation (C.1). The utilitarian problem, which assigns equal weight to both groups, corresponds to $z = 1/2$. The Rawlsian egalitarian problem, which maximizes the minimum utility between the two groups, can be interpreted as the limiting case $z \to 1$ when low-skilled workers are the worst-off group (i.e., $\widehat{\overline{U}}_L < \widehat{\overline{U}}_H$), as in this case the egalitarian objective concentrates entirely on their utility. The polar cases $z = 0$ and $z = 1$ correspond to a planner who targets exclusively high-skilled or low-skilled workers, respectively. These extremes admit a natural political economy interpretation: a planner with $z = 0$ effectively represents high-skilled residents only, as would arise, for instance, from franchise restrictions, elite capture of municipal institutions, or disproportionate lobbying by property or business interests. Whereas a planner with $z = 1$ reflects an objective entirely shaped by the welfare of a working-class electorate.

# D Optimal Upzoning Design

The framework developed in section 3 is not specific to transport infrastructure. The simulated annealing algorithm can be applied to any high-dimensional discrete policy choice that enters the quantitative urban model through its fundamentals, provided the planner can rank candidate policies by an economic objective $W$ that the QUM is able to evaluate. In this appendix, I illustrate this point by replacing the metro-design problem by a counterpart problem on the supply side of the land market: the planner no longer chooses where to lay metro lines but where to upzone, that is, where to relax local housing supply restrictions and allow new housing units to be built. Holding fundamentals, transport technology and the structural parameters fixed, the planner reallocates a fixed envelope of new units across locations and sampling welfare-improving allocations through the same simulated annealing logic as in the transport application.

Figure A5
Difference between $z = 0$ and $z = 1$ sampled metro networks




Notes: This figure shows the change in frequency of metro stations occurring in the simulated annealing runs between the $z = 0$ (targeting high-skilled workers) and the $z = 1$ scenarios (targeting low-skilled workers). The economic objective is

$\max: (1 - z) \times \widehat{U}_H(M; \theta) + z \times \widehat{U}_L(M; \theta)$. Blue indicates a positive change in frequency, while red indicates a negative change. The size of the dots is proportional to the absolute value of the frequency change.

## D.1 Optimization Environment

### D.1.1 Planner problem

The planner chooses an upzoning allocation $\Pi = (\bar{h}_1, \ldots, \bar{h}_{\mathbb{N}})$, where $\bar{h}_n \geq 0$ denotes the number of new housing units allowed in location $n$. The allocation is constrained by a fixed upzoning envelope $\bar{H}$, capturing the total amount of new construction the planner is allowed to authorize across the whole city,

$$\sum_{n \in \mathbb{N}} \bar{h}_n \;=\; \bar{H}, \qquad \bar{h}_n \geq 0 \;\forall n. \tag{D.1}$$

Given the baseline housing supply $H_n^S$ defined by equation (9), the counterfactual housing supply under upzoning allocation $\Pi$ is,

$$H_n^{S,\prime} \;=\; H_n^S \,+\, \bar{h}_n \;=\; H_n^S \, \widehat{H}_n^S(\Pi), \qquad \widehat{H}_n^S(\Pi) \;\equiv\; \frac{H_n^S + \bar{h}_n}{H_n^S}, \tag{D.2}$$

so the policy variable $\Pi$ enters the QUM through a multiplicative shock $\widehat{H}_n^S(\Pi)$ on the housing supply, with the prime in $H_n^{S,\prime}$ denoting the counterfactual policy. The matrix of commuting costs $d_{ni}$ is held at its baseline value: transport technology does not change in this experiment.

