Transport Infrastructures and Income Disparities Within Cities*

Morgan Ubeda†

March 3, 2021 – Last version here.

Abstract

This paper studies the effect of transportation networks on spatial inequalities and redistribution within metropolitan areas. To do so, I build and calibrate a spatial equilibrium model of a city that features non-homotheticicities and worker heterogeneity, allowing to capture rich patterns of workers sorting on commute costs and amenities. I then calibrate the model to the Paris urban area and use counterfactual simulations to study the effects of a) the Regional Express Rail and b) restricting car use in the city center. I find that on top of having a strong contribution to suburbanization and reducing welfare inequalities, the public transport network reduced income segregation in the area. Turning to the prospective effects of banning cars in the city center, the model predicts a reduction of the income disparities between Paris and its suburbs, at the cost of a substantial welfare loss.

Keywords commuting, amenities, income sorting, stratification

*I want to thank Pierre-Philippe Combes, Sonia Paty, Gilles Duranton, Jessie Handbury, Hans Koster and participants at the 14th UEA meeting in Philadelphia and 2019 Public Policy and Cities workshop in Lyon for their precious comments. I also thank Sylvie Charlot for helping me with accessing the restricted data. I gratefully acknowledge financial support from GATE, Université Lyon 2, the Fulbright Program and the Agence Nationale de la Recherche (ANR) research program ANR-18-CE41-0003-02. The data was accessed through CASD, who recieved financial support from ANR research program ANR-10-EQPX-17

†Université de Lyon, Université Lumière Lyon 2, GATE Lyon Saint-Étienne, UMR 5824, 93 Chemin des Mouilles 69131 Écully FRANCE. ubeda@gate.cnrs.fr

1 Introduction

In Europe, almost a quarter of greenhouse gas emissions comes from the transportation sector, of which 72% comes from road transport.[1]. In the US, these figures are respectively 28% and 59%.[2] To fight global warming an local pollution, cities around the world are trying to promote greener ways of commuting. In Paris for instance, the city plans to ban all thermal vehicles before 2030.[3] On the other hand, segregation and spatial inequalities are another major challenge faced by cities all over the world. With the French Riots and the more recent Yellow Vests movement, anti-gentrification protests in London, Leipzig, Brooklyn or Seattle, the negative consequences of spatial inequalities can be seen in all major cities in the world.

Most of the literature on sorting has emphasized the role of amenities (Lee and Lin, 2018, Koster and Rouwendal, 2017, Glaeser et al., 2018, Garcia-López et al., 2018, Couture and Handbury, 2020, Couture et al., 2018), school spending (Epple et al., 2001, Calabrese et al., 2006, Rothstein, 2006, e.g.) and place-based policies (e.g. González-Pampillón et al., 2019) on income stratification. [4] Yet, comparisons between cities show that polycentric cities that concentrate economic activity in peripheral sub-centers are less segregated than their monocentric counterparts (Garcia-López and Moreno-Monroy, 2018). This seems to indicate that job location and commuting are to be taken into account to explain spatial income distributions, so that one can wonder if it is possible to fill two needs with one deed and use transport improvements to reduce segregation.

To assess the stratifying and redistributive effect of transportation policies, I rely on the calibration and simulation of a quantitative spatial equilibrium model of a city. More precisely, I extend the model of Ahlfeldt et al. (2015) to introduce workers heterogeneity, both in terms of observable skill classes and unobserved

[1] *Greenhouse gas emissions from transport in Europe.* European Environment Agency, https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12, accessed on August 21, 2020.

[2] *Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990–2018.* EPA 430-R-20-02.

[3] *Plan Climat*, available in French and English at https://www.paris.fr/pages/nouveau-plan-climat-500-mesures-pour-la-ville-de-paris-5252, accessed on August 21, 2020.

[4] Two exceptions are the concurrent studies of Tsivanidis (2019) and Gaigné et al. (2019). Tsivanidis (2019) uses a similar model to estimate the welfare effects of the TransMilenio rapid bus network in Bogotá. However, he does not focus on income stratification, nor does he consider the effects of legislation on car use. Gaigné et al. (2019) on the other hand do not model transport mode choice, and thus do not explore the effects of precise policies.

talent/productivity. Within skill group, the income sorting of workers is governed by non-homotheticities in the preferences for housing, stemming from Stone-Geary preferences. These preferences imply that the willingness to pay rents in return for higher amenities and shorter commutes increases with income. I then calibrate the model to the Paris region and use model simulations to evaluate two transport policies: a) the Regional Express Rail (RER) and b) banning cars in the city.

The case of Paris is interesting for two reasons. First, it is a major European city, comparable in size, segregation and inequalities to other major metropolitan areas. Second, the impact of the RER has been studied previously using convincing IV strategies (Mayer and Trevien, 2017, Garcia-López et al., 2017). It is thus possible to benchmark the model against some known results in the literature.

I find that the Regional Express Rail had a negative effect on spatial income inequalities. Overall, the RER train system decreases the between-municipality coefficient of variation of mean incomes by 1.9%, and the income premium of Paris with respect to its suburbs (10 to 15 kilometers away from the geographical center of the city) by 3.7%. Further, it has sizable positive welfare effects. Indeed, it accounts for 3.32% of the welfare of low-skilled workers and 2.56% of the welfare of high-skilled workers, reducing welfare inequalities by 0.8%.

Turning to the car ban counterfactual, when focusing on commuting costs the model predicts that the policy would foster suburbanization, leading to a decrease of the income premium of the city relative to the rest of region. Comparing with the close suburbs, the income premium of the city would drop by 14% (1500 €). This would however come at the cost of a substantial welfare loss, of 2.6% for low-skilled and 3.2% for high-skilled workers. However, these effects depend on the amenity gains from the policy. It would require a more than 10% increase in amenities in Paris from pollution reduction and regained floor space for the policy to break even and start having a positive welfare effect. At that point, the sorting effects of the policy would be reversed: the income premium of the city would increase by 7%. As richer workers bid for floor space in this high-amenity center, the effect on welfare inequalities also flips, and the policy starts benefiting more the affluents.

The paper also provides new within-city estimates of agglomeration effects on total factor productivity, skill bias of agglomeration effects and residential amenities spillovers. Agglomeration effects and residential amenity spillovers are estimated using model-based instruments, as introduced by Allen et al. (2020). I find agglomeration effects comparable in size to previous results using between-cities designs (Ahlfeldt and Pietrostefani, 2019), but substantially lower than other within-city estimates as Ahlfeldt et al. (2015) and Tsivanidis (2019). Keeping city structure constant, the elasticity of TFP to total city population is 0.04.

The first contribution of this paper is to quantify the effects of transport policies on spatial income disparities. Several studies have shown the decentralizing effects of public transit and road infrastructures on employment and population gradients (e.g. Mayer and Trevien, 2017, Garcia-López et al., 2017, Garcia-López, 2012, Gonzalez-Navarro and Turner, 2018, Baum-Snow, 2007), but none has quantified the effects of public transports or road accessibility on income segregation.

Second, the paper extends the existing results on the effects of the Regional Express Rail on suburbanization. Mayer and Trevien (2017) use an IV strategy to estimate the causal impact of being connected to the Parisian Regional Express Rail network on a subset of municipalities. They conclude to a sizeable suburbanization effect on both employment and residential populations, with a stronger effect for high skilled workers. I add to their results by estimating the impact of the RER network on income sorting in the area.

Finally, the present paper also contributes to the literature on within-city quantitative spatial equilibrium models. Several recent studies in urban economics use a similar structural approach (Couture and Handbury, 2020, Almagro and Domínguez-Iino, 2019, Gaigné et al., 2019, Tsivanidis, 2019). I provide several robustness checks for the model fit, and make the case that this class of models can be used as stand-alone tools for policy evaluation. Indeed I estimate the model without targeting any particular policy and show that the model-based estimates are in line with reduced-form results on the RER network from Mayer and Trevien (2017). This lends credibility to using the model in cases where no natural experiment is available. Moreover, I estimate the housing consumption parameters that govern workers sorting on expenditure micro-data without targeting income disparities, and show that the model is able to fit the income sorting patterns in the data with those theoretically consistent parameters. Further, while I estimate amenities as structural residuals of the model as in Albouy (2016) and Ahlfeldt et al. (2015), I show that model-based amenities strongly correlate with observed amenities.

The remainder of the paper is organized as follows. Section 2 presents the model and discusses the mechanisms that lead to income sorting. Section 3 describes the estimation and calibration of the parameters of the model and local amenities. Section 4 discusses results. Finally, section 5 concludes.

2 Model

This section outlines the model and discusses workers sorting. The general structure of the model is similar to Ahlfeldt et al. (2015), with the addition of workers heterogeneity and Stone and Geary preferences.

2.1 Workers behaviour

A city or urban area is composed of S municipalities, denoted by i or j, each endowed with some land L_j. There are H workers in the city. Each worker has to choose in which municipality to live and in which municipality to work. Workers are perfectly mobile and receive their income from supplying labour to firms in their workplace. Firms use labour and floor space to produce a final good costlessly traded with the rest of the world.

There are two sources of heterogeneity in the model. First, workers are endowed with an observed type e, corresponding to their education level. Second, within skill classes, workers differ in their individual skills and abilities, denoted $l \in \mathbb{R}_+$. Following the canonical literature on the estimation of agglomeration economies (e.g. Combes et al., 2008) workers heterogeneity withing observed skill classes is modeled in terms of efficient labour supply differences. More precisely, a worker with ability l is assumed to supply l units of efficient labour. Therefore, given wages per efficient labour unit w_{je} for education e in municipality j, a worker with ability l simply receives an income of lw_{je}. The distribution of skills in the city for each type e is fix and denoted $\mathcal{F}_e$.

In what follows, education level indices are omitted when they are not necessary.

Conditional on her place of residence $i = 1, \ldots, S$ and her workplace $j = 1, \ldots, S$, agent n with ability l receives a wage lw_j, that she spends on a quantity x_{ijn} of the numéraire good and a quantity f_{ijn} of floor space. The numéraire is not subject to transport costs, and is therefore distributed at a constant price (normalized to unity) everywhere in the city. The budget constraint of n is thus

$$lw_j = Q_i f_{ijn} + x_{ijn}, \tag{1}$$

where Q_i is the residential floor space rent in municipality i.

Regarding workers preferences, I focus on the sorting of workers on the basis of local amenities, which precludes the use of homothetic preferences. Following Gaigné et al. (2019) and Tsivanidis (2019) and departing from the Cobb-Douglas specification in Ahlfeldt et al. (2015), I assume that workers have Stone and Geary preferences

$$U_{ijmn} \equiv z_{ijmn} B_{ijm} t_{ijm}^{-\tau_m} \left(\frac{x_{ijn}}{1-\beta} \right)^{1-\beta} \left(\frac{f_{ijn} - \underline{f}}{\beta} \right)^{\beta}, \tag{2}$$

where $B_{ijm} = B_i T_j$ are the local amenities perceived when living in i, working in j. They include B_i the proper residential amenities in i and the niceness of the workplace j besides its offered wage, T_j. Second, $t_{ijm}^{-\tau_m}$ is the utility cost of commuting between i and j using transport mode m, with t_{ij} the travel time and τ_m a mode-specific disutility parameter. The random variable z_{ijmn} captures idiosyncratic preferences of n for the commute ij and transport mode m, and $\beta \in (0,1)$ and $\underline{f} \geq 0$ are parameters that govern workers preferences for housing. $\underline{f}$ has a natural interpretation as an incompressible floor space consumption.

Stone and Geary preferences have many interesting properties. First, whenever $\underline{f} > 0$, the (indirect) marginal rate of substitution between floor space costs Q_i and local amenities B_{ij} is increasing with income. This induces a relatively higher willingness to pay for high amenity levels for rich households than for poor households. It provides a parsimonious and theoretically sound foundation for income sorting on the basis of amenities. When $\underline{f} = 0$ preferences are simple Cobb-Douglas.

Second, Stone and Gary preferences imply that the share of total income spent on housing is decreasing with income. This decrease is consistent with data on the housing consumption of French households. Indeed, our analysis of Expenditure Survey data in section 3.1 reports downward Engel curves ranging from 50% to 18% and shows that Stone and Geary preferences fit these curves well (cf Figure 2, section 3.1). This is in line with previous evidence using French data from Combes et al. (2018, p. 32, Table 6) who estimate that the share of housing in French households expenses is significantly decreasing in income. By maximizing (2) subject to the budget constraint (1), the individual demand for the private good (3), the individual demand for floor space (4), and the indirect utility of n when she chooses the commute ij (5) are respectively:

$$x_{ijn}^*(l) = (1-\beta)(lw_j - Q_i \underline{f}) \tag{3}$$

$$f_{ijn}^*(l) = \beta \frac{lw_j}{Q_i} + (1-\beta)\underline{f} \tag{4}$$

$$V_{ijmn}(l) = z_{ijmn} B_{ij} t_{ijm}^{-\tau_m} (lw_j - Q_i \underline{f}) Q_i^{-\beta}, \tag{5}$$

to the extent that these quantities are positive. When $lw_j - Q_i\underline{f} \leq 0$, i.e. when the worker cannot afford the incompressible floor space consumption in i by working in j, I set the indirect utility of the commute to $V_{ijmn}(l) = 0$.

In what follows, I assume that the idiosyncratic preference shock can be broken down into two components, $z_{ijmn} = \zeta_{ijmn}\xi_{ijn}$. The first term ζ_{ijmn} is the transport mode preference shock for worker n conditional on choosing commute ij, whilst the second one ξ_{ijn} captures idiosyncratic commute-specific preference shocks. Regarding the timing of the model, I assume that workers first learn about z_{ijn} and choose a commute (i.e. simultaneously decide on a workplace and a residential location) accordingly. After they choose their commute, they learn about the transport mode shock ζ_{ijmn} and decide on which transport mode to choose.

The model is then solved by backward induction. Conditional on having chosen commute ij, workers have to decide on a transport mode m. Assuming that ζ_{ijmn} are independently and identically Fréchet distributed, with scale parameters a_m and shape parameter $\theta > 1$, workers expected utility over transport modes conditional on ij is

$$\xi_{ijn}v_{ij}(l) \equiv \mathbb{E}\left[\max_m V_{ijmn}(l)\right] = \xi_{ijn}B_{ij}t_{ij}^{\frac{1}{\theta}}(lw_j - Q_i\underline{f})Q_i^{-\beta}, \tag{6}$$

with $t_{ij}^{\frac{1}{\theta}}$ the expected transport utility:

$$t_{ij}^{\frac{1}{\theta}} = \left[\sum_m a_m t_{ijm}^{-\tau_m\theta}\right]^{\frac{1}{\theta}}. \tag{7}$$

As in Ahlfeldt et al. (2015), I assume that ξ_{ijn} the idiosyncratic preference shocks for commutes are independent draws from Fréchet distributions with shape parameter $\epsilon > 1$. Standard discrete choice theory (cf. Ahlfeldt et al., 2015, for a detailed exposition) then yields the probability for a worker with skill l to choose commute ij:

$$\pi_{ij}(l) \equiv \Pr[v_{ijn}(l) > v_{kmn}(l),\ km \neq ij] = \frac{v_{ij}(l)^\epsilon}{\sum_{i=1}^S \sum_{j=1}^S v_{ij}(l)^\epsilon} \equiv \frac{v_{ij}(l)^\epsilon}{v(l)^\epsilon}, \tag{8}$$

with $v(l)$ the *ex ante* expected utility of a worker with skill level l. When $lw_j < Q_i\underline{f}$ however, utility is zero and so is the numerator of the choice probability. Since $\epsilon > 1$, these choice probabilities are still smooth and differentiable for any $w_j \in \mathbb{R}_+$ and any $Q_i \in \mathbb{R}_{++}$, as long as there is at least one commute in the city in which the worker can realize a positive utility.[5]

[5]Thereafter, I will implicitly assume $lw_j > Q_i\underline{f}$ when writing down choice probabilities. If a worker

The total probability to reside in i for a worker with skills l, $\pi_i^R(l)$ (respectively working in j, $\pi_j^M(l)$) is the sum over workplaces j (respectively dwelling places i) of the bilateral probabilities:

$$\pi_i^R(l) = \frac{\sum_{j=1}^S v_{ij}(l)^\epsilon}{v(l)^\epsilon}, \qquad \pi_j^M(l) = \frac{\sum_{i=1}^S v_{ij}(l)^\epsilon}{v(l)^\epsilon}. \tag{9}$$

Finally, the conditional probability to live in i when working in j is denoted $\pi_{ij|j}(l)$ and the probability to work in j conditional on living in i is denoted $\pi_{ij|i}(l)$:

$$\pi_{ij|j}(l) \equiv \frac{\pi_{ij}(l)}{\pi_{Mj}(l)}, \qquad \pi_{ij|j}(l) \equiv \frac{\pi_{ij}(l)}{\pi_{Ri}(l)}. \tag{10}$$

Armed with these choice probabilities, that describe the spatial distribution of workers conditional on wages, rents and amenities, we can now discuss sorting.

