The Production Function for Housing: Evidence from France

Pierre-Philippe Combes[*†]
University of Lyon and Sciences Po

Gilles Duranton[*‡]
University of Pennsylvania

Laurent Gobillon[*§]
Paris School of Economics-CNRS

Revised: February 2021

ABSTRACT: We propose a new non-parametric approach to estimate the production function for housing. Our estimation treats output as a latent variable and relies on the first-order condition for profit maximisation with respect to non-land inputs by competitive house builders combined with their zero-profit condition. Differences in the demand for housing across locations lead to differences in land prices and, in turn, differences in non-land input investments. For parcels of a given size, we compute housing production by summing across the marginal products of non-land inputs. We implement our methodology on newly built single-family homes in France. After taking care of estimation concerns, we find that the production function for housing is reasonably well, though not perfectly, approximated by a Cobb-Douglas function and close to constant returns after taking care of estimation concerns. We estimate an elasticity of housing production with respect to non-land inputs of about 0.65.

Key words: housing, production function.

JEL classification: R14, R31, R32

*We thank the editor, three anonymous referees, David Albouy, Nate Baum-Snow, Marcus Berliant, Felipe Carozzi, Tom Davidoff, Uli Doraszelski, Gabe Ehrlich, David Genesove, Joe Gyourko, Jean-François Houde, Stuart Rosenthal, Holger Sieg, Matt Turner, Tony Yezer, Oren Ziv, and seminar and conference participants for their comments and suggestions. We are also grateful to the *Service de l'Observation et des Statistiques (SOeS) - Ministère de l'Écologie, du Développement durable et de l'Énergie* for giving us on-site access to the main data, to the *Centre d'accès sécurisé aux données (CASD, ANR-10-EQPX-17 grant)* for remote secured access to supplementary confidential data, and to Guillaume Chapelle, Julian Gille, Clément Gorin, Morgan Ubeda, Benjamin Vignolles, and Alex Zanca for their help with the data.

†University of Lyon, CNRS, GATE-LSE UMR 5824, 93 Chemin des Mouilles, 69131 Ecully, France and Sciences Po, Economics Department, 28, Rue des Saints-Pères, 75007 Paris, France (e-mail: ppcombes@gmail.com; website: https://www.gate.cnrs.fr/ppcombes/). Also affiliated with the Centre for Economic Policy Research.

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

§PSE-CNRS, 48 Boulevard Jourdan, 75014 Paris, France (e-mail: laurent.gobillon@psemail.eu; website: http://laurent.gobillon.free.fr/). Also affiliated with the Centre for Economic Policy Research and the Institute for the Study of Labor (IZA).

1. Introduction

We propose a new non-parametric approach to estimate the production function for housing. Our estimation treats output as a latent variable and relies on the first-order condition for profit maximisation with respect to non-land inputs by competitive house builders combined with their zero-profit condition. Differences in the demand for housing across locations lead to differences in land prices and, in turn, differences in non-land input investments. For parcels of a given size, we compute housing production by summing across the marginal products of non-land inputs. We implement our methodology on newly built single-family homes in France. After taking care of estimation concerns, we find that the production function for housing is reasonably well, though not perfectly, approximated by a Cobb-Douglas function and close to constant returns. We estimate an elasticity of housing production with respect to non-land inputs of about 0.65.

A good understanding of the supply of housing is important for a number of reasons. First, housing is an unusually important good. It arguably provides an essential service to households and represents more than 30% of their expenditure in both France and the US (Combes, Duranton, and Gobillon, 2019). It is also an important asset. French households owned about 4.6 trillion dollar worth of housing in 2011 (Mauro, 2013). The value of the US residential stock owned by households was around 20 trillion dollar in 2007 (Gyourko, 2009). For both countries, this represents about 180% of their gross domestic product.

Housing and the construction industry also matter to the broader economy. The construction industry is arguably an important driver of the business cycle (e.g., Davis and Heathcote, 2005). The role of housing in the great recession has been studied by, among others, Chatterjee and Eyigungor (2015) and Kiyotaki, Michaelides, and Nikolov (2011). The broader effects of housing are not limited to the business cycle. Housing has also been argued to affect a variety of aggregate variables such as unemployment (Head and Lloyd-Ellis, 2012, Rupert and Wasmer, 2012) or economic growth (Davis, Fisher, and Whited, 2014, Hsieh and Moretti, 2019).

Finally, and most importantly, housing is also central to our understanding of cities. Different locations within a city offer different levels of accessibility to jobs and amenities. Housing production is crucial in transforming households' demand for locations into patterns of land use and housing consumption. Unsurprisingly, housing is at the heart of land use models in the spirit of Alonso (1964), Muth (1969), and Mills (1967) that form the core of modern urban economics. Related to this, the welfare consequences of land use regulations depend on the shape of the housing production function (Larson and Yezer, 2015). For instance, the consequences of minimum lot size requirements will depend on how easily substitutable land is in the production of housing.

Following Muth's (1969, 1975) pioneering efforts, there is a long tradition of work that estimates a production function for housing. Some of this work mirrors standard practice in productivity studies and regresses a measure of housing output on land and other construction inputs. Unfortunately, it is hard to separate the unit price of housing from the (quality-adjusted) amount of housing that a house offers. Then, a regression of the value of a house on land and non-land inputs is likely to have an error term correlated with the unit price of housing. Since we expect this price to determine the quantity of non-land inputs, the regression will not appropriately identify the production function for housing. This is a version of the unobserved price / unobserved quality problem that plagues the estimation of production functions (Ackerberg, Benkard, Berry, and Pakes, 2007, Syverson, 2011).

A popular alternative is to estimate the elasticity of substitution between land and other inputs directly by regressing the ratio of non-land to land inputs on the unit price of land. Because the price of land is equal to the value of a house minus the replacement cost of non-land inputs, this regression potentially suffers from severe reverse causation. With these important estimation caveats in mind, extant results are generally supportive of constant returns to scale in the production of housing and estimates for the elasticity of substitution between land and other inputs typically range between 0.50 and 0.75.[1]

To summarise, housing is highly heterogeneous and its unit price is unobserved. To produce housing, land, an immobile factor whose price varies greatly across locations, plays a particularly important role. These features call for specific estimation techniques, impose strong data constraints, and require careful attention to the sources of variation used for identification.

To meet our first challenge and separate the quantity of housing from its price per unit, we develop a novel estimation approach that relies on three main assumptions. First we assume a production function for housing that uses land and non-land inputs as primary factors. Since it cannot be directly observed, the quantity of housing is best thought of as a latent variable. Second, house builders maximise profit. They choose how much non-land inputs to use in order to build a house on a particular parcel of land given the price that households are willing to pay for housing

[1] Thorsnes (1997) is an interesting exception. He estimates an elasticity of substitution between land and other inputs statistically undistinguishable from one using high-quality data for which he observes both the price of land prior to construction and the price of the house when it is sold.

on this parcel. Third, we assume free entry among builders.

