The Costs of Agglomeration: Land Prices in French Cities

Pierre-Philippe Combes[*†]
Aix-Marseille University

Gilles Duranton[*‡]
University of Pennsylvania

Laurent Gobillon[*§]
Institut National d'Etudes Démographiques

November 2012

Abstract: We develop a new methodology to estimate the elasticity of urban costs with respect to city population using French land price data. Our preferred estimate, which handles a number of estimation concerns, stands at 0.041. Our approach also yields a number of intermediate outputs of independent interest such as a distance gradient for land prices and the elasticity of unit land prices with respect to city population. For the latter, our preferred estimate is 0.72.

Key words: urban costs, land prices, land use, agglomeration

JEL classification: R14, R21, R31

[*]We thank conference and seminar participants, Monica Andini, Morris Davis, Jan Eeckhout, Sanghoon Lee, François Ortalo-Magné, Henry Overman, Jean-Marc Robin, Nathan Schiff, and Daniel Sturm for their comments and suggestions. We also thank Pierre-Henri Bono and Giordano Mion for their help with the data. Finally, we are grateful to the *Service de l'Observation et des Statistiques (SOeS) - Ministère de l'Écologie, du Développement durable et de l'Énergie* for giving us on-site access to the data.

[†]Aix-Marseille University (Aix-Marseille School of Economics), CNRS, and EHESS, 2 Rue de la Charité, 13236 Marseille cedex 02, France (e-mail: `ppcombes@univmed.fr`; website: `http://www.vcharite.univ-mrs.fr/pp/combes/`). Also affiliated with the Centre for Economic Policy Research.

[‡]Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, USA (e-mail: `duranton@wharton.upenn.edu`; website: `https://real-estate.wharton.upenn.edu/profile/21470/`). Also affiliated with the Center for Economic Policy Research, the Spatial Economic Centre at the London School of Economics, and the Rimini Centre for Economic Analysis.

[§]Institut National d'Etudes Démographiques, 133 Boulevard Davout, 75980 Paris cedex 20, France (e-mail: `laurent.gobillon@ined.fr`; website: `http://laurent.gobillon.free.fr/`). Also affiliated with the Paris School of Economics, the Centre for Economic Policy Research, and the Institute for the Study of Labor (IZA).

1. Introduction

We develop a new methodology to estimate the elasticity of urban costs with respect to city population using French land price data. Our preferred estimate, which handles a number of estimation concerns, stands at 0.041. Our approach also yields a number of intermediate outputs of independent interest such as a distance gradient for land prices and the elasticity of unit land prices with respect to city population. For the latter, our preferred estimate is 0.72.

Reasonable estimates for urban costs and how they vary with city population are important for a number of reasons. First, following Henderson (1974) and Fujita and Ogawa (1982), cities are predominantly viewed as the outcome of a tradeoff between agglomeration economies and urban costs. Fujita and Thisse (2002) dub it the 'fundamental tradeoff of spatial economics'. The existence of agglomeration economies is now established beyond reasonable doubt and much has been learnt about their magnitude (see Rosenthal and Strange, 2004, Puga, 2010, Combes, Duranton, and Gobillon, 2011, for reviews). Far less is known about urban costs. High housing prices and traffic jams in Central Paris, London, or Manhattan are for everyone to observe. There is nonetheless little systematic evidence about urban costs. In this paper, we provide such evidence. This allows us to assert the existence of the fundamental tradeoff of spatial economics empirically.

Our estimate for the elasticity of urban costs with respect to population is close to the corresponding elasticity for agglomeration economies. That cities operate near aggregate constant returns to scale is suggestive that the fundamental tradeoff of spatial economics, despite its existence, is unlikely to have much power in determining the future evolution of city sizes. In turn, this finding may be an important reason for why cities of vastly different sizes exist and prosper.

Second, urban policies attempt to limit the growth of cities in many countries. These restrictive policies, which often take the form of barriers to labour mobility and zoning policies that limit new constructions, are particularly prevalent in developing countries (see Duranton, 2008, for a review). The underlying rationale for these policies is that the population growth of cities imposes large costs to already established residents by bidding up housing prices and crowding out the roads. Our analysis shows that in the French case, the costs of having larger cities are modest and of the same magnitude as agglomeration economies. This lends little support to the imposition of barriers to urban growth.

Finally, to obtain the elasticity of urban costs with respect to city population we must first estimate the elasticity of unit land prices with respect to population. This elasticity is interesting in its own right. Simple urban models in the tradition of Alonso (1964), Muth (1969), and Mills (1967) provide stark predictions for it. Our results suggest that these models provide surprisingly good approximations of the French urban reality after allowing cities to decentralise as they grow.

Tolley, Graves, and Gardner (1979), Thomas (1980), Richardson (1987), and Henderson (2002) are the main antecedents to our research.[1] To the best of our knowledge, this short list is close to exhaustive. Despite the merits of these works, none of their estimates has had much influence. We attribute this lack of credible estimate for urban costs and the scarcity of research on the subject to a lack of integrated framework to guide empirical work, a lack of appropriate data, and a lack of attention to a number of identification issues — the three main innovations of this paper.[2]

Urban costs take a variety of forms. In larger cities, housing is more expensive, commutes are longer, and the bundle of amenities that these cities offer may differ. We first develop a theoretical framework to show how we can estimate the elasticity of urban costs with respect to city population using land price data. Our starting point is a notion of spatial equilibrium. Residential mobility implies that urban (dis-)amenities and commuting costs are reflected into land prices. We can then monetise the costs of higher land prices using an expenditure function. Our theoretical model indicates that the elasticity of urban costs with respect to city population is the product of three quantities: the elasticity of unit land prices at the city centre with respect to population, the share of land in housing, and the share of housing in consumption expenditure. This last quantity can be readily obtained from a detailed evaluation made by the French ministry that oversees housing (CGDD, 2011). We follow their official estimate of 0.23. We can estimate the share of land in housing from our main source of data: a record of transactions for land parcels with a development or redevelopment permit which also contains a measure of construction costs. Our simple estimate of 0.25 is close to the results obtained in our companion paper (Combes, Duranton, and Gobillon, 2012) which provides a detailed investigation of the production function of housing.

[1]Thomas (1980) compares the cost of living for four regions in Peru focusing only on the price of consumption goods. Richardson (1987) compares 'urban' and 'rural' areas in four developing countries. Closer to the spirit of our work, Henderson (2002) regresses commuting times and rents to income ratio for a cross-section of cities in developing countries.

[2]As will become clear below, our work is also related to the large literature that estimates distance 'gradients' in cities following Clark's (1951) path-breaking work on urban densities (see McMillen, 2006, for a review). Most of this literature focuses on one city (or a small subset of cities) and usually considers distance gradients for population density or housing development. We look instead at land price gradients for a large sample of cities. There is also a literature that measures land values for a broad cross-section of urban (and sometimes rural) areas (Davis and Heathcote, 2007, Davis and Palumbo, 2008, Albouy and Ehrlich, 2012). We enrich it by considering the internal geography of cities and by investigating the determinants of land prices, population in particular, at the city level.

Estimating the elasticity of unit land prices at the centre of each city with respect to city population is more demanding and is the main focus of our empirical analysis. Guided by our model, we proceed in two steps. First, we use information about the location of parcels in each city to estimate unit land prices at its centre. Second, we regress these estimated prices at the centre on city population to obtain an estimate of the elasticity of unit land prices in each city with respect to city population. Our preferred estimate for this last quantity is 0.72. Multiplying this quantity by the share of land in housing and by the share in housing in consumption yields our preferred estimate for the elasticity of urban costs with respect to population of 0.041. Estimating the elasticity of land prices with respect to city population brings up a number of challenges. Most notably, land prices and city population are simultaneously determined. For instance, a high (unobserved) productivity in a city will cause both high land prices and a large population. To deal with this identification problem, our model suggests using amenities as instruments since amenities determine city population but do not otherwise intervene in the determination of land prices. We also confirm our results with a variety of other instruments external to the model.

2. Model

The model we develop in this section serves three purposes. First, it provides a simple and intuitive formula to compute urban costs. Second, it generates a specification that we later implement. Third, it highlights key identification concerns and suggests an empirical strategy to handle them.

Housing demand and housing supply

In the spirit of Alonso (1964), Mills (1967), and Muth (1969), we start with a monocentric circular city. City residents must commute to the central business district (CBD) to receive a wage W. The utility of a resident living at a distance D from the CBD is:

$$U(D) = \frac{M}{\beta^{\beta}(1-\beta)^{1-\beta}} \frac{h(D)^{\beta}\, x(D)^{1-\beta}}{v(D)}\,, \tag{1}$$

where M denotes the quality of amenities in the city, $h(D)$ is housing consumption, $x(D)$ is the consumption of a numéraire composite good, and $v(D)$ is the utility cost of commuting, which is increasing with distance.[3] The budget constraint is $Q(D)h(D) + x(D) = W$ where $Q(D)$ is the rental price of housing at a distance D from the CBD.

The first-order conditions for utility maximisation imply:

$$h(D) = \frac{\beta W}{Q(D)} \tag{2}$$

and $x(D) = (1 - \beta)W$. Substituting these two demand functions into equation (1) yields the following indirect utility:

$$U(D) = \frac{M W}{[Q(D)]^{\beta} v(D)} . \tag{3}$$

In equilibrium, the rental price of housing adjusts so that residents are indifferent across all residential locations between the CBD and the urban fringe. In particular, we have $U(D) = U(0)$ for $0 \leq D \leq \overline{D}$ where $\overline{D}$ denotes the urban fringe. Using this spatial equilibrium condition into equation (3) yields the following distance gradient for the rental price of housing in the city:

$$Q(D) = Q(0) \left[\frac{v(0)}{v(D)}\right]^{1/\beta} . \tag{4}$$

Turning to housing supply, we assume that housing is produced by competitive, profit-maximising builders using capital K, land L, and housing technology B. At any location, the production of housing is given by:

$$H = BK^{1-\alpha}L^{\alpha} . \tag{5}$$

For land located at a distance D from the CBD, the profit of a builder is $\pi(D) = Q(D)\, H(D) - R(D)L(D) - r^K K(D)$ where $R(D)$ is the rental price of land, r^K is the user cost of housing capital, and $L(D)$ is land available for development at distance D from the CBD. The first-order condition for profit maximisation with respect to capital and free entry among builders imply that capital usage is given by $K(D) = (1 - \alpha)R(D)L(D)/(\alpha\, r^K)$. In turn, land available for development at distance D from the CBD is $L(D) = 2\pi\theta\, D$ where $\theta \leq 1$ is the fraction of the land around the CBD

[3] This utility function calls for a few remarks. First, housing enters utility in a Cobb-Douglas manner. Although this formulation is consistent with the findings of Davis and Ortalo-Magné (2011), we return to it below. Second, urban models often introduce commuting as a monetary or a time cost. We model it as a utility cost. This follows a long tradition in transportation economics (Small and Verhoef, 2007). See Krueger, Kahneman, Schkade, Schwarz, and Stone (2009) for recent evidence. Eventually our specification for commuting costs boils down to an issue of the functional form for our estimating equations. We experiment widely with alternative functional forms in our empirical analysis. Third, we assume a constant numéraire price for non-housing goods. According to Combes, Duranton, Gobillon, Puga, and Roux (2013) for France and Handbury and Weinstein (2010) for the US, the elasticity of average retail prices with respect to population is around 0.01. Handbury and Weinstein (2010) also show that after accounting for the amenities provided by stores and for consumers buying more expensive varieties in larger cities, there are no significant systematic differences in retail prices across cities of different population sizes. This is consistent with our specification. Finally, Handbury and Weinstein (2010) also underscore the greater diversity of goods available in larger cities. We leave this aside here but note that taking this diversity effect into account would lower slightly our elasticity for urban costs.

that can be developed. For instance, $\theta = 1$ when the CBD is on a flat plain and $\theta = 0.5$ when it is on a linear coast. Using the last two expressions into equation (5) yields housing supplied at distance D from the CBD:

$$H(D) = 2\,\pi\,\theta\,B\,\left[\frac{1-\alpha}{\alpha}\,\frac{R(D)}{r^K}\right]^{1-\alpha} D\,. \tag{6}$$

Then, using again free entry among builders we get an expression that relates the rental price of housing to the rental price of land:

$$Q(D) = \frac{1}{B}\left(\frac{r^K}{1-\alpha}\right)^{1-\alpha}\left[\frac{R(D)}{\alpha}\right]^{\alpha}. \tag{7}$$

Urban costs

While we impose more structure to our model to be able to interpret our results further, we can already write our main formula for urban costs. To estimate urban costs, we consider an exogenous increase in population and ask what the monetary cost of this increase is for a resident. More precisely, we want to compute the elasticity of the expenditure needed to reach a given level of utility with respect to population.[4]

Let $e(Q(D), U(D))$ denote the expenditure function, i.e., the minimum expenditure to attain utility $U(D)$ in location D given the rental price of housing $Q(D)$. From the consumer problem of a resident located at distance D from the CBD, we can derive the Hicksian demand functions for housing and the composite numeraire and calculate that:

$$e(Q(D), U(D)) = \frac{Q(D)^\beta}{M} U(D) = \frac{Q(0)^\beta}{M} U(0). \tag{8}$$

Using equation (7) valued at $D = 0$, it is easy to obtain ρ the elasticity of expenditure with respect to city population, N:

$$\rho = \alpha\,\beta\,\phi\,, \tag{9}$$

where $\phi \equiv \partial R(0)/\partial N \times N/R(0)$ is the elasticity of the rental price of land at the CBD with respect to population.