2.2 The sorting of workers

When $\underline{f} > 0$, workers exhibit direct sorting both at the workplace and in their residential location choice. High ability workers are willing to forego more consumption than low ability workers for an increase in residential amenities or a decrease in travel times. They are also willing to forego more wage per unit of efficiency for an increase in workplace amenities or a decrease in travel times.

More precisely, use the conditional residential choice probability $\pi_{ij|j}$ to define the rate of substitution between rents and some commute characteristic X_{ij} as the variation in rents in i necessary to keep the share of j workers living in i stable when the X_{ij} increase/decrease as

$$\left.\frac{\mathrm{d}Q_i}{\mathrm{d}X_{ij}}\right|_{\mathrm{d}\pi_{ij|j}(l)=0}(l) = -\frac{\partial_{X_{ij}}\pi_{ij|j}(l)}{\partial_{Q_i}\pi_{ij|j}(l)}. \tag{11}$$

Direct computation of these quantities yields the following proposition:

Proposition 2.1. *Whenever* $\pi_{ij|j}(l) > 0$, *we have* $\left.\frac{\mathrm{d}Q_i}{\mathrm{d}B_i}\right|_{\mathrm{d}\pi_{ij|j}(l)=0}(l) > 0$ *and* $\left.\frac{\mathrm{d}Q_i}{\mathrm{d}t_{ij}^{1/\theta}}\right|_{\mathrm{d}\pi_{ij|j}(l)=0}(l) > 0$. *Further, they are increasing in l if and only if* $\underline{f} > 0$.

These elasticities are always positive, showing that all workers need to be compensated by a decrease in rents when amenities decrease or travel times increase.

gets too poor relative to floor space prices in the city, so that they cannot reach their incompressible floor space demand in any municipality, then it is simply assumed that they opt out from the city and leave.

When $\underline{f} = 0$, i.e. when preferences are Cobb-Douglas, these elasticities boil down to $1/\beta$: every worker, rich or poor, skilled or unskilled, will keep her probability to choose a municipality constant when her rent increases by $1/\beta$% in exchange for a 1% increase in amenities or decrease in expected commuting times. In this case residential choice probabilities are independent of talent and wages: everything else equal, skilled and unskilled households make the same residential choices.

Whenever $\underline{f} > 0$ however, this elasticity is strictly increasing in l. This means that when amenities in i increase (or travel times between i and j improve), more productive and thus richer workers can accept a stronger increase in rents while keeping their probability to live in i constant. This is the basic direct sorting effect that is induced by non-homotheticities in housing demand, and that drives differences in residential location choices between rich and poor workers in the model, which mimicks the classical Alonso-Muth single-crossing property.

Turning to workplace choice, define in a similar fashion the rate of substitution between wages and commute characteristics, conditional on residential locations, as

$$\left.\frac{\mathrm{d}w_j}{\mathrm{d}X_{ij}}\right|_{\mathrm{d}\pi_{ij|i}(l)=0}(l) = -\frac{\partial_{X_{ij}}\pi_{ij|i}(l)}{\partial_{w_j}\pi_{ij|i}(l)}. \tag{12}$$

Then the following proposition follows

Proposition 2.2. *Whenever $\pi_{ij|j}(l) > 0$, we have $\left.\frac{\mathrm{d}w_j}{\mathrm{d}T_j}\right|_{\mathrm{d}\pi_{ij|i}(l)=0}(l) < 0$ and $\left.\frac{\mathrm{d}w_j}{\mathrm{d}t_{ij}^{1/\theta}}\right|_{\mathrm{d}\pi_{ij|i}(l)=0}(l) < 0$. Further, they are increasing in l if and only if $\underline{f} > 0$.*

Whenever the commute has a positive probability to be selected, this quantity is strictly between zero and negative one, and monotonically decreasing with skills. All workers are willing to forego some income for an increase in their workplace quality (or a decrease in travel times), but for poorer workers the percentage increase needed to compensate a reduction in wages tends to infinity. This elasticity is also increasing in incompressible costs, so that everything else equal workers living in more expensive municipalities are less willing to forego wages for workplace niceness.

2.3 Aggregation

From individual choice probabilities, aggregate quantities at the municipal level can be computed as follows:

- Bilateral population in commute ij is given by summing residential probabilities over skill levels
$$H_{ij}^R = \bar{H} \int_0^\infty \pi_{ii}(l) \mathrm{d}\mathcal{F}(l), \tag{13}$$
- Total effective labour flow on ij is given by summing the supply from all skills l
$$H_{ij}^M = \bar{H} \int_0^\infty l \pi_{ij}(l) \mathrm{d}\mathcal{F}(l), \tag{14}$$
- Total income of residents in i is given by summing wages over workplaces and skill levels
$$W_i = \bar{H} \sum_j w_j \int_0^\infty l \pi_{ij}(l) \mathrm{d}\mathcal{F}(l). \tag{15}$$

with $\bar{H}$ the total population of the city. Finally, total labour supply in j is denoted $H_j^M \equiv \sum_i H_{ij}^M$ while total residential population in i is denoted $H_i^R \equiv \sum_j H_{ij}^R$.

Moreover, from the Fréchet preference shock the expected utility is given by (cf. Ahlfeldt et al., 2015, for a proof)

$$\mathbb{E}(U|l) = \left[\sum_{i=1}^S \sum_{j=1}^S \left[\tilde{B}_{ij}(l w_j - Q_i \underline{f}) Q_i^{-\beta} \right]^\epsilon \right]^{\frac{1}{\epsilon}}, \tag{16}$$

so that the total welfare of workers is

$$\mathbb{E}(U) = \int \mathbb{E}(U|l) \mathrm{d}\mathcal{F}(l). \tag{17}$$

2.4 Production

Production in each municipality is assumed to be Cobb-Douglas over workforce $\bar{H}_j^M$ and floor space F_j^M, with a share of floor space α:

$$y_j = A_j (\bar{H}_j^M)^{1-\alpha} (F_j^M)^\alpha, \tag{18}$$

where A_j is a total factor productivity (TFP) term that varies between municipalities. Labor supply is assumed to be a CES aggregate of total efficient labor units for high-skilled workers H, H_j^H and low-skilled workers L, H_j^L, with an elasticity of substitution σ and high skill bias A_j^S:

$$\bar{H}_{Mj} = \left[A_j^H (H_j^H)^{\frac{\sigma-1}{\sigma}} + (1 - A_j^H)(H_j^L)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{19}$$

Further, firms pay a rent Q_j per unit of floor space and a wage index ϖ_j per unit of aggregate labor. Under these assumptions, the profit of firms in j is thus

$$A_j(H_{Mj})^{1-\alpha}(F_{Mj})^{\alpha} - Q_jF_{Mj} - \varpi_j\bar{H}_{Mj}. \tag{20}$$

The first order conditions of profit maximization give the demand for commercial floor space, given workforce:

$$F_{Mj} = \frac{\alpha}{1-\alpha}\frac{\varpi_j\bar{H}_{Mj}}{Q_j}. \tag{21}$$

Moreover, the zero profits condition has to hold if profit maximizing firms operate in municipality j:

$$A_j = \left(\frac{Q_j}{\alpha}\right)^{\alpha}\left(\frac{\varpi_j}{1-\alpha}\right)^{1-\alpha}. \tag{22}$$

Finally, from the assumption of CES labor aggregate, the wage index ϖ_j is

$$\varpi_j = \left[A_j^H(w_j^H)^{1-\sigma}(1-A_j^H)(w_j^L)^{1-\sigma}\right]^{\frac{1}{\sigma-1}}. \tag{23}$$

2.5 The market for floor space

We assume that floor space is produced by a competitive development sector under CRS technology, using elastically supplied capital and land that is completely inelastically supplied. This implies an elastic supply of floor space, with a price elasticity inversely proportional to the share of land in the construction technology of the construction sector.

Formally, F_i the total floor space in i, available for both commercial and residential use, is supplied by a competitive development sector. Following Combes et al. (2017), developers use land L_i with rental price R_i and capital K_i with rental price P (common to all locations) as inputs to a CRS Cobb-Douglas technology:

$$F_i = C_iK_i^{1-\mu}L_i^{\mu}.$$

Developers treat land available for construction as given and fixed, $L_i = \bar{L}_i$, [6] and maximize their profit by choosing how much capital to invest for land development in i. Profit maximization gives the following supply function:

$$F_i = \tilde{L}_iQ_i^{\tilde{\mu}},$$

[6] Assuming that the supply of land is fixed does not seem to be a strong assumption in an urban context, where alternative uses of land such as agriculture are not a concern.

where $\tilde{L}_i \equiv \bar{L}_i C_i^{1/\mu}(\frac{1-\mu}{P})^{(1-\mu)/\mu}$ is a measure of land in i corrected by the easiness to build in i and $\tilde{\mu} \equiv \frac{1-\mu}{\mu}$ is the rent elasticity of floor space supply.

On the demand side, the demand of floor space from firms is given, as a function of workforce, by equation (21). For residents, total demand can be computed by aggregating the individual demand in (4) over skills and commute probabilities:

$$F_{Ri} = \beta \frac{W_i}{Q_i} + (1-\beta) H_i^R,$$

where W_i and H_{Ri} are total income and residential populations respectively, as per equations (15), and (13). Therefore, the market clearing condition is given by equating supply to both these demands:

$$\tilde{L}_i Q_i^{\tilde{\mu}} = \frac{\alpha}{1-\alpha} \frac{\bar{\omega}_i \bar{H}_i^M}{Q_i} + (1-\beta) \underline{f} H_i^R + \beta \frac{W_i}{Q_i}. \tag{24}$$

2.6 Agglomeration effects and spillovers

Local TFPs are allowed to depend on local workforce density:

$$A_i = \tilde{A}_i \left[\sum_j \exp(-\rho^A d_{ij}) \frac{\bar{H}_j^M}{L_j} \right]^{\lambda^A}, \tag{25}$$

where λ^A is the elasticity of TFP to total workforce in the city, while ρ^A is a spatial decay parameter measuring the reach of productivity spillovers.

High-skilled bias is allowed to depend on density in a similar way:

$$\frac{A_i^S}{1-A_i^S} = \tilde{A}_i^S \left[\sum_j \exp(-\rho^S d_{ij}) \frac{\bar{H}_j^M}{L_j} \right]^{\lambda^S}. \tag{26}$$

Finally, residential amenities depend on a local market potential that agregates total residential income around every location:

$$B_i = \tilde{B}_i \left[\sum_j \exp(-\rho^B d_{ij}) \frac{\bar{W}_j}{L_j} \right]^{\lambda^B}. \tag{27}$$

2.7 Equilibrium

Assume one type of workers to ease notations.

Definition 2.1 (Equilibrium)**.** An equilibrium of the model, conditional on parameter values $\{\beta, \underline{f}, \epsilon, \alpha, \rho, \eta, \delta, \lambda\}$, exogenous amenities (b_i), exogenous total factor productivity shifters (a_i), land areas $(\tilde{L}_i)$ and total city population H, is a set $\{(H_{ij}^M, H_{ij}^R)\}$ of number of workers and skill flow per commute so that:

1. the profit maximization condition for firms (22) holds
$$A_i = \left(\frac{Q_i}{1-\alpha}\right)^{1-\alpha}\left(\frac{w_i}{\alpha}\right)^{\alpha};$$
2. the market for floor space clears according to equation (24)
$$\tilde{L}_i Q_i^{\tilde{\mu}} = \left(\frac{(1-\alpha)A_i}{Q_i}\right)^{\frac{1}{\alpha}} H_{Mi} + (1-\beta)\underline{f} H_{Ri} + \beta\frac{W_i}{Q_i};$$
3. amenities are given by equation (27);
$$B_i = b_i\left[\sum_{j=1}^{S}\exp(-\rho t_{ij})W_j\right]^{\eta};$$
4. TFPs are given by equation (25);
$$A_j = a_j\left[\sum_{k=1}^{S}\exp(-\delta t_{jk})\left(\frac{H_{Mk}}{L_k}\right)\right]^{\lambda}.$$
5. Flows are in equilibrium:
$$H_{ij}^M = H\int_0^\infty l\pi_{ij}(l)\mathrm{d}\mathcal{F}(l),$$
$$H_{ij}^R = H\int_0^\infty \pi_{ij}(l)\mathrm{d}\mathcal{F}(l).$$

Proposition 2.3 (Equilibrium existence)**.** *Assume that floor space supply elasticity is strictly positive, $\tilde{\mu} > 0$, then an equilibrium exists for this economy.*

Table 1: List of parameters, estimation methods and sources.

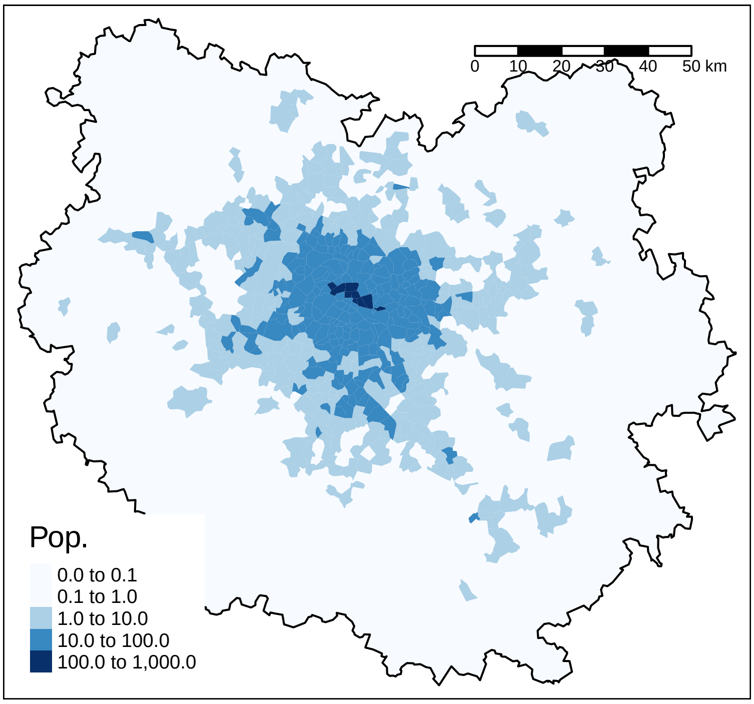
Quantity	Description	Method	Source*	Sect.
ϵ	Taste shock dispersion	Calibrated	Ahlfeldt et al. (2015)	3.1
τ	Utility cost of commuting	Estimated on commuting flows	DADS, TT, DVF	3.1
$\underline{f}$	Subsistence floor space quantity	Estimated on spending data	Expenditure survey	3.1
β	Floor space preference parameter	Estimated on spending data	Expenditure survey	3.1
$w_1, \ldots, w_J$	Local wages	Estimated on individual wages	DADS	3.2
$\mathcal{F}^e$	Talent distribution	Estimated on individual wages	DADS	3.2
ρ, λ	Spillovers	Estimated, model-based instruments	DADS, TT, DVF	3.6
α	Floor space share in prod.	Calibrated from macrod ata	National Accounts	3.3
σ	Skill complementarity	Calibrated from litterature	Wingender (2015)	3.3
$\tilde{\mu}$	Building supply elasticity	Calibrated from litterature	Combes et al. (2017)	3.4
$A_1, \ldots, A_J$	TFP	Residuals, zero profits condition	DADS, DVF	3.3
$B_1, \ldots, B_J$	Residential amenities	Residuals, location choice	DADS, TT, DVF	3.5
$T_1, \ldots, T_J$	Workplace niceness	Residuals, location choice	DADS, TT, DVF	3.5

*: See text in appendix B for a description of the data.

3 Data and calibration

For the rest of the paper, I calibrate the model on the Urban Areas of Paris, in 2015 (represented in Figure 1). It is by far the biggest Urban Area in the country, and the one that exhibits the highest levels of spatial inequalities both in terms of rents and wages. It has been a major commercial and cultural hub for most of the country's history, and thus offers important historical and cultural amenities.

Figure 1: Residents per km^2



For the delineation of the city, I use the National Statistical Institute *Aires Urbaines* 2010, that are constructed by sequential aggregation of municipalities around employment centers based on commuting flows. There has recently been a renewed interest in the literature about methods for delineating Urban Areas (see e.g. de Bellefon et al., 2020). For the purpose of the present paper, because the adjustment of the rent gradient is a key mechanism driving workers sorting as a response to changes in commuting costs, it is important that the limitations to urban sprawl imposed by the urban area boundaries do not influence the results. As illustrated in Figure 1, the definition used here allows for a wide buffer of low density areas (less than 1 h/km^2) around the city center, which means that relaxing commuting costs should not introduce artificial land scarcity.

Table 1 lists the parameters and fundamentals of the model, and the source and methods used to estimate or calibrate them.