By analogy with the metro planner's problem, the upzoning planner solves,

$$\max_{\Pi \in \mathcal{H}} W(\Pi; \theta) + \iota_\Pi, \tag{D.3}$$

where $\mathcal{H} = \{\Pi \in \mathbb{R}_+^{\mathbb{N}} : \sum_n \bar{h}_n = \bar{H}\}$ is the simplex of admissible upzoning allocations scaled by $\bar{H}$, $W(\Pi; \theta)$ is the expected utility computed at the counterfactual spatial equilibrium under housing supply $H^{S,\prime}$, and $\iota_\Pi$ is a Gumbel idiosyncratic shock with inverse-temperature $\sigma$ capturing the planner's residual preferences. As in section 3, the probability ratio of two candidate allocations is,

$$\frac{\pi(\Pi'; \theta)}{\pi(\Pi; \theta)} \;=\; \frac{\exp(\sigma \, W(\Pi'; \theta))}{\exp(\sigma \, W(\Pi; \theta))}, \tag{D.4}$$

which makes the same simulated annealing logic immediately applicable.

### D.1.2 Perturbation function

The state space is the upzoning vector $\Pi$. Three local moves form the perturbation function $\Psi$, mirroring the three perturbations used in the transport application (deviation, new line, removal). For all moves, the total envelope $\bar{H}$ is preserved exactly. To avoid clashes with $\delta$ (the gamma-function constant in the welfare expression) and $\xi$ (already used to denote the historical transport network), I denote the move sizes by $q_\Pi$ and $q_c$:

(i) *Reallocation.* A donor location $n$ with $\bar{h}_n > 0$ and a receiver location $n' \neq n$ are drawn at random, and a quantity $q_\Pi = q_{\Pi,\text{frac}}\,\bar{H}$ is transferred from $n$ to $n'$, with $q_{\Pi,\text{frac}} = 0.05$. This move smoothly redistributes upzoning between locations and is the upzoning analog of a metro line "deviation".

(ii) *Add upzoned location.* A previously non-upzoned location $n^*$ (with $\bar{h}_{n^*} = 0$) is selected at random, and a fixed chunk $q_c = q_{c,\text{frac}}\,\bar{H}$ is withdrawn proportionally from the currently upzoned locations and assigned to $n^*$, with $q_{c,\text{frac}} = 0.10$. This move expands the spatial support of upzoning, mirroring the creation of a "new line".

(iii) *Remove upzoned location.* An upzoned location $n$ is selected at random; its allocation $\bar{h}_n$ is reset to zero and redistributed proportionally across the remaining upzoned locations. This contracts the spatial support of upzoning, mirroring the "removal" of a metro line.

The three moves are drawn with probabilities $\mathbf{p} = (40/90,\ 30/90,\ 20/90)$, as in the metro application.

### D.1.3 Welfare via exact hat algebra

For a candidate allocation $\Pi'$, computing the economic objective $W(\Pi';\theta)$ requires solving a counterfactual spatial equilibrium. As in the transport application, I exploit exact hat algebra to avoid resolving the equilibrium from fundamentals at every iteration. Because only the housing supply changes, commuting costs are held at their baseline value ($\widehat{d}_{ni} = 1$) and there is no agglomeration in this canonical illustration ($\widehat{B}_n = \widehat{A}_i = 1$), so the policy enters the canonical QUM of section 2 exclusively through the multiplicative

housing supply shock,

$$\widehat{H}_n^S(\Pi) \;=\; \frac{H_n^{S,\prime}}{H_n^S} \;=\; \frac{H_n^S + \bar{h}_n}{H_n^S}. \tag{D.5}$$

Using "hat" expression with $\widehat{x} = x'/x$, every equation can be expressed in terms of change relative to the baseline equilibrium.

The relative change in wages, from the firm's first-order conditions (7)–(8) with $\widehat{A}_i = 1$, can be expressed as,

$$\widehat{w}_i \;=\; \widehat{Q}_i^{-\frac{1-\alpha}{\alpha}}, \tag{D.6}$$

which is function of changes in rent ($\widehat{Q}_i$).

The relative change in location choice probabilities, derived from equation (2), can be expressed as,

$$\widehat{\lambda}_{ni} \;=\; \frac{\left(\widehat{w}_i\right)^{\epsilon}\left(\widehat{Q}_n^{\beta}\right)^{-\epsilon}}{\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}}\lambda_{kl}\left(\widehat{w}_l\right)^{\epsilon}\left(\widehat{Q}_k^{\beta}\right)^{-\epsilon}}, \tag{D.7}$$

which is function of changes in living cost ($\widehat{Q}_n$) and changes in wages ($\widehat{w}_i$).