The first-order condition for profit maximisation by competitive house builders pins down the *marginal value product* of non-land inputs. Then, under free entry, the difference between the price of a house and the cost of the non-land inputs used to produce it should be equal to the price of the land parcel. We can use this zero-profit condition to eliminate the price of housing from the first-order condition and obtain a partial differential equation which links the *marginal product* of non-land inputs to the quantity of housing produced and the expenditure on both factors. For a given parcel size, this partial differential equation can be solved to obtain a non-parametric estimate of the amount of housing as a function of non-land inputs. Because our estimation is conditional on parcel size, the production function for housing is only partially identified.[2,3]

The second challenge is to find appropriate data. Our methodology requires information about the price of land parcels, their size, and the cost of construction. The unique data we use satisfy these requirements. They consist of several large annual cross-sections of land parcels sold in France with a building permit for a single-family home for which we know the subsequent building cost.

Given our approach and the data at hand, the third challenge is to use an appropriate source of variation. Although our estimation technique is non-standard, it remains that the supply of housing should be identified from the variation in the demand for housing across parcels. We develop a novel procedure inspired by instrumental variable approaches, which relies on exploiting the variation in systematic determinants of the demand for housing, like the city of a parcel or its location within this city, while conditioning out supply factors, such as the nature of the soil or the wage of construction workers.

We obtain three main results. First, we find that the elasticity of housing production with respect to non-land inputs is roughly constant at 0.65. As a first-order approximation, housing is produced

[2] To estimate production functions, Gandhi, Navarro, and Rivers (2020) jointly use the first-order condition for profit maximisation and the production function to eliminate unobserved persistent firm heterogeneity in productivity. This leads them to derive a partial differential equation similar to ours. For partial identification of the production function of housing, we only rely on the integration of this differential equation. For full identification, we make further assumptions about returns to scale in production. By contrast, Gandhi *et al.* (2020) make assumptions about the dynamics of productivity, insert the related equation into the production function and estimate the resulting specification that includes both the current and lagged values of inputs.

[3] Our approach consists in eliminating the unobserved price of output and rely about information on input prices and quantities. An alternative solution to this problem is to impose further assumptions about the structure of demand as in Klette and Griliches (1996) or De Loecker (2011). The production function can then be recovered from an extended productivity regression. Because standard assumptions about demand and industry structure made for manufacturing goods are questionable in our context, this type of approach is not appropriate here.

under constant returns to scale and is Cobb-Douglas in land and non-land inputs. This said, we can nonetheless formally reject that the housing production function is Cobb-Douglas and constant returns. We can also reject more general functional forms such as the CES. Finally, we find evidence of an elasticity of substitution slightly below one between land and non-land inputs.

In the recent literature, we note the work of Yoshida (2016). He develops a new approach to account for capital depreciation in housing and shows that standard estimates of the elasticity of substitution between land and other inputs can be sensitive to how depreciation is accounted for. Albouy and Ehrlich (2018) estimate a cost function for the production of housing at the city level. Their objective is to explore the determinants and implications of differences in housing productivity across cities. While our focus is to obtain a better measure of the amount of housing, Albouy and Ehrlich (2018) measure it simply using standard hedonics in an intermediate step. Finally, our work is most closely related to Epple, Gordon, and Sieg (2010) and subsequent work by Ahlfeldt and McMillen (2020) who also treat housing as a latent variable. We postpone a detailed comparison with their work until later.

2. Treating housing production as a latent variable

We first introduce the simplest version of our model with homogenous factors. We then consider factor heterogeneity before discussing some of our assumptions further and making a detailed comparison with Epple *et al.* (2010).

2.1 *Base model*

House builders produce a quantity of housing H using land T and non-land inputs K, which we refer to as (housing) capital for convenience.[4] The production technology $H(K,T)$ is strictly increasing and concave in K.

Land is exogenously partitioned into parcels, which differ in their area T and in their other characteristics noted x. These characteristics, which include location, determine the demand for housing on a parcel of land. A unit of housing on a parcel fetches a price, $P(x)$, which reflects the willingness to pay of residents to live there. This price is taken as given by competitive house builders. A builder, who develops a parcel of size T and characteristics x purchased at the endoge-

[4]Non-land inputs are essentially construction labour and materials, which both get frozen into housing through the construction process. This is consistent with the usual definition of capital.

nously determined price R, seeks to maximise profits $\pi = P(x)H(K,T) - K - R$ with respect to K after normalising the price of capital to one. Below, we extend this model to allow for supply differences between parcels, for the price of capital to vary across locations, for heterogeneity in the composition of capital, and for endogenous parcel size.

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

$$P(x)\frac{\partial H(K,T)}{\partial K} = 1\,. \tag{1}$$

The optimal amount of capital that satisfies this condition is given implicitly by $K^* = K^*(P(x),T)$. Because the production function for housing $H(.,.)$ is concave in K, K^* is unique given T. Applying the implicit function theorem to equation (1), the concavity of $H(.,.)$ also implies that $\partial K^*/\partial P > 0$. Hence, there is a bijection between the price of housing and the profit-maximising level of capital for any T and we can write $P(x) = P(K^*,T)$.

Then, free entry implies that the profit from construction is dissipated into the price of a parcel:

$$R = P(K^*,T)H(K^*,T) - K^* \equiv R(K^*,T)\,. \tag{2}$$

Note that the price of a parcel in equilibrium is uniquely defined for any K^* and T.

While in the data described below we observe K^*, T, and $R(K^*,T)$, we cannot separately measure $P(x)$ and $H(K^*,T)$. We only observe the product $P(x)\,H(K^*,T)$. To eliminate the unit price of housing $P(x)$, we can insert equation (1) into (2) and obtain the following partial differential equation:

$$\frac{\partial H(K^*,T)}{\partial K^*} = \frac{H(K^*,T)}{K^* + R(K^*,T)}\,. \tag{3}$$

For consistency with our empirical work below, we can rewrite this expression using natural logarithms on the left-hand side:

$$\frac{\partial \log H(K^*,T)}{\partial \log K^*} = \frac{K^*}{K^* + R(K^*,T)}\,, \tag{4}$$

In words, at the competitive equilibrium, the elasticity of housing production with respect to capital is equal to the share of capital in the cost of building a house.[5]

[5]The equality between the elasticity of output with respect to an input and this input's share in cost is often used in the literature since the cost share of inputs is usually readily available from firm data and can be used to estimate the output elasticity. See Klein (1953) and Solow (1957) for two early applications and Syverson (2011) for a more recent discussion. In a different spirit, De Loecker and Warzynski (2012) and followers use the difference between an input's cost share and its output elasticity under imperfect competition to recover price mark-ups . We differ from the literature by what we do next as we integrate 'individual' cost shares to recover output and separate it from prices.