3.1 Workers preferences

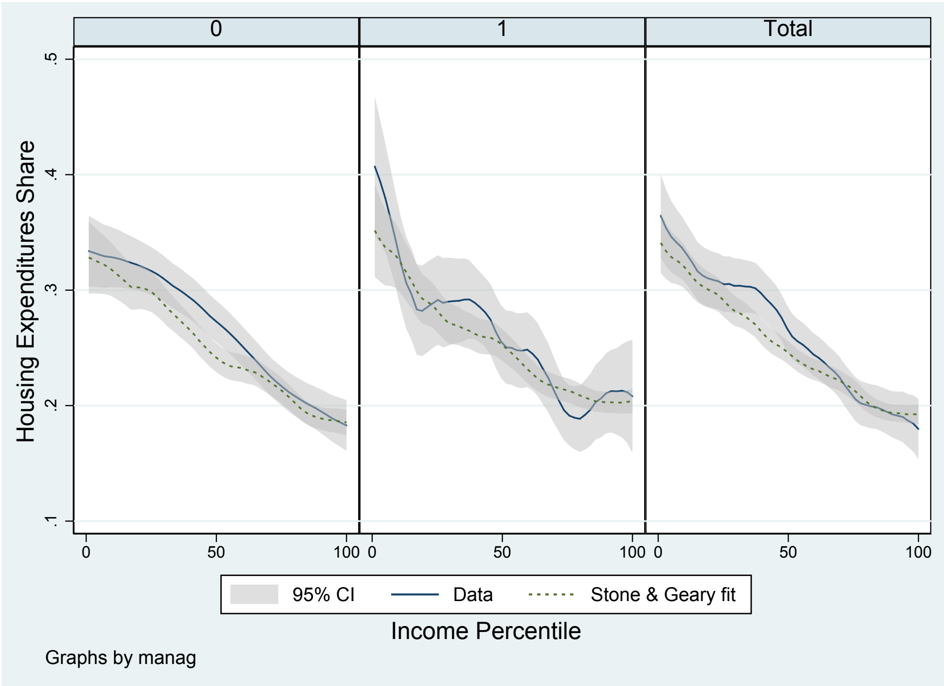
The taste shock dispersion parameter ϵ corresponds is the elasticity of commute choice to changes in real incomes. Several papers have estimated this quantity. Ahlfeldt et al. (2015) find values ranging from 6.6 to 6.8 depending on specifications using German historical data. Preferred estimate of Couture and Handbury (2020) is equal to 3, estimated on US data, while Monte et al. (2018) estimate a parameter equal to 3.3, still on US data. On Dutch data, Almagro and Domínguez-Iino (2019) reports coefficients ranging from 1.6 to 7 depending on household type. I settle for the mid-point and calibrate $\epsilon = 5$.

Housing consumption The Stone and Geary demand parameters, β and $\underline{f}$, are estimated separately for high and low skilled workers on housing expenditure from the Expenditure Survey data. This dataset gives monthly expenditures on housing, total floor space of the dwelling unit, monthly rent for the housing unit, monthly income and the number of workers for a sample of French households in 2006 and 2011. I restrict the sample to households from the Paris Area with at least one working member. From the Stone-Geary specification of utility, the share of income dedicated to housing is given by (4):

$$s_n = \frac{Q_n f_n^*}{w_n} = \beta + (1 - \beta)\underline{f}\frac{Q_n}{w_n}.$$

In the expenditure data, I compute income by active workers w_n by dividing the total salary of the household by the number of working members, and I do the same with household rents to obtain rents per worker. Keeping only households with above minimal wage workers and expenditure shares below one, the resulting Engel curve, pooling high and low skilled workers, is plotted in Figure 2. It is downward slopping, and the Stone-Geary specification estimated below is able to fit this relationship quite well.

Figure 2: Engel curve: data and Stone-Geary fit



Expenditure Survey data vs. Stone-Geary fit (predicted income). Unskilled workers (0), Skilled workers (1) and Total.

Table 2: Estimates of housing preference parameters.

	Raw income		Predicted income	
	L. Skill	H. Skill	L. Skill	H. Skill
β	0.149	0.140	0.129	0.132
	(0.0096)	(0.0137)	(0.0096)	(0.0134)
$\underline{f}$	21.62	29.32	23.13	27.15
	(1.290)	(2.542)	(1.321)	(2.889)
r2	0.431	0.529	0.392	0.332
N	505	329	505	329

Standard errors in parenthesis. Raw income and predicted income are defined in the text.

To estimate $\underline{f}$ and β for each worker type, I run the following regression

$$s_n = \beta + (1-\beta)\underline{f}\frac{Q_n}{w_n} + \tilde{\beta}_n,$$

with $\tilde{\beta}_n$ an individual error term capturing idiosyncratic variations in the marginal propensity to spend on housing. In columns (1) and (2) of Table 2 I report OLS estimates of β and $\underline{f}$ from this equation. The marginal propensity to spend, β, is estimated to be 0.149 for low-skilled workers and 0.140 for high-skilled workers. Incompressible floor space demand $\underline{f}$ is estimated to be higher for high-skilled workers (29.32 sq. meters) than for low-skilled workers (21.62 sq. meters). Note that these estimates pertain to the minimum floor space consumption per worker. With an average of 0.656 inactive household member per worker in the sample, these estimates thus correspond to 12 sq. meters per person for low-skilled workers and 17.7 sq. meters per person for high-skilled workers. As pointed out by Tsivanidis (2019), the slope of the Engel curve, and thus the estimate of $\underline{f}$ in this regression, could be overestimated if incomes are volatile.

As workers cannot adjust housing consumption instantaneously, shocks to w_n the year of the survey would inflate or deflate both Q_n/w_n and s_n, leading to an inflated estimate of $(1-\beta)\underline{f}$. To test the sensibility of the parameters to this issue I construct predicted incomes $\hat{w}_n$ by regressing individual incomes on 4-digits occupation codes and estimate the same OLS regression using these predicted incomes. The results of the regression on predicted income are reported in columns (3) and (4) of Table 2. With predicted incomes, the marginal propensity to spend is estimated at 0.129 and 0.132 respectively for high and low-skilled workers, and incompressible floor space is respectively 23.13 and 27.15. For the rest of the analysis, I set β and $\underline{f}$ to these estimated values. They are close to the raw income estimates, and the following results are not sensitive to using either one of them.

Transport costs From the specification of the transport mode choice problem in equation (7), the probablitiy to choose to take the car versus public transport conditional on living in i and working in j is

$$\Pr(c|ij) = \frac{a_c t_{cij}^{\tau_m \theta}}{a_c t_{cij}^{\tau_m \theta} + t_{pij}^{\tau_m \theta}}$$

where a_c is a parameter and t_{cij} and t_{pij} are travel times between i and j respectively by car and by public transport. I estimate a_c and $\tau_m\theta$ by OLS, regressing log odd-

Table 3: Transport mode choice.

	Low skill	High skill
Car time	-2.245***	-2.163***
	[-2.279,-2.211]	[-2.208,-2.119]
Public transport time	1.798***	1.638***
	[1.771,1.824]	[1.603,1.673]
Constant	-0.183***	0.0827
	[-0.242,-0.123]	[-0.00108,0.166]
r2	0.358	0.343
N	34328	18870

95% confidence intervals in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ratios for each location-destination-type pair on the log of travel times. The estimated coefficients are reported in Table 3. Estimated travel time elasticities are higher for travel times by car (2.245 for low skilled and 2.163 for low skilled) than for public transports (1.798 for low skilled and 1.638 for high-skilled). The magnitude of these parameters is in line with estimated travel time disutility parameters in the literature. Further, the estimated intercept is $a_c = -0.183$ for low-skilled workers, and 0.08 (not significantly different from zero at 5%) for high skilled workers, reflecting a higher preference for cars over public transports for high-skilled workers, irrespective of travel times. The expected utility of commuting between i and j, t_{ij}, is then (up to a multiplicative constant), given by equation (7). I then estimate the ratio of location choice dispersion to mode choice dispersion, ϵ/τ, by regressing commute flows on the expected disutility of travel times and origin and destination fixed-effects, running one separate regression per worker type:

$$\ln(\hat{s}_{ij}) = \delta_i + \delta_j - \frac{\epsilon}{\theta} \ln(t_{ij}) + e_{ij}.$$

As is common in the estimation of bilateral trade frictions, I report both OLS and PPML estimates to accomodate zeroes in the flows data. Table 4 reports the estimated elasticity of location choice to expected travel times disutility. The estimated parameter is slightly lower for High-skilled workers (1.238) than low-skilled workers (1.590). In the simulations, I set these parameters to their PPML estimates.

Table 4: Estimates of the dispersion parameter of mode choice

	(1) OLS Low skill	(2) PPML Low skill	(3) OLS High skill	(4) PPML High skill
Expected t.t. (ϵ/τ)	1.149***	1.590***	0.871***	1.238***
	(0.00345)	(0.00896)	(0.00425)	(0.0178)
r2	0.633		0.603	
N	93385	484416	60854	484416

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.2 Municipal wages

In order to simulate the model, one needs to recover the distribution $\mathcal{F}_l^H$ and $\mathcal{F}_l^L$ of individual productivities l within each worker type (high-skilled H and low-skilled L) separately from the wages paid by firms per unit of labor in each municipality. To do so, I regress the log of individual wages $\ln(w_n^e)$ for worker n with type $e = H, L$ on a set of individual variables X_n and workplace-type fixed effects $\ln(w_j^e)$:

$$\ln(w_n^e) = \ln(w_j^e) + \theta X_n + \epsilon_n \tag{28}$$

Once municipal wages $\ln(w_j^e)$ are estimated, I attribute everything else to individual productivities: $ln(l_n) = \hat{\theta} X_n + e_n$. In Table 5, I report descriptive statistics of the variations of local wages and individual productivities at the municipal level. While municipal effects have a higher explanatory power of municipal variations in wage than individual characteristics, we see that there appears to be positive sorting of the more productive workers withing the most productive cities, as pointed out by the 0.22 correlation between average municipal individual productivities and municipal wages.

3.3 Local productivities and technology parameters

The CES aggregator for unskilled and skilled labor (19) implies (log) labor demand ratios

$$\frac{w_j^H}{w_j^L} = \left(\frac{H_j^H}{H_j^L}\right)^{-1/\sigma} \frac{A_j^H}{1 - A_j^H}. \tag{29}$$

I set σ the elasticity of substitution between high and low skilled workers equal to 1.5, which is the consensus value in the litterature (see e.g. Wingender, 2015, and

Table 5: Contribution of individual and local effects to mean wage at the municipal level.

	Standard Deviation	Correlation with (log) Mean wage	$\ln(w^e_j)$
(log) Mean wage	0.115	1.00	0.89
$\ln(w^e_j)$ (Local)	0.091	0.89	1.00
Mean individual effects	0.053	0.64	0.22

Standard deviation and correlation coefficients within skill groups, between municipalities. First row is mean municipal wage, second is the estimated fixed-effect, the last one is municipal averages of individual productivities $\ln(l_n)$.

references therein.). I then compute the skill-bias of labor demand in each municipality by inverting (29)

$$\frac{A^H_j}{1 - A^H_j} = \frac{w^H_j}{w^L_j} \left(\frac{H^H_j}{H^L_j} \right)^{1/\sigma}$$

From the Cobb-Douglas technology on the upper nest of the production function $y_j = A_j \bar{H}_{Mj}^{1-\alpha} F^{\alpha}_{Mj}$, the production parameter α is equal to the share of floor space in firms costs. I compute this parameter in three ways. First, using aggregate national accounts, second by explicitly using commercial floor space data aggregated at the city level to compute the ratio of floor space expenditures over wage bill, and finally by regressing floor space expenditures on wage bill at the municipal level.

National accounts from *INSEE* report that the share of capital (GFCF) represents 30% of value added, while building and land make up 55% of capital expenditures of French firms. Normalizing so that floor space and labor shares sum to one, we get a share of developed land of $\alpha = 0.55 \times 0.3/(0.7 + 0.55 \times 0.3) = 19\%$.

Because of changing definitions of taxable commercial and professional floor space, land registers are not very reliable in their reporting of non-residential surfaces. Still, using non-residential floor space from those files and average rents per squared meter from the building transactions data, I obtain a share of floor space in firms costs of 26%. Finally, using those same data to regress floor space costs on total wage bill at the municipal level gives a coefficient of 0.28, implying a share $\alpha = 22\%$. Overall, the calibrated parameter from national accounts data is in line with raw correlations in the micro data, and I calibrate $\alpha = 20\%$.

Given this parameter, wage index $\bar{w}_j$ and rents Q_j for each municipality j, I compute TFPs from the zero profits condition

$$A_j = \left(\frac{\bar{w}_j}{1-\alpha}\right)^{1-\alpha}\left(\frac{Q_j}{\alpha}\right)^{\alpha}.$$

3.4 Construction sector

For the construction sector technology, I calibrate the share of land in the production of floor space μ to the estimates in Combes et al. (2017). For the city of Paris, they report elasticities to non-land inputs of 0.54 (Table 3). This gives a supply elasticity of $\tilde{\mu} = \mu/(1-\mu) = 1.17$. Estimated long-term elasticities of housing supply in the literature for constrained cities range between 1 and 4 depending on the nature of the housing market. Saiz (2010) reports unweighted mean elasticities across US Metropolitan Areas (MSAs) of 2.5, while Harter-Dreiman (2004) reports ranges of elasticities of $[1-2.1]$ for constrained housing markets and $[2.6-4.3]$ for unconstrained cities, still in the US. This puts the calibrated elasticity for Paris in the range of long-run elasticities estimated for constrained housing markets in the US.

Given $\tilde{\mu}$ and households preference parameters, adjusted land areas $\tilde{L}_i$ are computed for all i to solve the floor space market clearing equation in (24):

$$\tilde{L}_i = Q_i^{-\tilde{\mu}}\left(\frac{1-\alpha}{\alpha}\frac{\bar{w}_i\bar{H}_{Mi}}{Q_i} + (1-\beta)\underline{f}H_{Ri} + \beta\frac{W_i}{Q_i}\right).$$

3.5 Amenities

Given individual preference parameters $\epsilon, \beta, \tau, \underline{f}$, local wages w_j^e for every municipality and type and floor space rents Q_i for every municipality, we can compute total income at the workplace and at the residential place for each worker type

$$W_i^{Re} = H\sum_j w_j^e \int_0^\infty l\pi_{ij}(l)\mathrm{d}\mathcal{F}^e(l), \qquad e = \{H, L\}$$
$$W_i^{Me} = Hw_j^e \sum_j \int_0^\infty l\pi_{ji}(l)\mathrm{d}\mathcal{F}^e(l), \qquad e = \{H, L\}.$$

I calibrate amenities B_i^e and labour supply shifters T_j^e to the unique values that match predicted total income to total income in the data for all i, j and e, conditional on preference parameters, wages and rents.

3.6 Spillovers and agglomeration effects

Having recovered TFPs A_j, skill bias A_j^S and residential amenities B_i^e, for $e = \{H, L\}$, I estimate the spillover parameters using a model-based instrument approach à la Allen et al. (2020). The main concern with naive non-linear least square estimation of the spillover parameters ρ and λ in (25), (26) and (27), is that of the simultaneity of populations and exogenous amenities. Taking (25) as an example, the equation that we would like to estimate is

$$A_i = \left[\sum_j \exp(-\rho^A d_{ij}) \frac{\bar{H}_j^M}{L_j} \right]^{\lambda^A} \tilde{A}_i,$$

where $\tilde{A}_i$, the error term, is composed of exogenous variables that influence local productivity (e.g. natural advantages, access to rivers, slope, altitude, as well as unobserved infrastructures), besides agglomeration effects that are captured by the term in brackets. Because workers tend to move to high-TFP places to enjoy higher wages, $\tilde{A}_i$ is likely positively correlated with $\bar{H}_{Mj}$ the endogenous labor supply, so that we can expect naive estimates of λ to be biased upward.

This problem is not new, and there is a large body of literature concerned with designing ways to mitigate this endogeneity issue. The canonical approach (cf. Combes and Gobillon, 2015) is to instrument populations by long lagged values of itself, while controlling for geographical features that are likely to be part of $\tilde{A}_j$. The reasoning behind those instruments is that technological change has been such over the years that determinants of productivity that attracted populations centuries ago are not relevant anymore (at least conditional on controls), and only affect productivity in so far as they anchored populations. The identifying assumption that underlies this approach is thus that once we control for persistent geographical features that may be relevant for today's firms (such as climate or access to water) the factors that drove historical localization of manufacturing are not directly relevant to the localization of modern-day industries.

However, while the identifying assumptions are plausible in the case of productivity, their application to estimating residential amenity spillovers is more problematic. Indeed, in the case of residential amenities, one can argue that most of the natural features, views, monuments and historical prestige that drove the localization of residents a few centuries ago are still relevant to the location choice of modern workers. Especially within cities, where fine geographical features can make all the difference between a cold swamp and a nice riverside, it would be hard to argue that we have access to detailed enough control variables to make historical populations a valid instrument.

To recover these spillover parameters, I thus follow Allen et al. (2020) and use *model-based instruments*, constructed using exogenous variables that would have been used as controls in a traditional IV approach. The advantage of this approach is that we do not need an exhaustive list of control variables, as long as they have some relevance to location choice.