Expression (9) indicates that the elasticity of urban costs with respect to city population is the product of three terms: the elasticity of the rental price of land with respect to population, ϕ, times the share of land in housing, α, times the share of housing in expenditure, β. This expression is

[4] We are interested in the elasticity of the expenditure function with respect to population, not in the equivalent variation associated with an increase in population. The reason is that the equivalent variation would monetise the overall utility change from an exogenous increase in population and thus account for both the increase in urban costs and agglomeration economies. Because we want to isolate urban costs, we focus solely on the expenditure function.

intuitive. It states that the population elasticity of urban costs is equal to the population elasticity of the rental price of land at the CBD multiplied by the share of land in expenditure (which is itself the product of the share of land in housing and the share of housing in consumption).

Equilibrium rental price of land

Inserting equation (4) into (7), and substituting again into (7) valued at $D = 0$ yields the distance gradient for the rental price of land:

$$R(D) = \left(\frac{v(0)}{v(D)}\right)^{1/\alpha\beta} R(0)\,. \tag{10}$$

City population is obtained by summing population across all developed locations. In turn, population in a location is an equilibrium outcome which depends on housing supply and housing demand at this location:

$$N \equiv \int_0^{\overline{D}} N(D)\mathrm{d}D \;=\; \int_0^{\overline{D}} \frac{H(D)}{h(D)}\mathrm{d}D\,, \tag{11}$$

where $N(D)$ is population at distance D from CBD.

In expression (11) we can replace $H(D)$ by its expression in equation (6) and $h(D)$ by its expression in equation (2) using equation (7) to substitute for $Q(D)$. In the resulting expression, we can replace $R(D)$ using the distance gradient given by equation (10) before inverting it to obtain the rental price of land at the CBD:

$$R(0) = \alpha\beta\frac{W\,N}{\theta\,T(\overline{D})} \qquad \text{with} \quad T(\overline{D}) \equiv \int_0^{\overline{D}} 2\pi D\left(\frac{v(0)}{v(D)}\right)^{1/\alpha\beta} \mathrm{d}D\,. \tag{12}$$

This equation shows that the rental price of land at the CBD of a city depends positively on wages and population and negatively on the share of land that can be developed and an accessibility term $T(\overline{D})$ which depends on commuting costs and the distance between the CBD and the urban fringe.

An important feature of equation (12) is that the elasticity of rental price of land at the CBD with respect to city population is equal to one. This proportionality between the rental price of land at the CBD and city population arises because population density at any location D, $H(D)/h(D)$, is proportional to the rental price of land in D. In turn, the rental price of land in D is proportional to the rental price of land at the CBD because of the spatial equilibrium within the city. Then, population density can be summed across all locations to obtain total city population. This preserves the proportionality between population and the rental price of land at the CBD. This proportionality is not an exotic property of our framework. It also commonly arises in monocentric models where commuting costs are linear in distance and enter as an expenditure in the consumer programme (Fujita, 1989, Duranton and Puga, 2012).[5]

There is an important reason why the rental price of land at the CBD may increase less than proportionately with city population: We expect cities to decentralise as they grow. This should lessen the effect of population on land prices and push towards a lower population elasticity for the rental price of land. To make this point more explicitly we model decentralisation in a simple fashion by distinguishing between N, the population of the monocentric city, and $\overline{N}$, the population of the entire urban area. We assume that:

$$N = \overline{N}^{\gamma}\,, \tag{13}$$

with $0 < \gamma < 1$. The rest of the metropolitan population $\overline{N} - N$ works outside the CBD and lives outside the fringe $\overline{D}$ of the monocentric city. Equation (13) is obviously a reduced form. Modelling employment decentralisation from first principles is beyond the scope of this paper.[6] Here, we take as a given that metropolitan areas with greater population are more decentralised. We provide empirical evidence to that effect below and show that equation (13) provides a parsimonious but powerful description of employment decentralisation within French metropolitan areas. A deeper exploration of why the distribution of population within French metropolitan areas is potently described by equation (13) is left for further research.

On the other hand, the rental price of land at the CBD may increase more than proportionately with city population because of agglomeration economies. In their presence, larger cities are richer cities with stronger demand for housing and thus higher land prices. More formally, we assume that the wage in a city of unit population is A and increases according to:

$$W = AN^{\sigma}, \tag{14}$$

[5]In our model, the proportionality between $R(0)$ and N occurs for any specification of commuting costs and stems from the iso-elasticity of the demand for housing and constant returns in construction. In the class of model reviewed in Fujita (1989) and Duranton and Puga (2012), these two conditions are generally not satisfied. In these models, linear commuting costs play instead a key role. An increase in the distance to the CBD leads to a proportional increase in commuting costs. This should be offset by a lower expenditure on land for residents to remain indifferent across locations. In turn, the change in expenditure on land is equal to the change in the rental price of land times the quantity of land used per resident. That is, in equilibrium, $R'(D)h(D)$ is equal to the constant marginal increase in commuting costs. Since population density is inversely proportional to land use per resident, integrating over population density to obtain total city population as we do in equation (11) is equivalent to integrating over the distance gradient for the rental price of land $R'(D)$ from the CBD to the urban fringe. This immediately implies a proportionality between the rental price of land at the CBD and city population since $\int R'(D)\mathrm{d}D = R(D) - R(0)$. The caveat is that proportionality is observed for the differential land rent, not total land rent as in the model developed here.

[6]Existing models of the diffusion of jobs outside CBDs and the emergence of secondary sub-centres are notoriously cumbersome to handle (see for instance the older survey by White, 1999).

with $\sigma > 0$. We note that the existence and magnitude of agglomeration economies are well established (Rosenthal and Strange, 2004, Puga, 2010, Combes *et al.*, 2011). Existing estimates for the elasticity of wages with respect to population, σ, in France are between 0.015 and 0.03 (Combes, Duranton, Gobillon, and Roux, 2010).

There are a number of additional reasons why the rental price of land may increase more than proportionately with city population. First, the supply of housing in a city may not be as flexible as described by equation (6). A lower elasticity of housing supply with respect to current rental prices of land increases the population elasticity of the rental price of land.[7] Second, the price elasticity of housing may not be equal to one as indicated by equation (2). If anything the expenditure share of housing is likely to be higher in locations where it is more expensive. This argument suggests a price elasticity of housing below unity. This would, in turn, imply a higher population elasticity of the rental price of housing at the CBD in equation (12). Third, city population could generate a negative congestion externality on commuting, slow down traffic, and thus lead to steeper gradients and higher land prices.

Among these various possible extensions to our basic framework, we only retain agglomeration and decentralisation. First, agglomeration effects and urban decentralisation are both first-order features of contemporary French cities. Second, these two extensions provide a good empirical fit with the data. Third, we could not find robust persuasive evidence about the other issues mentioned above. For instance, as we show below, there is no evidence for steeper land price gradients in larger cities. As we show below, this in the data we use. This said, recall that our chief objective is to obtain an estimate for ϕ, the elasticity of land prices at city centres with respect to city population. The structure we impose here is useful to interpret our results but does not affect urban costs as stated in equation (9).

Inserting equations (13) and (14) into (12) yields:

$$R(0) = \alpha\beta \frac{A\,\overline{N}^{(1+\sigma)\gamma}}{\theta\,T(\overline{D})}\,. \tag{15}$$

This equation expresses the rental price of land at the CBD as a function of city productivity, population, physical expansion and commuting costs.

[7] To see this, consider an extreme case where the existing stock of housing does not depend on distance: $H(D) = H$. In this case, the term in $R(D)^{1-\alpha}$ in equation (6) is replaced by a constant. This leaves only a term in $R(D)^{\alpha}$ at the denominator of the integral in equation (11). This term in $R(D)^{\alpha}$ comes from (7) and the free entry condition among builders. Then, the exponent on N in equation (12) becomes $1/\alpha$. We also note that zoning regulations which make the supply of housing more land intensive distort α but have no other effect on the proportionately between the rental price of land and population.

Estimating equations

To recover the elasticity of the rental price of land with respect to city population, our empirical strategy proceeds in two steps. In the first step, we estimate the logarithm of the rental price of land at the CBD as suggested by equation (10). In a second step, we regress the estimator obtained in the first step on the logarithm of population and other variables as suggested by equation (15).

To derive a specification, we need to make a functional form assumption for commuting costs, $v(D)$. For our main estimations, we assume:

$$v(D) = (a + D)^{\tau} \quad \text{with} \quad \tau < 2\,\alpha\,\beta. \tag{16}$$

In this expression, the constant a can be thought of as a small distance internal to the CBD. Empirically, we only measure location at the municipal level and we need to add an internal distance for the residents of the central municipality.[8] This constant also makes sure that the utility of a resident at the CBD is well defined. Equilibrium stability also requires $\tau \geq 2\sigma/(1+\sigma)$.

After indexing parcels with i and cities with j, assuming that the internal distance to the CBD is small enough ($a << D$), inserting equation (16) into (10), taking logs, and assuming a measurement error on R_i, we obtain the following first-step specification:

$$\ln R_i \approx \ln R_{j(i)}(0) - \frac{\tau}{\alpha\beta} \ln D_i + \epsilon_i \tag{17}$$

where ϵ_i is the i.i.d. measurement error on the rental price of land for parcel i.[9]

After denoting city land area $S_j \equiv \pi\theta_j\overline{D}_j^2$ and assuming again that the internal distance to the CBD is small enough, inserting equation (16) into (15) yields our main second-step specification:

$$\ln R_j(0) \approx C_1 + (1+\sigma)\gamma \ln \overline{N}_j - \left(1 - \frac{\tau}{2\alpha\beta}\right) \ln S_j + \eta_{1j}\,, \tag{18}$$

$$\text{with} \quad C_1 \equiv \ln(\alpha\beta - \tau/2) - \frac{\tau}{2\alpha\beta} \ln \pi \quad \text{and} \quad \eta_{1j} \equiv \ln A_j - \frac{\tau}{2\alpha\beta} \ln \theta_j\,. \tag{19}$$

Equations (17) and (18) are the two steps of our estimation. In the first, land rent for each parcel is regressed on a city effect and distance to the CBD. This city effect measures the rental price of land at the CBD and is used as our second-step dependent variable to estimate the elasticity of the

[8] Alternatively, we expect the location with highest rental price of land to provide the best accessibility, not perfect accessibility as not all jobs are located at the CBD in reality.

[9] We can see two sources of measurement error. First, we expect some heterogeneity in prices coming from the fact that some buyers or sellers may be in a hurry to buy or sell, have idiosyncratic preferences, different abilities to bargain, etc. It could also be that the area of a parcel does not reflect the area that is actually constructible. This could occur, for instance, because of a steep slope, or because part of the parcel is not residentially zoned, etc. In this case, parcel area is mis-measured and this affects our computation of the price per unit of land.

rental price of land with respect to city population. The other important explanatory variable used in the second step is the land area of the city. As shown by the first part of equation (19), the constant of the second-step regression is a constellation of the key parameters of the model, α, β, and τ. The error term of the second-step regression is made explicit in the second part of equation (19). It contains both city productivity A and the share of the city that can be developed θ. Because these two terms are unobserved (or imperfectly observed) in the data, we treat them as possible missing variables in the regression.

Before dwelling deeper into this issue, we note that we can let the urban fringe be determined endogenously. Aside from housing, land can also be used for agriculture. Agricultural land rent around city j is $\underline{R}_j$. The urban fringe is thus such that $R(\overline{D}_j) = \underline{R}_j$. Using equation (10), we obtain $\overline{D}_j = (R_j(0)/\underline{R}_j)^{\alpha\beta/\tau} - 1$. Substituting into (18) and simplifying using $R(0) >> \underline{R}$ implies:

$$\ln R_j(0) \approx C_2 + (1+\sigma)\gamma\frac{\tau}{2\alpha\beta}\ln \overline{N}_j + \eta_{2j}\,, \tag{20}$$

$$\text{with} \quad C2 \equiv \frac{\tau}{2\alpha\beta}(\ln(\alpha\beta - \tau/2) - \ln\pi) \quad \text{and} \quad \eta_{2j} \equiv \frac{\tau}{2\alpha\beta}(\ln A_j - \ln\theta_j) + (1 - \frac{\tau}{2\alpha\beta})\ln \underline{R}_j\,. \tag{21}$$

The key difference between equations (18) and (20) is that in equation (18) land area is held constant whereas in equation (20) it is treated as an endogenous variable that can be substituted away. Hence, the exponent $(1+\sigma)\gamma$ on population, $\overline{N}$, in equation (18) measures the elasticity of the rental price of land in the CBD with respect to an increase in population keeping the urban fringe constant whereas the corresponding coefficient $(1+\sigma)\gamma\frac{\tau}{2\alpha\beta}$ in equation (20) corresponds to the same elasticity when the urban fringe is allowed to adjust. The former elasticity should be larger than the latter (and we can verify empirically that this is the case). Given the nature of our data (fixed boundaries for urban areas) and existing restrictions on housing development in France, our empirical analysis below puts greater emphasis on the estimation of regression (18). We nonetheless also estimate regression (20) and note that when we allow for an endogenous urban fringe, the second-step error term in equation (21) is more complicated than in equation (19) since it also contains unobserved agricultural land rent.

Before turning to our estimation strategy, one final interpretation issue must be discussed. Our model above assumes that amenities, M, remain constant as population grows. We can easily extend our model to allow for amenities to change endogenously as cities grow in population by assuming $M(\overline{N}) = \underline{M}\,\overline{N}^{\nu}$ where $\underline{M}$ is the natural level of amenities and ν is the elasticity of (endogenous) amenities with respect to population.

At the spatial equilibrium across cities, we have $U(D) = \underline{U}$. Using the fact that expenditure equals wage, $e(Q(D),U(D)) = W$, we can rewrite equation (8) as:

$$\underline{M}\,\overline{N}^{\nu} = \frac{Q(0)^{\beta}}{W}\underline{U}. \tag{22}$$

Making use of equations (7) and (14) we find: $\nu = \alpha\,\beta\,\phi - \sigma$. That is, the difference between our estimate of the elasticity urban costs and that of wages is also, at the spatial equilibrium between cities, an estimate of the elasticity of endogenous city amenities with respect to city population.