Consider that for parcels of size $T \in [\underline{T}, \overline{T}]$, parcel characteristics vary so that the price of housing is distributed over the interval $\left[\underline{P}, \overline{P}\right]$. The optimal level of capital in housing K^* then covers the interval $\left[\underline{K}, \overline{K}\right]$ where $\underline{K} = K^*(\underline{P}, T)$ and $\overline{K} = K^*(\overline{P}, T)$. The solution to the differential equation (4) for a given value of the optimal amount of capital inputs K^* in this interval is readily obtained by integration and can be written as:

$$\log H(K^*, T) = \int_{\underline{K}}^{K^*} \frac{K}{K + R(K, T)} \mathrm{d} \log K + \log Z\left(T\right) . \tag{5}$$

where $Z(T)$ is a positive function equal to $H(\underline{K}, T)$. Equation (5) enables the computation of the (unobserved) number of units of housing on a parcel of size T knowing the (observed) prices of parcels of the same size and the (observed) amounts of capital invested to build on them.

The intuition behind this result is relatively straightforward. Parcels differ in desirability and thus in their unit price for housing. This price is not observed but it appears in both the optimal capital investment rule described by the first-order condition (1) and in the zero-profit condition (2). We can use the latter equation to substitute for the price of housing in the former and derive differential equation (3), or its log equivalent in equation (4). We then readily obtain $\log H$ in equation (5) by integration over $\log K$.

To illustrate the workings of equation (5) and check the consistency of our approach, consider first a Cobb-Douglas production function. In this case, the price of land, R, and the capital used to build a house, K, are proportional. This implies that the term within the integral is constant. As a result, $\log H$ is proportional to $\log K$. That is, we retrieve a Cobb-Douglas form. To take another example, assume now that the production function has a constant elasticity of substitution between land and capital $\sigma = 2$. Then, profit maximisation implies that capital should increase with the square of parcel price given parcel size. Integrating the share of capital as in equation (5) implies that the production of housing is proportional to $(\sqrt{K} + z)^2$ where z is a constant. This functional form is indeed the generic functional form for a CES production function with an elasticity of substitution equal to two when a factor (land) is held constant.

2.2 *Factor heterogeneity*

An important assumption of our model so far is that land and capital are both homogeneous factors. We worry that parcels may differ in their ability to supply housing, that the price of housing capital differs across locations, and that housing capital is heterogeneous. We address these three concerns in turn.

Starting with land, consider first a simple illustrative example where all parcels are of unit size and the demand for housing is the same everywhere with $P(x) = P = 1$. Assume that housing is produced according to $H(K,y) = \frac{1}{a} y^{1-a} K^a$. The shifter y measures the ease of construction, which differs across parcels. Then, capital is given in equilibrium by $K(y) = y$ and parcel prices capitalise the ease of construction, $R(y) = \frac{1-a}{a} y = \frac{1-a}{a} K$. Using equation (5) to estimate the value of housing while ignoring y would wrongly imply that the production of housing is linear in K instead of being proportional to K^a.

More generally, assume that parcels are now characterised by two sets of characteristics, x and y. The characteristics x still affect the price that residents are willing to pay for housing, $P(x)$, while the characteristics y affect the production of housing directly, which is now given by $H(K,T,y)$. The analogue to the first-order condition (1) is $P(x) \partial H(K,T,y)/\partial K = 1$. The zero-profit condition also implies that y affects the price of land directly: $R(K^*,T,y) = P(x)H(K^*,T,y) - K^*$. The partial differential equation analogous to equation (4) is then:

$$\frac{\partial \log H(K^*,T,y)}{\partial \log K^*} = \frac{K^*}{K^* + R(K^*,T,y)}, \tag{6}$$

where $K^* = K^*(P(x),T,y)$ and $R(K^*,T,y)$ both depend on y and there is no longer a unique mapping between R and K^* given T. Hence, if we ignore y when we integrate equation (6), the computed quantity will depend on a mixture of values for y and will not recover the true H, even in cases where y is uncorrelated with P as in our simple example above.

A first possible solution to this problem is to note that equation (6) can still be integrated for a given y. Hence, if y is observed, we can still identify $H(K,T,y)$ for each T and y. For instance, it may be cheaper to build on a flat terrain. In this case, we can still identify how H varies with K depending on parcel size and the slope of the terrain. A limitation of this approach is that we may run out of statistical power as there are potentially many supply characteristics to consider.

A second solution is to obtain parametric predictions of K^* and R that both explicitly rely on demand factors x and condition out potential supply factors y in a first step. With $K^* = K^*(P(x),T,y) \equiv K^*(x,T,y)$ and $R = R(K^*(P(x),T,y),T,y) \equiv R(x,T,y)$, we can regress K^* and R on the parcel characteristics x that only affect the demand for housing, parcel area T, and the characteristics y that affect the supply of housing. We then reconstruct investment in housing capital and parcel price for a set value of y, say its average $\overline{y}$, as $\widehat{K} = \widehat{K}^*(x,T,\overline{y})$ and $\widehat{R} = \widehat{R}(x,T,\overline{y})$. Equation (6) then becomes,

$$\frac{\partial \log H(\widehat{K},T,\overline{y})}{\partial \log \widehat{K}} = \frac{\widehat{K}}{\widehat{K} + \widehat{R}}\,, \tag{7}$$

which can be readily integrated like equation (4).

While we postpone further discussion of how we implement this approach and extend it for the possible additional endogeneity of T, we note that it relies on the same principle as standard instrumental variables approaches. We exploit the (conditional) variation of exogenous variables x to predict K and R under the exclusion restriction that x does not play any role in $H(.)$ after conditioning out supply characteristics, y.[6]

Turning to capital, a first issue is that the price of construction material, or perhaps more likely, the price of construction labour (also embedded in K) may differ across parcels. Instead of being normalised to one everywhere, the unit price of capital is now noted r. The analogue to the first-order condition analogous to equation (1) is $P(x)\partial H(K,T)/\partial K = r$. With the zero-profit condition, $R(K^*,T,r) = P(x)H(K^*,T) - rK^*$, the partial differential equation analogous to equation (4) is now:

$$\frac{\partial \log H(K^*,T)}{\partial \log K^*} = \frac{rK^*}{rK^* + R(K^*,T,r)}\,, \tag{8}$$

where we observe rK^* and T in the data but cannot separate r from K^*.

Like with heterogeneous parcels, we cannot in general integrate this expression irrespective of r since this variable affects the price of land independently of K^*. Importantly, differences in price for housing capital do not create any problem in the Cobb-Douglas case where the value of land R only depends linearly on rK^*. In this particular case, the cost ratio in equation (8) is a constant, as already noted.