In short, the method goes as follows:

1. Regress model fundamentals (A_i, B_i, A_i^S) on a set of exogenous local characteristics Z_i (e.g. topographical characteristics), and predict $(\hat{A}_i, \hat{B}_i, \hat{A}_i^S)$ using Z_i.

2. Simulate the model using $(\hat{A}_i, \hat{B}_i, \hat{A}_i^S)$ and a first guess λ, ρ for the agglomeration and spillover parameters as inputs. Denote $(\hat{\hat{H}}_j^M)$ the simulated workforce and $(\hat{\hat{W}}_i)$ the predicted incomes.

3. Estimate $\hat{\lambda}$ and $\hat{\rho}$ from the structural equations (25), (26) and (27), using the simulated variables $(\hat{\hat{H}}_{Mj})$ and $(\hat{\hat{W}}_i)$ as instruments, *whilst controlling for* Z_i.

The validity of the instruments comes from the fact that, by construction, the predicted values are not correlated to components of (A_i, B_i, A_i^S) that are not controlled for by Z_i in the last step of the procedure.

Note that identification does not come from non-linearities of the model only. Indeed when running the actual IV regression in step 3 of the procedure we only control for i's own Z_i, whereas each equilibrium value $\hat{\hat{H}}_j^M$ is a combination of the whole Z and distances. Identification is thus achieved by using the model to weight distant values of Z and use them as instruments. When we suspect that one of the variables might have direct spillover effects on the productivity of neighboring municipalities, such as access to a river, it is thus important to control for the direct effect of distance to this amenity by directly including it in Z.

As a first guess, I use the values estimated by Ahlfeldt et al. (2015) and set the TFP parameters to $\lambda^A = 0.07$ and $\rho^A = -0.35$ and the amenities parameters to $\lambda^B = 0.35$ and $\rho^B = -0.8$ for both the high skilled and low skilled amenity indices. For the skill bias, I simply set both parameters λ^S and ρ^S to zero.

Table 22 in appendix reports the results of the first stage regressions, where I regress amenities, TFP, skill bias and workplace niceness on exogenous amenities. The

Table 6: Estimates of spillover parameters for productivity.

	(1) TFP	(2) TFP	(3) TFP IV	(4) Bias	(5) Bias	(6) Bias IV
main						
λ (Intensity)	0.0401***	0.0466***	0.0416***	0.0815***	0.0967***	0.0799***
	(0.00200)	(0.00243)	(0.00262)	(0.00613)	(0.00745)	(0.00846)
ρ (Decay)	-0.503***	-0.554***	-0.508***	-0.539***	-0.593***	-0.683**
	(0.0409)	(0.0416)	(0.0597)	(0.0695)	(0.0691)	(0.153)
Controls	No	Yes	Yes	No	Yes	Yes
F first stage			594.5			594.5
# of moments			16			16
J stat			5.698			5.849
N	696	696	696	696	696	696

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Agglomeration effects are measured in terms of total workforce, defined as the CES aggregate for labor. "Controls" include the variables in Table 22 in Appendix. Columns with no first-stage F-stat and overidentification tests are NLS regressions, columns with those statistics are model-based IV regressions. First stage F statistic reports the F-test from regressing $\bar{H}^M$ on its model-based counterpart.

included explanatory variables are mean altitude, maximum slope, distance to rivers, and a dummy equal to one if the centroid of the municipality is less than 5km away from a river. For amenities, I also include a dummy variable for listed buildings in the municipality. R-squared are 9% for skill bias, 13% for TFP, 11% and 10% for high and low skill amenities, and 11% and 13% for for labor supply shifters. All the F-stats are above 15. Altitude and slope are the main explanatory variables for TFP and skill bias, and they have the expected negative sign. For residential amenities, the listed buildings dummy is the only significant predictor. Existing studies in Europe (Koster et al., 2016, Koster and Rouwendal, 2017, Garcia-López et al., 2018) report that historic amenities are a strong driver of household location choice and sorting.

Table 6 shows the results of the naive non-linear least squares and GMM estimates of the agglomeration effects for TFP and skill bias. Regarding TFP, the GMM estimate of the elasticity is $\lambda^A = 0.040$, equal to the OLS estimate up to the third digit. The spatial decay is estimated at $\rho^A = -0.454$, lower than the OLS estimates. The magnitude of the estimated λ^A parameter is in line with the recommended value from Ahlfeldt and Pietrostefani (2019) of 0.04. It is also in line with the between-cities estimates of the effects of density on TFP in France from Combes et al. (2010) (IV

Table 7: Estimates of spillover parameters for amenities.

	(1)	(2)	(3)	(4)	(5)	(6)
	B0	B0	B0 IV	B1	B1	B1 IV
main						
λ (Intensity)	0.230***	0.179***	0.127	0.428***	0.427***	0.351***
	(0.0165)	(0.0196)	(0.0750)	(0.0163)	(0.0195)	(0.0507)
ρ (Decay)	-0.984***	-1.123***	-1.133	-0.711***	-0.755***	-0.682**
	(0.154)	(0.248)	(1.328)	(0.0527)	(0.0587)	(0.155)
Controls	No	Yes	Yes	No	Yes	Yes
F first stage			433.8			433.8
# of moments			16			16
J stat			30.48			23.30
N	696	696	696	696	696	696

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Lambda measures agglomeration and tau measures its spatial decay. Spillovers are measured in terms of total income per unit of land. “Controls” include the variables in Table 22 in Appendix. Columns with no first-stage F-stat and overidentification tests are NLS regressions, columns with those statistics are model-based IV regressions. First stage F statistic reports the F-test from regressing $\bar{W}^R$ on its model-based counterpart.

estimates on TFP ranging from 0.031 to 0.048 depending on specifications). Skill bias parameters are much higher, with a GMM estimate of $\lambda^S = 0.078$ and $\rho^S = -0.571$.

Similarly, Table 7 reports the estimated spillover effects for residential amenities, where the variable generating spillovers is total residential income per land unit. For low-skilled residential amenities, the OLS spillover is $\lambda^{B,L} = 0.23$. Adding controls, the effect drops to 0.179. Instrumenting total incomes further reduces the estimated coefficient to 0.127, although it is now imprecisely estimated and not significantly different from zero. Regarding the decay parameter, it increases from -0.984 without controls to -1.123 when introducing control variables, and stays stable when instrumenting. For high-skilled workers, OLS estimates of the spillover effects are 0.428 without controls and 0.427 when introducing control variables, while the GMM estimate is $\lambda^{B,H} = 0.351$. The decay parameter is quite stable to introducing control variables and when instrumenting (resp. 0.755 with and 0.711 without controls, and 0.682 when instrumenting). The strength of spillovers is thus approximately twice as high for high-skilled than for their low-skilled counterpart across specifications, which is in line with previous evidence that high-skilled workers value consumption amenities more than their low-skilled counterparts (e.g. Couture and Handbury, 2020).

Table 8: Variance of amenities.

(a) Low skilled

	St. Dev.	Correlation Tot.	End.
Total	1.07	1.00	0.67
Endogenous	0.42	0.67	1.00
Residual	0.85	0.93	0.35

(b) High skilled

	St. Dev.	Correlation Tot.	End.
Total	1.47	1.00	0.82
Endogenous	0.98	0.82	1.00
Residual	0.87	0.77	0.26

Variance decomposition of the estimated amenities. Total references the amenity index, endogenous is the predicted endogenous amenities based on GMM estimates of the spillover parameters, and residual is the part of amenities that is total minus endogenous. Everything is in logs.

Table 8 presents a decomposition of the variance of estimated amenities into estimated endogenous amenities (predicted from equation (27) with the GMM estimates of λ and ρ) and residuals. First, we see that high-skilled amenities have a higher variance and are better explained by the estimated endogenous component than low-skilled amenities, consistent with the stronger spillover parameters. Second, in both cases endogenous and residual amenities are positively correlated, as would be expected given the reverse causation between amenities and residential incomes.

In the model simulations of the next section, I set the spillover and decay parameter to their GMM estimates.

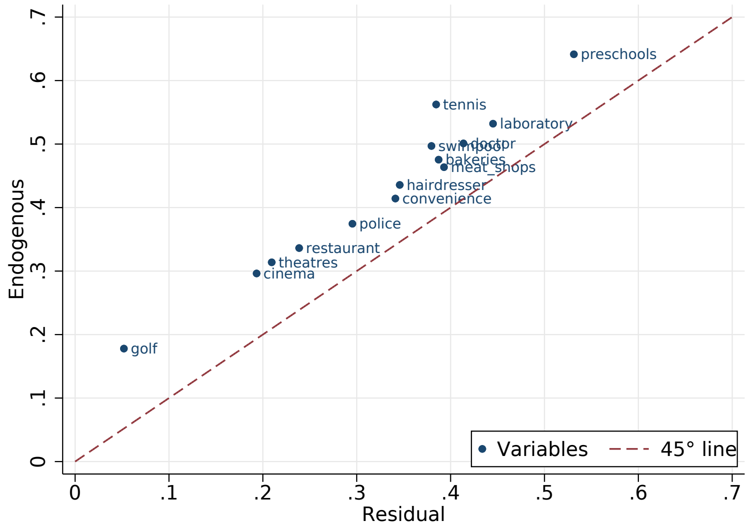
3.7 Correlation between model-based and observed amenities

In this section, I look at the correlation between the amenities computed above and a set of observed variables, as a way to check the validity of the model specification. I use the *Base Permanente des Équipements* dataset, a public dataset with the location of a wide range of endogenous amenities, from restaurants to swimming pools and general practitionners. A complete list of the variables used and associated descriptive statistics are reported in Table 23 in appendix.

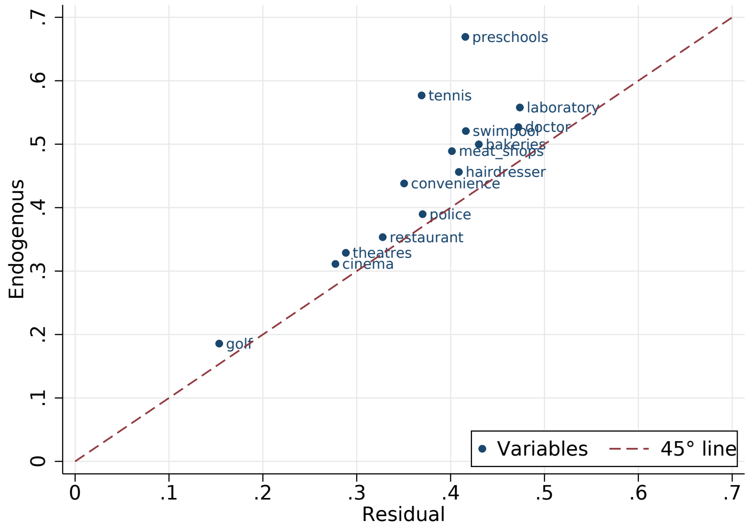
To explore the relationships between observed and model-implied amenities, Figure 3 reports raw correlations between observed and theoretical amenities both for high and low-skilled workers. The number of preschools, medical laboratories and doctors are the variables that correlate the most with endogenous amenities,

Figure 3: Raw correlations between observed and estimated amenities

(a) Low skilled



(b) High skilled



Note: each dot is an observed amenity. The y axis reads the correlation coefficient between this amenity and the estimated endogenous amenities, while the x axis reads its correlation with the residual (i.e. exogenous) amenities.

whilst cinemas, restaurants and theatres have the lowest correlation. Consistent with the endogenous nature of these amenities, their correlation is stronger with the estimated endogenous component of amenities (predicted from eq. (27) with the GMM parameter estimates from the previous section).

All these observed amenities also correlate positively with residual (”exogenous”) amenities. However all the dots lie above the 45° line, indicating that this correlation is systematically weaker than with endogenous amenities. This shows that the model-based endogenous amenities index captures most of the effect of observed amenities. Moreover, it is not surprising to see a positive correlation between the amenities residual and observed amenities, as we expect that high exogenous amenities anchor neighborhoods into high population, high endogenous amenities status.

To further explore the relationship between model-based amenities and observed amenities, Table 10 reports the results of linear regressions of model-based amenities on observed amenities. Due to the high degree of colinearity between observed amenities, I do not include the whole list of amenities included in Figure 3. Instead, I include the number of tennis courts and horse-riding clubs to represent sports and outdoors activities, the number of cinemas for consumption amenities, as well as the number of doctors and preschools.[7]

The R-squared of the regression of high-skilled amenities on these five observed variables (column (2)) is 0.56, which shows that the estimated high-skilled amenity index correlates strongly with observed amenities. Further, including them in a regression with the estimated high-skilled residential spillover index only raises the model's R-squared from 0.68 to 0.74, indicating that the estimated endogenous component of amenities indeed captures most of the effect of observed endogenous amenities.

For low-skilled workers, R-squared are globally lower but follow the same pattern. Consistent with the lower estimated spillover parameters in the previous section, the regression of low-skilled amenities on observed amenities yields a R-squared of 0.47 (column (5)). Including them in a regression of low-skilled model-based amenities on the estimated low-skilled residential spillovers raises the R2 from 0.45 to 0.55.

[7] Including the whole set of observed amenities included in Figure 3 only raises the R2 in column (2) to 0.61 and the one in column (5) to 0.50.

Table 10: Regression of model-implied amenities on observed amenities.

	High Skilled B			Low Skilled B		
	(1)	(2)	(3)	(4)	(5)	(6)
Endo. H-S	1.233***		0.896***			
	(37.96)		(21.76)			
Endo. L-S				1.709***		1.006***
				(23.81)		(11.24)
tennis		1.360***	0.621***		0.538***	0.175
		(8.28)	(4.74)		(4.06)	(1.38)
horses		2.186***	1.060***		0.547	0.0189
		(6.30)	(3.89)		(1.96)	(0.07)
cinema		0.335**	0.0714		0.0814	-0.0454
		(2.66)	(0.73)		(0.80)	(-0.48)
doctor		0.0411***	0.0547***		-0.00676	0.000505
		(3.31)	(5.71)		(-0.68)	(0.05)
preschools		0.517***	0.0499		0.632***	0.417***
		(7.44)	(0.86)		(11.25)	(7.57)
Constant	-15.11***	-11.09***	-14.15***	-12.54***	-10.95***	-11.97***
	(-112.37)	(-214.72)	(-96.86)	(-130.93)	(-262.71)	(-121.72)
r2	0.675	0.569	0.744	0.450	0.469	0.551
N	696	696	696	696	696	696

t-statistic in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Regression of model-based amenities on a set of observed endogenous amenities. In columns (1) to (3), the dependent variable is high-skilles model-based amenities. In columns (4) to (6), low-skilled model-based amenities.

4 Simulations

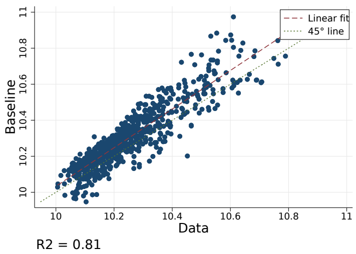
4.1 Model fit and overidentification tests

In this section, I report the results of a baseline simulation of the model with the estimated and calibrated parameter values from the previous section.

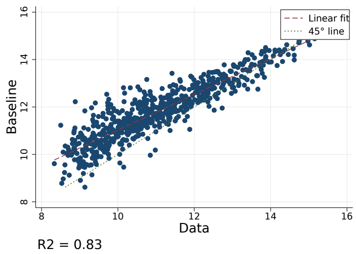
Simulations are computed using a fixed-point algorithm that reproduces a dynamic setup with myopic workers. With agglomeration effects and spillovers, these types of models are not guaranteed to yield a unique equilibrium. This simulation process thus chooses the equilibrium that is the "closest" to the baseline, in the sense that it is the equilibrium that the economy would reach under the following adjustment process. At each iteration t, taking previous period commute flows as fixed, wages $(w_j^e)^t$ and rents Q_i^t are computed that clear the floor space and labor markets, as per equation (22), (24) and (23). Endogenous amenities and TFPs are computed according to those new rents and wages. A new mass of workers is then computed for each commute using the bilateral choice probabilities. I then update incomes, rents and spillovers again, and the operation is repeated until the repartition of workers stabilizes.[8]

Figure 4: Overidentification checks: data vs. non-targeted variables

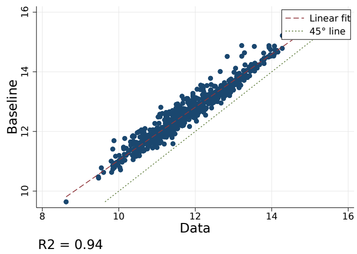
(a) Mean income



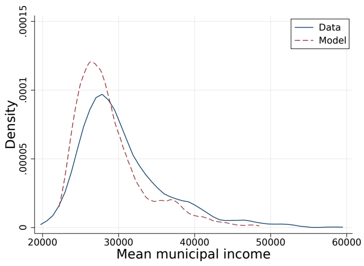
(b) Commercial floor space



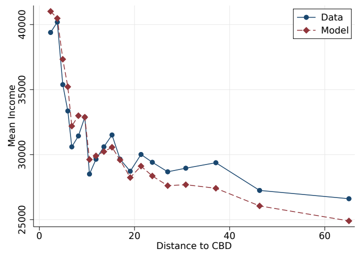
(c) Residential floor space



(d) Mean income



(e) Mean income



Note: Each observation is a municipality. (x axis) vs. model baseline simulations (y axis) for mean income (a), the total floor space of commercial building (b), and the total floor space of residential building (c). Panel (d) reports actual (solid) and predicted (dashed) density of mean incomes accross municipalities. Panel (e) reports bin-scatters of actual (solid) and predicted (dashed) mean income as a function of distance to the city center.