Estimation issues

While our model proposes a straightforward and intuitive formula for urban costs and suggests a simple two-step approach for estimation, it also raises two key identification issues.

The first is about functional forms. There is empirical support for the assumptions we make regarding housing demand (CGDD, 2011, Davis and Ortalo-Magné, 2011) and the production function for housing (Combes *et al.*, 2012). Unfortunately far less is known about commuting. To solve our model, we make specific assumptions about how commuting enters preferences, impose functional forms, and assume a particular geography. These choices lead to a specification where log land rent is regressed on log distance to the CBD. Although log-log specifications may be a popular choice when estimating urban gradients, log-level specifications are also popular.[10]

To respond to this specification challenge, we first verify that in the data, French cities are monocentric with respect to their land gradients and that our log-log specification is reasonable. In robustness checks, we estimate our first step with alternative functional forms for distance, with alternative definitions for the city centre, and with alternative non-monocentric geographies.

The simultaneous determination of the rental price of land at the CBD, city population, and land area is our second empirical challenge. To see this simultaneity problem more clearly, we now allow for city population to be endogenous by assuming free mobility and a reservation level of utility $\underline{U}$ outside the city.[11]

[10]For instance, $v(D) = e^{\tau\,D}$ is a natural alternative to the power specification for commuting costs used in equation (16). This alternative specification implies that the log of the rental price of land is proportional to the linear distance to the CBD instead of being proportional to its log. Then inserting $v(D) = e^{\tau\,D}$ into equation (12) leads to an integral of the form $D\,e^{D}$ for which the solution takes the form $(D-1)e^{D}$. In turn, this implies using both log area and its square root in the second step estimation corresponding to equation (15).

[11]Because 23% of the French population live in rural areas, we assume that the reservation utility is determined there. Assuming full urbanisation would imply further interdependencies across cities in equilibrium. To preview the discussion of the identification issues below, assuming full urbanisation would broaden the scope of possible instruments for city population and suggest using contemporaneous variables from other cities. We do not pursue this further because such instruments are empirically weak and generally unpersuasive.

This implies $\underline{U} = U(D) = U(0)$. Using equation (3) valued at $D = 0$, and equations (7), (14), (16), and (20) we obtain:

$$\ln \overline{N}_j \approx C_3 + \beta\, c \ln B_j + c \ln M_j - (\alpha\, \beta - \tau/2)\, c\, \ln \underline{R}_j + \eta_{3j}\,, \tag{23}$$

$$\text{with} \quad C3 \equiv f(\alpha, \beta, \tau, r^K, \underline{U}),\; c \equiv \frac{2}{\gamma\,[\tau(1+\sigma) - 2\sigma]},\, \eta_{3j} \equiv (1-\tau/2) c \ln A_j + \tau/2\, c\, \ln \theta_j\,, \tag{24}$$

where $f(.)$ is a function of its arguments. Equation (23) shows formally that city population and the rental price of land at the CBD as specified in equation (18) are simultaneously determined. Productivity and the share of the urban area that can be developed enter the error term of both equations. That is, two missing variables determine both the rental price of land and population.

By the same token we can also derive a corresponding expression for S_j using equations (18) and (21):

$$\ln S_j \approx \ln(\alpha\, \beta - \tau/2) + (1+\sigma)\gamma \ln \overline{N}_j + \ln A_j - \ln \underline{R}_j\,. \tag{25}$$

Aside from confirming that the simultaneity problem between the rental price of land at the CBD and city population also extends to city land area, this equation is interesting because it shows that we can obtain an independent estimate of $(1+\sigma)\gamma$ by regressing land area on population.

To deal with the simultaneity concern highlighted by equations (23) and (25), we use an instrumental variables strategy. Urban population and land area both enter equation (18) and their equilibrium values are otherwise determined by equations (23) and (25). This suggests that any variable that explains population and/or land area but is not otherwise present in equation (18) constitutes a valid instrument. We refer to such instruments as instruments 'internal' to the model.

From equations (23) and (25) we can see that amenities (M), housing technology (B), and the rental price of land at the urban fringe ($\underline{R}$) determine both population and land area but do not intervene directly into the determination of land prices in equation (18). To be valid, any instrument we use must be uncorrelated with the error term in equation (18) which contains unobserved city productivity (A) and share of land that can be developed (θ). In addition, instruments must be sufficiently strong predictors of city population and land area.

Because it is hard to think of variables that measure the housing production technology but are unrelated to the production technology for other goods and because measures of the rental price of land at the fringe of the city are likely to be correlated with the share of developable land, we restrict attention to amenities in our search for internal instruments. More specifically, we first rely on a measure of weather (January temperatures) and two measures of tourism (number of hotel rooms and share of one star rooms). There is strong evidence about the importance of weather to explain location choices in the US (Rappaport, 2007) and Europe (Cheshire and Magrini, 2006). Our use of tourism variable follows Carlino and Saiz (2008) who argue that the number of tourism visits is a good proxy for the overall level of consumption amenities of US cities and provide evidence to that effect. Hence, we expect urban areas with more hotel rooms, in particular more hotel rooms in upper categories, to be more attractive all else equal.

We show below that these instruments are often strong enough predictors of population and land area. Because they may be correlated with unobserved productivity and the unobserved share of developable land, these instruments may nonetheless fail to be uncorrelated with the error term in equation (18). While we acknowledge that weather may affect productivity directly, this effect is likely to be of limited importance since January weather in France is rarely bad enough to disrupt economic activity. Better January weather is also likely to be correlated with coastal locations (with the Mediterranean coast being warmer in the South and the Atlantic coast benefitting from the mitigating effects of the Gulf Stream). In turn, coastal locations may affect productivity. To preclude this correlation, we also control for coastal locations. More generally we use a broad range of control variables to reduce as much as possible the unobserved component of productivity and land development and to preclude possible correlations between the instruments and the error term. These supplementary controls include a broad range of geographical, socio-economic, land development, and geological characteristics of cities. Obviously these controls will also proxy directly for the unobserved productivity and share of developed land.

To confirm the findings obtained with our internal instruments, we also experiment extensively with 'external instruments', that is predictors of city population and land area which are not explicitly considered by our model. Following previous work (Ciccone and Hall, 1996, Combes *et al.*, 2010) we use long historical lags of population. We also construct predictors of city population and land area based on the sectoral and occupational composition of economic activity as suggested by urban theory (Henderson, 1974, Duranton and Puga, 2005). We verify that despite a different rationale, different reasons why they may fail to satisfy their exclusion restriction, and low correlations with the internal instruments, our external instruments essentially yield the same answer as the internal instruments. Further details of our instrumentation strategy and remaining estimation issues are discussed below.

3. Data

We use land price data. Using a record of transactions for land parcels with information about construction costs offer considerable advantage over more widely available records of property transactions for which disentangling between the value of land and the value of the property built on it is a considerable challenge.

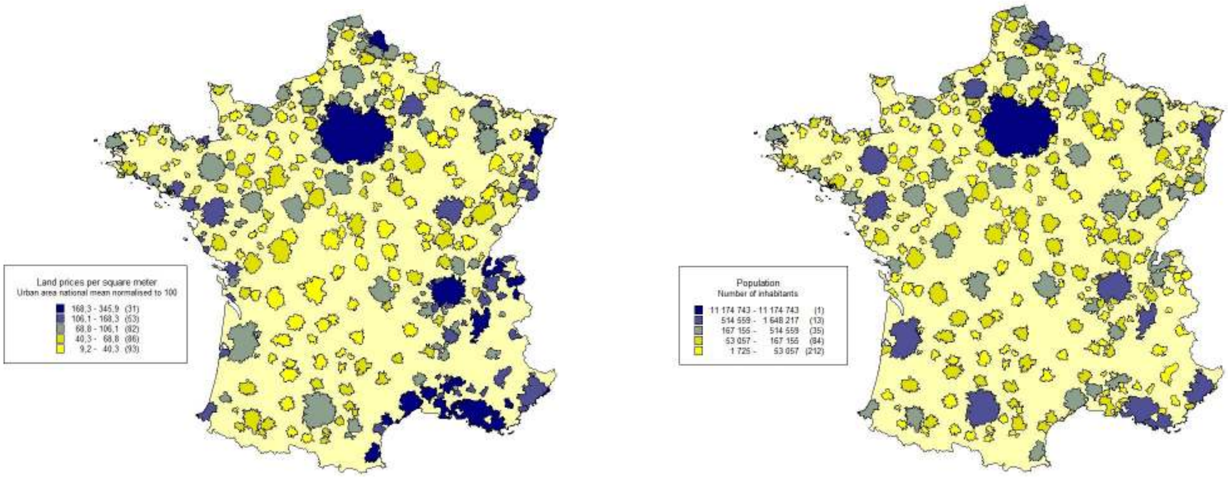
The data are extracted from the 2008 Survey of Developable Land Prices (*Enquête sur le Prix des Terrains à Bâtir*, EPTB) in France. For 2008, there are initially 82,586 observations. They correspond to 61% of all building permits for detached houses delivered this year. Appendix A provides further details about the origin of these data. For each observation we know the price, the municipality, a number of parcel characteristics, and the cost of construction incurred by the buyer. The municipality identifier allows us to determine in which urban area the parcel is located. Because French municipalities are tiny (on average, they correspond to a circle of radius 2.2 kilometres), we can approximate the location of parcels well by using the centroid of their municipality.

To measure distance to the centre of an urban area, our preferred metric is the log of the Euclidian distance between the centroid of a parcel's municipality and the barycentre of the urban area to which it belongs. To determine this barycentre, we weight all municipalities in an urban area by their employment.

The information about each transaction includes how the parcel was acquired (purchase, donation, inheritance, other), whether the parcel was acquired through an intermediary (a broker, a builder, another type of intermediary, or none), and some information about the house built including its cost. We also know the area of a parcel, its road frontage, whether it is 'serviced' (i.e., has access to water, sewerage, and electricity), whether there was already a building on the land when acquired, and whether there is a demolition permit for this building.

For 74,204 of the initial 82,586 observations, the transaction was a purchase. We ignore other transactions such as inheritances for which the price is unlikely to be informative. We also restrict our attention to transactions that were completed in 2008 since land prices are more difficult to interpret with transactions that took place well before a building permit was granted. That leaves us with 52,113 observations. We further limit our analysis to urban areas in mainland France to end up with 27,850 observations.

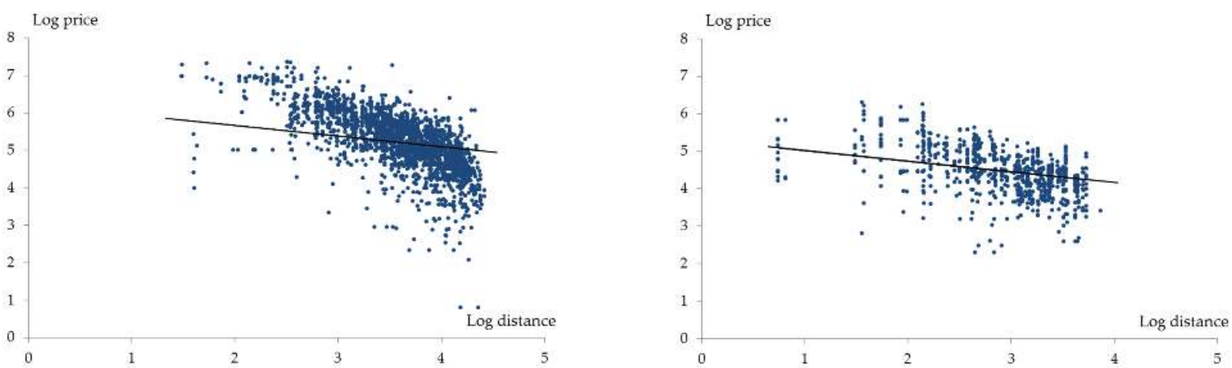
Figure 1: Mean land prices per square metre and population in French urban areas



Panel (a): Mean land prices

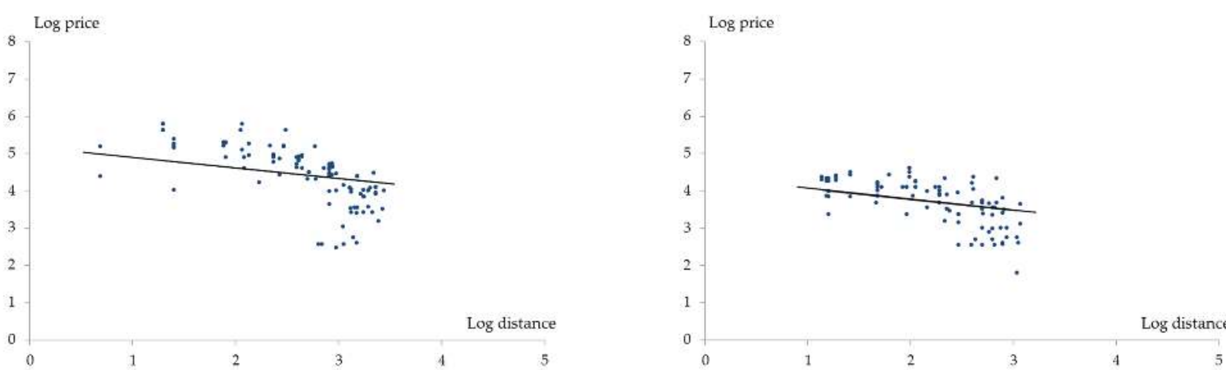
Panel (b): Population

Figure 2: Land prices per m^2 and distance to their barycentre for four urban areas



Panel (a): Paris

Panel (b): Bordeaux



Panel (c): Dijon

Panel (d): Chalon-sur-Saône

Table 1: Descriptive Statistics (parcel level)

Variable	Mean	St. Error	1st decile	Median	9th decile
Price (€ per m^2)	101.8	121.9	21.6	78.7	202.5
Price of construction (€ per m^2 of land)	169.8	137.4	62.6	150.0	293.2
Parcel area (m^2)	1,103	1,200	477	838	2,000
Road frontage / square-root parcel area	0.68	0.50	0.23	0.68	1.07
Serviced parcel	0.64	0.62	0	1	1
Prior building	0.05	0.29	0	0	0
Prior building to be demolished	0.02	0.20	0	0	0
Bought through an agency	0.24	0.55	0	0	1
Bought through a builder	0.17	0.49	0	0	1
Bought through another intermediary	0.17	0.49	0	0	1
Population (urban area, '000, 2007)	1,049	3,617	28	186	1,164
Land area (urban area, km^2)	2,042	4,360	198	859	3,875
Population growth (urban area, %, 1999-2007)	5.9	6.1	0.0	5.9	11.6
Distance to the barycentre (km)	13.0	14.7	2.8	10.1	25.9
Distance to the closest centre (km)	11.9	14.4	2.5	8.6	24.6

Notes: 27,850 observations corresponding to 46,697 weighted observations except row (4) (25,854 and 43,146).