The counterfactual residential and workplace employment can be expressed as,

$$\widehat{R}_n R_n \;=\; \sum_{i\in\mathbb{N}} \lambda_{ni}\,\widehat{\lambda}_{ni}\,R_{\mathbb{N}}, \tag{D.8}$$

$$\widehat{L}_n L_n \;=\; \sum_{n\in\mathbb{N}} \lambda_{ni}\,\widehat{\lambda}_{ni}\,R_{\mathbb{N}}, \tag{D.9}$$

which are functions of changes in commuting probabilities and the (fixed) total population $R_{\mathbb{N}}$ in the closed-city setting.

The relative change in residential housing demand, derived from equation (12), can be expressed as,

$$\widehat{H}_n^R H_n^R \;=\; \beta\left[\sum_{i\in\mathbb{N}} \frac{\lambda_{ni|n}^R\left(\widehat{w}_i\right)^{\epsilon}}{\sum_{l\in\mathbb{N}}\lambda_{nl|n}^R\left(\widehat{w}_l\right)^{\epsilon}}\,\widehat{w}_i w_i\right]\frac{\widehat{R}_n R_n}{\widehat{Q}_n Q_n}, \tag{D.10}$$

and the relative change in commercial housing demand can be expressed as,

$$\widehat{H}_n^L H_n^L \;=\; \frac{1-\alpha}{\alpha}\,\widehat{L}_n L_n\,\frac{\widehat{w}_n w_n}{\widehat{Q}_n Q_n}. \tag{D.11}$$

Finally, the floorspace market-clearing condition closes the system. The augmented housing supply, scaled by the upzoning shock $\widehat{H}_n^S(\Pi)$, must equal the total housing de-

mand at every location,

$$\widehat{Q}_n^{(1-\mu)/\mu} H_n^S \widehat{H}_n^S(\Pi) \;=\; \widehat{H}_n^R H_n^R \;+\; \widehat{H}_n^L H_n^L. \tag{D.12}$$

The system is iterated in $\widehat{Q}_n$ until the floorspace market clears at every location, taking the upzoning shock $\widehat{H}_n^S(\Pi)$ as the only exogenous policy input.

Finally, the change in expected utility induced by the upzoning policy follows directly from the population mobility condition (5), since the gamma-function constant $\delta$ cancels in the ratio,

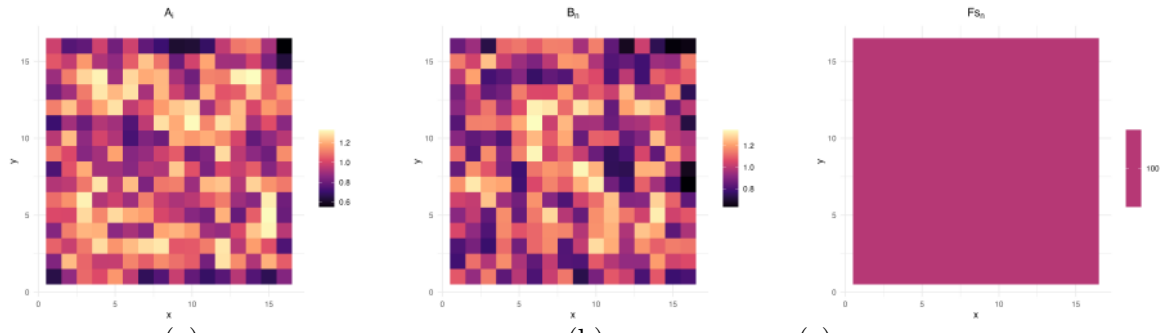
$$\widehat{\overline{U}}(\Pi;\theta) \;=\; \Big[\sum_{k\in\mathbb{N}}\sum_{l\in\mathbb{N}} \lambda_{kl}\,(\widehat{w}_l)^{\epsilon}\,(\widehat{Q}_k^{\beta})^{-\epsilon}\Big]^{1/\epsilon}. \tag{D.13}$$

The planner's economic objective is then the welfare gain $W(\Pi;\theta) = \widehat{\overline{U}}(\Pi;\theta)$ expressed in percentage utility points, as in section 3.

