The Costs of Agglomeration: Land Prices in French Cities

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

Gilles Duranton[*‡]
University of Toronto

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

Work in progress (do not cite): September 2011

ABSTRACT: We develop a new methodology to estimate the elasticity of urban costs with respect city population using unique French data. Our preferred estimate, which handles a number of estimation concerns, stands at 0.03. Our approach also yields a number of intermediate outputs of independent interest such as city specific distance gradients for land prices, the share of land in construction, and the elasticity of unit land prices with respect to city population. For the latter, our preferred estimate is 0.82.

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

JEL classification: R14, R21, R31

[*]We thank conference and seminar participants, Morris Davis, Sanghoon Lee, François Ortalo-Magné, and Henry Overman for their comments and suggestions.

[†]Aix-Marseille School of Economics, 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.

[‡]Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; website: http://individual.utoronto.ca/gilles/default.html). Also affiliated with the Centre for Economic Policy Research, the Rimini Centre for Economic Analysis, the Centre for Economic Performance and the Spatial Economic Research Centre at the London School of Economics.

[§]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 Centre for Economic Policy Research, PSE, and CREST.

1. Introduction

We develop a new methodology to estimate the elasticity of urban costs with respect city population using unique French data. Our preferred estimate, which handles a number of estimation concerns, stands at 0.03. Our approach also yields a number of intermediate outputs of independent interest such as city specific distance gradients for land prices, the share of land in construction, and the elasticity of unit land prices with respect to city population. For the latter, our preferred estimate is 0.82.

Reasonable estimates for urban costs and how they vary with city population are important for a number of reasons. Most crucially, 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 land prices and traffic jams in Central Paris, London, or Manhattan are for everyone to observe. There is nonetheless little systematic evidence about urban costs beyond these extreme examples. 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 larger than, but close to, the corresponding elasticity for agglomeration economies. That cities operate near aggregate constant returns to scale is suggestive of the fundamental tradeoff of spatial economics being unlikely to have much power in determining city sizes. In turn, this finding may be an important reason for why cities of vastly different sizes exist and prosper.

To estimate the elasticity of urban costs with respect to city population we 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 this elasticity. Our results suggest that these models provide surprisingly good approximations of the French urban reality after allowing for some amount of urban decentralisation as cities grow.

Tolley, Graves, and Gardner (1979), Thomas (1980), Richardson (1987), and Henderson (2002) are the main antecedents to our research. 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.[1] 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 no attention being paid so far to a number of identification issues, the three main innovations of this paper.

Urban costs take a variety of forms. In larger cities, land is more expensive; roads are more congested; and the bundle of amenities that these cities offer may differ. Some of these costs, like higher land costs, are monetary costs. Others, like the disutility from longer commutes or the fear of crime, are much harder to measure. Inspired by previous literature, we first develop a theoretical framework to show how we can estimate the elasticity of urban costs with respect to city

[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.

population using land price data and what will be included in this estimate. Our starting point is a notion of spatial equilibrium. Mobility within and between cities imply 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. The main result of our theoretical model is 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 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.

To estimate our second key quantity, the share of land in housing, our main source of data is a record of transactions for land parcels with a development or redevelopment permit which also contain construction costs as stated by the buyer. Such data 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. Using information about unit land prices and unit construction costs, we can estimate a cost function for housing and obtain the share of land in housing. We experiment with different approaches. They all yield comparable results and our preferred estimate for the share of land in housing is 0.18.

The estimation of our third fundamental quantity, the elasticity of unit land prices at the centre of each city with respect to city population, is more demanding. We proceed in two steps. First, we use information about the location of parcels in each city and other parcel characteristics from the same record of transactions for land parcels to estimate unit land prices at the centre of each city. Second, we regress these estimated (log) prices at the centre of each city on log city population to obtain an estimate of the elasticity of unit land prices at the centre of each city with respect to city population. Our preferred estimate for this third quantity is 0.82. Multiplying this population elasticity of unit land prices by the share of land in housing and by the share of housing in expenditure yields our preferred estimate for the elasticity of urban costs with respect to population of 0.033.

Estimating the elasticity of unit land prices in city centres with respect to city population brings up a number of challenges. Most notably, land prices and city population are simultaneously determined. A greater population in a city is expected to make land more expensive. But it may also be that cities where land is less expansive attract more population. To investigate this identification problem, we develop a number of instrumental variable strategies. The first relies on instruments derived from our theoretical model, where amenities determine city population but do not intervene directly in the determination of land prices. That is, provided a variable can be viewed as a pure consumption amenity that determines the attractiveness of a city but does not explain its productivity, it is a valid instrument for city population. We also use a number of other instruments external to the model. Following previous literature (e.g., Ciccone and Hall, 1996, Combes, Duranton, Gobillon, and Roux, 2010), we use long historical lags of population which predict current population well but are not obviously otherwise correlated with contemporaneous land prices after controlling for land characteristics. Finally, we also develop new instruments for city size building on the idea that the occupational mix of cities is also good predictor of their population size. Importantly, these IV strategies all yield comparable results even though they rely on different sources of variation.

Aside from the aforementioned literature on agglomeration and previous attempts at measuring urban costs, 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. Coulson (1991) and McMillen (1996) are two exceptions. They focus on distance gradients for land prices but they do so for only one city. There is also a literature that measures land values for a broader cross-section or urban (and sometimes rural) areas (Davis and Heathcote, 2007, Davis and Palumbo, 2008). We enrich this type of approach both by considering the internal geography of cities and by investigating the determinants of land prices, population in particular, at the city level.

The rest of this paper is organised as follows. Section 2 develops a theoretical framework to guide our empirical analysis. Section 3 discusses the data we use. Section 4 develops a methodology to estimate unit and aggregate land values in French urban areas. Section 5 provides estimates about land values at the centre of French urban areas. Section 6 estimates the elasticity of these land values with respect to population size. Section 7 turns to urban costs and their elasticity to population size. Finally, section 8 concludes.

2. Model

To inform our estimation strategy we first develop a simple model of a monocentric circular city in the spirit of Alonso (1964), Mills (1967), and Muth (1969), which we extend to allow for employment decentralisation and the choice of location between many cities. Our main goal is to derive a relationship between the rental price of land in cities and their population and other characteristics. The key equation in this respect is the accounting identity that integrates the residents over the entire city to obtain city population:

$$N = \int_0^{\overline{D}} N(D)\mathrm{d}D \;=\; \int_0^{\overline{D}} \frac{H(D)}{h(D)}\mathrm{d}D\,, \tag{1}$$

where $N(D)$ is the equilibrium level of population at distance D from the central business district (CBD), $H(D)$ is the supply of residential housing, $h(D)$ is individual housing consumption, and $\overline{D}$ denotes the urban fringe. In equation (1), we expect both the supply of residential development, $H(D)$, and individual housing demand, $h(D)$, to be affected by land rents at D.

Starting with housing demand, we assume that city residents must commute to the CBD to receive a wage W. The utility of a resident living at a distance D from the CBD is:

$$U(D) = \frac{M}{\beta^\beta(1-\beta)^{1-\beta}}\frac{h(D)^\beta\, x(D)^{1-\beta}}{v(D)}\,, \tag{2}$$

where M denotes the quality of amenities in the city, $x(D)$ is her consumption of a numéraire composite good, and $v(D)$ is the utility cost of commuting, which is increasing with distance.[2]

This resident maximises her utility (2) subject to the budget constraint: $Q(D)h(D) + x(D) = W$ where $Q(D)$ is the rental price of housing at a distance D from the CBD. The first-order condition for utility maximisation with respect to housing consumption implies:

$$h(D) = \frac{\beta W}{Q(D)} . \tag{3}$$

The first-order condition for utility maximisation with respect to the composite good leads to $x(D) = (1-\beta)W$. Substituting these two conditions into equation (2) yields the following indirect utility:

$$U(D) = \frac{M W}{[Q(D)]^{\beta} v(D)} . \tag{4}$$

In equilibrium, the rental price of housing adjusts so that residents are indifferent across all residential locations between the CBD and the urban fringe. Formally this implies $U(D) = U(0)$ for $0 \leq D \leq \overline{D}$. Using this spatial equilibrium condition into equation (4) yields the following distance gradient for the rental price of housing within the city:

$$Q(D) = Q(0) \left[\frac{v(0)}{v(D)}\right]^{1/\beta} . \tag{5}$$

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

$$H = BK^{1-\alpha} L^{\alpha} . \tag{6}$$

For land located at a distance D from the centre, 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 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 (6) yields housing supplied at distance D from the CBD:

$$H(D) = 2\,\pi\,\theta\, B \left[\frac{1-\alpha}{\alpha}\; \frac{R(D)}{r^K}\right]^{1-\alpha} D\,. \tag{7}$$

[2]According to Davis and Ortalo-Magné (2011), our specification for how housing consumption enters the utility function is empirically reasonable. Regarding commuting, urban models often introduce it as a monetary cost or a time cost whereas we model it as a utility cost. This follows a long tradition in transportation economics (Small and Verhoef, 2007). We further discuss the effect of this modelling choice below. In addition, we also assume a constant numéraire price for non-housing goods. According to Combes, Duranton, Gobillon, Roux, and Puga (2009) and Handbury and Weinstein (2010), the elasticity of average retail prices with respect to population is around 0.01 for France and the US, respectively. Handbury and Weinstein (2010) also show that after accounting for the fact that consumers in larger cities buy more expensive varieties and the amenities provided by stores, there are no significant differences in retail prices across cities of different population sizes. This is consistent with our specification.

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

$$Q(D) = \frac{1}{B}\left(\frac{r^K}{1-\alpha}\right)^{1-\alpha}\left[\frac{R(D)}{\alpha}\right]^{\alpha}. \tag{8}$$

Inserting equation (5) into (8), and substituting again into (8) valued at $D = 0$ yields the distance gradient for the rental price of land:

$$R(D) = \left(\frac{v(0)}{v(D)}\right)^{1/\alpha\beta} R(0)\,. \tag{9}$$

We can now return to equation (1) and replace $H(D)$ by its expression in (7) and $h(D)$ by its expression in (3), using (8) to substitute for $Q(D)$. In the resulting expression, we can replace $R(D)$ using the distance gradient given by (9) before inverting it to express the rental price of land at the CBD as a function of population:

$$R(0) = \frac{\alpha\beta\, W\, N}{2\pi\,\theta}\frac{1}{\int_0^{\overline{D}} D\left(\frac{v(0)}{v(D)}\right)^{1/\alpha\beta}\mathrm{d}D}\,. \tag{10}$$

An important feature of equation (10) is that the elasticity of rental price of land at the CBD with respect to city population is equal to one. That is, the rental price of land at the CBD is predicted to be proportional to city population. Such proportionality arises because population density at any location D, $H(D)/h(D)$, is proportional to the rental price of land, which in turn 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 and this preserves the proportionality with population to the rental price of land at the CBD. Of course, by (9), this proportionality extends to the rental price of land everywhere within the urban fringe.