Table 2: Descriptive Statistics (parcel means by population classes of urban areas)

City class	Paris	Lyon, Lille, Marseille	>200,000	≤200,000
Price (€ per m^2)	238.8	187.7	112.0	70.5
Price of construction (€ per m^2 of land)	251.5	189.3	177.5	152.6
Parcel area (m^2)	845	1203	1051	1163
Road frontage / square-root parcel area	0.66	0.55	0.63	0.63
Serviced parcel	0.56	0.52	0.66	0.66
Prior building	0.13	0.07	0.06	0.04
Prior building to be demolished	0.09	0.03	0.02	0.02
Bought through an agency	0.41	0.34	0.23	0.22
Bought through a builder	0.19	0.12	0.17	0.17
Bought through another intermediary	0.13	0.14	0.19	0.16
Population (urban area, '000, 2007)	11,837	1,645	540	81
Land area (urban area, km^2)	14,518	2,880	1,990	511
Population growth (urban area, %, 1999-2007)	5.9	5.6	7.0	5.1
Distance to the barycentre (km)	41.3	19.4	15.3	7.3
Distance to the closest centre (km)	40.2	18.1	13.9	6.5
Weighted number of land parcels	2,878	2,022	17,862	23,935
Weighted number with non-missing frontage	2,756	1,829	16,468	22,086
Number of urban areas	1	3	39	302

Notes: 27,850 parcel transactions corresponding to 46,697 weighted observations except row (4) (25,854 and 43,146). The numbers in column 3 are for all French urban areas with population above 200,000 excluding Paris, Lyon, Lille, and Marseille.

The two panels of figure 1 map mean unit land prices and population by French urban areas. To illustrate the reality of the data within particular urban areas, the four panels of figure 2 plot land prices per square metre and the distance to the employment weighted barycentre of the urban area for transactions occurring in Paris (the largest urban area), Bordeaux (a large urban area), Dijon (a mid-size urban area), and Chalon-sur-Saône (a small urban area). The gradient for each plot is the same and corresponds to the gradient for the country estimated below. Tables 1 and 2 provides further descriptive statistics. These figures and tables all underscore that land prices per square metre exhibit a lot of variation which appears to correlate with city population. In our data, mean land prices in Paris are more than three times those in small French urban areas.

The plots of figure 2 are also helpful to minimise the worry that parcels sold with a building permit form a highly selected sample of existing parcels in a urban area. In all cities we observe transactions close to the centre, in close suburbs, and remote suburbs. This is because French land use regulations encourage in-filling and try to limit expansions of the urban fringe.[12] The transactions we observe also cover a broad spectrum of parcels characteristics. This is because we use a systematic and compulsory survey based on administrative records. Unlike land transactions recorded by private real estate firms, ours are not biased towards large parcels.

Finally, we also use a wide variety of city level characteristics described in Appendix A.

4. Land values in French urban areas

Following equation (17), we first estimate

$$\ln P_i = \ln P_{j(i)}(0) - \delta \, \ln D_{k(i)} + T_i \, b + \epsilon_i \,, \tag{26}$$

where P_i is unit land *price* for parcel i, $k(i)$ is the municipality where i is located, $j(i)$ is the urban area where the municipality $k(i)$ is located. The fixed effects $P_j(0)$ capture unit land prices at the centre of j. The coefficient δ is the distance gradient of land prices when varying (log) distance D between the municipality where a parcel is located and the centre of the urban area it belongs to. According to our model, $\delta = \tau/(\alpha\,\beta)$. Finally, T_i is a vector of parcel characteristics.

There are two minor differences between specification (17) of our model and the regression (26) that we implement. The first is that because of data constraints we must use land price instead of

[12] French municipalities need to produce a planning and development plan (*plan local d'urbanisme*) which is subject to national guidelines and requires approval from the central government. Existing guidelines for municipalities or groups of municipalities insist on the densification or re-development of already developed areas to save on the provision of new infrastructure (usually paid for by higher levels of government) relative to expansions of the urban fringe.

land rent data. This affects the interpretation of the city fixed effect but plays otherwise no role in this first step of the estimation. The second difference is that we condition out a number of parcel characteristics. Parcel characteristics such as their shape or size are expected to matter. Our model ignores this type of heterogeneity. It would be straightforward but cumbersome to consider it in our model.

Table 3 reports summary results regarding the estimation of regression (26). Column (1) uses only ten parcel characteristics to explain their price per square metre. These characteristics are the seven listed in rows (4)-(10) of table 1, a dummy for missing information for road frontage, log parcel area, and its square. Most of these characteristics have a significant impact with the expected sign. For instance, the price of a (relatively wider) parcel at the last decile in terms of road frontage is 24% higher than the price of a (narrow) parcel at the first decile. A serviced parcel is 54% more expensive than a parcel with no access to basic utilities. A parcel is 6% more expensive when a building is present. This low value may be explained by the fact that an existing construction is sometimes a negative characteristic when a new building is planed. Indeed, when a building has been demolished or when the authorisation to be demolished is given, the price per square metre is 64% higher. Parcels sold by real estate agencies, builders, or other intermediaries are also significantly more expensive since real estate professionals are likely to specialise in the sale of more expensive parcels. Finally, smaller parcels fetch a higher price per square metre. Altogether, parcel characteristics explain 48% of the variance. Among parcel characteristics, log parcel area and its square matter most. Without them the R^2 falls to 10%.

Column (2) corresponds to a specification with 345 urban area fixed-effects. A large share of these effects are significantly different from the mean urban area effect, with 48% being above and 39% below. This variation is also underscored by the difference between the first decile of the urban area fixed effect, at 3.3, and the last decile, at 5.1, corresponding to prices per square metre about six times as high. With an R^2 of 52%, urban area fixed effects explain more than half of the variance in land prices, which is a first sign of the large role of urban areas in explaining land prices.

Column (3) adds log distance to the urban area barycentre as an explanatory variable to measure the distance gradient for land prices. The coefficient of -0.28 on this variable is highly significant. This elasticity implies that a parcel at the first decile of distance (2.8 kilometres from the centre) is 85% more expensive than a parcel at the last decile (25.9 kilometres from the centre). The R^2

Table 3: Summary statistics from the first step estimation regressions, 345 urban areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parcel log-area	-0.61^a				0.93^a	1.22^a	0.99^a	1.23^a
Parcel squared log-area	-0.03^a				-0.12^a	-0.14^a	-0.12^a	-0.14^a
Other parcel charac. signif. (max=8)	6			8	7	6	6	6
Mean urban area fixed effect		4.3	4.9	4.4	3.8	3.0	3.7	2.6
First decile		3.3	3.9	3.4	3.0	1.9	2.9	1.7
Last decile		5.1	5.9	5.4	4.7	4.2	4.6	3.5
% above mean (signif.)		47.6	48.1	46.8	22.0	24.1	22.1	22.1
% below mean (signif.)		39.4	41.7	41.3	20.5	33.4	22.8	25.3
Distance effect			-0.28^a	-0.26^a	-0.18^a		-0.26^a	-0.03^a
Mean distance effect by urban area						-0.26		
First decile						-0.48		
Last decile						-0.01		
% above mean (signif.)						32.1		
% below mean (signif.)						31.4		
R^2	0.48	0.52	0.57	0.62	0.79	0.81	0.80	0.81
Observations	26,177	26,177	26,177	26,177	26,177	26,177	26,177	26,177

Notes: The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. All columns are estimated with OLS. Columns (1) to (6) use log distance to the urban area barycentre. Columns (7) and (8) are identical to (5) but use log distance to the closest centre (among the two largest municipalities) and linear distance to the urban area barycentre, respectively. The eight other parcel characteristics are those listed in rows 4-10 of table 1 and a dummy for missing information for road frontage (which occurs for 7.2% of the observations).

increases to 57%, which confirms that location within an urban area also matters. In column (4), we enrich the specification of column (3) and re-introduce 8 parcel characteristics. They make little changes to the results.

Column (5) adds log parcel area and its square. This slightly lowers the land price gradient and the dispersion between city fixed effects. This regression is our preferred first-step specification and we use its output for our main estimations below. We think it is better to use a more demanding specification which includes log parcel area and its square. In addition, parcel area is likely to reflect local zoning and development patterns within a city (given the presence of city fixed effects). In robustness checks, we also use the fixed effects obtained from the other columns of this table. A side result of column (5) is that the overall price of land parcels per square metre is bell-shaped with respect to their land area. However, the area that maximises land price per square metre is small which suggests that prices per square metre generally decline with parcel area.

In column (6) we retain the same explanatory variables but allow for the coefficient on distance to vary by city. There is variation across cities in their land price gradient. However, allowing for a separate land gradient for each city has only a marginal effect on the R^2. We also show below that these land gradients are poorly explained by the explanatory variables we use at the second step of the regression. In column (7) we replicate the specification of column (5) but consider two centres for each urban area (corresponding to the two most populated municipalities) and use distance to the closest of two as explanatory variable. In column (8), we also duplicate our preferred specification of column (5) but use linear instead of logarithmic distance.

5. The elasticity of unit land prices with respect to population

We now report our main results for the second-step estimation where we use the urban area fixed effects estimated with our preferred first-step estimation as dependent variable. These fixed effects are the empirical counterparts to the rental price of land at the CBD in the theoretical model. We regress them on a number of urban area characteristics. Population is our explanatory variable of interest. We start with OLS results before turning to our main IV results. The end of this section provides a number of further robustness checks.

Following equation (18) we estimate:

$$\ln \hat{P_j}(0) = X_j\, \varphi + \xi_j\,, \tag{27}$$

where the dependent variable, land prices at the centre, is estimated in the first step, and X_i is a vector of city characteristics which includes population, land area, and other controls.

OLS *results*

Table 4 present results for second-step OLS estimations regressing urban area fixed effects on urban area characteristics. Column (1) uses only log population as explanatory variable for 278 French urban areas.[13] The estimation in column (1) can be thought of as a rudimentary estimation of

[13]There are 352 urban areas in France. For 7 of them, we have no observation so we can estimate a fixed effect for only 345 of them as reported in table 3. Then for 54 of them (mostly small urban areas), we miss some of our instruments. We also discard small urban areas with only one or two municipalities for which the fixed effect is not precisely estimated. There are 13 of them. We end up with 278 urban areas. For the sake of comparison with our IV results, we mostly report OLS results for 278 urban areas. We verify below that our results are not sensitive to our choice of sample.

Table 4: The determinants of unit land values at the centre, OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.310^a	0.650^a	0.613^a	0.601^a	0.616^a	0.629^a	0.583^a	0.579^a
	(0.024)	(0.030)	(0.027)	(0.029)	(0.030)	(0.028)	(0.031)	(0.034)
Land area		-0.466^a	-0.440^a	-0.427^a	-0.449^a	-0.467^a	-0.399^a	-0.415^a
		(0.032)	(0.029)	(0.031)	(0.030)	(0.032)	(0.029)	(0.035)
Population growth			3.254^a	3.079^a	3.135^a	3.350^a	3.744^a	3.322^a
			(0.394)	(0.415)	(0.443)	(0.394)	(0.395)	(0.447)
Controls	no	no	no	geog.	econ.	urb.	geol.	all
R^2	0.38	0.65	0.72	0.72	0.72	0.72	0.75	0.76

Notes: 278 observations in each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions include a constant.
Geog.: dummy variable for being on the sea and a peripherality index (employment-weighted mean distance to all other urban areas). Econ.: Income per capita and the share of workers with a university degree. Urb.: Share of urbanised and agricultural land. Geol.: ruggedness, three dummies for the erodibility of soils (low, intermediate, high), two dummies for hydrogeological classes, and two dummies for classes of soil (unconsolidated deposits and aeolian deposits). All: all previous controls.

equation (20). We find that population explains nearly 40% of the variance among first-step urban area fixed effects and the estimated elasticity, at 0.310, is highly significant. This regression confirms a strong association between land prices at the centre of urban areas and population size.

Columns (2) adds log land area and corresponds to the simplest estimation of equation (18). Adding land area makes the association between land prices and population even stronger with a coefficient of 0.650. On the other hand, land prices are negatively associated with land area. These signs are consistent with the predictions of our model and intuition. Population and land area account for 65% of the variance of urban area fixed effects.

Column (3) enriches the specification of column (2) to control for population growth over 1999-2007. Recall that, as dependent variables, we use urban area fixed effects estimated from data on land prices, not from data describing the rental price of land. Theoretically, land prices are best viewed as the net discounted value of all future land rents. Future growth in the rental price of land differs across urban areas and is unknown. However, we know contemporaneous population growth and we can use the fact that population growth is serially correlated. This suggests using contemporaneous population growth (1999 to 2007) as a control in our second step regression.