In calibrating the model, I do not directly target mean incomes and I do not use the data on floor space. They are therefore good candidates to test the specification of the model.

[8]In practice, convergence is declared when the maximum absolute relative deviation between two consecutive iterations of skill flows and workers flow of each type is lower than $10^{-}5$, i.e. when $\max\{|(H_{ij}^{Fe})^{t-1} - (H_{ij}^{Fe})^t|/(H_{ij}^{Fe})^t, \quad i,j = \{1,...,J\}, F = \{M,R\}, e = \{H,L\}\} < 10^{-}5$.

Although the model perfectly fits total populations, wages and rents conditional on the other observed variables in the model, it is not possible to calibrate the model to perfectly match total incomes and total populations at the same time, i.e. to perfectly predict mean incomes. Instead, the relationship between populations and income is determined by the non-homotheticities in workers preferences, their choice of workplace and the share of high-skilled workers in a given municipality. I find that the model captures mean incomes variations quite well (Figure 4 Panel (a), R2=.81), and although the model predicts less spatial disparities than observed, the shape of the distribution of mean incomes between municipalities (Figure 4 Panel (d)) is well captured. Finally, the specification of distance disutility allows to closely replicate the gradient of mean incomes in the city (Figure 4 Panel (e)).

Turning to commercial and residential floor space, they are well fitted with squared correlations of respectively 0.83 and 0.94, indicating that the demand functions are well calibrated. In Appendix A.6, I report maps of actual and predicted mean incomes and residential floor space. The model is able to closely replicate the spatial patterns of these variables.

4.2 Suburban train network (RER)

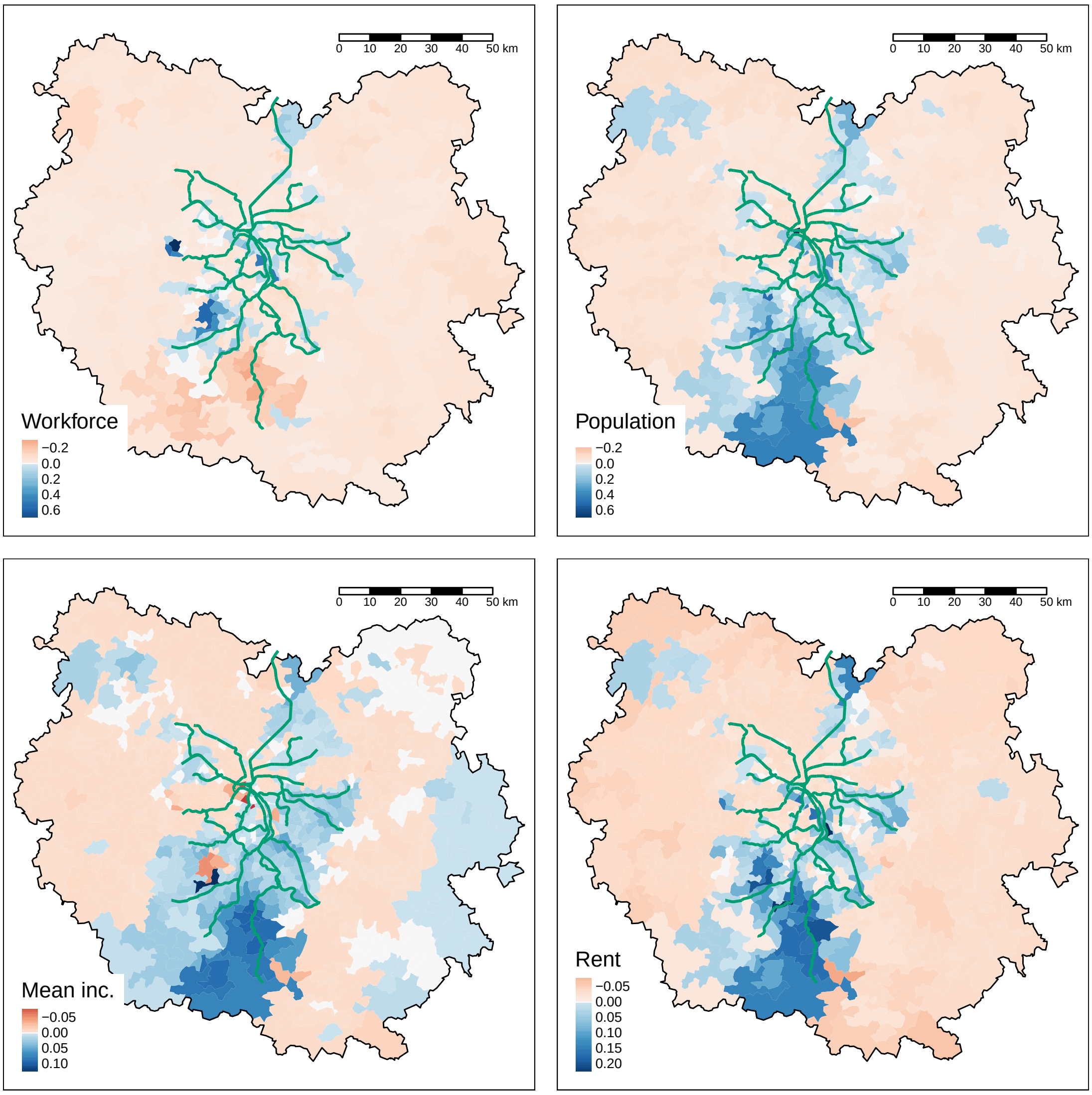
In this section, I look at the effect of the public transport network on the structure of the city by simulating a counterfactual Paris in which suburban trains from the Regional Express Rail (RER) network are removed.

The RER is a suburban rail network made of radial lines connecting Paris to its suburbs. In 1965, a plan to turn the mono-centric area of Paris into a poly-centric region was devised by the French government, that revolved around developing new sub-centers — the *New Towns* — that would house secondary business districts and residential areas. The RER rail system was devised as a set of radial lines that would cross the region to connect those sub-centers to Paris, complemented by a set of new metropolitan highways.

Inaugurated in 1977, the RER network was initially composed of two lines, one crossing the region from north to south and the other one from east to west, and was later extended to four lines, with a fifth one constructed in 2015.

With more than 500 Km of lines, the Regional Express Rail is the backbone of the Parisian transport network. Indeed, Table 17 in appendix reports summary statistics of travel times by public transport between pairs of municipalities in the Region, with and

Figure 5: Contributions of the suburban train (RER) to municipal outcomes.



Note: Maps of the contribution of the RER network to the number of workers (Workforce), the number of residents (Population), average income of residents (Mean inc.) and rents (Rent) in the Paris metropolitan area. Contributions are computed for variable y as $(y_{\text{Baseline}} - y_{\text{NoRER}})/y_{\text{Baseline}}$. Descriptive statistics in table 18.

without allowing the use of the RER. On average, the RER lowers travel times between all pairs of municipalities in the region by 22%, and travel times to the city center by 20%. Moreover, its effect on travel times is higher for municipalities located between 10 and 60 kilometers from the city center, because municipalities located further away are not connected to the network while those located closer to the center can use the faster metro lines.

Table 11: Effect of the RER network on commute costs to the CBD.

	(1) Unskil. Pop.	(2) Skil. Pop.	(3) Unskil. Emp.	(4) Skil. Emp.	(5) Rents	(6) Mean Inc.
RER=1	0.0885***	0.100***	0.0379***	0.0417***	0.0443***	0.0115***
	(0.0118)	(0.0134)	(0.00965)	(0.00786)	(0.00394)	(0.00174)
Constant	-0.0246***	0.000346	-0.0345***	-0.0299***	-0.0146***	0.00252**
	(0.00578)	(0.00656)	(0.00473)	(0.00385)	(0.00193)	(0.000851)
Observations	696	696	696	696	696	696
R^2	0.075	0.075	0.022	0.039	0.154	0.059

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Regressions of the difference in (log) outcome between baseline and no-RER simulations on a dummy equal to one if the municipality was eventually connected to the RER network. Columns (1) and (2) report results on population for low and high skilled workers respectively, columns (3) and (4) on employment, column (5) on rents and column (6) on mean income at the residential location.

General effects of the RER Maps in Figure 5 show the current contribution of the RER network to municipal outcomes. Contribution for variable y is measured as $1 - y_{\text{NoRER}}/y_{\text{Baseline}}$, where y_{NoRER} is the result of a simulation of the model with travel times computed by omitting the Regional Express Rail. It appears from these maps that the RER not only has a decentralizing effect on employment, populations, incomes and rents, but that this effect is heterogeneous conditional on distance. The biggest impact on populations and rents is measured for the most southern municipalities of the area, which are connected to the last stops of the D line of the RER, and otherwise poorly served by the rail network.

Quantitatively, Table 11 gives the effect of the RER network on the connected and non-connected municipalities in the region. It is computed by regressing the log of the difference between baseline outcomes and outcomes without the RER on a dummy equal to one when the municipality is connected to the RER. On average, unconnected municipalities get a loss of unskilled population of 2.5%, while connected municipalities get an additional increase of 8.9%. The effect on skilled population is stronger, with an average gain of 10% for connected municipalities. Further, the network have an effect on total incomes and income sorting in the area, with an increase of average incomes of 0.25% for unconnected municipalities and an additional 1.15% for connected ones.

The effect on employment is globally weaker than on populations, with a reduction for unconnected municipalities of 3.5% and 3.0% respectively for low and high-skilled workforce, and respective additional increases of 3.8% and 4.2%. These effects on population and employment have a substantial impact on the market for floor space, with a rent decrease if 1.5% on average in unconnected municipalities and an additional increase of 4.4% in connected municipalities.

Table 12: Aggregate effects of the RER network on incomes in the region.

	Mean	Total SD	Between SD	(%)	C−P
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	31129.84	20903.03	3544.28	16.96	10603.74
Effect (%)	0.09	-0.02	-1.93	-1.91	-3.68

Column 1 is the average income over all workers. Column 2 is the total standard deviation of income over individual workers. Column 3 is the between-municipality standard deviation. Column 4 is the ratio between the two times 100. Column 5 is mean income in the center minus mean income 10 to 15 kilometers away from the center.

Sorting The impact of the RER network on workers location translates into marked effects on income disparities, as reported in Table 12. First, average income in the area increases by 0.1% while the total standard deviation of incomes drop by 0.02, pointing to a small reduction of income inequalities. Turning to spatial inequalities, the between-municipalities SD drops by 1.93%. Further, the income premium of Paris with respects to its suburbs (10 to 15 kilometers away from the CBD) drops by 3.68%, which shows that the RER network lowers income inequalities between the center and the periphery.

Redistribution Finally, I compute the contribution of the Regional Express Rail to workers welfare and welfare inequality. Although the effect of the RER on location choices is stronger for high-skilled than low-skilled workers, its total welfare effect is higher for the latter group. As a result, it leads to a decrease in welfare inequalities. Indeed, the network accounts for 3.32% of the total welfare of low-skilled workers and 2.56% of that of high-skilled workers, thus reducing welfare inequalities by 0.78%.

Table 13: Welfare of high and low-skill workers, with and without the Regional Express Rail.

	Low skill	High skill	Ratio
Baseline	109.86	594.01	5.41
Counterfactual	106.22	578.79	5.45
Effect (%)	3.32	2.56	-0.78

Comparison with reduced-form results Mayer and Trevien (2017) evaluate the effect of the introduction of the regional rail system (RER) between 1970 and 1990 in the Paris area. The present simulation does not exactly replicate their setting, as my counterfactual simulation uses the current network without the RER, not the network as it was in 1970. Further, I study re-organization effects in a closed city when their measure takes into account both growth and relocation. They measure an effect of 8.8% on employment, and a positive but unstable effect on populations, stronger for high-skilled workers. Overall, the model-implied effects above are in line with the estimates they report, and in line with the broader literature on the effects of the RER network (e.g. Garcia-López et al., 2017).

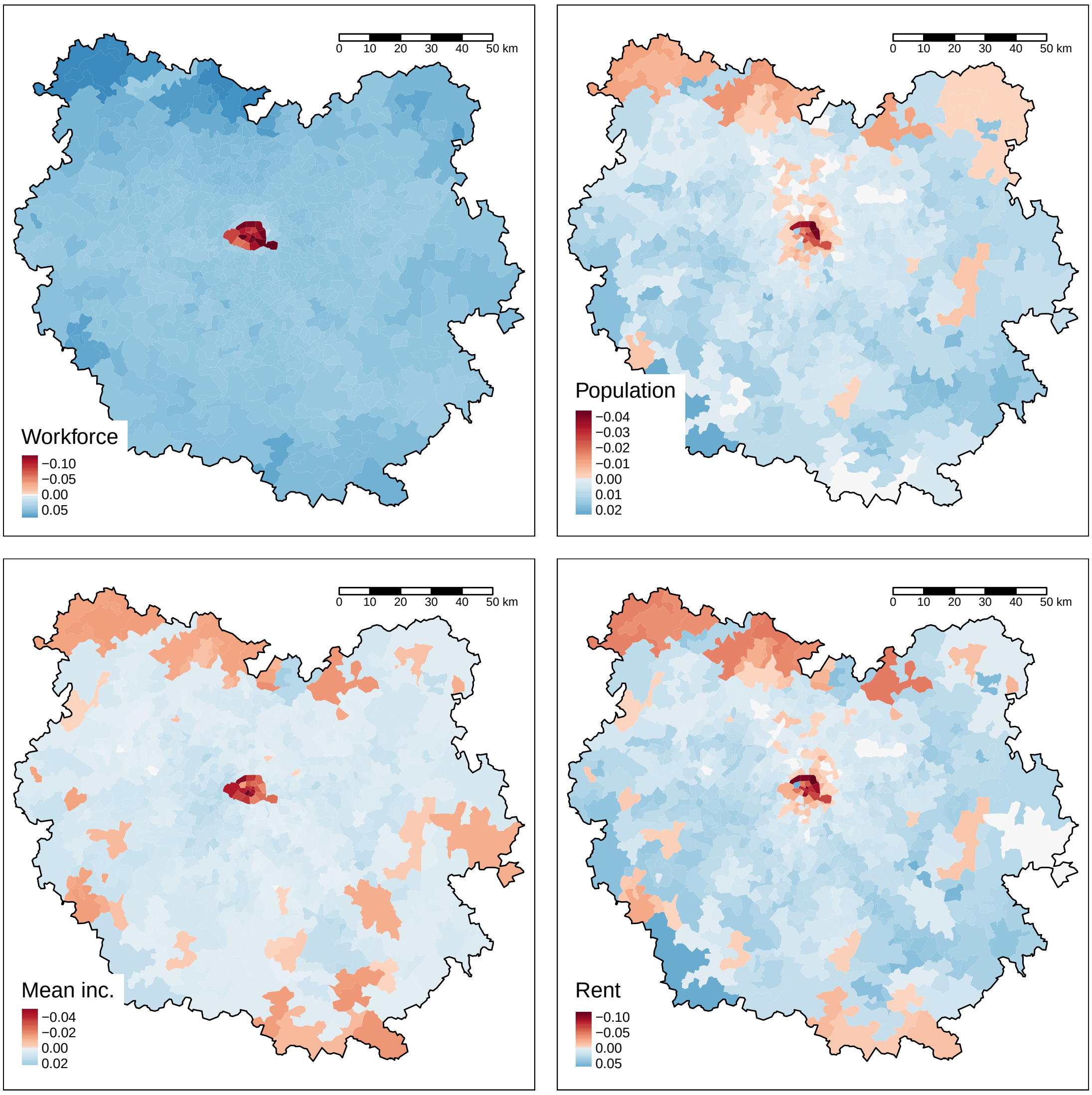
4.3 Banning cars from Paris

In this section, I turn to counterfactual simulations where commuting by car is banned when commuting from or to the city of Paris, so that all commuters within Paris, between Paris and the suburbs or vice-versa have to take public transports for commuting.

In 2017, following engagements taken as part of the 2015 COP21 agreements, the Council of Paris signed a document planning to ban thermal vehicles from its streets, with a plan of zero diesel cars by 2024 and zero thermal vehicles by 2030.[9] This announcement has been a source of debate in the region, with some opponents pointing out that the measure, given the current costs of electric vehicles, might penalize suburban workers. To assess this statement, I consider the somewhat more excessive situation of a complete ban on cars, thermal or not, within the city of Paris. This is implemented by setting the probability to travel by car to zero for every trip

[9] *Plan Climat*, available in French and English at `https://www.paris.fr/pages/nouveau-plan-climat-500-mesures-pour-la-ville-de-paris-5252`, accessed on August 21, 2020.