There are two possible solutions to the issues raised by heterogeneous prices for housing capital. They mirror those proposed above for heterogeneous land. First, we can reasonably assume that differences in the price of housing capital consist mainly of differences in construction costs across cities and, to a lesser extent perhaps, differences between more or less central locations within

[6]Our approach also resembles that of Olley and Pakes (1996) and followers. Like them, we worry that a shock y, unobserved to us but not to the house builder, affects both the choice of input and output. In Olley and Pakes (1996), capital evolves over time following new investments impacted by the shock y. Since the investment function $I = I(y,K)$ increases with y given K, it can be inverted to obtain $y = y(K,I)$. Then, y can be replaced by this last expression in the production function, $H(K,T,y) = H(K,T,y(K,I))$. Hence, the key parameters of the production function are estimated thanks to an economic proxy, investment, for unobserved shocks. In our case, we have $K^*(x,T,y)$ and $R(x,T,y)$ and we use the variations of observed proxies for x and y to predict K^* and R. The two main differences are that, first, we implement our approach by predicting the inputs $\widehat{K}^*(x,T,\overline{y})$ and $\widehat{R}(x,T,\overline{y})$ for a given $y = \overline{y}$ instead of using a control function and, second, we use exogenous proxies for y instead of an endogenous variable as in Olley and Pakes (1996).

cities.[7] We can thus integrate equation (8) separately for different sub-samples of observations likely to face the same price of housing capital, either because they belong to the same city or to the same type of location, central or suburban. Alternatively, we can also use information about the wages of construction workers by city, condition them out of parcel prices and housing capital, and then use the residualised values instead of the actual ones in our estimation.

Finally, we turn to the possible heterogeneity of the composition of capital itself. In our context, it is natural to assume two types of capital. The first produces 'raw floorspace' while the second produces 'housing quality'. House builders maximise their profits with respect to both K_1 and K_2. In the data, we only observe total capital, $K = K_1 + K_2$. In Appendix A, we show that, despite the heterogeneity of capital, our approach still recovers the quantity of housing for a given T.

A complication arises when the use of one type of capital, say K_1 used to produce floorspace, is subject to local regulatory constraints such as a cap. As we show in Appendix A, the elasticity of housing with respect to capital is then of the form $K^* / \left(K^* + R(K^*, T, \overline{K}_1)\right)$ where $\overline{K}_1$ is a cap on floorspace capital and $K^* = \overline{K}_1 + K_2^*$. Like with heterogeneous land or heterogeneous prices of capital, there is an extra argument entering the price of parcels so that the cost share cannot be integrated directly to recover the quantity of housing. To handle this issue, we distinguish between new constructions subject to stricter vs. less strict land use regulations. We postpone further discussion of land use regulations.

2.3 *Further threats*

We assume competitive developers operating with non-increasing marginal returns to capital on exogenously determined parcels. Because of the ease of entry in this industry, our assumption of competitive builders strikes us as reasonable.[8] Then, our results below strongly support our assumption of decreasing marginal returns to capital. Section 7 relaxes the assumption of exogenous parcel size and provides evidence against the endogenous determination of parcel size.

---

[7] Another source of differences in construction costs across parcels may arise from economies of scale associated with being able to build many houses at the same location at the same time vs. building only one house. While this is a concern, we show in the next section that it is unlikely to be a first-order issue here since most new houses in France are in-fills that are built individually or in very small numbers. The caveat is nonetheless that we estimate the production function for one house, not for builders who may build several houses jointly.

[8] A search on the French yellow pages (`http://www.pagesjaunes.fr/`) yields 1,783 single-family house builders for Paris (largest urban area with population above 12 million), 111 for Rennes (10th largest urban area with population 654,000), and still 38 for Troyes (50th largest urban area with population 188,000). (Search conducted on 21st May 2013 looking for 'constructeurs de maisons individuelles' – builders of single family homes – typing 'Ile-de-France' to capture the urban area of Paris, 'Rennes et son agglomération', and 'Troyes et son agglomération' for the other two cities.)

We also assume in our model that housing is homogeneous. This assumption is core to our approach since it allows us to think in terms of units of housing that can be measured and compared across houses purchased by different households.[9] In defense of this assumption, we note first that we mainly consider the construction cost of houses before they get fully customised. In a robustness check, we use available information about the degree of completion of houses. Should housing heterogeneity matter, we expect it will matter more for houses at a more advanced state of completion. In another robustness check, we conduct separate estimations for different socio-economic groups of buyers who may compete on different segments of the market.

Our model of housing construction is static and ignores a range of dynamic issues. First, housing prices may diverge over time across locations. Below, we extend our model to two periods and show how differences in expectations about future prices across locations affect our estimations. Second, we ignore that housing development is, to a large extent, an irreversible decision. Under uncertainty, irreversible development implies that the price of vacant land includes an option value to develop it. This option value of waiting is highest for marginal parcels at the urban fringe. We omit this issue here because most parcels in the data are more centrally located as we show below.

Finally, we assume that the price of housing on a parcel does not depend on the intensity of its development. This is a simplification because single-family homes are indivisible (by definition) and the price per unit of housing may decline with the quantity of housing offered by a house, at least beyond a certain quantity. In this case, house builders are no longer price takers since the unit price of housing is also determined by the quantity of housing built: $P(x,H)$. In turn, the first-order condition analogous to equation (1) contains a term in $\partial P/\partial H$ which can no longer be eliminated using the zero-profit condition.

2.4 *Comparison with Epple* et al. *(2010)*

We now compare our approach with that of Epple *et al.* (2010). Appendix B provides formal derivations. Like us, Epple *et al.* (2010) develop a non-parametric estimation of the housing production function using restrictions from theory and treat output as a latent variable. The first substantive difference is that our approach may be viewed as the primal of Epple *et al.*'s (2010) dual approach. We rely on the first-order condition for profit maximisation after eliminating the unobserved unit

[9]The polar opposite implies that each house is a uniquely differentiated variety over which residents have unique idiosyncratic preferences. Our approach is unable to deal with such an extreme case where each and every resident values each and every house differently since the notion of a common unit of housing is no longer well defined.

price of housing using a zero-profit condition. From the resulting partial differential equation, we can then provide an explicit solution for housing quantities. Epple *et al.* (2010) rely instead on duality theory and Hotelling's lemma to recover the supply function of housing before computing the production function. This is a less direct route which relies on a more intricate differential equation with no closed-form solution. Second, we make no restriction regarding returns to scale. Epple *et al.* (2010) impose instead constant returns to scale in the production of housing. Empirically, we show below that we can reject constant returns. In Appendix B, we rework the approach of Epple *et al.* (2010) in the absence of constant returns to scale. Third, and most importantly, we tackle various forms of factor heterogeneity, which Epple *et al.* (2010) do not consider.

Other differences with Epple *et al.* (2010) are more cosmetic. We exploit information about capital K, parcel prices R, and land areas T, whereas Epple *et al.* (2010) use house values $P\,H$ instead of capital together with parcel prices and land areas. This difference is superficial because both approaches impose the same zero-profit condition $P\,H = K + R$. Hence, capital can be immediately obtained by subtracting parcel prices from house values. Finally, we implement our approach on very different data: newly built houses for an entire country instead of assessed land values for all houses in a single city, Pittsburgh. For comparison purpose, we estimate a variant of their approach on our data below.