Such proportionality between the rental price of land at the CBD and city population 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, 2011). The general intuition is the same as here. An increase in distance to the CBD leads to a proportional increase commuting costs and this should be offset by a lower expenditure on land for residents to remain indifferent across locations at the spatial equilibrium. 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. Slightly more formally, in equilibrium it must be that $R'(D)h(D)$ 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 (1) 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.[3]

However, 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

[3] With the caveat that proportionality is observed for the differential land rent, not total land rent as in the model developed here.

should lessen the effect of population 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

$$N = \overline{N}^{\gamma}, \tag{11}$$

with $0 < \gamma < 1$ while the rest of the metropolitan population works outside the CBD and lives outside the fringe $\overline{D}$. Equation (11) is obviously a reduced form. Modelling employment decentralisation from first principles is beyond the scope of this paper. 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). Here, we take as a given that metropolitan areas with greater population are more decentralised (as evidenced by Glaeser and Kahn, 2001, for the US). We provide empirical evidence to that effect below and show that (11) provides a parsimonious but powerful description of employment decentralisation within French metropolitan areas. A deeper exploration of why the distribution of employment within French metropolitan areas is so potently described by (11) 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 we expect production in the city to be subject to agglomeration economies. Larger cities are thus richer cities with stronger demand for housing. More formally, we can assume that the wage in a city of unit population is A and increases according to,

$$W = AN^{\sigma}, \tag{12}$$

with $\sigma > 0$. We also assume $\sigma < \alpha\beta/(1 - \alpha\beta)$ to satisfy a stability condition when we allow for mobility in and out of the city. 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 *et al.*, 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 (7). For cities built a long time ago, a lower elasticity of housing supply with respect to current rental prices of land would increase the population elasticity of the rental price of land. To see this, consider an extreme case where the existing stock of housing does not depend on distance: $H(D) = H$. Then, the exponent on N in equation (10) becomes $(1 + \sigma)/\alpha$.[4] Second, the price elasticity of housing may not be equal to one as indicated by equation (3). 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 one. This would, in turn, imply a higher population elasticity of the rental price of housing at the CBD in equation (10). Third, city population could generate a negative externality on commuting. As can be seen easily from (10),

[4]In equation (7), the term in $R(D)^{1-\alpha}$ is replaced by a constant which leaves only a term in $R(D)^{\alpha}$ at the denominator of the integral in equation (1). This term in $R(D)^{\alpha}$ comes from (8) and the free entry condition among builders. It remains within the integral because, even if the city is flat, new housing (re-)development occurs and (8) still applies. Zoning regulations that make the supply of housing more land intensive distort α but on their own they have otherwise no effect on the proportionately between the rental price of land and population.

a congestion term in N that multiplies $v(D)$ would directly lead to a higher exponent on N on the right-hand side of this expression.

In what follows we retain the specifications in (11) for decentralisation and (12) for agglomeration but do not enrich our model further. The first reason is that our two extensions provide a good empirical fit with the data. The second is that we could not find robust persuasive evidence about the last series of issues mentioned above. For instance increased congestion in larger cities should lead to steeper distance gradients for the rental price of land. As we show below, there is no evidence for this in the data we use.

Finally, to derive an empirical specification from equation (10), we need to impose some functional form assumptions for $v(D)$. For our main estimations, we assume:

$$v(D) = (1 + D)^{\tau} . \tag{13}$$

In this expression, we add one to the distance to the CBD to make sure that the utility of a resident at the CBD is well defined. This can be thought as 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.

Using (11), (12), and (13) into (10) and taking advantage of the fact that by choice of units $\overline{D}$ can be made large relative to one leads to

$$R(0) \approx C_1 \, \overline{N}^{(1+\sigma)\gamma} S^{-1+\frac{\tau}{2\alpha\beta}} , \tag{14}$$

where $S \equiv \pi\theta\overline{D}^2$ is the city land area and $C_1 \equiv (\alpha\beta - \tau/2)A \,/\, (\pi\,\theta)^{\frac{\tau}{2\alpha\beta}}$ is a constellation of parameters.

A natural alternative to the specification for commuting costs used in equation (13) would be $v(D) = e^{\tau\, D}$. This alternative functional form leads to a different specification for the distance gradient for the rental price of land where the log of the rental price of land is proportional to the linear distance to the CBD instead of being proportional to its log. As we show below, neither specification obviously dominates the other. For what we want to do this alternative specification also leads to a more complicated formulation of $R(0)$ relative to that in equation (14). In particular, it leads to non-linear regressions which are more difficult to handle.[5] In any case, we perform supplementary estimations using linear distance rather than log distance to assess the robustness of our empirical findings.

To allow for the urban fringe to be determined endogenously, we follow the literature and assume that, aside from housing, land can also be used for agriculture. Agricultural land rent around the city is $\underline{R}$. The urban fringe is thus such that $R(\overline{D}) = \underline{R}$. Using equation (9), we obtain:

$$\overline{D} = \left(\frac{R(0)}{\underline{R}} \right)^{\frac{\alpha\beta}{\tau}} - 1 . \tag{15}$$

Replacing into (14) and simplifying using again the fact that $R(0) >> \underline{R}$ leads to

$$R(0) \approx C_2 \, \overline{N}^{(1+\sigma)\gamma \frac{\tau}{2\alpha\beta}} , \tag{16}$$

[5]Inserting $v(D) = e^{\tau\, D}$ into equation (10) leads to an integral of the form $D\, e^{D}$ for which the solution takes the form $(D-1)e^{D}$.

where $C_2 \equiv [(\alpha\beta - \tau/2)A \,/\, (\pi\,\theta)]^{\frac{\tau}{2\alpha\beta}}\, \underline{R}^{1-\frac{\tau}{2\alpha\beta}}$ is a constellation of parameters.

The key difference between equations (14) and (16) is that in (14) land area is held constant whereas in (16) it is allowed to adjust. Hence, the exponent $(1+\sigma)\gamma$ on N in equation (14) 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 (16) corresponds to the same elasticity when the urban fringe is allowed to adjust. Since $\tau < \alpha\beta$ empirically, the former elasticity should be larger than the latter. 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 equation (14).

Finally, to allow for city population to be endogenous as well, we can assume free mobility and a reservation level of utility $\underline{U}$ outside the city.[6] This implies $\underline{U} = U(D) = U(0)$. Using (4), (8), (12), and (13) we obtain

$$N = \left[\left(\frac{r^K}{1-\alpha} \right)^{(1-\alpha)\beta} \left(\frac{R(0)}{\alpha} \right)^{\alpha\beta} \frac{\underline{U}}{MAB} \right]^{1/\sigma} . \tag{17}$$

Importantly, we note that amenities, M, and the productivity in housing, B, determine population in equation (17) but not play any direct role in (14) and (16) to determine $R(0)$.

Urban costs

We consider an exogenous increase in population and ask what is the monetary cost of this increase 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. 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 $Q(D)$ and ρ is its elasticity with respect to population in the metropolitan area, $\overline{N}$. Note that 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 utility change from an exogenous increase in population and thus include both the increase in urban costs and agglomeration economies. Because we want to isolate urban costs, we focus solely on the expenditure function.

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

$$e(Q(D),U(D)) = \frac{Q(D)^\beta}{M} U(D) = \frac{Q(0)^\beta}{M} U(0). \tag{18}$$

Using equation (8), it is easy to obtain ρ, the elasticity of expenditure with respect to population which is given by:

$$\rho = \alpha\,\beta\,\phi\,, \tag{19}$$

[6] Because only 77% of the French population live in urban areas, we assume that the reservation utility is determined by rural areas. 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 conceptually unpersuasive.

where ϕ is the elasticity of the rental price of land at the CBD with respect to population. This expression 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 housing in expenditure, β. In a city with fixed boundaries equation (14) indicates that $\phi = (1+\sigma)\gamma$. In a city where the urban fringe can adjust as the city grows we have $\phi = (1+\sigma)\gamma\tau/(2\alpha\beta)$ by (16).

Before turning to the data, one final issue must be discussed. Our model above assumes that amenities, M, remain constant as population grows. We can easily extend our model to allow amenities to change endogenously as cities grow in population by assuming $M(\overline{N}) = \underline{M}\,\overline{N}^{\nu}$ where $\underline{M}$ is the intrinsic 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 (18) as

$$\underline{M}\,\overline{N}^{\nu} = \frac{Q(0)^{\beta}}{W}\underline{U}. \tag{20}$$

Making use of (8) and (12) we find

$$\nu = \alpha\,\beta\,\phi - \sigma. \tag{21}$$

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.

3. Data

The data are extracted from the French Survey of Developable Land Prices (*Enquête sur le Prix des Terrains à Bâtir*, *EPTB*) which 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 the universe of all building permits for an individual house. Parcels are drawn randomly in each strata which corresponds to a group of municipalities (in France, there are around 36,000 municipalities). The sampling scheme was designed to survey about two third of all building permits. Some French regions paid for a larger sample. For 2008, there are initially 82,586 observations. They correspond to 61% of all building permits delivered this year.

For each observation we know the price, the municipality, some information about the landowner and how the parcel was acquired, and a number of parcel characteristics. The municipality identifier allows us to determine in which urban area the parcel is located.

The information about the transaction includes how the parcel was acquired (purchase, donation, inheritance, other), whether the parcel was acquired through an intermediary (a broker, a builder, other, or none), and some information about the house to be built including its expected cost. Among the parcel characteristics, we know its area, 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. Note that there is a non-trivial fraction of missing values for the road frontage (7.2% of the observations in our sample). As a consequence, we construct a dummy for missing road frontage.

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.

The two panels of figure 1 map mean unit land prices and population by French urban areas. To illustrate the content of the data within particular urban areas, the four panels of figure 2 plot prices per square metre and 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 exact definition for the measure of distance used in this figure is discussed below together with a number of alternatives. Tables 1 and 2 provides further descriptive statistics for all our main variables. 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.

We also use population, overall and sectoral employment, and occupation data that we extracted from the 1990, 1999, and 2006 population censuses at the French statistical institute (INSEE). Other data about municipalities such as their land area and some geographical characteristics come from the 1988 municipal inventory, also from INSEE. Soil data come from the European soil database. We refer to Combes *et al.* (2010) for further description of these data. Historical population data are also described in Combes *et al.* (2010). We also make use of tourism data from INSEE ('key figures' from their department of tourism) and climate measures from Mitchell, Carter, Jones, Hulme, and New (2004).

4. Methodology

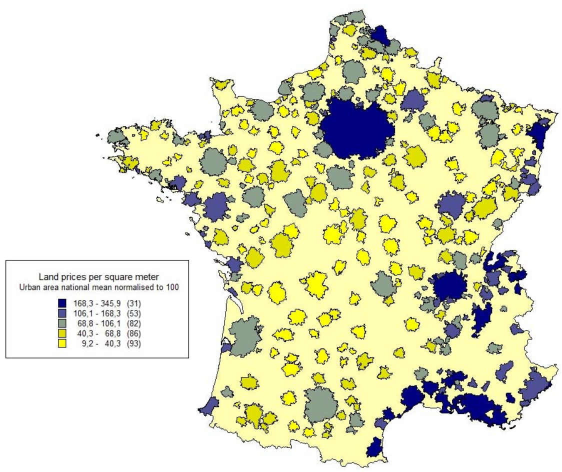
While our ultimate goal is to estimate the elasticity of urban costs with respect to population, ρ, we first focus on the estimation of the elasticity of the rental price of land, ϕ. As suggested by equation (19), our main quantity of interest ρ can be obtained by multiplying ϕ by the share of land in expenditure.