Figure 6: Effects of banning cars in Paris.

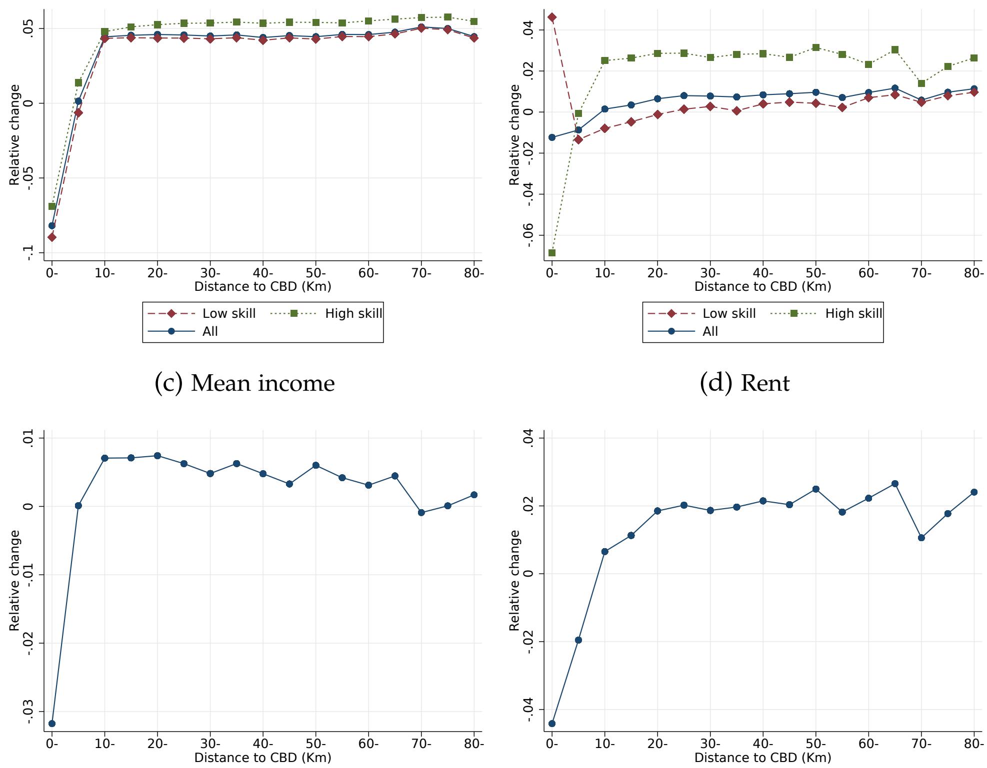


Note: Maps of the effects of banning commuting by cars to and from Paris on the number of workers (HM), the number of residents (HR), average income of residents (meaninc) and rents (rent) in the Paris metropolitan area. Inelastic floor space supply. Descriptive statistics in table 18.

from or to the city of Paris.

I focus on short-term effects where residential and commercial floor space are fixed to their baseline levels. Section A.5 in appendix shows the results from an alternative simulation where floor space is elastic and landlords are allowed to convert between commercial and residential floor space.

Figure 7: Effect of banning cars in Paris.



(a) Employment

(b) Population

(c) Mean income

(d) Rent

Note: Average effect of banning cars in Paris on employment, population, mean income and the between-municipality coefficient of variation of mean incomes in bins of 5km from the city center (first district of Paris). Inelastic floor space supply.

Figure 6 shows maps of the effect of a car ban on the Grand Paris region. Contrary to the predictions of a monicentric model, the model predicts that increasing the cost of accessing the city center would penalize the center itself, as firms and workers relocate outside of the city walls. This leads to a decrease in rents in the center, allowing for lower income residents to relocate within the city.

General effects Figure 7 Panel (a) shows the effect of the car ban on employment, population, income and rent gradients. The center of Paris experiences a loss in employment of 7%, that relocate 10km away from the center. Regarding populations, the center experiences a loss in high-skilled residents of 6%. For low-skilled workers, the pattern is u-shaped: their number increases by 4% in the center, decreases in the close suburbs and increases again in the more remote locations of the area.

Table 14: Aggregate effects of banning in Paris on incomes in the region.

	Mean	Total SD	Between SD	(%)	C−P
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	31093.20	20784.18	3538.56	17.03	8727.20
Effect (%)	-0.21	-0.55	1.77	2.33	-14.66

Column 1 is the average income over all workers. Column 2 is the total standard deviation of income over individual workers. Column 3 is the between-municipality standard deviation. Column 4 is the ratio between the two times 100. Column 5 is mean income in the center minus mean income 10 to 15 kilometers away from the center.

Sorting Turning to income sorting, the outflow of high skilled workers and the lower rents in the centrer translate into a decrease of the income premium of the city, as incomes within 5 kilometers fall by 3%, and incomes in the suburbs rise by 0.5%. This effect corresponds to a drop of the mean income premium of the center relative to the close suburbs (from 10 to 15 kilometers) of €1500. This amounts to a 14.7% reduction from the baseline income gap of €10227. In terms of total spatial income heterogeneity, this does however translate into an increase of 1.8% of the between-municipality standard deviation of mean incomes. Because the total standard deviation of mean incomes decreases, relative segregation (measured as the ratio between the two) increases by 2.3%.

Table 15: Welfare of high and low-skill workers, effects of banning cars in the city center.

	Low skill	High skill	Ratio
Baseline	109.86	594.01	5.41
Counterfactual	107.13	576.46	5.38
Effect (%)	-2.49	-2.95	-0.48

Welfare Finally, the policy creates a welfare loss of 2.49% for low-skilled workers and 2.95% for high-skilled workers, slightly reducing welfare inequalities. These effects are quite substantial, as they are roughly of the same size as the positive effects of the Regional Express Rail.

Increasing amenities in the center It should be noted, however, that these estimates do not take into account the direct effects of banning cars on local amenities in the center in terms of air quality improvement, noise reduction and alternative uses of streets — e.g. terraces. Predicting the magnitude of these effects is challenging, as we lack data on similar measures that would allow to estimate the elasticity of local residential amenities to banning cars from the city.

Therefore, I assess the potential importance of these effects by running alternative simulations where I artificially increase exogenous amenities in the city by 5%, 10% and 15%. In these simulations, I assume that the effects of the policy is proportional to the baseline valuation of residential amenities in the center, so that the relative increase in amenity valuation is identical for high and low-skilled workers. Further, like in the baseline treatment where exogenous amenities stay constant, endogenous amenities still adjust according to the spillover equations (27).

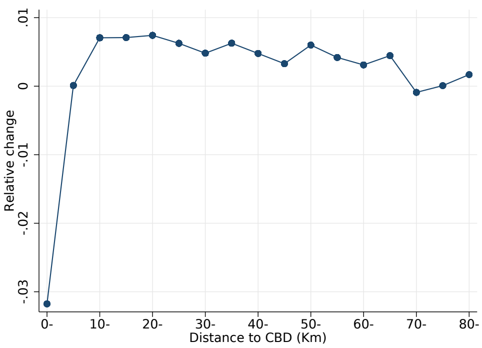
Table 16: Welfare of high and low-skill workers, effects of banning cars in Paris. Alternative assumptions on the effects of the policy on amenities in the city.

	Low skill	High skill	Ratio
No effect on amenities			
Baseline	109.86	594.01	5.41
Counterfactual	107.13	576.46	5.38
Effect (%)	-2.49	-2.95	-0.48
Amenities increase by 5%			
Baseline	109.86	594.01	5.41
Counterfactual	108.33	583.98	5.39
Effect (%)	-1.40	-1.69	-0.30
Amenities increase by 10%			
Baseline	109.86	594.01	5.41
Counterfactual	109.47	591.41	5.40
Effect (%)	-0.36	-0.44	-0.08
Amenities increase by 15%			
Baseline	109.86	594.01	5.41
Counterfactual	110.56	598.77	5.42
Effect (%)	0.63	0.80	0.17

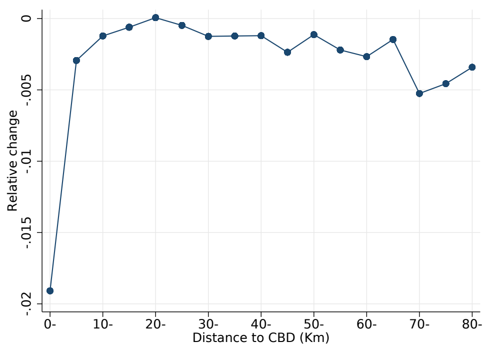
In Table 16, I report the welfare effects of banning cars in Paris under those three alternative assumptions on the effects of the policy on residential amenities in the city. It would take an increase of exogenous amenities comprised between 10% and 15% to reverse the total welfare effect of the policy and make it positive. At a 15% increase in amenities, the welfare effect of the policy on low-skilled workers is 0.63% and the effect on high-skilled workers is 0.80%. As a result, when the welfare effects of the policy become positive, its effects on inequalities are reversed and it benefits more high-skilled workers.

Figure 8: Effect of banning cars in Paris on the income gradient in the region. Alternative assumptions on the effects of the policy on amenities in the city.

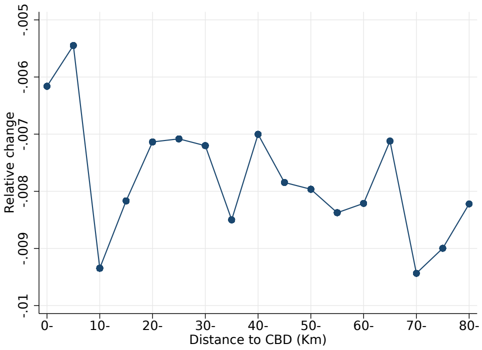
(a) No effect on amenities



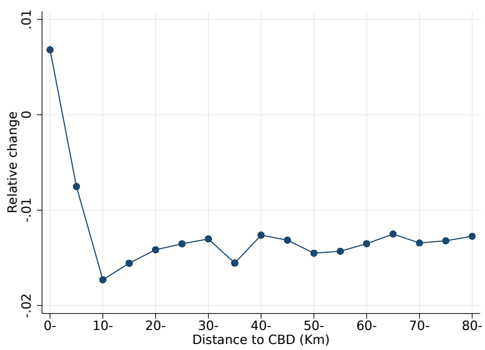
(b) Amenities increase by 5%



(c) Amenities increase by 10%



(d) Amenities increase by 15%



Note: Average effect of banning cars in Paris on mean income in bins of 5km from the city center (first district of Paris), under alternative assumptions on the magnitude of the increase in amenities in the city due to the policy. Inelastic floor space supply.

Turning to spatial inequalities, a 10% increase in amenities is enough to reverse the effects of the policy. As documented in Panel (c) of Figure 8, at that point effect of the policy on the income gradient is almost zero. Further, for a higher increase in amenities the mean income premium of the center over the suburbs increases substantially. For a 15% increase in amenities, mean income increases by 0.5% in the center and decreases by 1.5% in the periphery (10 kilometers away). As reported in Table 21 in appendix, this corresponds to a 7.8% increase in the income gap between the center and the periphery. Although inequalities between the center and the periphery increase, segregation as a whole drops by 2.77%. This is because when richer workers flow in or out of the inner city, they locate in high incomes-high amenity suburban municipalities. When they move back into the center following an increase in amenities, the richest suburban locations become poorer while the poorest ones stay stable, reducing spatial inequalities.

Plausible values for the increase in amenities To better inform the plausible effects of the policy, I turn to back of the envelope calculations of the potential increase in amenities due to the reduction in air pollution, using existing estimates from the literature.

First, Chay and Greenstone (2005) reports an elasticity of house value to suspended particulates between 0.2% and 0.35%, while official estimates from the City of Paris state that 35% of the suspended particulates emissions in the city come from road traffic. Extrapolating from the estimates in Chay and Greenstone (2005), we should anticipate a 7% to 12% in house values.

Similarly, Sullivan (2016) find a semi-elasticity of house values to NO2 of 0.3%. The City of Paris estimates that 60% of the baseline concentration in the city of $40\mu g/m^3$ comes from road traffic. Applying the estimate from Sullivan (2016), we would get a 6% increase in house value from reduced pollution.

Looking at the results of model simulations, this 6% to 12% increase in house values would necessitate a 5% to 10% increase in amenities. If in the upper part of this range, the reduction in pollution alone would therefore put the policy in a regime where it has no impact on spatial income inequalities at large (-0.25%) and close to no impact on inequalities between the center and the periphery (0.33%). Further, the negative welfare effects would be moderate, and of comparable size for low skilled (-0.36%) and high skilled workers (-0.44%).

5 Conclusion

In this paper, I use an equilibrium model of a metropolitan area to evaluate the impact of transportation infrastructures on spatial disparities within cities. Calibrated on the Paris region, the model is able to closely replicate the spatial repartition of economic activities and the income gradient in the city. Further, simulated effects of public transports on local employment and population are in line with existing reduced-form results.

Simulating away the Regional Express Rail, the model shows that it hat sizable effects on income sorting between the city and its suburbs, and reduces the total income inequalities in the area. Further, it does bring a higher welfare gain to low-skilled workers than high-skilled workers, reducing welfare inequalities. On the other hand, looking at an increase in travel time costs through banning cars in the city I find a sizable reduction in spatial inequalities between the city and its suburbs, at the cost of a total welfare loss for both skilled and unskilled workers.

The total welfare loss of banning cars is of a similar magnitude to the welfare gains from the Regional Express Rail, a very important transportation infrastructure. These effects depend however on the direct amenity gains in the city center from the policy. It would require a more than 10% increase in amenities in Paris from pollution reduction and regained floor space for the policy to break even and start having a positive welfare effect, at which point it would start increasing segregation. Back of the envelope calculations based on estimated effects of pollution on housing values suggest that the policy should be close to this break even point. As these amenity effects are determinant in the total effect of the policy, further work should quantify them.

References

Ahlfeldt, G. M. and Pietrostefani, E. (2019). The economic effects of density: A synthesis. *Journal of Urban Economics*, 111:93–107.

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.

Albouy, D. (2016). What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities. *Review of Economics and Statistics*, 98(3):477–487.

Allen, T., Arkolakis, C., and Takahashi, Y. (2020). Universal gravity. *Journal of Political Economy*, 128(2):393–433.

Almagro, M. and Domínguez-Iino, T. (2019). Location Sorting and Endogenous Amenities: Evidence from Amsterdam.

Baum-Snow, N. (2007). Did highways cause suburbanization? *Quarterly Journal of Economics*, 122(2):775–805.

Calabrese, S., Epple, D., Romer, T., and Sieg, H. (2006). Local public good provision: Voting, peer effects, and mobility. *Journal of Public Economics*, 90(6-7):959–981.

Chay, K. Y. and Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, 113(2):376–424.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742.

Combes, P.-P., Duranton, G., and Gobillon, L. (2017). The Production Function for Housing: Evidence from France.

Combes, P.-P., Duranton, G., and Gobillon, L. (2018). The Costs of Agglomeration: House and Land Prices in French Cities. *The Review of Economic Studies*, 63.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. *Agglomeration Economics*, I(February):15–65.

Combes, P. P. and Gobillon, L. (2015). The Empirics of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 5A, chapter 5, pages 247–348. Elsevier.

Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2018). Income Growth and the Distributional Effects of Urban Spatial Sorting.

Couture, V. and Handbury, J. (2020). Urban revival in America. *Journal of Urban Economics*, 119.

de Bellefon, M. P., Combes, P. P., Duranton, G., Gobillon, L., and Gorin, C. (2020). Delineating urban areas using building density. *Journal of Urban Economics*, (November 2019).

Epple, D., Romer, T., and Sieg, H. (2001). Interjurisdictional sorting and majority rule: an empirical analysis. *Econometrica*, 69(6):1437–1465.

Gaigné, C., Koster, H. R., Moizeau, F., and Thisse, J.-F. (2019). Who Lives Where in Cities? Amenities, Commuting and Income Sorting.

Garcia-López, M. à. (2012). Urban spatial structure, suburbanization and transportation in Barcelona. *Journal of Urban Economics*, 72(2-3):176–190.

Garcia-López, M. À., Hémet, C., and Viladecans-Marsal, E. (2017). How does transportation shape intrametropolitan growth? An answer from the Regional Express Rail. *Journal of Regional Science*, 57(5):758–780.

Garcia-López, M. À. and Moreno-Monroy, A. I. (2018). Income segregation in monocentric and polycentric cities: Does urban form really matter? *Regional Science and Urban Economics*, 71(May):62–79.

Garcia-López, M.-À., Pasidis, I., and Viladecans-Marsal, E. (2018). Amphitheaters, cathedrals and operas: The role of historic amenities on suburbanization. CEPR Discussion Papers 13129, C.E.P.R. Discussion Papers.