3. Data

The observations in our data are transacted land parcels with a building permit that are extracted from the French Survey of Developable Land Price (*Enquête prix des terrains à bâtir*). This survey is conducted every year in France since 2006 by the French Ministry of Ecology, Sustainable Development, and Energy. The sampling frame is drawn from Sitadel, the official land registry, which covers the universe of all building permits for detached houses. The survey selects building permits for owner-occupied, single-family homes. Permits for extensions to existing houses are excluded. For a small fraction of parcels (less than 3% in 2006), there is also a demolition permit. Our study period covers transactions from 2006 to 2012.[10] Originally, about two thirds of the transactions with permits were sampled. The survey became exhaustive in 2010. It is mandatory

[10]It is important to keep in mind that, unlike in the US, there was no housing burst in France during this period and that the heterogeneity of housing price fluctuations across locations was far from being as extreme as in the US.

and the response rate, after one follow-up, is above 75%. Annually, the number of observations ranges from 48,991 (in 2009) to 127,479 (in 2012).

While it is possible to get new houses built in many ways in France, the arrangement we study covers a large fraction of new constructions for single-family homes.[11] Households typically first buy a constructible land parcel, obtain a building permit, and get a house built through a general contractor, through an architect, or directly by themselves. Only about 20% of new houses are 'self-built' as French law requires the use of a general contractor or an architect for constructions above 100,000 Euros. Getting a new house by first buying land before subsequently signing a contract with a builder is fiscally advantageous as it avoids paying stamp duties on the structure. This arrangement also greatly reduces financing constraints for house builders and lowers their risks. Finally, this arrangement is consistent with the model above where the house buyer pays first for the land parcel $R(K^*,T)$ and then pays the builder for the structure $P\,H(K^*,T) - R(K^*,T) = K^*$.

For each transaction, we know the price of the parcel, its size, whether it is 'serviced' (i.e., has access to water, sewerage, and electricity), its municipality, how it was acquired (purchase, donation, inheritance, other), some information about its buyer, whether the parcel was acquired through an intermediary (a broker, a builder, another type of intermediary, or none), and some information about the house built, including its cost (but with no breakdown between material and labour) and its floor area.

The notion of 'building costs' is potentially ambiguous but we know whether the reported cost reflects the cost of a fully decorated house, the cost of a serviced house prior to decoration (i.e., excluding interior paints, light fixtures, faucets, kitchen cabinets, etc), or only the cost of the bare-bone structure. The second category, ready-to-decorate houses, represents the large majority of our observations (72%). We only consider parcels that were purchased and ignore inheritances and donations. We also appended a range of municipal and urban area characteristics to our main data. They are described in Appendix C.

Table 1 provides descriptive statistics for all our main variables. The first interesting set of facts pertains to the variation in parcel size, total construction costs, and parcel price per square metre. A parcel at the top decile is about four times as large as a parcel at the bottom decile. Interestingly, for

[11]The consultancy Développement-Construction reports between 120,000 and 160,000 new single-family homes per year during the period (http://www.developpement-construction.com/). These magnitudes essentially match our number of observations after accounting for the sampling frame and the response rate and allowing for a lag between land transactions and the completion of houses.

Table 1: Descriptive Statistics

Variable	Mean	St. deviation	1st decile	Median	9th decile
Entire country:					
Parcel area	1,156	947	477	883	2,079
House area	123	35	87	117	161
Construction cost	127,552	55,002	78,442	115,000	190,667
Parcel price	63,371	57,648	19,672	50,000	120,000
Parcel price per m^2	80	86	14	58	166
Urban areas:					
Parcel area	1,030	812	434	806	1,843
House area	126	36	89	120	167
Construction cost	132,979	58,714	80,778	119,000	200,000
Parcel price	75,481	63,891	27,465	59,841	140,000
Parcel price per m^2	101	99	23	76	204
Greater Paris:					
Parcel area	839	673	329	665	1,492
House area	134	42	90	126	188
Construction cost	151,319	73,713	89,173	132,850	236,605
Parcel price	141,711	102,709	69,155	124,406	220,000
Parcel price per m^2	237	193	67	182	466

Notes: The sample contains 386,177 observations for the entire country, 218,767 observations for urban areas, and 17,178 observations for Greater Paris. Parcel and house areas are in square metres. Parcel prices and construction costs are expressed in 2012 Euros, using the French consumer price index.

construction costs, the corresponding inter-decile ratio is only about 2.4 whereas for parcel prices per square metre, it is nearly 12. The second interesting feature of the data highlighted in table 1 is that this variation does not only reflects differences across urban areas or between urban and rural areas. Even when we only consider transactions from Greater Paris, we still observe considerable variation in parcel price per square metre.

A first important reason for the variation in parcel prices is that new constructions in French urban areas are, in their large majority, in-fills that occur everywhere in their urban area, from more expensive central locations to cheaper peripheral ones. To illustrate this feature, figure 1 represents the probability distribution function of the relative distance of new constructions to the centre of their urban area. Consistent with the preponderance of in-fills, another data source, the Survey of Commercialisation of New Dwellings (Enquête sur la Commercialisation des Logements Neufs) indicates that less than 10% of building permits for single-family homes are for houses part of a group of five or more.

Beyond the systematic variation explained by the location of parcels within and across urban

Figure 1: Probability distribution function of the relative distance for new constructions, French urban areas 2006-2012

Notes: All years of data used. 218,767 observations. For each new construction, we compute the distance between the centroid of its municipality and the population-weighted centroid of its urban area and divide by the greatest observed distance for any new construction in this urban area. Less than 2% of the observations are beyond 95% of the maximum distance to the centre. The modal distance is at about 40% of the maximum distance to the centre of the urban area.

areas, the price of a parcel also reflects features such as idiosyncratic buyer preferences, their degree of optimism about future housing prices, and the bargaining ability and eagerness of buyers and sellers.[12] In addition, the market for constructible land in many parts of France is thin. Comparing construction costs across competing builders is perhaps easier than assessing the value of a parcel of land. This lack of information about land prices may add to the variation in parcel prices.

Regardless of its exact source, this idiosyncratic variation is an important reason why the implementation of the approach described below uses predicted parcel prices obtained from a kernel non-parametric regression. Kernel smoothing on a grid allows us to reduce the noise for particular transactions while obtaining quasi-continuous series for land prices and capital, which are needed for our non-parametric estimation.