Our model suggests the following empirical strategy. Based on equations (9) and (13), we can first estimate

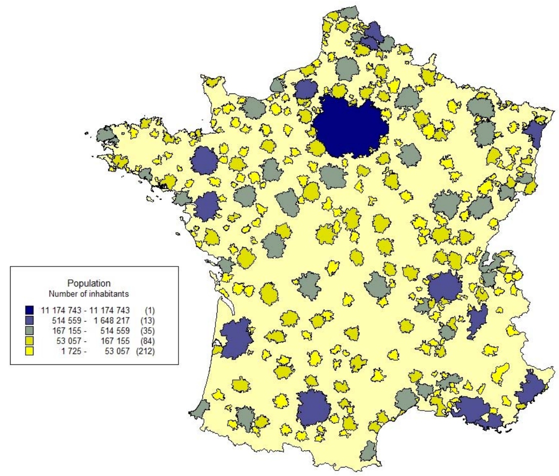
$$\ln P_i = \gamma_{j(i)} + \delta_{j(i)} \ln D_{k(i)} + T_i\, b + \epsilon_i \tag{22}$$

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 γ_j capture unit land prices at the centre of j. The coefficients δ_j capture the distance gradients of land prices when varying the (log)

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

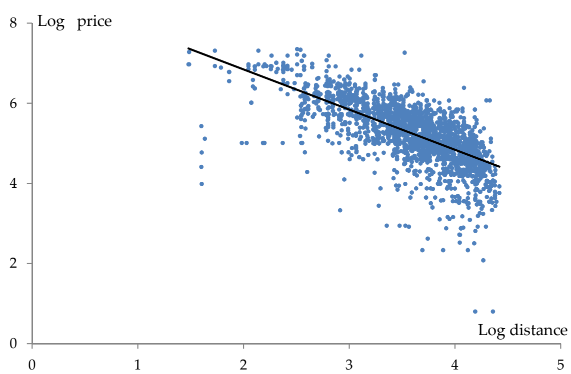


Panel (a): Mean land prices

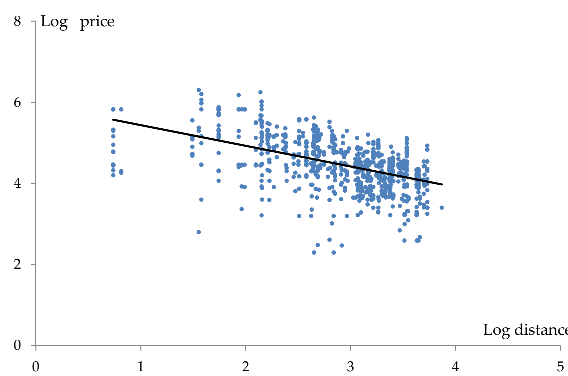


Panel (b): Population

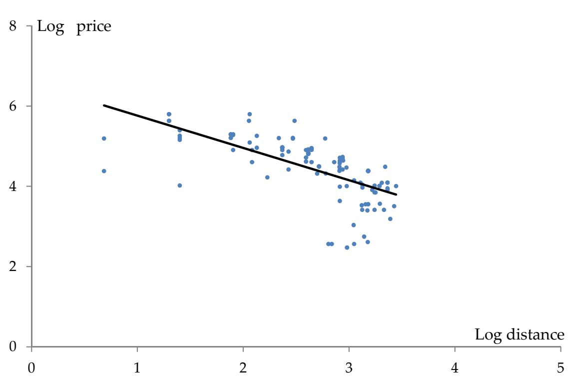
Figure 2: Land prices per m^2 and distance to their barycentre for four urban areas



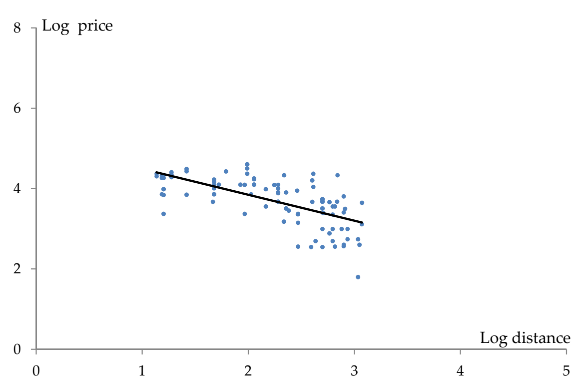
Panel (a): Paris



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)	1103	1200	477	838	2000
Road frontage / square-root 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 46697 weighted observations except row (3) (25854 and 43146).

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 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. The numbers in column 3 are for all French urban areas with population above 200,000 excluding Paris, Lyon, Lille, and Marseille.

distance D between the municipality where a parcel is located and the centre of the urban area it belongs to. Finally, T_i is a vector of parcel characteristics.

There are two differences between the theoretical specification that arises from equation (9) and the corresponding empirical implementation described by (22). The first is that we use land price instead of land rent data. We return to this issue below. The second 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.

We can also express the urban area fixed effects γ which correspond land price at the centre and the coefficient for the distance gradient for land prices δ as functions of the characteristics of their urban area:

$$\gamma_j = X_j\psi + \eta_j\,, \tag{23}$$

$$\delta_j = X_j\varphi + \kappa_j\,. \tag{24}$$

(Here for simplicity, the regressors are supposed to be the same in the two equations). Depending on the vector of urban area characteristics X and more specifically depending on whether X includes land area, regression (23) can be viewed as implementing equation (14) or equation (16) from the model.

Turning to equation (24), we have $\delta_j = -\frac{\tau_j}{\alpha\beta}$ according to equation (9). Our model above assumes no further dependence of the commuting cost parameter τ_j on the characteristics of urban area j. However, the possible dependence of the commuting cost parameter on urban area characteristics has implications regarding the estimation of (23) as made clear below. In addition, this is an issue of obvious empirical interest and policy relevance.

A multi-step approach

It is possible to estimate (22) by OLS to recover estimators of the urban area fixed effects denoted $\widehat{\gamma}_j$ and the urban area distance gradients for land prices denoted $\widehat{\delta}_j$. Equations (23) and (24) can be rewritten as:

$$\widehat{\gamma}_j = X_j\psi + \eta_j + \zeta_j \tag{25}$$

$$\widehat{\delta}_j = X_j\varphi + \kappa_j + \xi_j \tag{26}$$

where $\zeta_j = \widehat{\gamma}_j - \gamma_j$ and $\xi_j = \widehat{\delta}_j - \delta_j$ are sampling errors. The coefficients ψ and φ can then be estimated by OLS. Their standard errors can be recovered using the same type of FGLS procedure as in Combes, Duranton, and Gobillon (2008). This procedure is described in Appendix A.

This said, the parameters ψ and φ can also be estimated consistently in one step only. When injecting (23) and (24) into (22), we get:

$$\ln P_i = X_{j(i)}\psi + (X_{j(i)}\varphi)\ln D_{k(i)} + T_i b + (\eta_{j(i)} + \kappa_{j(i)}\ln D_{k(i)} + \epsilon_i) \tag{27}$$

and the whole set of parameters can be estimated by OLS. Note however that in that case, it is burdensome to compute the standard errors of the estimated parameters. This is because the error term is now $\eta_{j(i)} + \kappa_{j(i)} \ln D_{k(i)} + \epsilon_i$, and the covariance structure is less straightforward. We nonetheless explore with one-step estimations below.

Estimation issues: First step

While most of the important problems regarding our estimation pertain to the second step, it worth commenting briefly on a number of issues that arise when estimating our first step, equation (22) with OLS.

First, as already argued in the introduction using a record of transactions for land parcels with information about construction costs offer considerable advantage over records of property transactions for which disentangling between the value of land and the value of the property is a considerable challenge. One may worry that parcels sold with a building permit may form a highly selected sample of existing parcels in a urban area. To alleviate this concern, note first that 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. Moreover, French land use planning regulations imply that the transactions we observe cover a broad spectrum of parcels in terms of location, size, etc.[7]

A key first-step explanatory variable, distance is a second source of concern. First, there are several ways to measure empirically distance to the CBD, D. In addition, using different functional forms for the utility function (2) leads to different functional forms for the estimating equation (22) as already mentioned. To address these two issues, we experimented extensively with measures of distance and functional forms in the estimation of regression (22).

Our preferred distance variable measures the log of the great circle distance between the centroid of a parcel's municipality and the barycentre of the urban area to which it belongs where we weight all municipalities in an urban area by their employment. We prefer this measure to many alternatives because we believe it captures the concept of 'accessibility' that is core to our model better than measures of distances to specific, and possibly arbitrary, 'central points'. This said, our preferred measure of distance is highly correlated with the distance to the centroid of the municipality with most employment in the urban area (usually the core municipality). The correlation between the two is 0.92. Our preferred measure of distance is also highly correlated with a measure of distance to the densest municipality of the urban area. The correlation between the two is 0.90. Consistent with these high correlations we can easily replicate our results with these two alternative measures of distance.

We also compute a measure of distance that allows for urban areas to have two centres and use the log of the distance to the closest centre (where centres are again defined by density of employment). The correlation between this measure of distance and our preferred one is still high at 0.85. Finally, we also experimented with functional forms and considered a linear version of our

[7] French municipalities need to produce a planning and development plan ('plan local d'urbanisme') which is subject to national guidelines and requires central approval. Existing guidelines allow municipalities or groups of municipalities to expand their urban fringe but tend to 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). This implies that many parcels in the data are well inside the urban fringe, sometimes even close to the centre of the urban area.

preferred measure of barycentric distance instead of a log. The correlation between the two is also high at 0.80. We replicate our results with these alternative measures below.

Our use of municipal centroids for the location of parcels instead of their exact co-ordinates is of no practical importance either. French municipalities are tiny: An average French municipality corresponds to a circle of radius 2.2 kilometres.

We also use a number of parcel characteristics such as their size, shape, whether they are 'serviced' (i.e., have access to utilities), and the intermediaries involved in the transaction.[8] One might worry that some of these characteristics might be determined simultaneously with prices. For instance, with parcels that already have a building sitting them, builders may only incur demolition costs when these parcels are expensive, i.e. when they have good unobservable characteristics. While such concerns are certainly valid, they do not appear to affect our results. We can replicate our main findings without using parcel characteristics.

Estimation issues: Second step

When estimating regression (23), the first important issue to note is that population — our explanatory variable of interest — and land area may be simultaneously determined with land prices. For instance, one might expect a lower population in more expensive urban areas. Adding to this, the productivity of urban areas, A, enters the theoretical equation (14) but constitutes a missing variable in the corresponding regression (23). This missing variable is expected to be correlated with population and land area as suggested by equations (15) and (17). To deal with this endogeneity concern, we develop a number of instrumental variables strategies.