The coefficient on population growth is positive and highly significant. This confirms that higher contemporaneous growth, which is likely correlated with higher future growth, is strongly associated with higher future land rents and therefore higher current land prices. A higher population growth by one percentage point during the period is associated with land prices which are more than 3% higher. Put differently, a one standard deviation increase in population growth rate (i.e., 0.047) is associated with a 15% increase in land prices. We also find that adding population growth raises the share of variance explained even further to 72% but leaves the coefficients on population and land area roughly unchanged. This is unsurprising given the weak link between city population and city population growth.

Columns (4) to (8) enrich the specification of column (3) by adding supplementary control variables that aim to proxy for unobserved productivity and the unobserved share of land used around the CBD. Column (4) considers two geographical characteristics: a dummy for being on the sea and a peripherality index that captures how far an urban area is relative to the others. Column (5) considers two important socio-economic characteristics: income per capita and the share of workers with a university degree. Column (6) considers two important land development characteristics: the share of land that is urbanised and the share of land that is used for agricultural purposes (and could perhaps be urbanised). Column (7) considers a number of geological characteristics of soils. Finally column (8) includes all these supplementary controls together.

Interestingly, geographic, socio-economic, and development variables have virtually no effect on the coefficient on population and only minimal effects on the coefficients on land area and population growth. They also leave the R^2 of the regression virtually unchanged. The addition of geological characteristics has a larger but still modest downward effect on the coefficients on population. Geological characteristics also lower the magnitude of the coefficient on land area. While for the coefficient on population the difference between column (3) and column (8) is statistically significant, it remains economically small (about 10%). Overall, these results offer a first indication that the coefficients on population and land area are not sensitive to the inclusion of a wide range of proxies for unobserved productivity and the share of land developed around the CBD. In turn, this suggests that the worries caused by the endogeneity of population and land area may be unwarranted, perhaps because the French population is geographically poorly mobile. To investigate this issue more in depth, we now turn to our IV estimations.[14]

[14]These results are consistent with extant results from the agglomeration literature which only finds a minor bias caused by the endogeneity of population. See for instance Combes *et al.* (2010).

Table 5: Unit land values at the centre, population and area instrumented

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.721^a	0.673^a	0.614^a	0.680^a	0.683^a	0.702^a	0.673^a	0.685^a
	(0.076)	(0.064)	(0.089)	(0.063)	(0.064)	(0.063)	(0.062)	(0.066)
Land area	-0.563^a	-0.504^a	-0.425^a	-0.499^a	-0.502^a	-0.538^a	-0.501^a	-0.504^a
	(0.096)	(0.083)	(0.183)	(0.083)	(0.083)	(0.083)	(0.082)	(0.087)
Population growth	3.017^a	3.117^a	3.238^a	3.090^a	3.083^a	3.055^a	3.115^a	3.080^a
	(0.430)	(0.412)	(0.428)	(0.412)	(0.413)	(0.416)	(0.510)	(0.414)
Overidentification p-value	0.68	0.26	0.36	0.76	0.99	0.84	0.27	0.95
First-stage statistic	9.8	13.3	6.4	13.2	13.0	13.4	10.2	11.8
Endogeneity p-value	0.13	0.24	0.41	0.02	0.01	0.05	0.10	0.01
Number of hotel rooms	Y	Y	Y	N	N	Y	Y	N
Share of 1-star rooms	Y	Y	N	N	N	N	N	N
Temperature in January	Y	N	Y	N	N	Y	N	Y
Bartik occupations 1999	N	Y	N	N	N	N	Y	N
Bartik industry 1999	N	N	Y	N	N	N	N	N
Urban population 1831	N	N	N	Y	Y	N	N	Y
Urban density 1881	N	N	N	Y	N	N	N	N
Henderson industry 1999	N	N	N	Y	Y	Y	Y	N
Henderson occupations 1999	N	N	N	N	Y	N	N	Y
1st Shea part. R^2, population	0.13	0.18	0.10	0.18	0.18	0.19	0.20	0.17
1st part. Fisher, population	264.9	291.7	279.5	181.2	156.7	293.2	295.5	148.9
1st Shea part. R^2, area	0.10	0.13	0.07	0.13	0.13	0.13	0.13	0.11
1st part. Fisher, area	108.7	103.4	100.4	118.0	98.5	100.5	101.3	100.5

Notes: 278 observations in each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions are estimated with LIML and include a constant but no supplementary control.

IV *results*

Table 5 reports results for a series of IV regressions using a number of instruments for population and land area. Because the first-stage statistics are sometimes below the critical thresholds of Stock and Yogo (2005) for weak instruments, we estimate all our IV regressions with LIML instead of TSLS.[15] A related worry is that our instruments may be particularly weak for one of our two endogenous variables, population and land area. We report measures of partial R^2 at the bottom of the table to show that this is not the case.

In column (1), we instrument population and land area with temperatures in January, the number of hotel rooms, and the share of hotel rooms in lower categories. As argued above, January

[15]Limited information maximum likelihood (LIML) is a one-stage IV estimator. Compared to two-stage least squares (TSLS), it provides more reliable point estimates and test statistics with weak instruments.

temperature is our preferred measure of climate. We also experimented extensively with a variety of other climate variables. They lead to similar results but tend to be weaker as instruments. We also expect cities that are more attractive in terms of amenities to attract more tourists and hence have more hotel rooms, particularly in the highest categories. Because this is the specification that sticks most closely to our model, we take column (1) as our preferred specification.

Relative to their corresponding OLS values in column (3) of table 4, the IV coefficients on population and land area in column (1) are larger in magnitude. However, for population the difference is less than a standard deviation whereas for land area it is only slightly more than a standard deviation. These differences are not statistically significant. They are also economically small. For the coefficient on population the difference between IV and OLS is only about 10%. We also easily fail to reject our test of overidentifying restrictions.

Although using amenities to instrument for population and land area is natural and is backed by our model, our amenities variables might be correlated with land prices through a number of other channels. For instance, good amenities might adversely affect productivity by favouring leisure. Such correlation with unobserved productivity would violate our exclusion restriction. To deal with this issue we consider a range of other possible instruments for city population and land area. These instruments are not explicitly considered by our model and we refer to them as instruments external to our model.

Following Ciccone and Hall (1996) and much of the recent agglomeration literature we use long historical lags of population variables: urban population in 1831 and urban density in 1881. These measures constitute strong predictors of contemporaneous population patterns in France (Combes *et al.*, 2010). To be valid as instruments, historical populations should only affect contemporaneous population through the persistence of where people live. In particular, the natural advantages of cities which may have attracted people in 1831 should no longer be present. Otherwise they would be part of contemporaneous productivity of urban areas and violate the exclusion restriction. Our case for these instruments thus rests on the extent of the changes in the French economy since 1831 and 1881.

Our second group of external instruments for population size and land area is novel. Following Henderson (1974), a large literature on urban systems has developed. A key idea from that literature is that cities specialise in their activities. It is also the case that different activities benefit differently from agglomeration effects. In turn, this implies that, depending on what they produce, cities reach different equilibrium sizes. For instance, textile cities tend to be small whereas banking cities tend to be large. More specifically, to construct this instrument, we compute mean urban area employment for a typical worker if each three-digit sector. Then for each urban area, we compute our instrument by multiplying the local share of employment in a sector by mean urban area employment for this sector before taking the sum across sectors for each city (see Appendix A for details).

In a twist to this idea, Duranton and Puga (2005) argue the main dimension along which cities now specialise is no longer by sector but by occupation following the greater possibilities offered to firms to separate their operations geographically. For instance, an urban area with a high proportion of accountants or lawyers will be predicted to be large while an urban area with a high proportion of blue-collar occupations will be predicted to be small. Following the same approach as with sectors, we compute mean urban area employment for a typical worker in each four-digit entry of the French standard occupational classification. Then, for each urban area, we compute our instrument by multiplying the local share of employment of an occupation by mean urban employment for this occupation before taking the sum across occupations for each city.

Finally, our third group of instruments relies on ideas first developed in Bartik (1991). We expect cities initially specialised in sectors or occupations that grow more nationally to end up larger at the end of the period. For instance, cities specialised in information technologies are expected to have grown whereas cities specialised in the production of steel are expected to have declined in population. More specifically, we construct this instrument by interacting the initial composition of economic activity of an urban area by sector with the growth of those sectors at the national level during the period. We also duplicate this using the occupational structure of cities and the growth of occupational groups. The resulting variables are predictors of growth during the period which are also correlated with population size at the end of the period since cities that grow faster end up larger.

Just like with our internal instruments, it is possible to imagine reasons why the exclusion restriction associated with each of these external instruments might fail. By using these instruments in different combinations we can nonetheless compare their answers and perform meaningful overidentification tests. A p-value above 10% indicates that the instruments at hand yield answers that are not statistically different. They are either jointly valid or all invalid in the same way. For this, they would need to have the same correlation with the error term of our regression.

This seems implausible. First, our instruments rely on different sources of variation in the data: weather, tourism activity, long population lags, occupational structure, and sectoral structure. Second, the correlations between them are generally low. For instance the correlation between January temperature and the other instruments we use is always below 0.20. The share of one star hotels has a correlation always below 0.30 with the other instruments except the 1999 Henderson instrument, etc.

The results for different combinations of internal and external instruments are reported in columns (2) to (8) of table 5. The coefficients on population, land area, and population growth are stable. The coefficient on population varies between 0.614 and 0.702 and is generally within one standard deviation of our preferred estimate of column (1). The coefficient on land area varies between -0.425 and -0.538 and also remains within a standard deviation of our preferred coefficient. The coefficient on population growth exhibits even more stability. We also note that the overidentification tests are always easily passed. The results for the endogeneity tests are more mixed. In some cases, the difference between the IV and OLS results is statistically significant. Although the differences remain economically small, this is indicative of a tendency for the coefficients on population and land area to be larger in magnitude under IV than under OLS. A first possible reason is that our instruments correct for measurement error. Recall also that our model suggests two missing variables: unobserved productivity which should be positively correlated with land prices at the centre and the fraction of land developed around the CBD which should be negatively correlated with land prices. Hence, our OLS coefficients could be upward-biased if the unobserved productivity bias dominates or downward-biased if the bias caused by the unobserved fraction of developed land dominates. Our IV results suggest the second bias may be more important. This said, the small difference between our OLS and IV results suggests that these biases are small.

Before checking the robustness of our results further, we return to our model. Recall that the coefficient on log population is $(1+\sigma)\,\gamma$. Estimates of the agglomeration parameter σ in the literature are low: 0.015 to 0.03 in France (Combes *et al.*, 2010). Since our preferred estimate for the coefficient on population is 0.721, these numbers suggest an estimate of 0.70 to 0.71 for γ, the decentralisation coefficient introduced in equation (13). To obtain a simple independent estimate of this coefficient, we can estimate equation (13) directly by regressing log population in the core municipality of urban areas on the log population of their urban areas.[16] We find a coefficient of 0.796 with an R^2 of 0.89 in a simple OLS univariate regression. Adding the full list of controls used in table 3 leads to a slightly lower coefficient of 0.728 (with an R^2 of 0.90). In both cases, these coefficients are highly significant and remarkably close to the indirect estimates of 0.70 to 0.71 implied by our main regressions.

We also note that equation (25), which relates log land area to log urban area population, provides us with another independent estimate of $(1+\sigma)\gamma$. When we regress log land area on log population alone, we obtain a coefficient of 0.795 and an R^2 of 0.66. Adding the full list of controls again leads to a slightly lower coefficient of 0.695 with an R^2 of 0.75. Again, these coefficients are highly significant and remarkably close to our indirect estimate of 0.721.

Further robustness checks

As argued above, our main concerns regard functional forms and the existence of missing variables correlated with land prices and with population or land area. To mitigate the latter concern, we verified in table 5 that statistically similar coefficients were obtained with different combinations of instruments. We verify here that these IV results are also robust to introducing supplementary control variables. In table 6 we duplicate all the specifications of table 5 but add the full list of geographical, socio-economic, development, and geological controls used in table 4.

When we compare table 6 (with controls) with table 5 (without), the first main conclusion is that the coefficients on population are about the same. On average, they are higher by 0.013 with controls than without. The maximum gap is 0.037 and occurs for our preferred set of instruments in column (1). These differences are small, both statistically and economically. For the coefficients on land area, we find that they are slightly smaller in magnitude with controls than without. Finally, the coefficient on population growth is about 10% larger. We also note that despite adding powerful controls, our instruments are not weakened. If anything, they are stronger. Overall these results confirm and reinforce the results obtained previously.

[16]Of course, we do not expect the boundaries of the core municipality of urban areas to coincide exactly with the monocentric city of our model because French municipalities are small and the centre of urban areas should attract workers well beyond this core municipality. However, a result of our model above is that as cities grow, population density should increase everywhere inside the monocentric city by the same proportion when the urban fringe is fixed. To see this, recall that density at any point located at distance D from the CBD is $d(D) = H(D)/(2\pi\theta D h(D))$. Using (2), (6), (7), (10), and (12) shows the result. We also note that French urban areas are defined by iterative aggregation of municipalities using a minimum threshold for the share of workers that commute to the rest of the urban area.