## D.2 Numerical example

To illustrate the optimal upzoning design, I reuse the same synthetic city as in section 2: a $16\times16$ grid with 500-metre cell spacing, walking-only transport at 5 km/h, the productivity $(A_i)$ and amenity $(B_n)$ draws shown in Figure A6, and a uniform baseline housing supply $H_n^S = 100$ at every location. Structural parameters are unchanged: $\alpha = 0.7$, $\beta = 0.25$, $\epsilon = 5$, $\mu = 0.4$, $\kappa = 0.01$.

Figure A6
Upzoning experiment: Spatial distribution of fundamentals

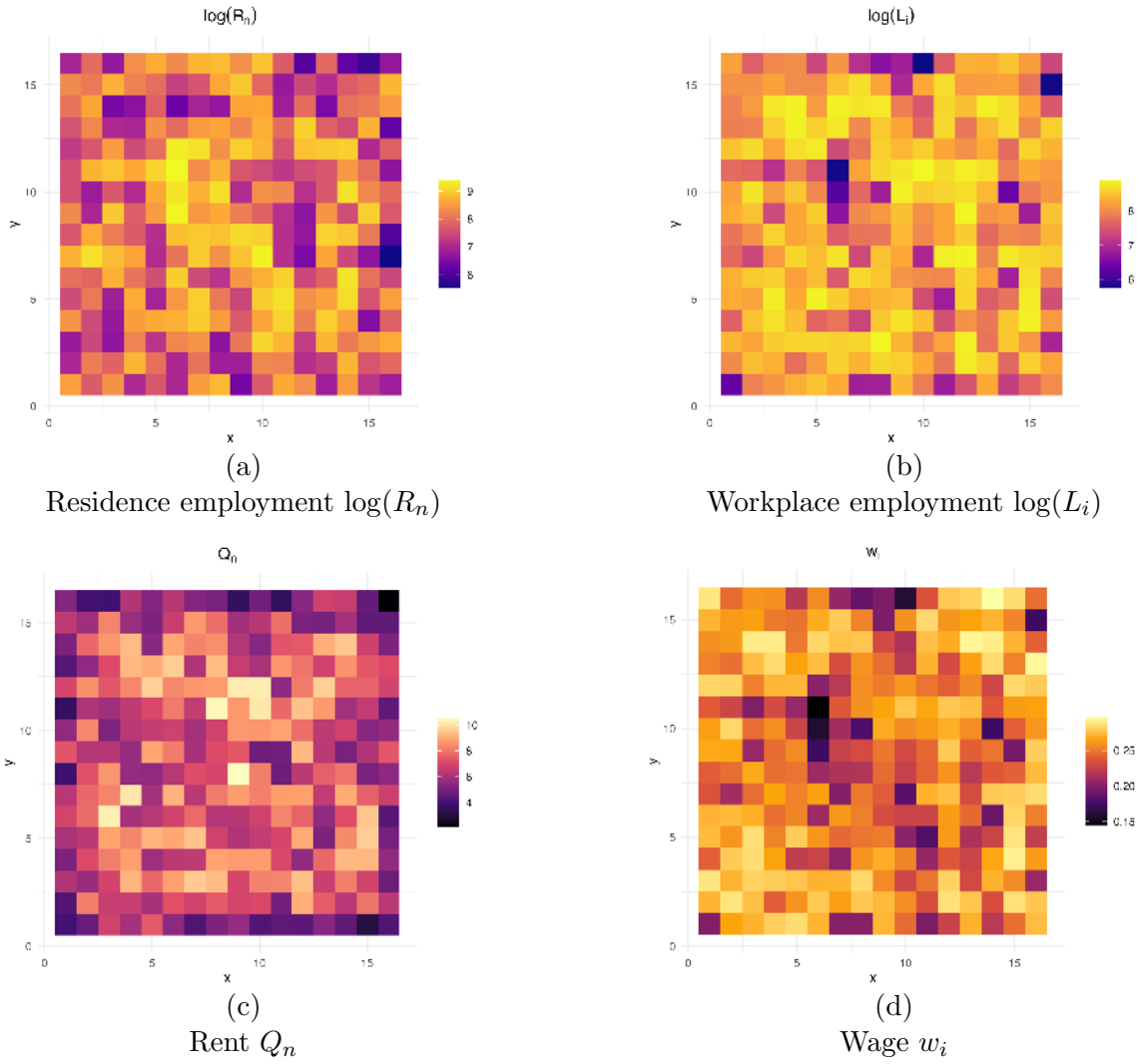


(a) Productivity $A_i$ (b) Amenities $B_n$ (c) Baseline housing supply $H_n^S$

Notes: This figure presents the spatial distribution of fundamentals used in the upzoning experiment. Panel (a) shows productivity $A_i$, Panel (b) shows amenities $B_n$, and Panel (c) shows the (uniform) baseline housing supply $H_n^S = 100$. The fundamentals are identical to those of the metro experiment in Section 2.

The initial spatial equilibrium $\{R_n, L_i, Q_n, w_i\}$ (Figure A7) serves as the baseline against which upzoning shocks are evaluated.

Figure A7
Upzoning experiment: Initial spatial equilibrium




(a) Residence employment $\log(R_n)$

(b) Workplace employment $\log(L_i)$

(c) Rent $Q_n$

(d) Wage $w_i$