Our first set of instruments is motivated by our model. Urban population and land area both enter equation (14) and their equilibrium values are otherwise determined by equations (15) and (17). This suggests that any variable that explains population and/or land area but is not directly present in equation (14) constitutes a valid instrument according to our model. From (15) we can see that any term that determines either the rental price of land at the urban fringe ($\underline{R}$) or population (since $R(0)$ depends on N) or is a possible instrument. Regarding the determinants of the rental price of land at the urban fringe, soil and terrain characteristics do not constitute plausible instruments since such characteristics also proxy for the share of developed land, θ, which appears in equation (14).[9] Equation (17) is more promising due to the presence of M, the level of amenities. Consumption amenities will affect the demand for an urban area without affecting land prices directly.[10]

[8] To measure size, we use area and its square. To measure shape we use the ratio of the road frontage to the square root of the area. This relationship might not be linear since both extremely wide and extremely narrow parcels may be less desirable. However, most parcels are long and narrow as opposed to short and wide. We also know whether there is already a construction on the parcel and if it can be taken down. We also have dummies for transactions involving directly the builder, a broker, or other intermediaries.

[9] As we argue below, the soil and terrain characteristics we use affect land development (θ). However it is hard to make the case that they are unrelated to agricultural productivity ($\underline{R}$). For instance terrain ruggedness determines land development but also arguably agricultural productivity.

[10] Productivity in housing, B, also appears in (17) but not in (14). However, this may be an artefact of our model where construction matters only for residential purpose not for production. In addition it is hard to think of variables that affect 'strongly' (in an IV sense) the productivity of the construction industry without affecting productivity elsewhere (A).

As instruments, we then rely on two measures of weather (January temperatures and precipitations) and two measures of tourism (number of hotel rooms and share of one star rooms). We refer to these instruments as instruments 'internal' to the model. There is strong evidence about the importance of weather to explain location choices in the US (Rappaport, 2007) and Europe (Cheshire and Magrini, 2006). While we acknowledge that weather may affect productivity directly, we note that this effect is likely to be of limited importance since January weather in France is rarely bad enough to disrupt economic activity. 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 also use instruments which are not explicitly considered by our model. We refer to these instruments as 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, a measure of market potential also from 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 in production which may have attracted people in 1831 should no longer be present. Otherwise they should be part of contemporaneous productivity A of urban areas and would 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. This said, there are nonetheless aspects of past natural advantages that may persist today such as proximity to large markets or direct access to the sea. For this reason, we include these variables as controls to preclude any further correlation between our historical instruments and contemporaneous population. See Combes *et al.* (2010) for further discussion.

Our third group of 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. 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. These ideas suggest the following instrument. For each four-digit occupation in the French standard occupational classification, we compute mean urban area employment. Then for each urban area, we compute our instrument by interacting the local share of employment of an occupations with mean employment at the urban area level for this occupation before summing across occupations. Simply put, an urban area with a high proportion of bankers will be predicted to be large while an urban area with a high proportion of blue-collar occupations will be predicted to be small.[11]

[11] We experimented with sectors as well as occupations. However, only the instruments derived from the occupational structure are strong enough for our purpose.

While the occupational structure of urban areas provides strong predictors of their population, the validity of our 'Henderson' instruments relies on the occupational structure of urban areas affecting land prices only through population. In particular, these instruments should be uncorrelated with urban area productivity. To some extent, we expect the occupational structure of urban areas to be driven by natural advantage. Coastal urban areas may have a productive advantage and employ a greater proportion of workers in ship building. That suggests again using the geographical characteristics of urban areas to preclude these correlations.

It is important to note that our instruments rely on different sources of variation in the data: weather, tourism activity, long population lags, and occupational structure. While these instruments are not completely orthogonal (e.g., warm January is correlated with tourism activity), we take these instruments to be sufficiently different for overidentification tests to be meaningful. For instance the correlation between January temperature and the other instruments we use below in table 7 is 0.20. The correlation between 1831 market potential and the other instruments is below 0.30. The share of one star hotels has a correlation 0.30 with the other instruments except the 1999 Henderson instrument, etc. Getting the same answer from different instruments would be reassuring since it is extremely unlikely that they are all correlated with land prices in the same way after conditioning out their effects on population.

The second important issue to note when estimating regression (23), is that we use an urban area fixed effect estimated from land price data and not from land rents. Land prices can be viewed theoretically as the net discounted value of all future land rents where growth in the rental price of land differs across urban areas. Future growth in land rents is unknown. However, we know contemporaneous population growth and we can use the fact that population growth is highly serially correlated. This suggests using contemporaneous population growth (1999 to 2006) as a control in our second step regression.

As made clear by our results below, population growth is an important determinants 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. Alternatively, there may be missing variables that determine both population growth and land prices in urban areas. A first response to this issue is to note that population growth is only weakly correlated with land area and population so that introducing population growth in the regression does not affect the magnitude of the other coefficients. We provide evidence to that extent below. A second response is to find appropriate instruments for population growth. Following Bartik (1991) it is standard practice to interact the initial composition of economic activity of the urban area by sector with the growth of those sectors at the national level during the period.

A third issue is that the distance gradient of the land price estimated in the first step, $\tau/(\alpha\beta)$, enters equation (14) and constitutes a missing variable in (23). We show below that distance gradients by urban areas do not appear to be systematically related to our two variables of interest, population and land area.

Table 3: Summary statistics from the first step estimation regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parcel log-area	-	-	-	-	-	1.23^a	-	1.20^a
Parcel squared log-area	-	-	-	-	-	-0.14^a	-	-0.14^a
Other parcel charac. signif. (max=8)	8	-	-	-	8	6	-	6
Mean urban area fixed effect	-	4.3	4.9	5.2	4.7	2.9	5.1	3.0
First decile	-	3.3	3.9	3.9	3.4	1.8	4.0	2.0
Last decile	-	5.1	5.9	6.3	6.0	4.1	6.2	4.1
% above mean (signif.)	-	48	48	37	35	24	47	23
% below mean (signif.)	-	39	42	42	44	33	44	30
Log-distance effect	-	-	-0.28^a	-	-	-	-0.39^a	-
Mean distance effect by urban area	-	-	-	-0.35	-0.33	-0.24	-	-0.26
First decile	-	-	-	-0.68	-0.59	-0.49	-	-0.44
Last decile	-	-	-	-0.05	-0.02	-0.01	-	-0.01
% above mean (signif.)	-	-	-	25	26	28	-	22
% below mean (signif.)	-	-	-	36	34	32	-	31
R^2	0.10	0.52	0.57	0.62	0.66	0.81	0.59	0.81
Observations	27,850	27,850	27,850	27,850	27,850	27,850	27,850	27,850

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 (3) and (6) but use log distance to the closest centre (among the two). The eight other parcel characteristics are those listed in rows 4-10 of table 1 and a dummy for missing information for road frontage.

5. Land values in French urban areas

Table 3 summarises results of the first step of the estimation, which corresponds to regression (22). Column (1) uses only eight parcel characteristics to explain their price per square metre. These characteristics are those listed in rows 4-10 of table 1 and a dummy for missing information for road frontage. All 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 is possibly explained by the fact that an existing construction can be 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. Finally, parcels sold by real estate agencies, builders, or other intermediaries are also significantly more expensive. Real estate professionals are likely to specialise in the sale of more expensive parcels. Altogether, parcel characteristics explain 10% of the variance.

Column (2) corresponds to a specification where only 345 urban area fixed-effects are introduced. A large share of these effects are significantly different from the mean urban area effect, with 48% being above and 39% below. This feature 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, with a mean effect at 4.3. Finally, with an R^2 of 52%, the 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 area effects in explaining land prices.

Column (3) adds the log distance to the urban area barycentre as an explanatory variable. The coefficient on this variable, the distance gradient for land prices, is very significant and is estimated at -0.28. 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 increases to 57%, which confirms that location within an urban area matters.

In column (4), we allow for the distance gradient for land prices to vary across urban areas. The R^2 increases further to 62%. We find that 61% of the distance gradients for land prices are significantly different from the mean of -0.35. We also note that considering distance gradients by urban area only slightly decreases the differences across urban areas in terms of their fixed effect. These results are barely affected when controlling for parcels characteristics in column (5). The R^2 increases further to 66%. In the complete regression of column (6) that also considers land area and its square, 81% of the variance in land prices is now explained by the model. This regression is our preferred first-step specification and we use its output for our main second-step estimations below. A side result of our estimation here is that the overall price of land parcels 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. Finally, we also note that controlling for all parcels characteristics, including their area, reduces the variation among urban area fixed effects and among the distance gradients for land prices.

Columns (7) and (8) replicate columns (3) and (6) but use distance to the closest of two centres rather than barycentric distance to measure parcel location. The results are very close to those obtained in columns (3) and (6) respectively except for, unsurprisingly, slightly steeper distance gradients for land prices. We note that the R^2 in column (6) and (8) are the same suggesting there is no obvious gain in allowing for multiple centres.

In results not reported here, we also ran the specifications of columns (2) to (8) with linear distance rather than log distance. The results are similar to those in table 3. If we perform a more detailed comparison for our preferred regression in column (6), we find that when using linear distance the R^2 and the coefficients on parcel characteristics are essentially the same. The coefficients for the distance gradients for land prices are of course different. Their mean of -0.025 suggests smaller effects of distance on prices when using linear distance instead of log distance. These smaller effects get reflected in slightly less variation in the urban area fixed effects. Only 32% of these are significantly different from the mean when using linear distance instead of 57% with log distance.[12]

[12]The greater amount of variation when estimating the first step with distance in log could be an indication of more noisy estimates. This appears not to be the case. The R^2 of the second stage regressions are similar when the first stage uses log distance instead of linear distance.

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

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) FGLS	(7) OLS	(8) 1 OLS
Population	0.371^a	0.672^a	0.642^a	0.621^a	0.463^a	0.381^a	0.397^a	0.454^a
	(0.032)	(0.049)	(0.047)	(0.051)	(0.058)	(0.067)	(0.056)	(0.050)
Land area	-	-0.404^a	-0.387^a	-0.356^a	-0.379^a	-0.402^a	-0.297^a	-0.351^a
		(0.054)	(0.051)	(0.055)	(0.058)	(0.058)	(0.052)	(0.052)
Population growth	-	-	3.615^a	3.099^a	3.006^a	4.001^a	3.262^a	3.257^a
			(0.681)	(0.723)	(0.739)	(0.772)	(0.623)	(0.536)
Geography	no	no	no	some	all	all	all	all
R^2	0.32	0.44	0.49	0.50	0.54	0.55	0.50	0.73
Observations	285	285	285	285	285	285	342	27,850

Notes: The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions include a constant. Standard errors clustered by urban area in column (8).
Some amenities: dummy for being on the sea, dummy for mountain or forest, and a peripherality index (employment-weighted mean distance to all other urban areas).
All amenities: ruggedness, two dummies for the erodibility of soils (high and low vs. intermediate), two dummies for hydrogeological classes, and two dummies for classes of soil (unconsolidated deposits and aeolian deposits).

6. 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 in the first step as dependent variables. These fixed effects are the empirical counterparts to the rental price of land at the CBD in the theoretical model. We regress them against a number of urban area characteristics, especially population. We start with OLS results before turning to our main IV results. The end of this section provides a number of further robustness checks.