Table 6: Unit land values at the centre, population and area instrumented, all controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.758^a	0.674^a	0.634^a	0.709^a	0.699^a	0.694^a	0.656^a	0.712^a
	(0.087)	(0.081)	(0.091)	(0.066)	(0.060)	(0.059)	(0.060)	(0.057)
Land area	-0.583^a	-0.477^a	-0.425^a	-0.421^a	-0.413^a	-0.494^a	-0.448^a	-0.438^a
	(0.111)	(0.111)	(0.123)	(0.093)	(0.093)	(0.071)	(0.079)	(0.085)
Population growth	3.432^a	3.426^a	3.428^a	3.603^a	3.595^a	3.443^a	3.438^a	3.582^a
	(0.465)	(0.446)	(0.442)	(0.464)	(0.415)	(0.449)	(0.444)	(0.464)
Overidentification p-value	0.44	0.08	0.05	0.95	0.72	0.77	0.22	0.47
First-stage statistic	10.4	10.5	8.7	14.7	14.1	27.9	18.2	16.5
Endogeneity p-value	0.02	0.12	0.23	0.01	0.01	0.02	0.09	0.01
Number of hotel rooms	Y	Y	Y	N	Y	Y	Y	N
Share of 1-star rooms	Y	Y	N	N	Y	N	N	N
Temperature in January	Y	N	Y	N	Y	Y	N	Y
Bartik occupations 1999	N	Y	N	N	N	N	Y	N
Bartik industry 1999	N	N	Y	N	N	N	N	N
Urban population 1831	N	N	N	Y	Y	N	N	Y
Urban density 1881	N	N	N	Y	N	N	N	N
Henderson industry 1999	N	N	N	Y	Y	Y	Y	N
Henderson occupations 1999	N	N	N	N	Y	N	N	Y
1st Shea part. R^2, population	0.17	0.19	0.16	0.27	0.33	0.34	0.35	0.37
1st part. Fisher, population	79.4	78.9	79.2	56.5	49.7	86.8	91.0	51.5
1st Shea part. R^2, area	0.11	0.11	0.09	0.15	0.15	0.24	0.20	0.18
1st part. Fisher, area	40.4	35.0	37.0	17.2	15.4	20.7	35.5	13.1

Notes: 278 observations in each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions are estimated with LIML and include a constant and the full set of supplementary controls described in 4.

Population growth is an important determinant of land prices in our second-step regression. Its coefficient is highly significant and it increases the explanatory power of the regression. Like with the level of population and perhaps more strongly so, we expect some simultaneity with land prices since it is reasonable to expect lower population growth in more expensive urban areas. Although the coefficient of population growth is not our main coefficient of interest, a biased coefficient on population growth may percolate to the coefficients on population and land area. A first answer here is to note that population growth is only weakly correlated with land area and population. Recall that introducing population growth in the regression does not affect the magnitude of the other two coefficients as shown in table 4. A second response is to find appropriate instruments for population growth. We can use our Bartik variables for sectors and occupations to instrument for population growth.

Table 7: Unit land values at the centre, growth instrumented

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.607^a	0.610^a	0.607^a	0.598^a	0.688^a	0.657^a	0.661^a	0.662^a
	(0.028)	(0.029)	(0.028)	(0.029)	(0.060)	(0.058)	(0.060)	(0.058)
Land area	-0.436^a	-0.438^a	-0.436^a	-0.429^a	-0.532^a	-0.469^a	-0.483^a	-0.484^a
	(0.030)	(0.030)	(0.030)	(0.031)	(0.084)	(0.076)	(0.079)	(0.076)
Population growth	3.780^a	3.512^a	3.743^a	4.581^a	3.825^a	3.473^a	3.417^a	3.449^a
	(0.821)	(1.106)	(0.811)	(0.970)	(0.832)	(0.708)	(0.728)	(0.711)
Overidentification p-value	–	–	0.76	0.06	0.50	0.44	0.44	0.43
First-stage statistic	81.0	39.2	41.6	29.9	7.5	8.0	9.1	7.3
Number of hotel rooms	N	N	N	N	Y	N	N	Y
Temperature in January	N	N	N	N	Y	Y	Y	Y
Bartik occupations 1999	N	Y	Y	N	N	N	N	N
Bartik industry 1999	Y	N	Y	N	Y	Y	Y	Y
Bartik occupations 1990	N	N	N	Y	N	N	N	N
Bartik industry 1990	N	N	N	Y	N	N	N	N
Urban population 1831	N	N	N	N	N	Y	Y	Y
Henderson industry 1999	N	N	N	N	Y	N	Y	N
Henderson occupations 1999	N	N	N	N	N	Y	N	Y
1st Shea part. R^2, population	-	-	-	-	0.21	0.22	0.26	0.22
1st part. Fisher, population	-	-	-	-	237.3	112.6	293.2	292.3
1st Shea part. R^2, area	-	-	-	-	0.13	0.15	0.19	0.15
1st part. Fisher, area	-	-	-	-	83.4	76.2	100.5	88.9
1st Shea part. R^2, growth	0.23	0.13	0.23	0.18	0.23	0.31	0.30	0.31
1st part. Fisher, growth	81.0	39.2	41.6	29.9	9.0	10.5	15.2	15.1

Notes: 278 observations for each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions are estimated with LIML and include a constant but no supplementary control.

Table 7 reports results for specifications where population growth is instrumented. In columns (1) to (4), we use our Bartik predictors of growth based on industry and occupational composition for 1990-1999 and 1999-2007 to instrument for population growth. We note that these instruments are strong. Relative to the OLS specifications of column (3) of table 4, the coefficients on population and land area are almost identical. The instrumented coefficients on population growth are higher than their corresponding OLS coefficients but the differences are not significant. In columns (5) to (8), we also instrument for population and land area using the same instruments as in the last two tables. The results remain very close to those of table 5. Finally in results not reported here, we also duplicated the specifications of table 7 but added the same supplementary controls as in table 6. The results were again unaffected. From this, we conclude that our results are unaffected by

Table 8: Unit land values at the centre, effect of first-stage specification, OLS regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.573^a	0.712^a	0.783^a	0.766^a	0.613^a	0.644^a	0.614^a	0.636^a
	(0.040)	(0.038)	(0.038)	(0.036)	(0.027)	(0.049)	(0.028)	(0.027)
Land area	-0.526^a	-0.613^a	-0.539^a	-0.545^a	-0.440^a	-0.336^a	-0.408^a	-0.457^a
	(0.049)	(0.041)	(0.041)	(0.038)	(0.029)	(0.053)	(0.030)	(0.029)
Population growth	3.210	3.900^a	3.645^a	3.520^a	3.254^a	3.503^a	3.400^a	3.437^a
	(0.439)	(0.553)	(0.552)	(0.517)	(0.394)	(0.710)	(0.402)	(0.391)
R^2	0.69	0.63	0.67	0.69	0.72	0.51	0.72	0.74
Observations	26,177	278	278	278	278	278	278	278

Notes: The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions include a constant but no supplementary control at the city level. Column (1) corresponds to a single-step regression including eight parcel characteristics, population, population growth, and land area with standard errors clustered by urban area. Columns (2) to (8) estimate the same OLS specification as in column (3) of table 4 but use as dependent variable the output of the corresponding column in table 3.

how we treat population growth.

Next we explore the sensitivity of our results to the specification of the first stage. So far we have used as dependent variable the urban area fixed effects estimated from our preferred first-step estimation of column (5) in table 3. In table 8, we repeat our main second-step OLS specification for all the first-step specifications used in table 3. Because column (1) of table 3 does not estimate city fixed effects, it has no corresponding second-step estimation. Given this, column (1) of table 8 reports the results of a single-step estimation that includes all our parcel characteristics from the first step and our three main second-step explanatory variables, population, land area, and population growth.

In column (1), the coefficient on city population when estimated in one step is only marginally lower at 0.573 than our preferred (two-step) OLS estimate of 0.613. The gap for the coefficient on land area is larger, -0.526 instead of -0.440, but still minor. Columns (2) to (4) correspond to first-step specifications which do not include log parcel area nor its square. Relative to our preferred OLS, we find coefficients that are somewhat larger in magnitude for both population and land area. While these differences do not affect our conclusions, we note that they are larger than other

Table 9: Unit land values at the centre, change of gradient specification, OLS all controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.579^a	0.587^a	0.590^a	0.568^a	0.531^a	0.499^a	0.622^a	0.665^a
	(0.034)	(0.035)	(0.035)	(0.035)	(0.062)	(0.051)	(0.054)	(0.083)
Land area	-0.454^a	-0.393^a	-0.391^a	-0.367^a	-0.391^a	-0.363^a	-0.612^a	-0.700^a
	(0.036)	(0.036)	(0.036)	(0.036)	(0.064)	(0.052)	(0.056)	(0.086)
Population growth	3.403^a	3.445^a	3.362^a	3.420^a	3.310^a	3.518^a	2.855^a	2.249^b
	(0.449)	(0.452)	(0.457)	(0.461)	(0.813)	(0.661)	(0.703)	(1.085)
R^2	0.77	0.77	0.77	0.76	0.57	0.58	0.61	0.44

notes: ^{a}, ^{b}, ^{c}: 278 observations in each regression. Significant at 1%, 5% and 10% respectively. All columns correspond to the same specification which include all supplementary controls. The urban fixed effects from the first step were all estimated conditioning out all parcel characteristics. In column (1), we take distance in levels at the first step. In column (2), we take log distance to the densest municipality of the urban area. In column (3), we take log distance to the municipality with the most employment. In column (4), we take log distance to the closest of two centres. In column (5), we take a log distance gradient specific to each urban area. In column (6), we take a gradient in level specific to each urban area. In column (7), we estimate land prices at 10 kilometers from the centre using log distance gradients specific to each city. In column (8), we repeat the same exercise with gradients in level.

differences in other robustness tests.[17]

Columns (5) to (8) correspond to first-step specifications that all include all parcel characteristics. Column (5) repeats our preferred OLS. Column (6) indicates that allowing for city-specific gradients makes little change to the results. Columns (7) and (8) show that allowing for multiple centres or taking distances in levels rather than logs makes again only a small difference to the coefficients on log population, log area, and population growth.

We now examine more systematically the effects of the specification of the distance gradient in the first stage. We take again our preferred first stage estimation with parcel characteristics but consider alternative measures of distance. Table 9 reports results corresponding to the OLS specification of column (8) in table 4 with all city-level controls to avoid repeating some of the results already reported in the previous table.

In column (1) of table 9, we use urban area fixed effects estimated controlling for distance in levels instead of log. This makes no difference to the coefficient on city population and only minimal differences to the coefficients on land area and population growth. In column (2), we

[17] We also experimented with IV estimations. When using the same dependent variable as in column (4) (where the first step controls for all parcel characteristics except log parcel area and its square), we estimate the population elasticity to be 0.935 (instead of 0.721) with our preferred specification of column (1) of table 5 and 0.766 (instead of 0.680) for the estimation of column (4) of the same table which uses a completely different set of instruments. Overall, the IV estimates for these two alternative dependent variables are slightly higher and exhibit more variation across the different sets of instruments.

Table 10: Econometrics variants

	(1) OLS	(2) One-step OLS	(3) FGLS	(4) WLS	(5) OLS	(6) One-step OLS	(7) FGLS	(8) WLS
Population	0.611^a	0.614^a	0.616^a	0.617^a	0.568^a	0.611^a	0.579^a	0.580^a
	(0.029)	(0.040)	(0.027)	(0.026)	(0.036)	(0.040)	(0.034)	(0.034)
Land area	-0.460^a	-0.487^a	-0.444^a	-0.445^a	-0.393^a	-0.478^a	-0.415^a	-0.416^a
	(0.028)	(0.049)	(0.030)	(0.029)	(0.034)	(0.041)	(0.036)	(0.036)
Population growth	3.804^a	3.126^a	3.187^a	3.168^a	3.405^a	2.325^a	3.222^a	3.211^a
	(0.378)	(0.437)	(0.393)	(0.387)	(0.432)	(0.473)	(0.449)	(0.440)
Controls	no	no	no	no	all	all	all	all
R^2	0.65	0.71	0.74	0.73	0.69	0.74	0.78	0.77
Obs.	345	26,177	278	278	345	26,177	278	278

notes: ^{a}, ^{b}, ^{c} : Significant at 1%, 5% and 10% respectively. Controls are the urban area characteristics described in table 4. The one-step estimations of columns (2) and (6) also control for the 10 parcels characteristics used in column (5) of table 3.

change the definition of the centre of an urban area and use the densest municipality. This makes virtually no difference. In column (3), we use the municipality with most employment as centre and it again makes no difference. In column (4), we allow for two centres and take the distance to the closest. This makes very little difference yet again.[18] In columns (5) to (8), we allow the distance to the centre to affect prices differently in different cities by estimating city specific gradients. We do this in log in column (5) and in levels in column (6). In columns (7) and (8), we repeat the same exercise to estimate this time land prices at 10 kilometres from the centre. While in columns (5) and (6), the coefficients on population and land area are slightly smaller in magnitude relative to a single gradient, in columns (7) and (8), they are slightly larger. This generally confirms our main results but also highlights a greater fragility of the results of specifications that use gradient specific to each urban area in the first step.

In table 10, we provide a final set of robustness checks. In column (1) we run our main OLS regression of column (3) of table 4 using our largest sample of 345 cities instead of the 278 for which our instruments are available. This makes virtually no difference. In column (2), we use the same dependent variables as in the two step estimation of column (3) of table 4 but perform our estimation in one step. Relative to the one-step estimation reported in table 8, the specification

[18]Consistent with these results, the pairwise correlations between our preferred distance metric and the four alternatives used in columns (1) to (4) are 0.80, 0.90, 0.92, and 0.85, respectively.

here includes a distance gradient. Again this makes no difference to our results despite the fact that our two-step regressions take cities as observations in our second step whereas this one-step estimation takes parcels as units of observations (and thus imposes a different weighting scheme).

An issue with our two-step estimations is that, as second-step dependent variable, we use an estimator instead of its true value. This may affect the standard errors. It is possible to take into account the specific structure of the covariance matrix of error terms in second step using feasible generalised least squares (FGLS). Alternatively, we also computed weighted least squares (WLS) with robust standard errors. This approach is often considered to be more robust. These two estimation techniques are briefly described in Appendix B. Columns (3) and (4) report results for the FGLS and WLS estimations corresponding to our main OLS specification which includes population, land area, and population growth as explanatory variables. The results are virtually the same. The standard errors are unchanged with FGLS, and if anything, are slightly lower with WLS. Finally columns (5) to (8) duplicate the specification of columns (1) to (4) adding our full set of controls for urban area characteristics. These results are best compared to the OLS specification of column (8) of table 4 which also includes the same controls. Again the results are extremely similar.