Glaeser, E. L., Kim, H., and Luca, M. (2018). Measuring Gentrification: Using Yelp Data To Quantify Neighborhood Change.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108(September):85–106.

González-Pampillón, N., Jofre-Monseny, J., and Viladecans-Marsal, E. (2019). Can urban renewal policies reverse neighborhood ethnic dynamics? *Journal of Economic Geography*, 20(2):419–457.

Harter-Dreiman, M. (2004). Drawing inferences about housing supply elasticity from house price responses to income shocks. *Journal of Urban Economics*, 55(2):316–337.

Koster, H. R. and Rouwendal, J. (2017). Historic Amenities and Housing Externalities: Evidence from the Netherlands. *The Economic Journal*, 127(605):F396–F420.

Koster, H. R. A., van Ommeren, J. N., and Rietveld, P. (2016). Historic amenities, income and sorting of households. *Journal of Economic Geography*, 16(1):203–236.

Lee, S. and Lin, J. (2018). Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income. *The Review of Economic Studies*, 85(1):663–694.

Mayer, T. and Trevien, C. (2017). The impact of urban public transportation evidence from the Paris region. *Journal of Urban Economics*, 102:1–21.

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

Padgham, M. (2019). dodgr: An R package for network flow aggregation. *Transport Findings*.

Rothstein, J. M. (2006). Good principals or good peers? Parental valuation of school characteristics, tiebout equilibrium, and the incentive effects of competition among jurisdictions. *American Economic Review*, 96(4):1333–1350.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, 125(3)(August):1253–1296.

Sullivan, D. M. (2016). The True Cost of Air Pollution: Evidence from House Prices and Migration. *Harvard Environmental Economics Program Discussion Paper*, (16):1–69.

Tsivanidis, N. (2019). The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio. *Working Paper*, pages 1–59.

Wingender, A. M. (2015). Skill complementarity and the dual economy. *European Economic Review*, 74:269–285.

A Appendix

A.1 Regional Express Rail

Table 17: Travel times in minutes, with and without the RER network.

Dist. to	To the whole region				To the city center		
center (km)	RER	No RER	Dif (%)	Con.(%)	RER	No RER	Dif (%)
0-	94.67	111.29	18.83	23.94	39.67	41.39	5.16
5-	108.71	131.16	21.33	26.32	57.46	66.12	16.50
10-	117.28	144.60	23.54	30.11	67.09	80.07	20.98
15-	123.99	153.98	25.04	32.95	73.28	94.95	32.09
20-	132.44	160.57	22.02	31.22	82.69	103.12	25.66
25-	144.52	178.59	23.55	33.43	93.74	117.73	26.78
30-	152.99	190.51	25.38	39.68	100.49	124.28	24.69
35-	154.29	185.83	20.48	26.69	102.67	119.15	16.28
40-	165.80	202.63	22.26	43.27	112.26	130.47	15.55
45-	173.98	215.38	22.69	50.27	118.83	142.74	16.58
50-	188.50	220.03	16.44	23.77	131.79	140.92	6.17
55-	184.25	222.37	20.44	28.27	130.00	142.52	9.35
60-	179.69	235.56	28.57	52.35	121.19	150.67	19.57
65-	206.64	226.41	9.63	16.45	145.95	145.47	-0.22
70-	226.19	249.21	10.28	16.82	164.07	165.85	0.89
75-	231.49	254.32	9.13	15.72	171.35	171.45	0.06
Total	138.52	168.61	22.11	31.19	86.18	102.34	19.63

Average travel times using public transports, in minutes. Each row is a distance bin (in km) to the city center. First four columns report the travel time toward the whole area, originating from a given bin. RER reports travel times with the network, "No RER" travel times without the network, and "Dif.(%)" the relative difference between the two. "Con.(%)" is the relative difference for the restricted sample of commutes where either the origin or destination municipality has a RER stop. Three last columns report the travel time to the city center, originating from a given bin.

Table 18: Effects of the Regional Express Rail on municipalities

	Mean	S.D.	Min.	Q1	Q2	Max.
Mean income	0.0051	0.0197	-0.0931	-0.0050	0.0107	0.1303
Unskilled population	-0.0111	0.1052	-0.2155	-0.0671	0.0118	0.8484
Skilled population	0.0138	0.1248	-0.3111	-0.0610	0.0609	0.6113
Unskilled employment	-0.0306	0.0816	-0.3149	-0.0555	-0.0239	0.7959
Rent	-0.0051	0.0457	-0.0929	-0.0326	0.0062	0.2471

Descriptive statistics of the contribution of the RER network to municipal outcomes.
Each observation is a municipality in the Paris region. $N = 696$.

A.2 Effects of the whole public transport network

Table 19: Welfare of high and low-skill workers, with and without public transports.

	Low skill	High skill	Ratio
Baseline	109.86	594.01	5.41
Counterfactual	85.40	465.62	5.45
Effect (%)	22.27	21.61	-0.84

Table 20: Aggregate effects the transport network on incomes in the Paris region.

	Mean	Total SD	Between SD	(%)	C−P
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	30945.86	20460.62	3555.16	17.38	8976.88
Effect (%)	0.68	2.10	-2.24	-4.44	12.22

Column 1 is the average income over all workers. Column 2 is the total standard deviation of income over individual workers. Column 3 is the between-municipality standard deviation. Column 4 is the ratio between the two times 100. Column 5 is mean income in the center minus mean income 10 to 15 kilometers away from the center.

A.3 Banning cars

Table 21: Aggregate effects of banning cars from Paris.

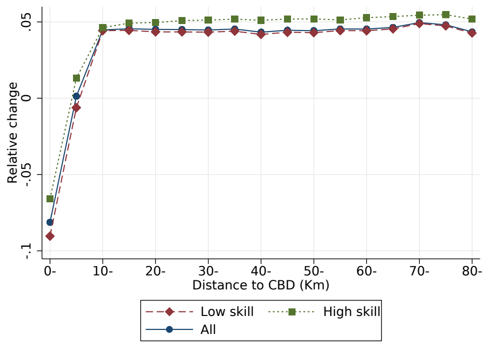
	Mean	Total SD	Between SD	(%)	C−P
No effect on amenities					
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	31093.20	20784.18	3538.56	17.03	8727.20
Effect (%)	-0.21	-0.55	1.77	2.33	-14.66
Amenities increase by 5%					
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	31025.42	20718.69	3481.36	16.80	9491.11
Effect (%)	-0.43	-0.86	0.12	0.99	-7.19
Amenities increase by 10%					
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	30966.18	20659.48	3428.68	16.60	10260.20
Effect (%)	-0.62	-1.15	-1.39	-0.25	0.33
Amenities increase by 15%					
Baseline	31158.68	20899.29	3477.11	16.64	10226.89
Counterfactual	30914.19	20604.64	3380.69	16.41	11025.90
Effect (%)	-0.78	-1.41	-2.77	-1.38	7.81

Column 1 is the average income over all workers. Column 2 is the total standard deviation of income over individual workers. Column 3 is the between-municipality standard deviation. Column 4 is the ratio between the two times 100. Column 5 is mean income in the center minus mean income 10 to 15 kilometers away from the center.

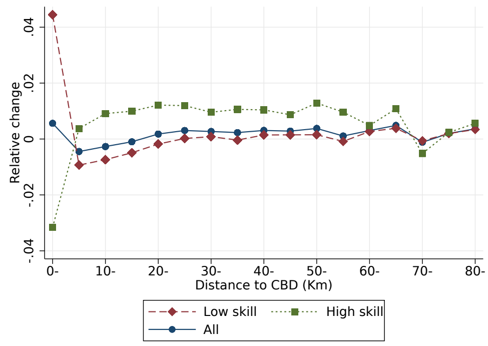
A.4 Banning cars: gradients with amenity gains

Figure 9: Effect of banning cars in Paris, 5% amenity increase in the city.

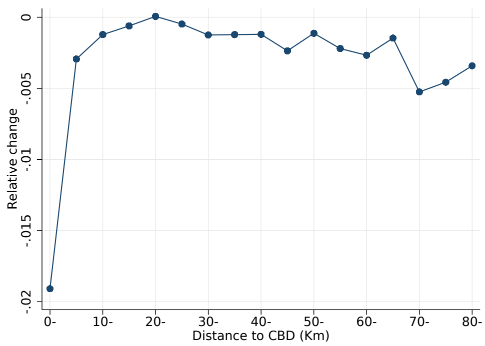
(a) Employment



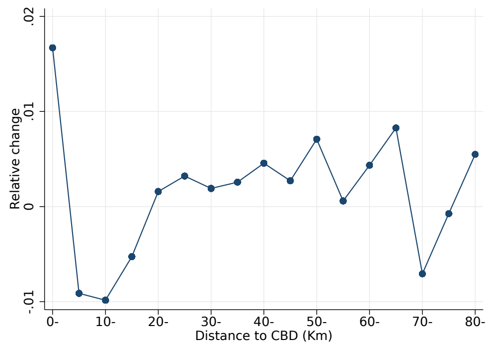
(b) Population



(c) Mean income



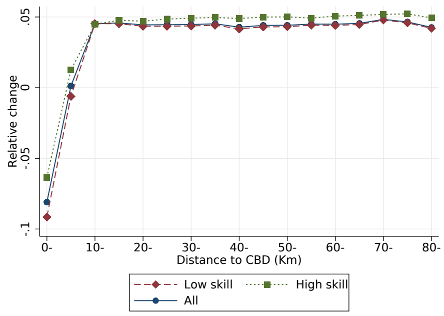
(d) Rent



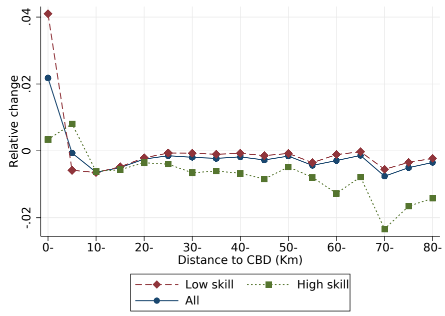
Note: Average effect of banning cars in Paris on employment, population, mean incomes and the between-city coefficient of variation of mean incomes, in bins of 5km from the city center (first district of Paris). Inelastic floor space supply.

Figure 10: Effect of banning cars in Paris, 10% amenity increase in the city.

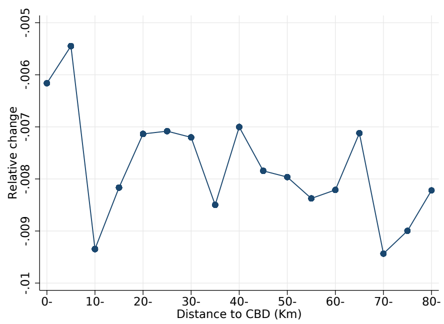
(a) Employment



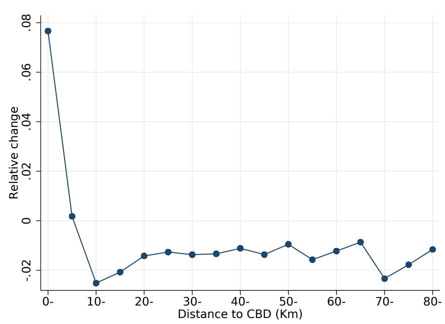
(b) Population



(c) Mean income

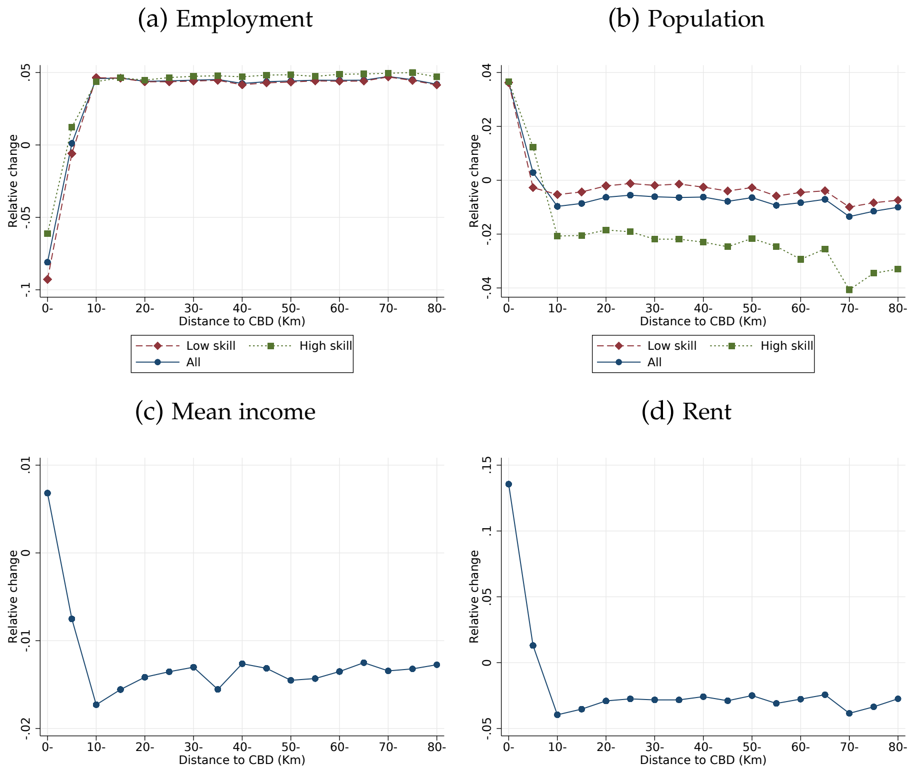


(d) Rent



Note: Average effect of banning cars in Paris on employment, population, mean incomes and the between-city coefficient of variation of mean incomes, in bins of 5km from the city center (first district of Paris). Inelastic floor space supply.

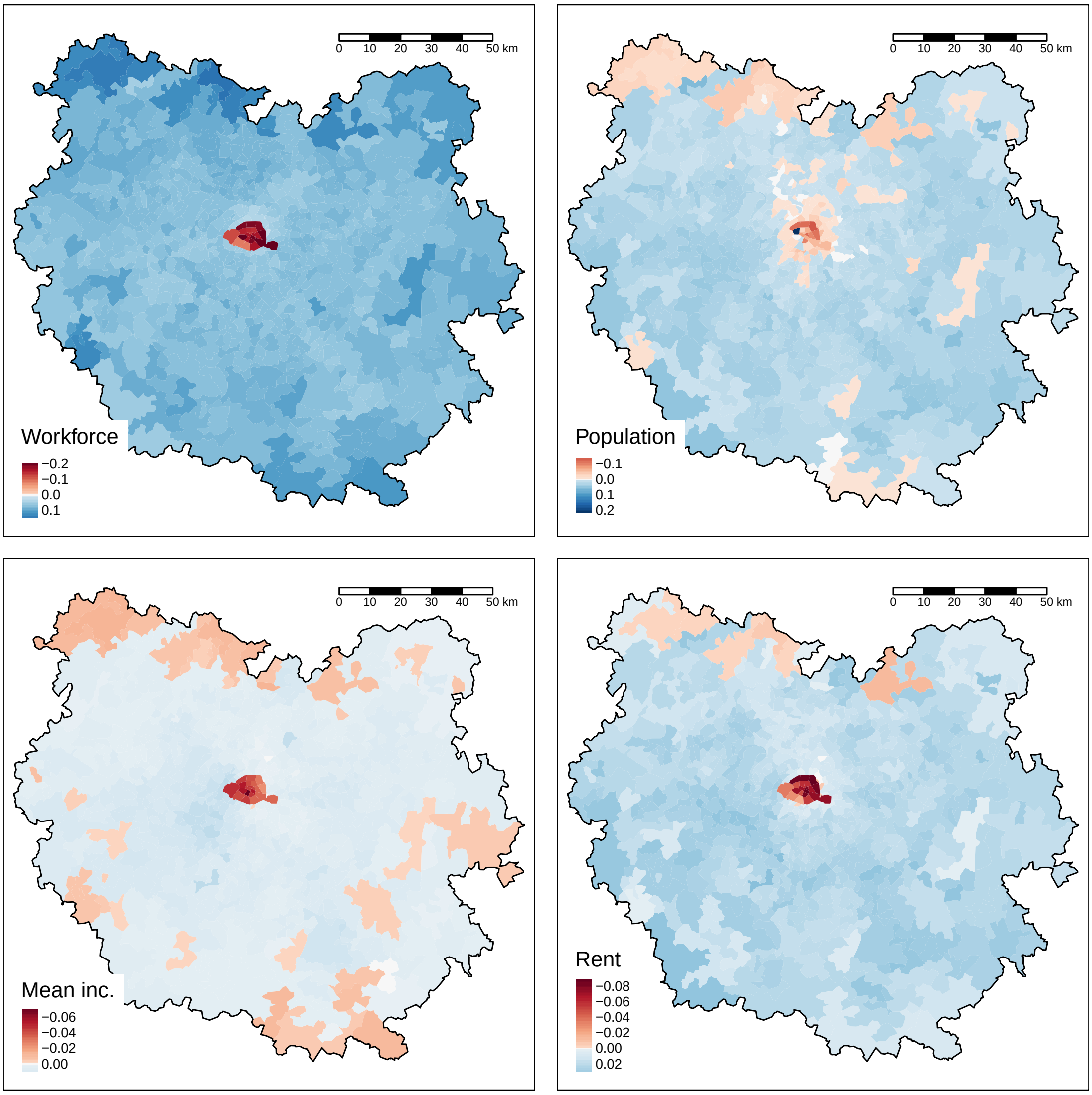
Figure 11: Effect of banning cars in Paris, 15% amenity increase in the city.