OLS *and* FGLS *results*

Table 4 present results for OLS and FGLS second-step estimations where we regress the urban area fixed effects on urban area characteristics. Column (1) uses only log population as explanatory variable for 285 French urban areas. This sample of urban areas is smaller than the full sample of 345 urban areas we use above Because some of the instruments we use below are missing for 57 urban areas and we have three urban areas with poorly estimated fixed effects. This restricts our IV results to 285 urban areas. In most cases, urban areas with missing instruments are small with few observations. For the sake of comparison, we run most of our OLS regressions on this sample. because it includes only population, this regression can be thought of as a rudimentary estimation of equation (16). We find that population explains about a third of the variance among first-step urban area fixed effects and the estimated elasticity, at 0.371, is highly significant. For now, this regression simply confirms a strong association between land prices at the centre of urban areas and population size after controlling for parcels characteristics and relative location within the urban area at the first step.

Columns (2) adds log land area and corresponds to the most basic specification of equation (14). Adding land area makes the association between land prices and population even stronger with a coefficient of 0.672. On the other hand, land prices are negatively associated with land area. These signs are consistent with an interpretation in terms of demand (population) and supply (area) as described by our model. These two variables account for 44% of the variance of urban area fixed effects. Column (3) enriches this specification further to control for population growth over 1999-2006. As argued above this helps bridge the gap between land prices available in the data and land rents used in the theoretical model. We expect higher contemporaneous growth, which is likely correlated with higher future growth, to imply higher land prices. We find that adding population growth raises the R^2 even further to 49% but leaves the coefficients on population and land area unchanged. The coefficient on population growth is positive and highly significant. Column (4) further adds some geographical characteristics that could affect the productivity of urban area such as a dummy for being on the sea, a dummy for mountains or forests, and a peripherality index that captures how far an urban area is relative to the others. These characteristics leave the results unchanged.

Column (5) is our most complete specification which adds another seven variables to control for soil and terrain characteristics. These controls are useful because they help us proxy further for the fraction of land that can be developed in an urban area, θ.[13] This term appears in equation (14) but is not directly observed in the data. We find that adding soil and terrain characteristics lowers the coefficient on urban area population to 0.463 and raises the R^2 to 0.54. This suggests that soil and terrain characteristics that limit development such as a rugged environment are associated with a higher population. We use our OLS estimates for this column as our benchmark to compare to our IV results below. Column (6) replicates column (5) but uses the FGLS estimator described in Appendix A to account for the fact that urban area fixed effects are estimated with some measurement error. This has only small effects on the coefficients on population, land area, and population growth. Column (7) also replicates column (5) but uses a larger sample of 342 urban areas instead of 285. This leads to slightly lower coefficients on population and land area relative to column (5). Finally, column (8) replicates again the estimation of column (5) but estimates everything in one step following the specification of equation (27) and uses our original 27,850 land transactions. This leads to results very similar to those of column (5).

IV *results*

Table 5 reports results for a series of IV regressions using four internal instruments for population and land area. Column (1) uses the number of hotel rooms and the share of one star room to

[13] These seven variables are all related to land use or regulations that will limit development. Ruggedness obviously limits development. Our two dummies for erodibility (low and high) determine land that will be used particularly intensively or be protected. Our two dummies for hydrogeological class capture the accessibility of water and risks associated with water (flood, and landslide) and thus land use. Finally, the two dummies for soils made of unconsolidated deposits and aeolian deposits (such as loess) will capture more fragile soils for which development will be more limited.

Table 5: Unit land values at the centre, IV regressions with internal instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.775^a	0.836^a	1.012^a	0.696^a	0.945^a	0.928^a	0.467^a	1.106^a
	(0.236)	(0.136)	(0.161)	(0.159)	(0.176)	(0.208)	(0.058)	(0.193)
Land Area	-0.590^b	-0.666^a	-0.878^a	-0.490^a	-0.678^a	-0.662^b	-0.372^a	-1.007^a
	(0.280)	(0.175)	(0.183)	(0.182)	(0.202)	(0.245)	(0.058)	(0.218)
Population Growth	3.080^a	3.051^a	2.974^a	-	3.153^a	3.159^a	5.067^a	8.297^a
	(0.767)	(0.780)	(0.850)	-	(0.774)	(0.766)	(1.850)	(2.810)
Geography	all	all	all	all	some	no	all	all
Overidentification p-value	-	0.74	0.29	0.74	0.82	0.33	0.87	0.45
First-stage statistic	6.0	11.5	13.2	10.3	8.1	5.0	25.0	4.4
Observations	285	285	285	285	285	285	285	285
Number of hotel rooms	Y	Y	Y	Y	Y	Y	N	Y
Share of 1-star rooms	Y	Y	Y	Y	Y	Y	N	Y
Temperature in January	N	Y	N	Y	Y	Y	N	Y
Precipitations in January	N	N	Y	N	N	N	N	Y
Bartik Industry 1990	N	N	N	N	N	N	Y	N
Bartik Industry 1999	N	N	N	N	N	N	Y	Y
1st Shea part. R^2, population	0.06	0.15	0.18	0.14	0.09	0.06	-	0.16
1st part. Fisher, population	118.7	79.0	80.1	78.7	251.1	219.5	-	50.9
1st Shea part. R^2, area	0.04	0.12	0.13	0.11	0.08	0.05	-	0.12
1st part. Fisher, area	64.8	51.5	50.6	48.2	165.1	131.5	-	32.3
1st Shea part. R^2, pop. gr.	-	-	-	-	-	-	0.16	0.12
1st part. Fisher, pop. gr.	-	-	-	-	-	-	25.0	11.3

Notes: 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.

instrument for population and land area. Column (2) adds January temperature to the instrument sets. Column (3) uses instead January precipitations. Column (4) uses the same set of instruments as column (2) but drops population growth from its controls. Column (5) also replicates the instrumentation strategy of column (2) but drops soil and terrain characteristics from its controls. Column (6) also drops the other geographical characteristics. Column (7) no longer instruments population and land area but instead instruments population growth using Bartik instruments which rely on the 1999 and 1990 sectoral structures of urban areas and national growth of sectors to predict population growth. Finally, column (8) instruments population, land area, and population growth using our four amenity instruments and the Bartik instrument for employment growth in 1999.

Several results are worth discussing. First, we note that in several instance such as columns (1), (6), and (8), our instruments have a low first stage statistic between 4.4 and 6.0. In columns (2), (3), (4), and (7), the first stage statistic is nonetheless above 10. To minimise the problems caused by weak instruments, we perform all our IV estimation using limited information maximum likelihood (LIML) instead of two-stage least squares. Because we instrument two endogenous

Table 6: Unit land values at the centre, IV regressions with external instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.865^a	0.910^a	0.844^a	0.888^a	0.951^a	0.988^a	0.487^a	0.802^a
	(0.167)	(0.179)	(0.110)	(0.117)	(0.095)	(0.155)	(0.079)	(0.111)
Land Area	-0.798^a	-0.760^a	-0.775^a	-0.871^a	-0.694^a	-0.755^b	-0.336^a	-0.625^a
	(0.210)	(0.220)	(0.158)	(0.166)	(0.114)	(0.202)	(0.079)	(0.148)
Population Growth	2.914^a	3.012^a	2.920^a	-	3.183^a	3.110^a	14.575^a	5.978^b
	(0.810)	(0.812)	(0.802)	-	(0.773)	(0.761)	(4.180)	(2.665)
Geography	all	all	all	all	some	no	all	all
Overidentification p-value	–	0.05	0.86	0.50	0.77	0.49	0.87	0.26
First-stage statistic	13.1	9.1	15.8	16.2	32.8	7.9	8.1	5.4
Observations	285	285	285	285	285	285	285	285
Urban Population 1831	Y	Y	Y	Y	Y	Y	N	Y
Market Potential 1831	Y	Y	Y	Y	Y	Y	N	Y
Urban density 1881	N	Y	N	N	N	N	N	Y
Henderson Occupations 1990	N	N	N	N	N	N	Y	Y
Henderson Occupations 1999	N	N	Y	Y	Y	Y	Y	Y
1st Shea part. R^2, population	0.14	0.14	0.32	0.32	0.33	0.11	-	0.32
1st part. Fisher, population	65.6	54.2	68.9	69.1	266.1	263.9	-	48.6
1st Shea part. R^2, area	0.09	0.09	0.15	0.16	0.26	0.08	-	0.18
1st part. Fisher, area	34.5	29.7	15.8	16.3	101.5	96.6	-	11.0
1st Shea part. R^2, pop. growth	-	-	-	-	-	-	0.06	0.10
1st part. Fisher, pop. growth	-	-	-	-	-	-	1.58	6.1

Notes: 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.

variables, population and land area, one might worry that our instruments may be particularly weak for one variable. The measures of partial R^2 at the bottom of the table show that this is not the case. Second and most importantly, the coefficients on population and land area both increase in magnitude when we instrument for these two variables. While our preferred OLS estimate for the population elasticity of land prices at the centre is 0.463, the IV estimates vary between 0.696 and 1.106. Whenever the IV regression is overidentified, the overidentification test is always easily passed. This indicates that the weather and tourism instruments yield answers that are not statistically different. Third, we find that the coefficient on population growth does not change when instrumenting for population and land area. In addition, the results for population and land area do not depend on the inclusion of population growth in the regression. At the same time, the coefficient on population increases a lot when instrumenting for this variable. The differences between OLS and IV for population and population growth are consistent with a situation where high land prices limit the arrival of new residents to urban areas.

Table 6 reports results for a series of IV regressions now using five external instruments for population and land area. This table follows roughly the same pattern as table 5. Column (1) uses urban population in 1831 and market potential for the same year. This last variable is the sum of all other urban areas' population in 1831 discounted by distance. It predicts current population since urban with a higher market potential in 1831 are larger today.[14] Column (2) adds urban density in 1881. Column (3) uses instead the Henderson instrument for 1999. Column (4) uses the same set of instruments as column (3) but drops population growth from its controls. Column (5) also replicates the instrumentation strategy of column (3) but no longer uses soil and terrain characteristics as controls. Column (6) also drops the other geographical characteristics. Column (7) no longer instruments population and land area but instead instruments population growth using the Henderson instrument for 1990 and for 1999 using the idea that past changes in the occupational structure of cities can predict their growth while at the same time not be otherwise directly related to land prices. Finally, column (8) instruments population, land area, and population growth using our all five external instruments.

The results are very similar to those of table 5 despite using a very different set of instruments. When instrumenting for population using these external instruments, the elasticity of land prices with respect to population is between 0.802 and 0.988 versus a range from 0.696 to 1.106 with the external instruments in table 5. This tighter range of coefficients in table 6 is arguably due to the somewhat greater strength of our external instruments relative to the internal instruments for population and land area. Our use of changes in the occupational structure to instrument for growth also leads to much higher coefficients on population growth as in table 5 even though these instruments are weaker than the Bartik instrument of table 5.