Two further measurement issues require discussion. First, the construction costs reported by surveyed households may not accurately reflect how much they actually paid. We do not think this is the case. We first note that contracts with house builders usually include a small number of installments and we expect households to remember the headline figure. We can investigate this issue further using data from the Survey of Costs of New Dwellings (Enquête sur le Prix de Revient des Logements Neufs). This is a detailed survey of builders that forms the basis of the French construction price index, which, in turn, is used to index rent increases for residential rented

[12]In a regression of log land price per square metre, the extensive parcel and location characteristics used in table 7 explain only 65% of the variance within each decile of parcel area.

properties, parking spots, and, until quite recently, commercial leases. From the second quarter of 2010 through the fourth quarter of 2012, we could match all 2,336 observations in this survey with our main data using the building permit identifier. Reassuringly, the correlation between the two measures of housing costs is 0.83, for both levels and logs. Finally, we note that our identification approach, which relies on predicting construction costs using exogenous characteristics of parcels, also corrects for measurement error in construction costs like a traditional instrumental variable estimation. The similarity in our results using actual and predicted construction costs is consistent with measurement issues for construction costs being minor.

Second, we expect parcel areas to be measured with a high degree of accuracy since they are an essential part of the official land registry. The data we use reports parcel size for both the transaction and the building permit. We retain the latter to account for land assembly. When the size of the permitted parcel is larger than the size of the purchased parcel, we increase parcel price proportionately. For only about 1% of the permits, the permitted parcel is more than 20% smaller than the purchased parcel. When we treat parcel size as endogenous below and thus possibly correct for measurement error, we find that it makes essentially no difference to our results.

In summary, we rely on data documenting parcel prices and construction costs for newly built single-family homes in France. These data stem from an unusual institutional setting where house buyers purchase empty parcels before construction to save on transaction costs and from a unique follow-up survey of buyers that can be, for some houses, matched to another survey of house builders. While we know of no other equivalent data elsewhere, we note that the same information could, in principle, be obtained from transaction prices for new homes and prior transaction prices for empty parcels or tear-downs. Then, construction costs can be obtained indirectly by subtracting the price of land from the price of new houses. Alternatively, information about construction costs, if available, may be used to recover land prices.[13] Nonetheless, the information extracted from these alternative sources of data is likely to be more noisy than the data at hand.[14]

[13] In the US, construction prices for new homes are broadly available from several data providers. Tear-downs can be detected from the same data. Land transactions are recorded by CoStar and have been used, among others, by Albouy, Ehrlich, and Shin (2018). A challenge is that CoStar records only transactions above a given threshold. These are often large parcels which get subdivided. The prices of these large parcels must then be transformed to recover the value of smaller parcels, possibly using information from nearby houses to infer what the size of the small parcels might be as per Gyourko and Krimmel (2020). Construction costs are available from real estate consultancy RS Means. They have been used for instance by Gyourko and Saiz (2006). They are however limited to a few standard house specifications at the level of entire metropolitan areas.

[14] Obtaining construction costs by subtracting parcel prices from house prices may create a spurious correlation due to measurement error when we regress construction costs on land costs, as in the traditional approach. This spurious correlation is not a problem in our approach which relies on a cost share.

4. Implementation

We now turn to the five practical steps we take to compute the production of housing and estimate the elasticity of housing production with respect to capital. First, we make the data comparable over time. Second, we estimate the price of parcels R for any pair of capital K and parcel size T using kernel smoothing. Third, we non-parametrically estimate the amount of housing $H(K,T)$ for a given T using equation (5) on a (K,T) grid. Fourth, we describe the shape of $H(K,T)$ by means of simple regressions. Finally, we extend our approach to condition out factor heterogeneity.

Because we use data for a six year period, we first make parcel prices and capital investments comparable over time. We do so by correcting for year effects, which we obtain from regressions of log parcel prices and log capital on year fixed effects.

Our second step is more consequential. In the data, the price of parcels is observed only for some values of capital and parcel size, not for the entire continuum as required by the integration in equation (5). To make up for these 'missing values', we estimate the price of land for any given level of capital K and parcel size T from actual observations with slightly larger and slightly lower values of K and T using a kernel non-parametric estimation.

The kernel we use is the product of two independent normals and the bandwidth is computed using a standard rule of thumb for the bivariate case (see Silverman, 1986). For a given value of (K,T), the estimated price of land is given by the following formula:

$$\tilde{R}\,(K,T) = \sum_i \omega_i R_i \quad \text{with} \quad \omega_i = \frac{L_{h_K}\,(K - K_i)\, L_{h_T}\,(T - T_i)}{\sum_i L_{h_K}\,(K - K_i)\, L_{h_T}\,(T - T_i)}\,, \tag{9}$$

where $L_h\,(x) = \frac{1}{h} f\left(\frac{x}{h}\right)$ with $f\,(\cdot)$ the density of the normal distribution with $h = N^{-1/6}\sigma\,(X)$. $\sigma\,(X)$ is the standard deviation for variable X computed from the data and N is the number of observations. This kernel estimator has the property of making $R(K,T)$ unique, which is requested by our model.

The correlation between actual parcel prices and those predicted using the rule-of-thumb bandwidth is 0.45. Using instead bandwidths that are half, a quarter, and a tenth of the rule-of-thumb bandwidth leads to correlations between actual and predicted parcel prices of 0.50, 0.57, and 0.66, respectively.[15] We verify below that our choice of bandwidth does not affect our results.

[15] While smaller bandwidths lead to a better fit at the observed values of capital and parcel size, they are potentially problematic for some points of our grid since they may not allow the use of enough observations around these points to obtain accurate predicted parcel prices or correct for measurement error in R.

We also note that we kernel-smooth over R before computing the cost share associated with K, $\tilde{R}$, and T. An alternative is to consider that the cost share $CS(K,R,T) = \frac{K}{K+R(K,T)}$ is computable only at the observed values of capital and parcel size. We can then directly estimate smoothed cost shares $\tilde{CS}(K,R,T)$ for other values of capital and parcel size using a kernel non-parametric regression. This alternative approach differs because it implies smoothing a function of R and K instead of smoothing R only. In practice however, the two approaches yield very similar results as we show below in a robustness check.

For our third step, the integral in equation (5) is computed using a trapezoidal approximation on a (K,T) grid. An estimator of the production function at a given node $\left(K_i^g,T^g\right)$ is:

$$\widehat{\log H}(K_i^g,T^g) = \sum_{j=2}^{i} \left(\frac{c_{j-1}+c_j}{2}\right)\left(\log K_j^g - \log K_{j-1}^g\right), \tag{10}$$

where $K_j^g, j = 1,...,J$ are the grid values of capital and:

$$c_j = \frac{K_j^g}{K_j^g + \tilde{R}(K_j^g,T^g)}. \tag{11}$$

The lower bound, K_1^g, in equation (10), which corresponds to $\underline{K}$ in model, is the lowest value of the capital we consider. We can potentially use any value of capital as lower bound but there is a trade-off. A smaller value will allow us to study the variations of the housing production function over a wider range of values for capital but this may come at the cost of being in a region where there are few observations. We restrict attention to observations above the first decile (and below the ninth decile) of capital to estimate the production of housing for each of the nine deciles of parcel size. This corresponds to 74.2% of all land values. With 900 uniformly distributed values of capital for each decile of parcel size, we thus generate a fine grid of 8,100 (K,T) points from 386,177 observations for the entire country in our data set.[16]