Extensions

Larger cities might experience systematically higher distance gradients for the rental price of their land because of congestion. We verified in table 9 that allowing for city specific gradients did not affect our second-step results. However, we could still face a problem because the land gradient enters equation (18) which underlies our main second-step regression. If the distance gradient is constant across cities, it is part of the constant of the regression. If it varies idiosyncratically across cities, it enters the error term. Finally, if it varies systematically with our explanatory variables, our results might be biased. To verify that this is not the case, we regress the city-specific gradients estimated in column (6) of table 3 on the same explanatory variables that we use above. Table 11 reports results for a variety of OLS and IV specifications.

A number of interesting conclusions can be drawn from the results of table 11. First, as can be seen from column (1), population, land area and city growth have little explanatory power. The R^2 for this specification is only 0.03 instead of 0.72 for the corresponding regression using

Table 11: Determinants of the distance gradients for land prices, OLS and IV regressions

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV	(8) IV
Population	0.022	0.040	0.033	−0.050	−0.013	0.005	0.001	0.006
	(0.028)	(0.038)	(0.041)	(0.066)	(0.069)	(0.077)	(0.060)	(0.082)
Land area	-0.102^a	-0.096^b	-0.090^b	−0.005	−0.065	−0.042	−0.085	−0.093
	(0.031)	(0.039)	(0.041)	(0.087)	(0.090)	(0.086)	(0.079)	(0.102)
Population growth	−0.051	−0.197	−0.203	0.094	0.028	0.180	0.017	−0.498
	(0.413)	(0.498)	(0.498)	(0.435)	(0.431)	(0.490)	(0.738)	(1.243)
Controls	no	all	all+gini	no	no	all+gini	no	all+gini
R^2	0.03	0.12	0.12	-	-	-	-	-
Overidentification p-value	-	-	-	0.94	0.70	0.89	0.57	0.49
First-stage statistic	-	-	-	13.4	11.8	21.9	8.0	7.9
Number of hotel rooms	-	-	-	Y	N	Y	N	N
Temperature in January	-	-	-	Y	Y	Y	Y	Y
Bartik industry 1999	-	-	-	N	N	N	Y	Y
Urban population 1831	-	-	-	N	Y	N	Y	Y
Henderson industries 1999	-	-	-	Y	N	Y	N	N
Henderson occupations 1999	-	-	-	N	Y	N	Y	Y
1st Shea part. R^2, population	-	-	-	0.19	0.17	0.26	0.22	0.23
1st part. Fisher, population	-	-	-	293.2	148.9	74.7	112.6	33.6
1st Shea part. R^2, area	-	-	-	0.13	0.11	0.21	0.15	0.15
1st part. Fisher, area	-	-	-	100.5	100.5	19.8	76.2	9.8
1st Shea part. R^2, pop. growth	-	-	-	-	-	-	0.31	0.15
1st part. Fisher, pop. growth	-	-	-	-	-	-	10.5	16.5

Notes: 278 observations in each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All IV regressions are estimated with LIML and all regressions include a constant. In columns (3), (6), and (8), we also control for a Gini index of city income inequalities since the monocentric model predicts that inequalities should affect the land price gradient (Fujita, 1989).

city fixed effects as dependent variable. Adding a full set of urban area characteristics in column (2) only raises the R^2 to 0.12. Second, the coefficient on log population is always small and never significant despite small standard errors. Third, the coefficient on land area is sometimes significant with OLS and never with IV. Cities with a larger area appear to have modestly steeper land gradients. Overall, while it is difficult to prove a negative result, the results of table 11 are nonetheless suggestive that population does not affect the distance gradients for land prices. This shows that our strategy is consistent. A better understanding of the drivers of the variation of these city distance gradients is left for future work.

We now turn to the estimation of equation (20) where land area is allowed to adjust with population whereas most of the regressions reported so far were concerned with equation (18)

Table 12: Unit land values at the centre, regressions without land area

	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV	(6) IV	(7) IV	(8) IV
Population	0.288[a]	0.346[a]	0.286[a]	0.360[a]	0.278[a]	0.422[a]	0.296[a]	0.370[a]
	(0.022)	(0.034)	(0.024)	(0.094)	(0.027)	(0.067)	(0.025)	(0.047)
Population growth	3.903[a]	3.526[a]	3.912[a]	3.557[a]	3.934[a]	3.698[a]	3.305[a]	3.322[b]
	(0.528)	(0.551)	(0.526)	(0.527)	(0.779)	(0.556)	(1.035)	(1.359)
Controls	no	all	no	all	no	all	no	all
R^2	0.48	0.64	-	-	-	-	-	-
Overidentification p-value	-	-	0.52	0.53	0.66	0.75	0.11	0.30
First-stage statistic	-	-	629.9	145.8	267.8.1	47.5	25.3	12.7
Number of hotel rooms	-	-	Y	Y	N	N	Y	Y
Bartik industry 1999	-	-	N	N	Y	Y	Y	Y
Urban population 1831	-	-	Y	Y	Y	Y	Y	Y
Urban density 1831	-	-	N	N	N	N	Y	Y
1st Shea part. R^2, population	-	-	0.82	0.53	0.66	0.26	0.79	0.52
1st part. Fisher, population	-	-	629.9	145.8	267.8	43.2	358.3	82.1
1st Shea part. R^2, population growth	-	-	-	-	-	-	0.27	0.16
1st part. Fisher, population growth	-	-	-	-	-	-	27.6	4.82

Notes: 278 observations in each regression. The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All IV regressions are estimated with LIML and all regressions include a constant.

where land area was included in the regression. As already argued, with land development being heavily restricted, we view the coefficient on population in the estimation of (18) as more relevant to think about urban costs in contemporary French cities. However, estimating (20) is interesting in its own right.

Table 12 reports results for two OLS and six IV regressions using urban area fixed effects as dependent variable and population as main explanatory variable. Column (1) is an OLS specification which includes only population and population growth as dependent variables. The coefficient on population is 0.288. Adding supplementary city controls in column (2) raises this coefficient slightly to 0.346. In columns (3) to (6), population is instrumented using two combinations of internal and external instruments and controlling or not for city characteristics. In columns (7) and (8), population growth is also instrumented, again with and without city controls. The IV coefficient on population ranges between 0.278 and 0.296 and is thus very close to the corresponding OLS coefficient of column (1). In columns (4), (6), and (8), city characteristics are included as controls and the coefficient on population is between 0.360 and 0.420. This is slightly higher than the corresponding OLS coefficient of column (2). Given that we use our set of controls only for robustness, we retain the IV coefficient of column (3) of 0.286 as our preferred estimate.[19]

The estimation of equation (18) also allows us to perform an 'overidentification' test to evaluate the overall consistency of our approach. We note from equation (20) that the coefficient on population is interpreted as $(1+\sigma)\gamma\ \tau/(2\alpha\beta)$ in our model whereas, when including land area in equation (18), the same coefficient now corresponds to $(1+\sigma)\gamma$. Taking our preferred estimate of 0.721 for the conditional elasticity of land prices with respect to population and 0.286 for the corresponding unconditional elasticity, we obtain an estimate of $\tau/(2\alpha\beta) = 0.286/0.721 = 0.397$. From equation (18), we also know that the coefficient on land area corresponds to $-1+\tau/(2\alpha\beta)$. Taking our preferred estimate of -0.563 for this coefficient implies $\tau/(2\alpha\beta) = 0.437$. Although obtained independently, these two numbers are remarkably close.[20]

6. The elasticity of urban costs with respect to population

The elasticity of urban costs with respect to urban area population is the product of the elasticity of unit land prices with respect to population times the share of housing in expenditure β times the share of land in housing α. Our preferred estimate for the first elasticity is 0.72. We also note that most of our other estimates are between 0.6 and 0.8.

A recent study by Davis and Ortalo-Magné (2011) estimates the share of housing in expenditure at 0.24 in the US. For France, a detailed evaluation made by the French ministry that oversees housing (CGDD, 2011) proposes a very similar number at 0.23. We retain this number since an independent estimate would be beyond the scope of this paper.

Turning to the share of land in housing, note that the first-order conditions for profit maximisation in housing development with respect to land implies that the *user cost* of land r^L is such that $r^L L = \alpha QH$ where α is the share of land in housing production and QH is the value of housing. The second first-order condition implies that the user cost of capital r^K is such that $r^K K = (1-\alpha)QH$.

[19]The coefficient on population can also be computed indirectly from our preferred specification of column (1) in table 5. Using the fact that the elasticity of land area with respect to population is 0.932 and that our preferred estimates for the elasticity of unit land prices with respect to population and land area are 0.721 and -0.563, we find the unconditional population elasticity of land prices to be $0.721 - 0.932 \times 0.563 = 0.193$. This number is fairly close to our preferred direct IV estimates of the same elasticity, which is 0.286.

[20]The first step estimation described by equation (17) provides an estimate of the coefficient on distance which corresponds to $-\tau/(\alpha\beta)$. Table (3) reports values between -0.18 and -0.26. These clearly lead to smaller estimates of $\tau/(2\alpha\beta)$. We attribute this difference to the uncertainty around the exact functional form for distance in the first step and mismeasurement of distance.

These two first-order conditions imply

$$\frac{\alpha}{1-\alpha} = \frac{r^L\,L}{r^K\,K}\,. \tag{28}$$

Then, we know from our data that the value of land accounts for about 40% of the value of housing. That is, $L/K \approx 0.66$. Because housing capital depreciates, we think of the user cost of capital as being equal to the interest rate plus the rate of housing depreciation. Taking values of 5% for the rate of interest and 1% for housing capital depreciation yields $r^K = 0.06$. For the US, Davis and Heathcote (2005) take a slightly higher value of 1.5% for depreciation.

Unlike capital, land does not generally depreciate but appreciates instead. According to our results, a 1% higher population causes 0.7% higher land prices. Assuming a growth rate in real incomes slightly above 1% suggests a rate of appreciation for the price of land in urban areas of about 2% per year (with 0.7% coming from population growth and 1.3% from income growth). This rate appreciation is slightly below the rate of 3% used by Davis and Heathcote (2007). In turn a 2% annual appreciation suggests a user cost of land $r^L = 0.05 - 0.02 = 0.03$. Inserting these numbers into equation (28) yields $\alpha = 0.25$.[21] Combining this estimate of 0.25 for the share of land in housing with 0.23 for the share of housing in expenditure and a population elasticity of unit land prices of 0.72 yields an elasticity of urban cost with respect to population of 0.041.

While one may take issues with any of the numbers we use, it is hard to reach an elasticity of urban costs that varies dramatically from our derived estimate of 0.041. For instance, taking our highest estimate of around 0.80 for the population elasticity of unit land prices would only raise our elasticity of urban costs to 0.046. Increasing the user cost of land from 3% to 4% would raise the elasticity of urban costs to 0.05. The unrealistic assumption of no appreciation in land prices over time (and thus a user cost of land of 5%) would only imply an elasticity of urban costs of about 0.06. Even in the extreme case of no land price appreciation, a share of housing of 35% instead of 23%, and an elasticity of land prices of 0.8, we only get an elasticity of urban costs of 0.10.

It is possible to make our estimate of the elasticity of urban costs lower as well. Assuming for instance a low elasticity of land prices of 0.5 or a low user cost for land of 0.02 would lead to an elasticity of urban costs of about 0.03. Taking extreme values of 0.02 for the user cost of land, 0.15 for the share of housing, and 0.4 for the elasticity of land prices yields an elasticity of 0.01. Overall,

[21] Combes *et al.* (2012) propose an alternative and more detailed approach to the estimation of a production function for housing which yields very similar results for the share of land and show that the production function for housing can be closely approximated by a Cobb-Douglas function with constant returns.

our preferred value for the elasticity of urban costs is 0.041. One could reasonably conceive any number between 0.03 and 0.05. Getting numbers outside of this tight range requires making more extreme assumptions.

Our preferred value of the elasticity of urban costs of 0.041 also needs to be contrasted with the already reported estimates for agglomeration effects in France which range from 0.015 to 0.03 (Combes *et al.*, 2010). Although the coefficient on urban costs is higher than that on agglomeration, the two numbers are very close.

We also note that allowing the land area to adjust following increases in population leads to a much lower estimate of the elasticity of land prices with respect to population of 0.286 instead of 0.721. In turn, this leads to an estimate of 0.016 for the elasticity of urban costs with respect to population. Urban costs are much lower when the physical growth of cities is not restricted.

7. Conclusion

This paper develops a new methodology to estimate the elasticity of urban costs with respect to city population. Building on an extension of the standard monocentric framework, our model derives this elasticity as the product of the three terms: the elasticity of the rental price of land in the centre of cities with respect to population, the share of housing in consumer expenditure, and the share of land in housing.

While we rely on an external estimate for the share of housing in consumer expenditure and obtain the share of land in housing directly from the data, we devote considerable attention to the estimation of the population elasticity of land prices. Implementing our approach on unique land price data for France we obtain an estimate of 0.041 for the population elasticity of housing costs. Our estimated elasticity of unit land prices with respect to city population of 0.72 is also of interest.

These findings have a number of interesting implications. The first is that we provide the first evidence on the cost side about what Fujita and Thisse (2002) dub the 'fundamental tradeoff of spatial economics'. While the existence of increasing urban costs associated with the scarcity of land and a loss of accessibility to central locations as cities grow was never really in doubt for any casual observer of cities, it is important to provide some quantitative estimates for them. In this respect, our estimates for urban costs are somewhat modest. Our preferred estimate for urban costs is close to but larger than existing estimates for agglomeration effects. This suggests that cities operate close to constant returns in the aggregate. This also implies that the cost of cities being oversized might be small as already suggested by Au and Henderson (2006). This also means that large deviations from optimal size might happen at a low economic cost for cities. Put differently, while our results provide evidence regarding the existence of the fundamental tradeoff of spatial economics, they also suggest that this tradeoff between agglomeration and urban costs may be of little relevance to understand the future evolution of the size of cities.