Finally, table 7 reports results for a series of IV regressions now using both internal and external instruments for population and land area. Columns (1) to (6) use various combinations of instruments for population and land area keeping a complete set of controls. The coefficient on population now fluctuates between 0.809 and 0.899. This range is even narrower than previously. This is due to the fact that the mixing internal and external instruments leads to stronger sets of instruments. The range for the coefficients on land area is equally narrow, from -0.772 to -0.662. As previously the coefficient on (uninstrumented) population growth remains extremely stable at around three. Columns (7) and (8) also instrument for population growth. Column (8) which uses a broad variety of instruments is our preferred specification with coefficients of 0.821 for population, -0.750 for land area, and 4.835 for population growth.

Further robustness checks

We now confirm our results further by running a number of robustness checks related to the use of different first step specifications, using linear instead of log distance in the first step, checking that the distance gradients for land prices are not systematically related to our main variables, and finally showing that estimating equation (16) (without using land area) yields estimates which are consistent with those of (14).

[14]Interestingly the correlation between 1831 urban population and 1831 (external) market access is negative.

Table 7: Unit land values at the centre, IV regressions with internal and external instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.893^a	0.833^a	0.834^a	0.812^a	0.816^a	0.879^a	0.781^a	0.821^a
	(0.167)	(0.126)	(0.102)	(0.120)	(0.110)	(0.125)	(0.107)	(0.103)
Land Area	-0.779^a	-0.754^a	-0.755^a	-0.676^a	-0.680^a	-0.732^a	-0.648^a	-0.750^a
	(0.209)	(0.147)	(0.134)	(0.152)	(0.133)	(0.143)	(0.132)	(0.139)
Population Growth	2.970^a	2.935^a	2.935^a	3.012^a	3.010^a	3.014^a	4.900^a	4.835^a
	(0.808)	(0.794)	(0.794)	(0.776)	(0.776)	(0.796)	(1.631)	(1.677)
Geography	all	all	all	all	all	all	all	all
Overidentification p-value	0.47	0.76	0.95	0.52	0.81	0.96	0.22	0.16
First-stage statistic	8.9	19.0	17.1	16.8	17.5	15.0	11.6	11.2
Observations	285	285	285	285	285	285	285	285
Urban Population 1831	Y	Y	Y	Y	Y	N	Y	Y
Market Potential 1831	Y	Y	Y	N	N	Y	N	Y
Number of hotels	Y	N	N	Y	Y	Y	Y	N
Share of 1-star hotels	N	N	N	N	N	Y	N	N
Temperature in January	N	Y	Y	N	Y	Y	Y	Y
Henderson Occupations 1999	N	N	Y	Y	Y	N	Y	Y
Bartik Industry 1999	N	N	N	N	N	N	Y	Y
1st Shea part. R^2, population	0.15	0.24	0.37	0.26	0.31	0.25	0.33	0.38
1st part. Fisher, population	108.6	44.5	52.3	118.8	89.5	59.4	71.6	41.9
1st Shea part. R^2, area	0.9	0.17	0.21	0.16	0.21	0.19	0.22	0.21
1st part. Fisher, area	46.8	28.2	13.3	29.1	21.6	36.8	17.5	10.7
1st Shea part. R^2, pop. growth	-	-	-	-	-	-	0.23	0.23
1st part. Fisher, pop. growth	-	-	-	-	-	-	13.3	12.3

Notes: 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.

To show that the results do not depend on the details of the first step, table 8 replicates our benchmark OLS estimation of column (5) of table 4 using a number of alternatives. Column (1) performs a one-step estimation that corresponds to this benchmark using our eight parcel characteristics as first step controls. Columns (2) to (8) replicate this benchmark OLS estimation using the first-step output from columns (2) to (8) of table 3. We can see that our preferred first-stage specification of column (6) of table 3 leads to a lower coefficient for population relative to less complete first step specifications. This is because many of our first-step controls that are associated with high prices are also correlated with the population size of urban areas. We can also see that, again, the difference between one- and two-step estimations is minimal.

To show that our results are not affected by our use of log distance in the first step of the estimation, table 9 reports results for a number of OLS and IV second-step regressions for which we use linear distance in the first step (but otherwise retain our preferred first-step estimation). Column (1) is a simple OLS regression that includes only population. It corresponds to column

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	0.708^a	0.708^a	0.765^a	0.604^a	0.577^a	0.463^a	0.746^a	0.487^a
	(0.050)	(0.051)	(0.050)	(0.075)	(0.073)	(0.058)	(0.051)	(0.054)
Land area	-0.672^a	-0.634^a	-0.527^a	-0.437^a	-0.456^a	-0.379^a	-0.457^a	-0.371^a
	(0.053)	(0.050)	(0.050)	(0.074)	(0.072)	(0.058)	(0.050)	(0.053)
Population growth	3.568^a	4.123^a	3.884^a	3.002^a	2.709^a	3.006^a	4.036^a	3.426^a
	(0.518)	(0.641)	(0.635)	(0.948)	(0.928)	(0.739)	(0.645)	(0.679)
Geography	all	all	all	all	all	all	all	all
R^2	0.45	0.63	0.67	0.52	0.53	0.54	0.67	0.56
Observations	27850	285	285	285	285	285	285	285

Notes: The superscripts *a*, *b*, and *c* indicate significance at 1%, 5%, and 10% respectively. Standard errors between brackets. All regressions include a constant. OLS estimates corresponding to the first-step specifications reported in table 3. Column (1) corresponds to a single-step regression including eight parcel characteristics, all geographical characteristics, population, population growth, and land area. with standard errors clustered by MSA.

(1) of table 4. Column (2) further includes land area and population growth. It corresponds to column (3) of table 4. Column (3) further includes all our geographical characteristics and thus corresponds to our benchmark estimation of column (5) of table 4. We can immediately see that while the coefficient on population is slightly lower when using linear distance in column (1) and (2), the results of column (3) are virtually identical to those obtained when using log distance in the first step of the estimation.

Column (4) of table 9 corresponds to column (2) of table 5 with internal instruments and yields very similar results. Column (5) of table 9 corresponds to column (3) of table 6 with external instruments and yields again very similar results. Column (6) of table 9 corresponds to column (2) of table 7 and mixes both internal and external instruments just like column (7) which corresponds to column (6) of table 7. In both cases using linear or log distance in the step leads to results which are statistically the same. Finally, column (8) corresponds to column (8) of table 7 and instruments for population, land area, and population growth. Again, using linear instead of log distance makes no noticeable difference.

Next, recall that larger cities might experience a systematically different distance gradient for the rental price of their land. In such case, looking at how the price of land at the centre changes with population is no longer sufficient to assess urban costs. Table 3 documents that the distance gradient of the rental price of land differs across cities. Recall that our preferred first-step specification in column (6) of table 3 implies that the mean elasticity of unit land prices with respect to distance to the barycentre of an urban area is -0.24, where 28% of urban areas have a significantly higher elasticity and 32% a significantly lower elasticity. However, this variation in the distance gradient for land prices does not appear to be correlated with our key variables of interest, population and land area.

To show this, table 10 reports results corresponding to the estimation of equation (24) using the distance gradient for land prices estimated in the first stage as dependent variable instead

Table 9: Unit land values at the centre, effect of linear distance in the first stage, OLS and IV regressions

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV	(8) IV
Population	0.279^a	0.555^a	0.459^a	0.750^a	0.750^a	0.809^a	0.815^a	0.770^a
	(0.028)	(0.039)	(0.049)	(0.131)	(0.097)	(0.109)	(0.109)	(0.088)
Land Area	-	-0.397^a	-0.364^a	-0.589^a	-0.645^a	-0.705^a	-0.616^a	-0.681^a
	-	(0.042)	(0.049)	(0.148)	(0.130)	(0.128)	(0.155)	(0.120)
Population Growth	-	3.880^a	3.335^a	3.350^a	3.279^a	3.265^a	3.312^a	4.651^a
	-	(0.562)	(0.627)	(0.655)	(0.665)	(0.683)	(0.680)	(1.426)
Geography	no	no	all	all	all	all	all	all
R^2	0.26	0.51	0.54	-	-	-	-	-
Overidentification p-value	-	-	-	0.60	0.26	0.58	0.75	0.079
First-stage statistic	-	-	-	11.4	16.9	18.8	14.8	11.0
Observations	285	285	285	285	285	285	285	285
Urban Population 1831	-	-	-	N	Y	Y	N	Y
Market Potential 1831	-	-	-	N	Y	Y	Y	Y
Number of hotel rooms	-	-	-	Y	N	N	N	N
Share of 1-star rooms	-	-	-	Y	N	N	N	N
Number of hotels	-	-	-	N	N	N	Y	N
Share of 1-star hotels	-	-	-	N	N	N	Y	N
Temperature in January	-	-	-	Y	N	Y	Y	Y
Henderson Occupations 1999	-	-	-	N	Y	N	N	Y
Bartik Industry 1999	-	-	-	N	N	N	N	Y
1st Shea part. R^2, population	-	-	-	0.15	0.29	0.24	0.24	0.38
1st part. Fisher, population	-	-	-	81.6	59.5	45.4	62.1	42.5
1st Shea part. R^2, area	-	-	-	0.12	0.16	0.17	0.19	0.21
1st part. Fisher, area	-	-	-	53.1	17.2	28.5	38.7	10.8
1st Shea part. R^2, pop. growth	-	-	-	-	-	-	-	0.24
1st part. Fisher, pop. growth	-	-	-	-	-	-	-	12.3

Notes: 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.

of the urban area fixed effect. We retain the first-step specification of column (6) of table 3 to generate these second-step results. To avoid presenting a long list of results that mirror all the second-step results presented so far with the urban areas fixed effects, we only reproduce the sample of specifications used in table 9 with three OLS and five IV estimations. The results from this table are clear. For no specification do the coefficients for population, land area, and population growth ever come close to being significant. Worse, depending on the details of the specification the coefficients om population and land area change sign. The extremely R^2 for the three OLS regressions are also consistent with a lack of systematic influence of population and land area on the distance gradient for land prices. While it is difficult to prove a negative result, the findings of table 10 are nonetheless suggestive that population and land area do not affect the distance

Table 10: 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	0.135	−0.032	0.195	−0.388	0.552	0.582	0.210	−0.066
	(0.131)	(0.224)	(0.288)	(0.767)	(0.543)	(0.603)	(0.599)	(0.490)
Land Area	-	−0.156	−0.169	0.737	0.310	0.453	−0.069	1.070
	-	(0.241)	(0.286)	(0.869)	(0.740)	(0.719)	(0.689)	(0.679)
Population Growth	-	2.742	3.112	3.583	3.930	4.108	3.380	1.530
	-	(3.206)	(3.637)	(3.630)	(3.651)	(3.678)	(3.571)	(7.918)
Geography	no	no	all	all	all	all	all	all
R^2	0.001	0.01	0.03	-	-	-	-	-
Overidentification p-value	-	-	-	0.065	0.090	0.070	0.047	0.063
First-stage statistic	-	-	-	11.4	16.5	18.5	14.7	10.8
Observations	288	288	288	288	288	288	288	288
Urban Population 1831	-	-	-	N	Y	Y	N	Y
Market Potential 1831	-	-	-	N	Y	Y	Y	Y
Number of hotel rooms	-	-	-	Y	N	N	N	N
Share of 1-star rooms	-	-	-	Y	N	N	N	N
Number of hotels	-	-	-	N	N	N	Y	N
Share of 1-star hotels	-	-	-	N	N	N	Y	N
Temperature in January	-	-	-	Y	N	Y	Y	Y
Henderson Occupations 1999	-	-	-	N	Y	N	N	Y
Bartik Industry 1999	-	-	-	N	N	N	N	Y
1st Shea part. R^2, population	-	-	-	0.15	0.29	0.24	0.24	0.37
1st part. Fisher, population	-	-	-	80.9	59.3	45.2	62.1	42.2
1st Shea part. R^2, area	-	-	-	0.12	0.16	0.17	0.18	0.20
1st part. Fisher, area	-	-	-	52.0	16.7	27.6	38.1	10.6
1st Shea part. R^2, pop. growth	-	-	-	-	-	-	-	0.23
1st part. Fisher, pop. growth	-	-	-	-	-	-	-	11.8

Notes: 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.

gradients for land prices. Understanding the drivers of the variation of these distance gradients for land prices is left for future work.