In the fourth step, we regress housing production as estimated in equation (10) on capital investment to describe how housing production varies with capital. For instance, under Cobb-Douglas and for any fixed T, the relationship between log housing and log capital is linear. We estimate

[16]We use a uniform grid. It would be possible to consider a greater density of grid points closer to actual transactions and fewer points in regions where there are fewer transactions. A procedure to determine an optimal grid accounting for both the precision with which the points on the grid are estimated and the approximation error in the integration, is beyond the scope of this paper. While we do not know what an 'optimal' grid looks like, we note that it is unlikely to matter in practice since we use the high density of transactions between the first and last decile of capital to create an extremely fine grid along the K axis with capital increments of less than 0.12%. This allows us to exploit precise estimates for our grid while keeping the approximation error from integration low.

confidence intervals of the estimated coefficients by bootstrap. At each iteration, we draw with replacement a random sample from the universe of all transacted land parcels. We recompute parcel prices at each point of our (K,T) grid using kernel non-parametric regressions before re-estimating housing production and regressing it on capital investment. Confidence intervals can be deduced from the re-estimated coefficients.

Finally, in the last step, we extend our approach to handle estimation issues related to factor heterogeneity. We describe our implementation in the case of parcel characteristics y that directly affect the production of housing. For heterogeneity in the price or in the composition of housing capital, we follow a similar approach. Consider for instance that some parcels are more costly to develop because of a steep slope or because their soil is harder to excavate. For a given unit price of housing, these parcel are worth less in equilibrium and our base approach will fail to recover the appropriate production of housing from the data. As argued above, we can compute housing production if we purge our variables of interest, K and R, from the effects of supply characteristics and rely only on the variation in demand for housing. Consider the regressions:

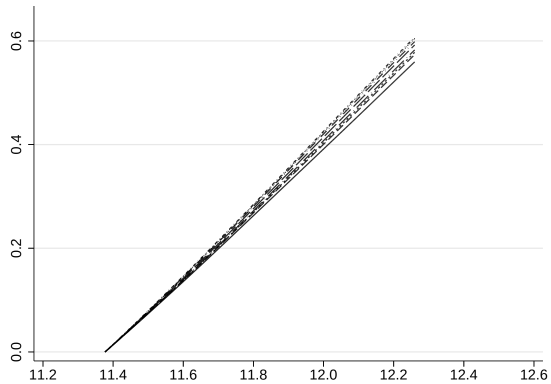
$$\log Z_i = X_i\, a^Z + Y_i\, b^Z + f^Z(T_i) + \epsilon_i^Z \quad \text{for} \quad Z = K,R\,. \tag{12}$$

In equation (12), X is the vector of characteristics that (are assumed to) affect the demand for housing, Y is a vector of characteristics that potentially affect the supply of housing, $f^Z(T)$ is a non-parametric function of T, and ϵ_i^Z is an error term.

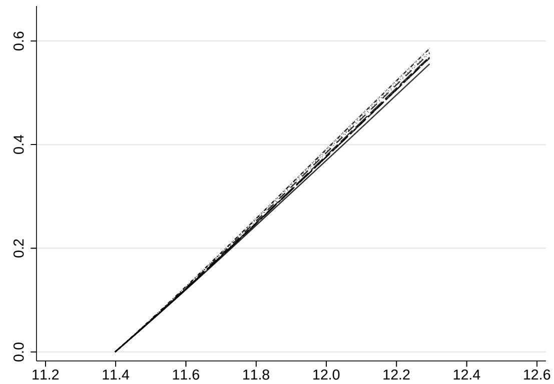
The vector X is the empirical counterpart of x in our framework above while Y is the empirical counterpart of y. To estimate $f^Z(T)$, we use indicator variables for every parcel size centile. Then, under the assumption that the residuals are normally distributed, we compute unbiased predicted values $\widehat{Z}_i = \exp(X_i \hat{a}^Z + \overline{Y}\,\hat{b}^Z + \hat{f}^Z(T_i) + (\hat{\sigma}^Z)^2/2)$ for $Z = K,R$ where $\overline{Y}\,\hat{b}^Z$ is the mean effect of supply characteristics and $\hat{\sigma}^Z$ is the estimator of the standard deviation of the error term in equation (12). These predicted values for capital and parcel prices depend only on demand characteristics X and not on supply characteristics Y. They are then used to estimate parcel values non-parametrically in equation (10). We postpone the important discussion of the variables included into X and Y.

Our model takes parcel size T as given. However, the location characteristics that affect the cost of construction may also affect parcel size. For instance, parcels may be larger where construction is more costly. This suggests applying the same approach to parcel size. However, we need to

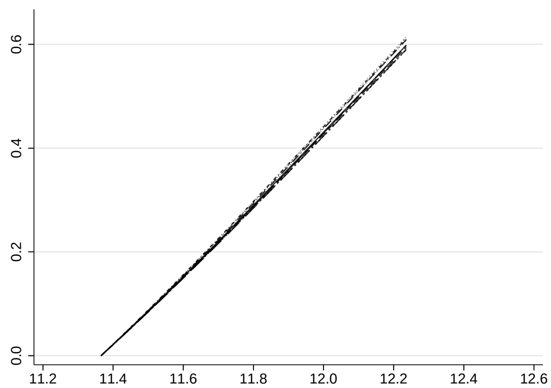
Figure 2: log housing production as a function of log capital investment, non-parametric estimates



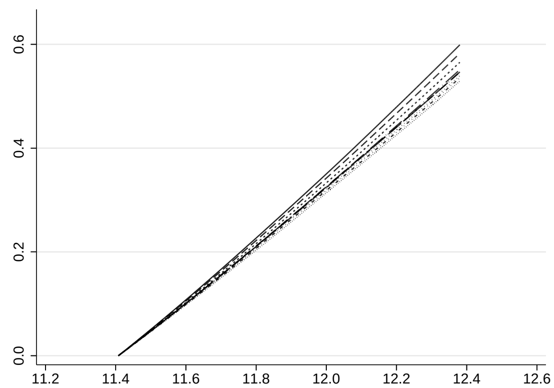
Panel (A) Entire country



Panel (B) All urban areas



Panel (C) Urban areas, 50,000-100,000



Panel (D) Urban areas, more than 500,000 (excl. Paris)

Notes: The log of housing production is represented on the vertical access and the log of capital investment is represented on the horizontal access. To ease the comparison across deciles of parcel size, we normalise log $H(\underline{K})$ to zero for all deciles. 386,177 observations for the entire country and 218,767 for urban areas.

amend equation (12) and consider instead,

$$\log Z_i = X_i\, a^Z + Y_i\, b^Z + \epsilon_i^Z \quad \text{for} \;\; Z = K,R,T\,, \tag{13}$$

where T is now an endogenous variable and it is no longer included as determinant of K^* and R.