Notes: This figure shows the initial walking-only spatial equilibrium used as the baseline for evaluating upzoning shocks. The fundamentals are those of Figure A6, and the structural parameters are calibrated as in Section 2.

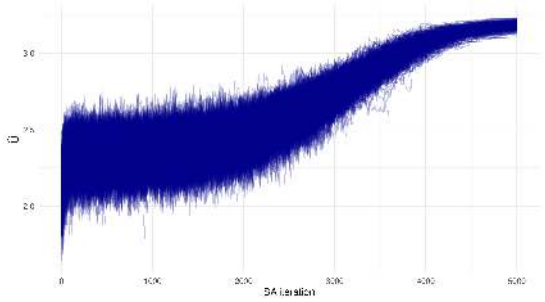
The upzoning envelope is set to $\bar{H} = 2{,}500$ new housing units, i.e. approximately a 10% increase in the city-wide housing stock relative to a baseline of $\mathbb{N} \times 100 = 25{,}600$ units. Each independent SA chain is initialized at iteration 1 by drawing $k_{\text{init}} = 16$ locations at random and distributing the envelope across them with weights drawn from $\mathcal{U}(0, 100)$. Subsequent iterations apply the perturbation function $\Psi$ described above. I run 500 independent chains of 4,000 iterations each, with an exponential cooling schedule $\sigma_1 = 1$ to $\sigma_K = 1{,}000$.

## D.3 Simulation results

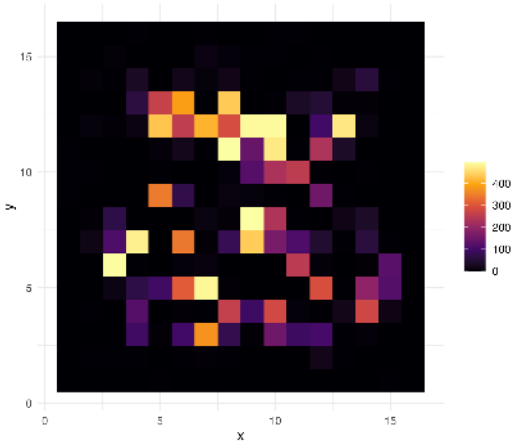
Panel (a) of Figure A8 shows the convergence of the 500 independent SA chains. Welfare gains converge to roughly 3.18% utility gain on average, confirming that the algorithm

successfully samples high-welfare upzoning allocations. Panel (b) shows the frequency with which each location is upzoned in the final iteration of the 500 runs; the most strategically attractive cells are upzoned in up to 450 (over 500) chains, while peripheral cells are upzoned in fewer than 50 (over 500).