We now turn to the estimation of equation (16) where land area is allowed to adjust with population whereas most of the regressions reported so far were concerned with equation (14) where land area was included in the regression. As already argued, with land development being heavily restricted through planning regulations, we view the coefficient on population in the estimation of (14) as more relevant to think about urban costs. However, estimating (16) is interesting in its own right.

First, it provides an opportunity to verify whether the coefficient on population remains higher with IV than with OLS. To confirm this, table 11 reports results for two OLS and six IV regressions

Table 11: 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	0.216^a	0.230^a	0.247^a	0.307^a	0.278^a	0.282^a	0.357^a	0.394^a
	(0.051)	(0.050)	(0.076)	(0.094)	(0.075)	(0.086)	(0.084)	(0.089)
Population Growth	-	3.234^a	-	-	3.286^a	3.291^a	5.086^a	14.057^a
	-	(0.793)	-	-	(0.779)	(0.781)	(1.927)	(3.552)
Geography	all	all	all	all	all	all	all	all
R^2	0.44	0.47	-	-	-	-	-	-
Overidentification p-value	-	-	0.70	0.08	0.91	0.94	0.16	0.30
First-stage statistic	-	-	102.7	36.2	101.1	65.2	18.6	6.5
Observations	285	285	285	285	285	285	285	285
Urban Population 1831	-	-	Y	N	Y	Y	Y	Y
Urban density 1881	-	-	N	N	N	N	Y	N
Number of hotels	-	-	Y	Y	Y	N	N	Y
Share of 1-star hotels	-	-	N	Y	N	N	N	N
Temperature in January	-	-	N	Y	N	N	N	N
Henderson Occupations 1999	-	-	N	N	N	N	N	Y
Bartik Industry 1999	-	-	N	N	N	Y	Y	N
1st Shea part. R^2, population	-	-	0.43	0.29	0.43	0.33	0.36	0.52
1st part. Fisher, population	-	-	102.6	36.2	101.1	65.3	55.2	74.9
1st Shea part. R^2, population growth	-	-	-	-	-	-	0.17	0.09
1st part. Fisher, population growth	-	-	-	-	-	-	20.5	6.94

Notes: 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.

using urban area fixed effects as dependent variable and population as main explanatory variable. We note that the IV coefficients for population in columns (3) to (8) are all higher than the OLS coefficient of columns (1) and (2). This is highly consistent with our prior findings. Returning to our preferred specification of column table 7, we can compute an elasticity of land prices with respect to population that does not condition out land area. Using the fact that the elasticity of population with respect to land area is 0.874, we find the unconditional population elasticity of land prices to be $0.802 - 0.874 \times 0.625 = 0.256$. This number is very close to the direct IV estimates of the same elasticity in table 11. In addition, the coefficient on population growth remains around three when not instrumented and much higher (but estimated without much precision) when instrumented. This is again consistent with our findings above.

The second reason why estimating equation (16) is interesting is that it allows us to perform an 'overidentification' test to evaluate the overall consistency of our approach. We note from equation (16) that the coefficient on population corresponds to $(1+\sigma)\gamma\ \tau/(2\alpha\beta)$ whereas when including land area in equation (14) the same coefficient now corresponds to $(1+\sigma)\gamma$. If we take 0.821 in column (8) of table 7 to be our preferred estimate of $(1+\sigma)\gamma$, this suggests a value of $\tau/(\alpha\beta)$ between 0.60 and 0.96 from the first row of IV estimates in table 11. From (14), we also know that the coefficient on land area corresponds to $-1 + \tau/(2\alpha\beta)$. Our preferred IV estimation of (14) in column (8) of table 7 has a coefficient of -0.75 for land area which suggests a value of 0.50 for $\tau/(\alpha\beta)$. This is slightly below the values suggested from the estimation of (16) but obviously of the same magnitude.[15]

We also note that estimates of σ in the literature are low including coefficients from 0.015 to 0.03 in France (Combes *et al.*, 2010). Since our preferred estimate for the coefficient on population is 0.821, these numbers suggest an estimate of 0.80 to 0.81 for γ, the decentralisation coefficient introduced in equation (11). To obtain a simple independent estimate of this coefficient, we can estimate equation (11) directly by regressing log population in the core municipality of urban areas against the population of their urban areas. Of course, we do not expect the boundaries of the core municipality of urban areas to coincide 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 by the same proportion when the urban fringe is fixed.[16] Following the logic of our model we can thus obtain an estimate of the decentralisation coefficient by regressing population in the core municipality against the population of the entire urban area. We find a coefficients of 0.799 in a simple OLS univariate regression. This coefficient is highly significant and remarkably close to the indirect estimate implied by our main regressions.

7. The elasticity of urban costs with respect to population

Recall that 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. We just estimated the first elasticity to be about 0.82.

A recent by Davis and Ortalo-Magné (2011) sets the share of housing in expenditure to 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: 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 $r^L L = \alpha Q H$ 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 is such that $rK = (1-\alpha)QH$. These two first order conditions imply

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

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

[15] This coefficient of $\tau/(\alpha\beta)$ is also estimated in the first step as minus the distance gradient for land prices. Table (3) reports coefficients between -0.24 and -0.39 which are again of the same magnitude.

[16] To see this we can define density at any point located at distance D from the CBD as $d(D) = H(D)/(2\pi\theta D)$ using (3), (7), (8), (9), and (10) shows the result.

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%.

Land, unlike capital does not generally depreciate. Instead, according to our results it appreciates by about 3% for each percentage point of population growth. We also know that the population of French urban areas increases by about 1% per year. That suggests a user cost of land $r^L = 0.05 - 0.03 = 0.02$.[17] Inserting these numbers into (28) yields $\alpha = 2/11 \approx 0.18$. Appendix D proposes 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 Cobb-Douglas function with constant returns. Combining this estimate of 0.18 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 0.82 yields an elasticity of urban cost with respect to population of 0.033.

This value of ρ around 0.033 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.

8. 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 we devote considerable attention to the estimation of the population elasticity of land prices and develop a new approach to estimate the share of land in housing. Implementing our approach on unique land price and construction cost data for France we obtain an estimate of 0.033 for the population elasticity of housing costs.

This finding has 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'. Even though the elasticity of wages with respect to city population is lower than the elasticity of urban costs, population growth leads to wage increases that, in absolute terms, dominate the rise in urban costs in small cities. This is because wages are much larger than urban costs in these cities. However in large cities, the converse holds and the increase in wages associated with a larger population size should eventually be dominated rising urban costs. Overall the net benefits from cities should be bell-shaped.

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

[17]This number is the same as Davis and Heathcote (2007) but for possibly different reasons. Population growth in French metropolitan areas is slightly slower than in the US on the one hand but land supply is arguably on average more constrained on the other.

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). From a more positive perspective, 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 size of existing cities. Finally, we show above that the difference between our estimates for urban costs and agglomeration effects can be interpreted, under some conditions, as the elasticity of how endogenous amenities increase or decrease with population size. Our results suggest that this elasticity could be extremely low. This finding is consistent with recent results from Albouy (2008).

References

Albouy, David. 2008. Are big cities really bad places to live? Improving quality-of-life estimates across cities. Working Paper 14472, 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.

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. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2):723–742.

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, 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 (MASS.): National Bureau of Economic Research, 15–65.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Sébastien Roux, and Diego Puga. 2009. The productivity advantages of large markets: Distinguishing agglomeration from firm selection. Mimeographed, University of Toronto.

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.

Coulson, N. Edward. 1991. Really useful tests of the monocentric city model. *Land Economics* 67(3):299–307.

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

Duranton, Gilles and Diego Puga. 2011. The growth of cities. Work in progress, University of Toronto.

Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. A new approach to estimating the production function for housing. *American Economic Review* 100(3):905–925.

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.

Glaeser, Edward L. and Matthew Kahn. 2001. Decentralized employment and the transformation of the American city. *Brookings-Wharton Papers on Urban Affairs* :1–47.

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.

McMillen, Daniel P. 1996. One hundred fifty years of land values in Chicago: A nonparametric approach. *Journal of Urban Economics* 40(1):100–124.

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.

Mitchell, Timothy D., Timothy R. Carter, Pilip D. Jones, Mike Hulme, and Mark New. 2004. A comprehensive set of high-resolution grids of monthly climate for europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Technical report, Tyndall Centre for Climate Change Research Working Paper 55.

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.

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. Correction of OLS standard errors

Caution. The notations in these appendices are not consistent with the rest of the paper.

A. Theoretical expressions

The model considered here is of the form:

$$y = X\beta + \varepsilon + \eta \,, \tag{A1}$$

where y is a $N \times 1$ vector corresponding to the dependent variable, X is a $N \times K$ matrix corresponding to the explanatory variables, β is a $K \times 1$ vector of parameters, ε is a $N \times 1$ vector of error terms, and η is a $N \times 1$ vector of sampling errors with known covariance matrix V.

In this section, the error terms in the vector ε are independently and identically distributed with variance σ^2. We explain how to take into account the specific covariance structure of error terms of model (A1) in the computation of OLS standard errors.