References

Albouy, David and Gabriel Ehrlich. 2012. Metropolitan land values and housing productivity. Working Paper 18110, National Bureau of Economic Research.

Alonso, William. 1964. *Location and Land Use; Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.

Au, Chun-Chung and J. Vernon Henderson. 2006. Are Chinese cities too small? *Review of Economic Studies* 73(3):549–576.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo (MI): W.E. Upjohn Institute for Employment Research.

Card, David and Alan B. Krueger. 1992. School quality and black-white relative earnings: A direct assessment. *Quarterly Journal of Economics* 107(1):151–200.

Carlino, Gerald A. and Albert Saiz. 2008. Beautiful city: Leisure amenities and urban growth. FRB of Philadelphia Working Paper No. 08-22.

Cheshire, Paul C. and Stefano Magrini. 2006. Population growth in european cities: weather matters-but only nationally. *Regional Studies* 40(1):23–37.

Ciccone, Antonio and Robert E. Hall. 1996. Productivity and the density of economic activity. *American Economic Review* 86(1):54–70.

Clark, Colin. 1951. Urban population densities. *Journal of the Royal Statistical Association Series A* 114(4):490–496.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2011. The identification of agglomeration economies. *Journal of Economic Geography* 11(2):253–266.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2012. The production function of housing: Evidence from France. Processed, University of Toronto.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. 2013. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica* forthcoming.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. Estimating agglomeration economies with history, geology, and worker effects. In Edward L. Glaeser (ed.) *The Economics of Agglomeration*. Cambridge (MA): National Bureau of Economic Research, 15–65.

Commissariat Général au Développement Durable. 2011. *Comptes du Logement: Premiers Résultats 2010, le Compte 2009*. Paris: Ministère de l'Ecologie, du Développement Durable, des Transports et du Logement.

Davis, Morris A. and Jonathan Heathcote. 2005. Housing and the business cycle. *International Economic Review* 46(3):751–784.

Davis, Morris A. and Jonathan Heathcote. 2007. The price and quantity of residential land in the United States. *Journal of Monetary Economics* 54(8):2595–2620.

Davis, Morris A. and François Ortalo-Magné. 2011. Household expenditures, wages, rents. *Review of Economic Dynamics* 14(2):248–261.

Davis, Morris A. and Michael G. Palumbo. 2008. The price of residential land in large US cities. *Journal of Urban Economics* 63(1):352–384.

Duranton, Gilles. 2008. Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics* 41(3):689–736.

Duranton, Gilles and Diego Puga. 2005. From sectoral to functional urban specialisation. *Journal of Urban Economics* 57(2):343–370.

Duranton, Gilles and Diego Puga. 2012. The growth of cities. Work in progress, University of Toronto.

Fujita, Masahisa. 1989. *Urban Economic Theory: Land Use and City Size*. Cambridge: Cambridge University Press.

Fujita, Masahisa and Hideaki Ogawa. 1982. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics* 12(2):161–196.

Fujita, Masahisa and Jacques-François Thisse. 2002. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge University Press.

Guérin-Pace, France and Denise Pumain. 1990. 150 ans de croissance urbaine. *Economie et Statistiques* 0(230):5–16.

Handbury, Jessie H. and David E. Weinstein. 2010. Is new economic geography right? Evidence from price data. Processed, Columbia University.

Henderson, J. Vernon. 1974. The sizes and types of cities. *American Economic Review* 64(4):640–656.

Henderson, Vernon. 2002. Urban primacy, external costs, and the quality of life. *Resource and Energy Economics* 24(1):95–106.

Krueger, Alan B., Daniel Kahneman, David Schkade, Norbert Schwarz, and Arthur A. Stone. 2009. National time accounting: The currency of life. In Alan B. Krueger (ed.) *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being*. Chicago: NBER and University of Chicago Press, 9–86.

McMillen, Daniel P. 2006. Testing for monocentricity. In Richard J. Arnott and Daniel P. McMillen (eds.) *A Companion to Urban Economics*. Oxford: Blackwell, 128–140.

Mills, Edwin S. 1967. An aggregative model of resource allocation in a metropolitan area. *American Economic Review (Papers and Proceedings)* 57(2):197–210.

Muth, Richard F. 1969. *Cities and Housing*. Chicago: University of Chicago Press.

Puga, Diego. 2010. The magnitude and causes of agglomeration economies. *Journal of Regional Science* 50(1):203–219.

Rappaport, Jordan. 2007. Moving to nice weather. *Regional Science and Urban Economics* 37(3):375–398.

Richardson, Harry W. 1987. The costs of urbanization: A four-country comparison. *Economic Development and Cultural Change* 35(3):561–580.

Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2119–2171.

Small, Kenneth A. and Erik T. Verhoef. 2007. *The Economics of Urban Transportation*. New York (NY): Routledge.

Stock, James H. and Motohiro Yogo. 2005. Testing for weak instruments in linear IV regression. In Donald W.K. Andrews and James H. Stock (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, 80–108.

Thomas, Vinod. 1980. Spatial differences in the cost of living. *Journal of Urban Economics* 8(1):108–122.

Tolley, George S., Philip E. Graves, and John L. Gardner. 1979. *Urban Growth Policy in a Market Economy*. New York: Academic Press.

White, Michelle J. 1999. Urban areas with decentralized employment: Theory and empirical work. In Edwin S. Mills and Paul Cheshire (eds.) *Handbook of Regional and Urban Economics vol. III*. North-Holland: Elsevier Science, 1375–1412.

Appendix A. Data description

Enquête sur le Prix des Terrains à Bâtir (EPTB). This survey is conducted every year in France since 2006 by the French Ministry of Ecology, Sustainable Development, Transports, and Housing. The sample is composed of land parcels drawn from Sitadel, the official registry which covers the universe of all building permits for a detached house. Households must be the project managers and houses must include only one dwelling. Permits for extensions to existing houses are excluded.

Parcels are drawn randomly in each strata (about 3,700 of them) which corresponds to a group of municipalities (about 36,000 in France). Two thirds of the permits are surveyed. Some French regions paid for a larger sample. In total, this corresponds to about 120,000 permits per year. The survey is mandatory and the response rate after one follow-up is expected to be around 75%. This means that around 90,000 households should be surveyed every year. For the year 2008, we have 82,586 observations.

Population census. We have access to data on employment and population at the municipality level for the 1990, 1999 and 2007 censuses. These data are used for several purposes such as the construction of our explanatory variable of interest (population of each urban area), an important control variable (population growth), and several supplementary control variables (share of workers with a university degree, peripherality index, etc). We also use population data to compute the barycentre of urban areas. Finally we use detailed information from the 1/4 sample of the 1990 census and the 1/20 sample of the 1999 census to construct measures of employment (by municipality of residence) by 4-digit occupational category and by 4-digit sector for each urban area (weighting by survey rates for the data to be representative of the whole population of occupied workers). The resulting aggregates are used to construct Bartik and Henderson instruments.

Bartik and Henderson instruments. To ease the exposition, we index a given census by t and the previous census by $t-1$. Denote N_{jpt} the population in urban area j in the four-digit professional occupation p at time t, N_{jst} the population in urban area j in the four-digit sector s, N_{jt} the population in urban area j, $\bar{N}_{pt}$ the average urban area population in the four-digit professional occupation p ($\bar{N}_{pt} = \sum_j N_{jpt}/n$ where n is the number of urban areas), and $\bar{N}_{st}$ the average urban area population in the four-digit sector s ($\bar{N}_{st} = \sum_j N_{jst}/n$). The Henderson 'occupation' instrument is:

$$H_{jt}^{occc} = \sum_p \left(\frac{N_{jpt}}{N_{jt}} \right) \bar{N}_{pt} \tag{A1}$$

A similar computation is applied to sectoral employment to construct the Henderson sectoral instrument. The Bartik sectoral instrument is:

$$B_{jt}^{sec} = \sum_s \left(\frac{N_{jst-1}}{N_{jt-1}} \right) \frac{\bar{N}_{st}}{\bar{N}_{st-1}} \tag{A2}$$

A similar computation is applied to construct the Bartik occupation instrument.

Income. Average household income and Gini indexes by urban area come from the 2007 survey *Revenus fiscaux localisés des ménages* of the French Institute of Statistics (INSEE) and *Direction Générale des Finances Publiques* of the Ministry of Economics and Finance.

Land cover data. The French Ministry of Ecology, Sustainable Development, Transports, and Housing provides information about land cover for every municipality. The raw information comes from the 2006 Corine Project (Coordination of Information on the Environment by the European Environment Agency). We first aggregate the data at the urban area level and then compute the surface that is or can be developed as the sum of the two 1-digit categories: 1, 'artificial surfaces' and 2, 'agricultural areas' from which we exclude the 2-digit category 2.4, 'Heterogeneous agricultural areas'. Other excluded 1-digit categories compise: 8, 'forest and semi natural areas'; 4, 'wetlands'; and 5, 'waterbodies'.

Historical population data. We use a file containing some information on population by municipality for 27 censuses covering the 1831-1982 period (Guérin-Pace and Pumain, 1990). Over 1831-1910, the data contain only information on "urban municipalities" which are defined as municipalities with at least 2,500 inhabitants. The population of municipalities varies over time. Municipalities appear in the file when their population goes above the threshold and disappear from the file when their population goes below the threshold. Data are aggregated at the urban area level to construct our historical instruments.

Tourism data. These data at the municipality level are constructed by the French Institute of Statistics (INSEE) since 2002 from the census and a survey of hotels. It contains some information on the number of hotels depending on their quality (from zero star to four stars) and the number of rooms in these hotels. We construct our instruments, the number of hotel rooms and the share of 1-star rooms, by aggregating the data for 2006 at the urban area level.

Climate measures The original data come from the ATEAM European project as a high-resolution grid of cells of 10 minutes (18.6 km) per 10 minutes. These data came to us aggregated at the département level. The value of a climate variable for a département was computed as the average of the cells whose centroid is located in that département. The climate variables include the average monthly precipation (in mm), the temperature (in C), the cloudiness (in % time) and the potential evapotranspiration (in hPa) for January and July. We attribute to each municipality the value of its département. The value of an urban area is computed as the average of its municipalities, weighting by the area.

Soil variables We use the European Soil Database compiled by the European Soil Data Centre. The data originally come as a raster data file with cells of 1 km per 1 km. We aggregated it at the level of each urban area. We refer to Combes *et al.* (2010) for further description of these data.

Appendix B. FGLS and WLS estimators

In this appendix, we explain how we construct weighted least squares (WLS) and feasible general least squares (FGLS) estimators used in some second-stage regressions. The model is of the form:

$$P = X\varphi + \zeta + \eta \,, \tag{B1}$$

where P is a $J \times 1$ vector stacking the fixed effects capturing unit land prices at the centre, $\ln P_j(0)$, with J the number of urban areas, X is a $J \times K$ matrix stacking the observations for urban area variables (area, population, population growth, etc.), ζ is a $J \times 1$ vector of error terms supposed to be independently and identically distributed with variance σ^2, and η is a $J \times 1$ vector of sampling errors with known covariance matrix V.

It is possible to construct a consistent FGLS estimator of φ as:

$$\widehat{\varphi}_{FGLS} = \left(X'\widehat{\Omega}^{-1}X\right)^{-1} X'\widehat{\Omega}^{-1}P\,, \tag{B2}$$

where $\widehat{\Omega}$ is a consistent estimator of the covariance matrix of $\zeta + \eta$, $\Omega = \sigma^2 I + V$. To compute this estimator, we use an unbiased and consistent estimator of σ^2 which can be computed from the residuals of an OLS residuals of equation (B1) and denoted $\widehat{\zeta + \eta}$:

$$\widehat{\sigma}^2 = \frac{1}{N-K}\left[\widehat{\zeta + \eta}'\widehat{\zeta + \eta} - tr\left(M_X V\right)\right]\,, \tag{B3}$$

where $M_X = I - X\left(X'X\right)^{-1} X'$ is the projection orthogonally to X. We thus use $\widehat{\Omega} = \widehat{\sigma}^2 I + V$ in the computation of (B2). A consistent estimator of the covariance matrix of the FGLS estimator is:

$$\widehat{V}\left(\widehat{\varphi}_{FGLS}\right) = \left(X'\widehat{\Omega}^{-1}X\right)^{-1}\,. \tag{B4}$$

As the FGLS is said not to be always robust, we also compute a WLS estimator in line with Card and Krueger (1992), using the diagonal matrix of inverse of first-stage variances as weights, denoted Δ. The estimator is given by:

$$\widehat{\varphi}_{WLS} = \left(X'\Delta X\right)^{-1} X'\Delta P\,, \tag{B5}$$

with a consistent estimator of the covariance matrix given by:

$$\widehat{V}\left(\widehat{\varphi}_{WLS}\right) = \left(X'\Delta X\right)^{-1} X'\Delta\widehat{\Omega}_w\Delta X\left(X'\Delta X\right)^{-1}\,,$$

where $\widehat{\Omega}_w = \widehat{\sigma}^2_w I + V$ with $\widehat{\sigma}^2_w$ a consistent estimator of σ^2 based on the residuals of WLS denoted $\Delta^{1/2}\widehat{(\zeta + \eta)}$ and given by:

$$\widehat{\sigma}^2_w = \frac{1}{tr\left(\Delta^{1/2}M_{\Delta^{1/2}X}\Delta^{1/2}\right)}\left[\Delta^{1/2}\widehat{(\zeta+\eta)}'\Delta^{1/2}\widehat{(\zeta+\eta)} - tr\left(\Delta^{1/2}M_{\Delta^{1/2}X}\Delta^{1/2}V\right)\right]\,. \tag{B6}$$