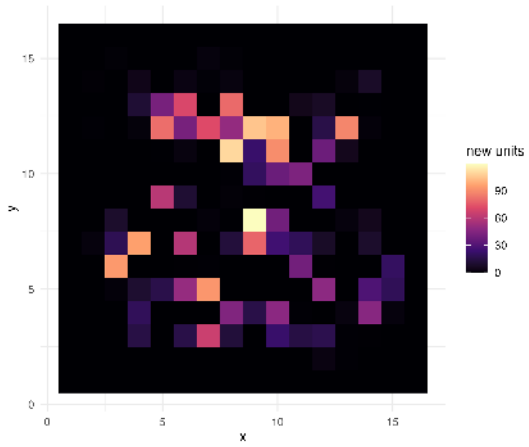
Figure A8
Upzoning experiment: Simulated annealing results



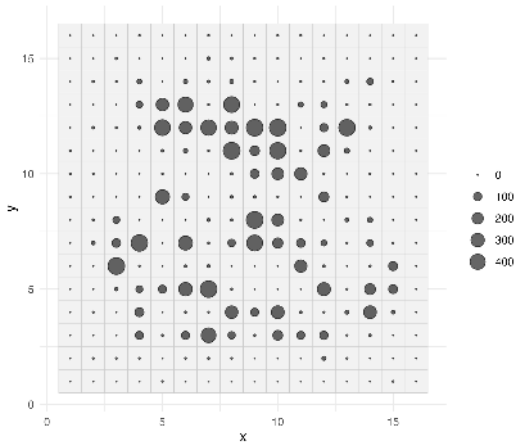

(a)
Welfare convergence ($\widehat{\overline{U}}$)




(b)
Upzoning frequency by location




(c)
Mean new units per cell (across runs)




(d)
Upzoning frequency (bubble map)

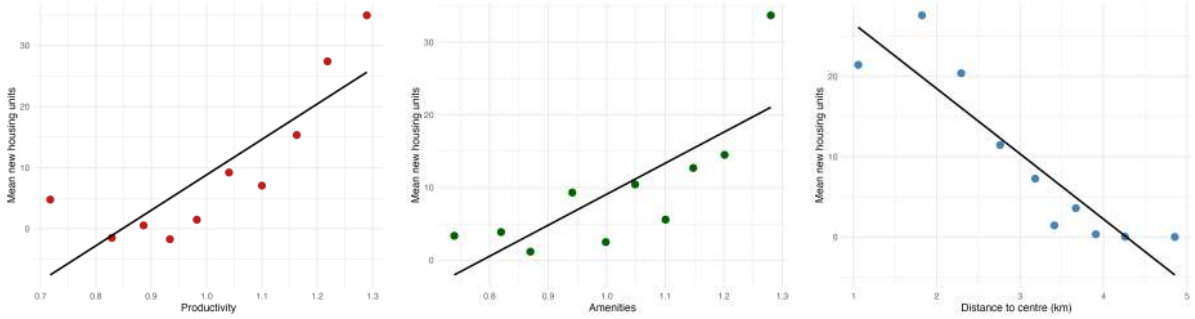
Notes: Panel (a) shows the trajectory of the welfare gain $\widehat{\overline{U}}$ along each of the 500 independent SA chains. Panel (b) shows the heatmap of the frequency with which each location is upzoned in the final iteration of the 500 chains. Panel (c) reports the mean number of new housing units allocated to each location across the 500 chains. Panel (d) is a bubble-map representation of the upzoning frequency, with marker size proportional to the number of chains that allocate at least one new unit to the location.