The OLS estimator of parameters is given by:

$$\widehat{\beta}_{OLS} = \left(X'X\right)^{-1} X'y\,. \tag{A2}$$

The OLS (biased) estimator of σ^2 when ignoring the sampling error is given by:

$$\widehat{\sigma}^2_{OLS} = \frac{1}{N-K}\widehat{\varepsilon+\eta}'\widehat{\varepsilon+\eta}\,. \tag{A3}$$

The OLS estimated covariance matrix of $\widehat{\beta}_{OLS}$ when ignoring the sampling error is given by:

$$\widehat{V}\left(\widehat{\beta}_{OLS}\right) = \widehat{\sigma}^2_{OLS}\left(X'X\right)^{-1}\,. \tag{A4}$$

For the k^{th} estimated coefficient $\widehat{\beta}^k_{OLS}$, the Student p-value is given by:

$$P^k_{OLS} = P\left(|T| > \left|\frac{\widehat{\beta}^k_{OLS}}{\sqrt{\widehat{V}\left(\widehat{\beta}^k_{OLS}\right)}}\right|\right)\,, \tag{A5}$$

where T follows a Student law with parameter $N-K$.
The R^2 obtained when ignoring the sampling error is given by:

$$R^2 = 1 - \frac{\widehat{\varepsilon+\eta}'\widehat{\varepsilon+\eta}}{(y-\overline{y})'(y-\overline{y})}\,, \tag{A6}$$

where $\overline{y}$ is the vector where each element equals the average of the elements of y. Note that for the R^2 to be meaningful, model (A1) must include a constant.
Importantly, an unbiased and consistent estimator of σ^2 is given by:

$$\widehat{\sigma}^2_{corr} = \frac{1}{N-K}\left[\widehat{\varepsilon+\eta}'\widehat{\varepsilon+\eta} - tr\left(M_X V\right)\right]\,, \tag{A7}$$

where $M_X = I - X\left(X'X\right)^{-1}X'$ is the projection orthogonally to X, $\widehat{\varepsilon+\eta} = y - X\widehat{\beta}_{OLS}$ is the vector of estimated residuals obtained when the model is estimated with OLS.
Denote by $\Omega = \sigma^2 I + V$ the covariance matrix of $\varepsilon+\eta$. An unbiased and consistent estimator of this covariance matrix is $\widehat{\Omega} = \widehat{\sigma}^2_{corr}I + V$. A consistent estimator of the variance of $\widehat{\beta}_{OLS}$ is given by:

$$\widehat{V}_{corr}\left(\widehat{\beta}_{OLS}\right) = \left(X'X\right)^{-1}X'\widehat{\Omega}X\left(X'X\right)^{-1}\,. \tag{A8}$$

For the k^{th} estimated coefficient $\widehat{\beta}^k_{OLS}$, the corrected Student p-value is given by:

$$P^k_{corr} = P\left(|T| > \left|\frac{\widehat{\beta}^k_{OLS}}{\sqrt{\widehat{V}_{corr}\left(\widehat{\beta}^k_{OLS}\right)}}\right|\right)\,. \tag{A9}$$

It is also possible to construct a corrected R^2 for the model as:

$$R^2_{corr} = 1 - \frac{\widehat{\varepsilon+\eta}'\widehat{\varepsilon+\eta} - tr\left(M_X V\right)}{(y-\overline{y})'(y-\overline{y}) - tr\left(M_X V\right)}\,. \tag{A10}$$

B. Stata procedure

The procedure **olscorr(string scalar y, string matrix x, string scalar cov)** allows to compute the quantities expressed above. The arguments of the functions are the following:

- **y**: string containing the name of the dependent variable.
- **x**: string containing the names of the explanatory variables. The names of the explanatory variables are separated by blanks. For instance x="x1 x2" corresponds to the case where there are two explanatory variables in the regression which names are x1 and x2.
- **cov**: prefix of columns corresponding to the covariance matrix V. The columns of the covariance matrix must be named as: cov1, cov2,...,covN.

Note that the Stata dataset must only contain observations participating to the estimations (currently, "if" statements are not allowed). The dataset however can contain additional columns that are not used in the estimations.
Also, the procedure deals with cases where there are some missing values. If some elements in y take a missing value, corresponding lines in y and X are skipped, as well as the corresponding lines and columns in V.
Note finally that the constant is not included by default as when using the **regress** instruction in Stata. For a constant to be included, the matrix of explanatory variables X must include a vector which elements are all equal to one.
The procedure creates the following variables in the dataset:

- **bols**: OLS estimator of β given by (A2).
- **sig**: square-root of the OLS estimator of σ^2 given by (A3).
- **stdols**: OLS estimator of the standard errors of $\widehat{\beta}_{OLS}$ derived from (A4).
- **pvalols**: OLS Student p-value of elements of $\widehat{\beta}_{OLS}$ derived from (A5).
- **R2**: R^2 obtained from OLS and given by (A6).
- **sigc**: square-root of the corrected estimator of σ^2 given by (A7).
- **stdolsc**: corrected estimator of the standard errors of $\widehat{\beta}_{OLS}$ derived from (A8).
- **pvalolsc**: corrected Student p-values of elements of $\widehat{\beta}_{OLS}$ derived from (A9).
- **R2**: corrected R^2 given by (A10).

Appendix B. FGLS estimator

A. Theoretical expressions

It is also possible to construct a consistent FGLS estimator of $\widehat{\beta}$ as:

$$\widehat{\beta}_{FGLS} = \left(X'\widehat{\Omega}^{-1}X\right)^{-1} X'\widehat{\Omega}^{-1}y \tag{B1}$$

A consistent estimator of its covariance matrix is given by:

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

For the k^{th} estimated coefficient $\widehat{\beta}^k_{FGLS}$, the Student p-value is given by:

$$P^k_{FGLS} = P\left(|T| > \left|\frac{\widehat{\beta}^k_{FGLS}}{\sqrt{\widehat{V}\left(\widehat{\beta}^k_{FGLS}\right)}}\right|\right) \tag{B3}$$

B. Stata procedure

The procedure **fgls(string scalar y, string matrix x, string scalar cov)** allows to compute the quantities expressed above where **x**, **y** and **cov**, are the same as in the previous section.
The procedure creates the following variables in the dataset:

- **bfgls**: FGLS estimator of β given by (B1).
- **stdfgls**: estimated standard errors of the FGLS estimator derived from (B2).
- **pvalols**: FGLS Student p-value of elements of $\widehat{\beta}_{FGLS}$ derived from (B3).

Appendix C. WLS estimator

In some cases, the error terms in the vector ε are independently distributed but their variance varies such that for the i^{th} element of ε the variance is given by σ^2/n_i where n_i is an integer. In particular, this occurs if the i^{th} error term is the average of n_i random variables independently and identically distributed with variance σ^2.

In that case, the model can be rewritten as:

$$\widetilde{y} = \widetilde{X}\beta + \widetilde{\varepsilon} + \widetilde{\eta} \tag{C1}$$

where $\widetilde{y} = Wy$ with $W = diag(\sqrt{n_1},...,\sqrt{n_N})$, $\widetilde{X} = WX$, $\widetilde{\varepsilon} = W\varepsilon$, $\widetilde{\eta} = W\eta$. In that case we have $V\left(\widetilde{\varepsilon}\right) = \sigma^2 I$ and $V\left(\widetilde{\eta}\right) = \widetilde{V}$ with $\widetilde{V} = WVW$. It is then possible to conduct the same correction of standard errors of estimated parameters for model (C1) as for model (A1), using $\widetilde{V}$ instead of V.

The procedure **wlscorr(string scalar y, string matrix x, string scalar w, string scalar cov)** allows to compute the quantities expressed above where **x**, **y** and **cov**, are the same as in the previous section. Here, **w** is a string containing the name of the variable in which the vector $(n_1,...,n_N)'$ is stored.
This procedure creates the same variables in the dataset as **olscorr** except than the names are not exactly the same: (bols, sig, stdols, pvalols, R2, sigc, stdolsc, pvalolsc, R2) becomes (bwls, wsig, stdwls, pvalwls, wR2, wsigc, stdwlsc, pvalwlsc, wR2).

Appendix D. Estimating a production function for housing

We develop here a novel approach to the estimation of the production function for housing. The approach developed recently by Epple, Gordon, and Sieg (2010) would not be suitable for our purpose since we need to use different observables, the value of land and the cost of housing capital instead of the value of land and the value of the house for Epple *et al.* (2010). In addition, our approach does not impose constant returns to scale but instead allows to test for it. There are two other differences with the approach of Epple *et al.* (2010): (i) our production function measures the production of the flow of housing services and not the production of houses and (ii) we implement our approach on a broad cross-section of cities using variation in the rental price of land to identify the production function for housing. More precisely, we rely on the fact that housing can be sold at different prices in different locations . As a result, different locations have different rental prices for land. That's the fundamental source of variation that we use.

To develop our approach we need to modify our model above. Parcel area, T, which differs across parcels now enters as an argument in the rental price of housing. In addition, we no longer assume a specific functional form for the production function for housing. For simplicity, we also ignore distance arguments so that builder's profit for a parcel of area T can be written as

$$\pi = QH(K,T) - r^K K - R(T)\,, \tag{D1}$$

where $H(K,T)$ is the production of housing services, r^K is the user cost of capital, $R(T)$ is the rental price of a parcel of land with area T, and Q is the rental price of a unit of housing.

The first order condition for profit maximisation with respect to capital is

$$Q\frac{\partial H(K,T)}{\partial K} = r^K\,. \tag{D2}$$

Free entry implies that all the rents from building are dissipated into the rental price of land so that

$$R^*(K^*,T) = QH(K^*,T) - r^K K^*\,. \tag{D3}$$

Eliminating the rental price of housing from the previous two expressions yields the following differential equation

$$\frac{\partial H(K^*,T)}{\partial K^*} = \frac{r}{r^K K^* + R^*(K^*,T)} H(K^*,T)\,. \tag{D4}$$

When taking T as given, the solution of this differential equation is well known and satisfies

$$\ln H(K^*,T) = \int_{K_0}^{K^*} \frac{r}{r^K K + R^*(K,T)} \mathrm{d}K + \ln Z(T)\,, \tag{D5}$$

where $Z(T)$ is a positive function of T. This implies

$$H(K,T) = Z(T) \exp\left(\int_{K_0}^{K^*} \frac{r}{r^K K + R^*(K,T)} \mathrm{d}K \right)\,. \tag{D6}$$

This expression tells us for any T how much housing services are produced from an amount of capital K. We can then regress the log of $H(K,T)$ on the log of K for all the T for which it is estimated. Finding the same coefficient would indicate that the share of capital is constant.

To implement this approach we would need to observe, for any given T, the rental price of land $R^*(K^*,T)$ and capital K^* for many parcels in many cities. While this would be extremely demanding in terms of data, a simpler approach is just to estimate non-parametrically $H(K^*,T)$ for a given band for T and repeat the exercise for all the bands. In our empirical implementation we take deciles of land area. Focusing only on transaction with a construction taking place on empty land, we use 18053 observations. We first estimate housing services within each decile using (D6) and regress the log of housing services on the log of capital in each of those deciles. The coefficient on capital is estimated at 0.838. The smallest coefficient is for the first decile and is equal to 0.793 whereas the largest coefficient is for the last decile and is equal to 0.869. More generally, this coefficient increases as larger parcels are considered. However this tendency is modest and to a first approximation this coefficient can be taken to be constant.

We could next assume constant returns and derive a share of housing equal to 0.162. However we can also estimate the share of housing directly. Rather than use the first-order condition with respect to capital as in (D4) it is possible to develop an expression corresponding to (D6) that gives the production of housing services as a function of land conditional on some amount of capital K being used. The log of this quantity can then be regressed on the log of land to obtain the share of land for each capital decile. We find an average of 0.150 for the coefficient of land across all deciles with a maximum of 0.185 for the first decile and a minimum of 0.137 for the seventh decile. Consistent with the results for capital there is a tendency for smaller houses to be more land intensive. While land and capital deciles do not match exactly, it is easy to see that the sum of the share of land and capital is really close to one, 0.988 on average across all deciles.