Note: Average effect of banning cars in Paris on employment, population, mean incomes and the between-city coefficient of variation of mean incomes, in bins of 5km from the city center (first district of Paris). Inelastic floor space supply.

A.5 Banning cars: elastic floor space supply

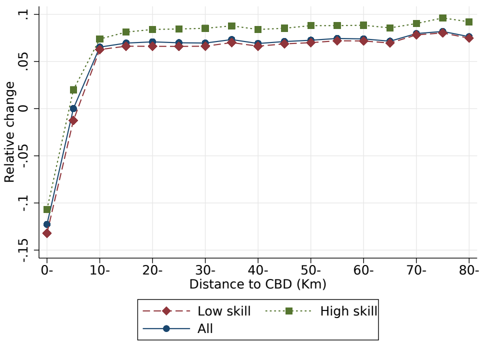
Figure 12: Effects of banning cars in Paris.



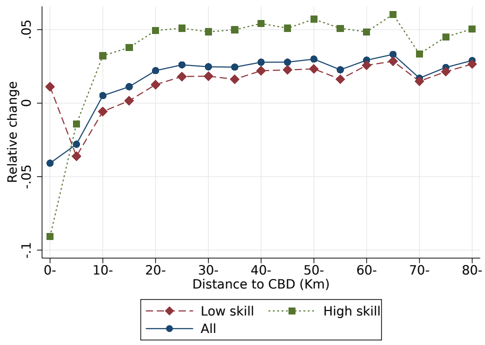
Note: Maps of the effects of banning commuting by cars to and from Paris on the number of workers (HM), the number of residents (HR), average income of residents (meaninc) and rents (rent) in the Paris metropolitan area. Elastic floor space supply.

Figure 13: Effect of banning cars in Paris.

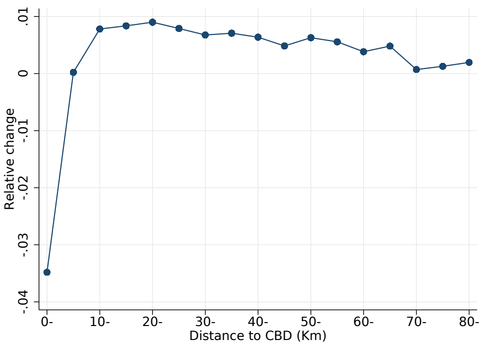
(a) Employment



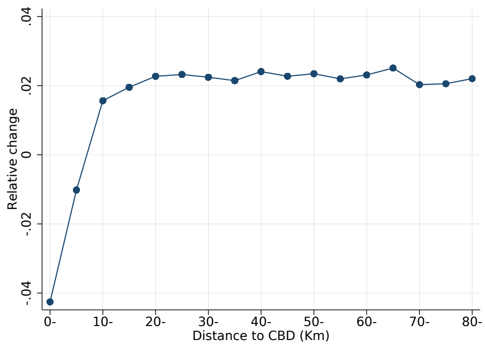
(b) Population



(c) Mean income



(d) Rent



Note: Average effect of banning cars in Paris on employment, population, mean incomes and the between-city coefficient of variation of mean incomes, in bins of 5km from the city center (first district of Paris). Elastic floor space supply.

A.6 Baseline maps

Figure 14: Mean income. Actual (left) vs predicted (right)

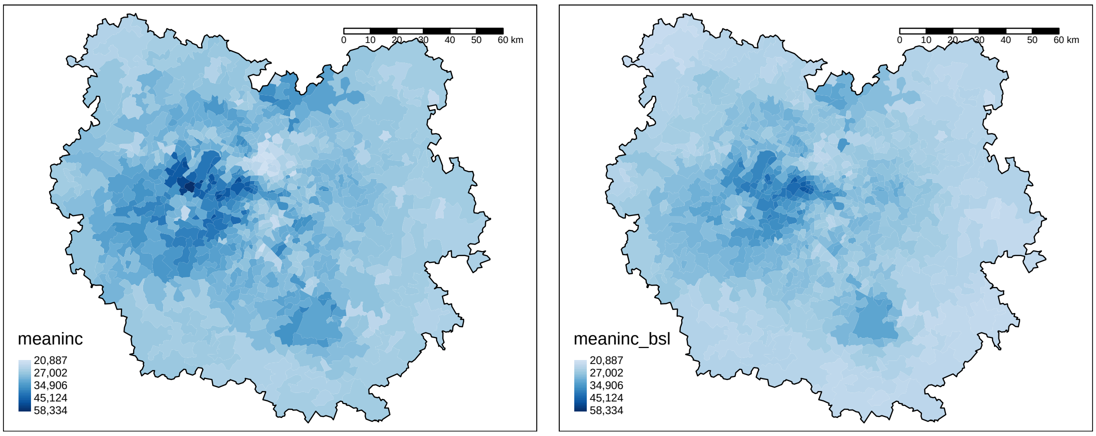
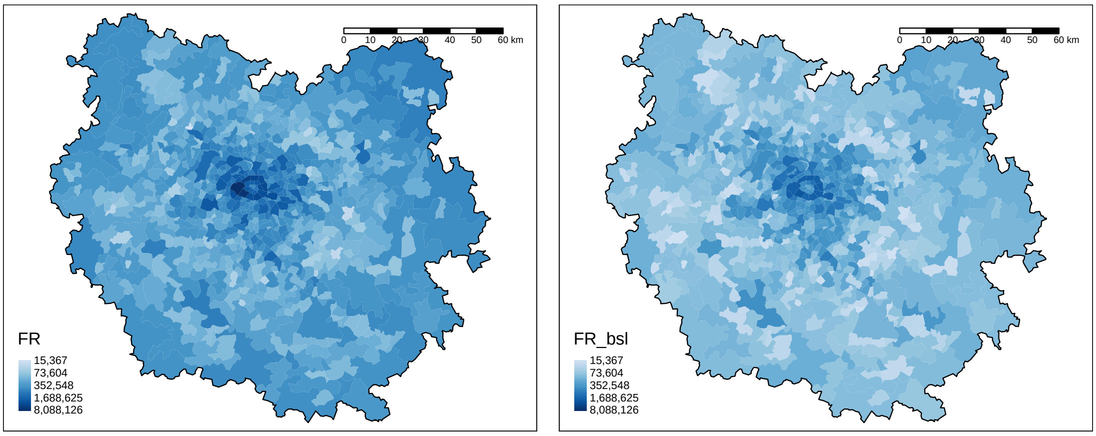


Figure 15: Residential floor space. Actual (left) vs predicted (right)



A.7 Model based IV

Table 22: Decomposition of fundamentals.

	(1)	(2)	(3)	(4)	(5)	(6)
	TFP	Skill bias	B (s.)	B (u.)	T (s.)	T (u.)
Mean altitude (log)	-0.0930***	-0.162**	-0.183	0.00874	-0.573***	-0.510***
	(-4.68)	(-3.01)	(-1.52)	(0.07)	(-3.98)	(-3.96)
Maximum slope (log)	-0.0143***	-0.0413***	-0.0161	0.0362*	-0.0444*	-0.0346*
	(-5.41)	(-5.78)	(-1.00)	(2.33)	(-2.31)	(-2.02)
Distance to river (log)	0.0119	0.0305	-0.0243	-0.0114	0.0429	0.0777
	(1.40)	(1.33)	(-0.47)	(-0.23)	(0.70)	(1.41)
Distance to river ¡ 5km	0.0440*	0.0789	0.202	0.0777	0.234	0.264*
	(2.21)	(1.47)	(1.67)	(0.66)	(1.62)	(2.04)
Listed building dummy			0.597***	0.602***	0.544***	0.590***
			(8.11)	(8.44)	(6.20)	(7.53)
Constant	10.58***	0.351	-2.646***	-3.901***	2.269***	1.856**
	(122.07)	(1.50)	(-5.01)	(-7.62)	(3.59)	(3.29)
r2_within	0.134	0.0889	0.117	0.102	0.116	0.127
F	26.70	16.86	18.26	15.75	18.08	20.06
N	696	696	696	696	696	696

Regression of model fundamentals on exogenous variables. Predicted values from these regressions are used as inputs for the model simulation that generates instruments for incomes and workforce.

Table 23: Observed amenities

	count	mean	sd	min	max
restaurant	696	14.79053	59.61473	0	766.6667
convenience	696	1.766946	5.252166	0	57.29167
theatres	696	.1270503	.5421777	0	7.070707
cinema	696	.0956781	.3891422	0	6.976744
meat_shops	696	.8831948	2.166097	0	24.65278
bakeries	696	1.915192	4.725206	0	40.40404
preschools	696	.6877911	1.008322	0	7.142857
hairdresser	696	3.530352	9.460522	0	111.4583
doctor	696	2.630785	5.836409	0	49.15966
laboratory	696	.2067682	.4275947	0	3.669725
police	696	.067342	.183323	0	1.834862
tennis	696	.2408132	.3004812	0	2.857143
golf	696	.0183569	.0662338	0	.6944444
hiking	696	.0154199	.0560326	0	.4385965
horses	696	.0516236	.1067405	0	1.587302
swimpool	696	.1071565	.2320357	0	2.020202

Description of observed amenities, in number per squared kilometer. Observation is a municipality. N=696.

B Data

B.1 Data sources

Workers Microdata (DADS): The *Déclarations Automatiques de Sécurité Sociale* are an administrative, restricted-access dataset on the universe of French workers. Sent by employers to the social security administration on a yearly basis to be used for the computation of social security contributions. They contain the salaries, hours worked, occupation, workplace and dwelling place of every French employee. They are exhaustive on the universe of French private payroll employees and available from 1993 to 2015. However, it is not a proper panel as individual IDs are scrambled every two years. Absent data on education, I use occupation categories, and treat grey matter, managers and professionals as high skill workers and the rest of the workforce as low skill workers.

Household Expenditure survey: The *Enquête Budget des Familles* is a representative survey of French households expenditures conduced by the National Statistical Institute. It contains household composition, housing expenditures, household income and housing surface area. For the estimations, I pool the 2006 and 2011 waves of the survey.

Building transactions (DVF): The *Demande de Valeurs Foncières* is an open dataset, exhaustive on the universe of building transactions in France starting in 2014.

Land registry files (FF): The *Fichiers Fonciers* from the French tax administrations are an exhaustive dataset on the universe of French properties. They report, for each property in France, its floor space area and its fiscal status, either as a dwelling or as a place of business.

Travel Times (TT): Average road travel times between municipalities are computed using extractions of the road network from the OpenStreetMap project and the *dodgr* R package (Padgham, 2019). For the public transport network, I use publicly available GTFS transit timetables and compute travel times between municipality centroids at 8 in the morning on a tuesday.

B.2 Geographical units

Because the estimation procedures used cannot handle geographical units with zero employment or zero residents, to ensure some precision in the estimation of local wages, rents, TFPs and amenities and to comply with legal regulations on exporting aggregates from restricted acces microdata in France, I pull municipalities toguether into groups so that each group has at least 10 workers and 10 residents of each type. To minimize the heterogeneity between municipalities in a same group, I use a procedure that tries to minimize the rent differential between merged municipalities. More precisely, I use the following iterative procedure:

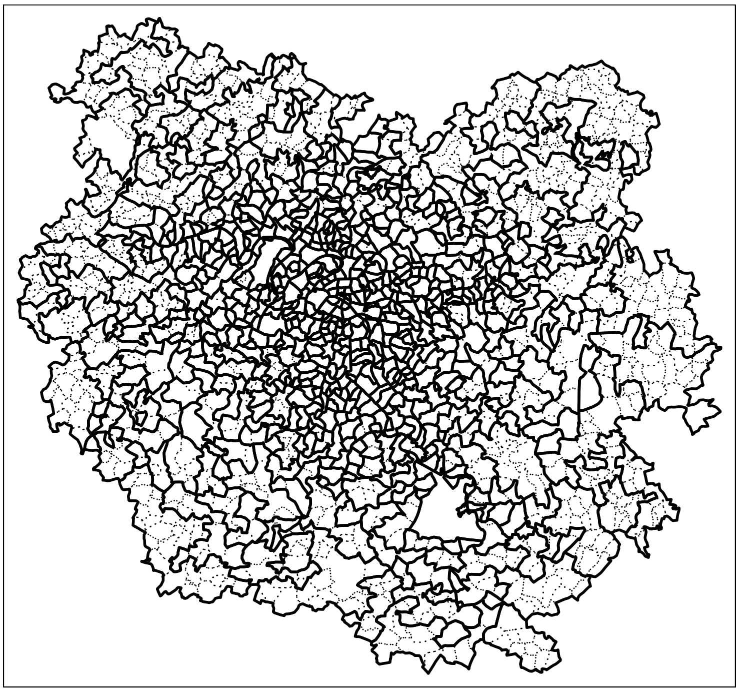
0. Create groups consisting of only one municipality. Make a list of the groups that do not meet the criterion.
1. If the list is empty, exit. Else, choose the first group of the list.
2. Amongst adjacent groups, find the one that has the closest average rent per squared meter and merge the two groups.
3. Place that group at the bottom of the list. Go to 1.

Rents are used to measure the distance between municipalities because it is the variable with the best coverage in the raw data, with no missing value at the municipal level. Second, rent is a good indicator of the general attractiveness of a location as it is strongly correlated with income and population. I therefore expect that pooling neighboring municipalities with similar rents will also minimize the within unit variation in populations, income and amenities. Overall, the procedures leaves central, highly densely populated areas unchanged and only pools peripheral, almost empty locations. These locations are highly homogenous in their emptiness and inexpensiveness, and they mainly serve as an outside option to allow workers to move out of the city center in counterfactual simulations as they have little weight in the estimation.

B.3 Travel times

Travel times by car are computed for all pairs ij using the road network extracted from OpenStreetMap. They are computed in minutes between each pair of municipalities, and are theoretical travel times based on the road network and speed limits.

Figure 16: Municipalities (dashed) and pooled units (solid) for the Urban Area of Paris.

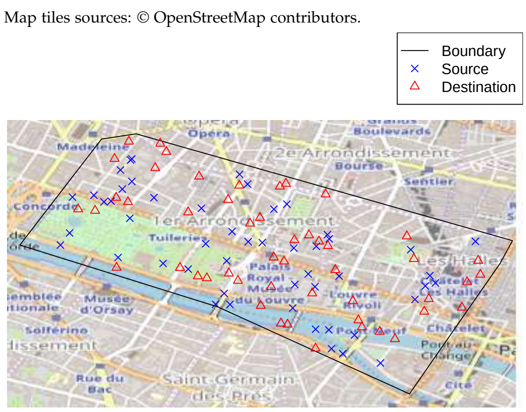


Congestion is not taken into account. Figure 17b plots one line of the travel time matrix for the Urban Area of Paris. Travel times in the Paris area range from 7 minutes to two and a half hours.

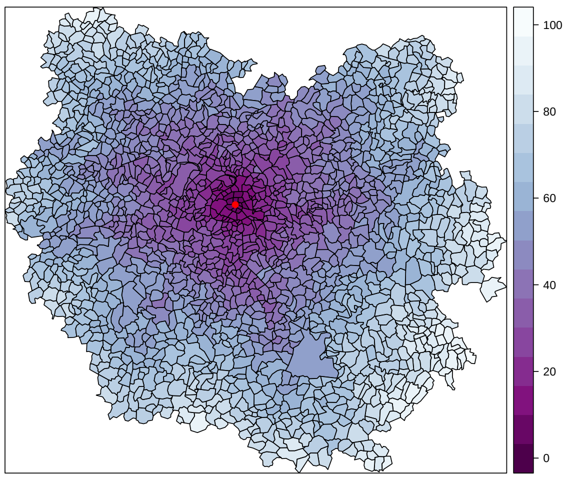
To approximate the average travel time between municipalities, I average travel times between randomly drawn pairs of points within each municipalities. For each municipality pair, 50 origins and 50 destinations are randomly drawn. Figure 17a shows such a sample of origins and destinations within the first district of Paris. Given the 50 sources and 50 destinations, the 2500 pairwise travel times between them are computed and their average is taken as the average travel time within the municipality.

Travel times by public transport are computed for all pairs of municipalities using the street network from OpenStreetMap to find entry points and GTFS data on the transport network in the region from RATP and SNCF. These two data sources are fed into the OpenTripPlanner API. The public transport travel times used in the paper correspond total travel times by all means of public transport available, including walking time and waiting time, assuming a departure time at 8 AM on a Tuesday.

Figure 17: Examples.



(a) Example of sample points



(b) Travel times in minutes from the 1st district of Paris (red dot)