To probe the structural drivers of optimal upzoning, Figure A9 reports binned scatter plots of the mean number of new units allocated to a location against the three location-specific factors: productivity ($A_i$), amenities ($B_n$), and distance to the city center ($d_{cc}$). For Panels (a) and (b), the mean allocation is residualized on $\log(d_{cc})$, longitude and latitude, in order to isolate the role of fundamentals beyond mechanical centrality effects.

Three patterns emerge. First, the optimal upzoning allocation is positively correlated

with amenities $B_n$: high-amenity locations attract more new units, consistent with the residential-utility channel: relaxing housing supply where workers most want to live yields the largest welfare gain, because the marginal entrant captures both the high-amenity premium and the rent relief afforded by the local supply shift. Second, the allocation is positively correlated with productivity $A_i$, reflecting the fact that targeting these locations are crucial to access high-wages. Third, housing allocations decrease with distance to the city center, reflecting that central locations enjoy the lowest commuting costs and therefore the highest commuting-market access, which makes a marginal housing unit there valuable for a much larger pool of potential commuters than a peripheral unit.

Figure A9
Upzoning experiment: Allocation vs. fundamentals

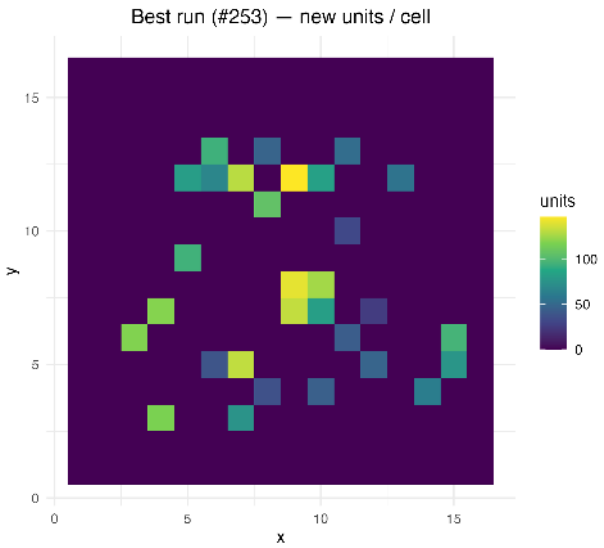



(a) vs. productivity $A_i$ (b) vs. amenities $B_n$ (c) vs. distance to centre

Notes: Binned scatter plots of the mean number of new housing units allocated to each location (across the 500 SA chains, at the final iteration) against location-specific factors. In Panels (a) and (b) the dependent variable is residualized on $\log(d_{cc})$, longitude and latitude to absorb mechanical centrality effects. Panel (c) reports the raw relationship.

Finally, Figure A10 shows the allocation of the best-performing chain, which achieves a welfare gain of ≈3.25%. The optimal allocation concentrates new units in ≈30 cells, with an average of ≈80 units per upzoned cell (peak ≈130 in a single cell located near the centre of the grid). This pattern matches the binned-scatter evidence in Figure A9 (units are concentrated in central, high-amenity, high-productivity cells), suggesting that the SA algorithm is consistent across the welfare distribution of sampled allocations.

Figure A10
Upzoning experiment: Best-performing allocation




Notes: This figure shows the upzoning allocation of the SA chain that achieves the highest welfare gain at the final iteration. The color scale corresponds to the number of new housing units allocated to each location. The total across all locations equals the upzoning envelope $\bar{H} = 2{,}500$ units.

Taken together, these results illustrate that the simulated annealing framework developed in section 3 extends naturally to the supply side of the land market. The same general-equilibrium logic that motivates an evaluation of transport networks, workers reallocating across residence and workplace, rents adjusting, welfare integrating all general equilibrium responses, applies to upzoning shocks, and the same algorithm samples welfare-improving allocations. Beyond this numerical example, the methodology can be applied to historical episodes of housing-supply liberalization or to forward-looking policy experiments where a city planner faces an upzoning envelope and must choose where to allow new construction.