5. Results

5.1 *Base results*

Before looking at formal estimation results, it is useful to visualise our non-parametric estimations. Each panel of figure 2 plots the estimated log production of housing, log H, as a function of capital investment, log K, for every decile of parcel size, T. This is the empirical counterpart to equation (5) after normalising $\log H(\underline{K})$ to zero for all deciles. Panel (A) represents the production function for housing for the entire country while panels (B), (C), and (D) do the same for all urban areas, small urban areas with population between 50,000 and 100,000 and large urban areas with population above 500,000 (bar Paris), respectively. We obtain similar patterns for other urban area size classes.

Although we must remain cautious with visual impressions, several remarkable features emerge from figure 2. First, as might be expected, housing production always increases with capital. More specifically, log housing is an apparently linear function of log capital with a slope of about 0.65. This is of course consistent with a Cobb-Douglas function with an elasticity of housing production with respect to capital of about 0.65. In turn, this value suggests an elasticity of housing production with respect to housing prices of $0.65/(1-0.65) \approx 1.86$.[17] Second, the relationship between $\log H$ and $\log K$ appears very similar across all deciles of parcel size. The last important feature of figure 2 regards the differences across panels. While the relationship between $\log H$ and $\log K$ is very much the same across the first three panels, the last panel for large urban areas is modestly different with more dispersion across deciles and a less steep slope on average.

We next turn to regressions to describe these features more precisely. Our first set of results is reported in panel (A) of table 2 where, for each decile of parcel size, we regress our non-parametric estimates of the log production of housing on log capital for observations located in urban areas. Each regression relies on 900 observations obtained after smoothing parcel prices as per equation (9) at values of capital between its first and last deciles. Each column of table 2 corresponds to a separate decile of parcel size. The estimated capital elasticity of housing for the first decile is 0.62. It is 0.64 for the second to the fifth decile, 0.65 for the seventh and eighth, and finally 0.66 for the last decile. These findings point at some apparent log supermodularity in the production of housing and reject Cobb-Douglas. This said, these differences in capital elasticity between deciles of parcel size are economically small. Finally, we note that our linear regressions provide a near perfect fit as the R^2 is always above 0.999.[18]

Panel (B) of table 2 replicates the regressions of panel (A) adding the square of log capital as explanatory variable. The estimated coefficient of the quadratic term is significant in all regressions

[17] We keep in mind that this is a micro-elasticity. The aggregate elasticity of housing supply also depends on the number of parcels to be developed and the choice of mode of development (single- vs. multi-family).

[18] Recall that we work with smoothed data, which condition out idiosyncratic variation. To be clear, this R^2 does not measure how well our regression fits the raw data but how well the functional form imposed by the regression fits the non-parametric estimate of the housing production function.

Table 2: log housing production in urban areas, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A)									
log (K)	0.624[a]	0.637[a]	0.639[a]	0.638[a]	0.642[a]	0.650[a]	0.653[a]	0.659[a]	0.661[a]
	(0.00092)	(0.00084)	(0.00082)	(0.00087)	(0.00098)	(0.0014)	(0.0017)	(0.0022)	(0.0027)
	[0.00090]	[0.00078]	[0.00081]	[0.00090]	[0.0011]	[0.0014]	[0.0018]	[0.0022]	[0.0026]
R^2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Observations	900	900	900	900	900	900	900	900	900
Panel (B)									
log (K)	0.114[a]	-0.023	-0.112[a]	-0.033	-0.017	0.085	0.266[a]	0.232[b]	0.257[b]
	(0.041)	(0.028)	(0.029)	(0.044)	(0.053)	(0.070)	(0.088)	(0.103)	(0.141)
	[0.040]	[0.028]	[0.031]	[0.041]	[0.052]	[0.069]	[0.087]	[0.107]	[0.130]
$[\log (K)]^2$	0.021[a]	0.028[a]	0.032[a]	0.028[a]	0.028[a]	0.024[a]	0.016[a]	0.018[a]	0.017[a]
	(0.0017)	(0.0012)	(0.0012)	(0.0019)	(0.0023)	(0.0030)	(0.0037)	(0.0044)	(0.0060)
	[0.0017]	[0.0012]	[0.0013]	[0.0017]	[0.0022]	[0.0029]	[0.0037]	[0.0045]	[0.0055]
R^2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Observations	900	900	900	900	900	900	900	900	900

Notes: OLS regressions with a constant in all columns. Bootstrapped standard errors with 100 bootstraps in parentheses and with 1,000 bootstraps in squared parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10% (for 1,000 bootstraps). Non-parametric estimates of housing production rely on 218,767 observations.

with a coefficient between 0.016 and 0.032. Together with the differences in coefficients across deciles of parcel size in the previous panel, this is another indication that the production function for housing is not strictly log linear in capital. Instead, housing is a log convex function of capital, but only modestly so. For the third decile of parcel size for which log convexity is strongest, our estimates imply that the capital elasticity of housing is only about 0.05 larger for houses built at the top decile of capital (where $\log K \approx 12.2$) relative to houses built at the bottom decile (where $\log K \approx 11.4$).[19]

Before turning to various forms of heterogeneity, we discuss four technical issues. First, table 2 reports two series of standard errors with 100 and 1,000 bootstraps. Taking 1,000 bootstraps instead of 100 does not make any substantive difference. Because these bootstraps are computationally intensive, we only report standard errors computed from 100 bootstraps in subsequent tables. Second, we verify that our results are not affected by our choice of bandwidth. Table 9 in Appendix D repeats the estimations of table 2 for bandwidths equal to a half, a quarter, and a tenth of the

[19] At this point, we only note two modest deviations from Cobb-Douglas but do not explore functional form issues further since our findings may be contaminated by unobserved factor heterogeneity. We return to functional form approximations in section 6.

rule-of-thumb bandwidth, respectively. The results are virtually identical. This shows that we need to consider extreme forms of under-smoothing before running into problems where 'holes' in the data are no longer properly smoothed away. Third instead of ignoring all parcels in the bottom and top deciles as in table 2, table 10 in Appendix D considers a broader range of observations where we sort parcels by ascending investment and eliminate 3% of aggregate land values in each tail. Despite a loss of precision, this table shows that the results of table 2 are generally robust to considering more extreme regions of the data where observations are sparser. Finally, table 11 in Appendix D duplicates table 2 but smooths the cost share $K^*/(K^* + R)$ directly instead of smoothing R prior to computing the cost share. This alternative leads to a minor change to the results as we estimate capital elasticities that are 0.02 to 0.03 larger relative to those of table 2. All the other features of table 2 remain.

5.2 *Factor heterogeneity: Results using predicted values for K^*, R, and T*

We now turn to factor heterogeneity. In section 2, we proposed two strategies to avoid biases caused by factor heterogeneity. The first is to use predicted values of K^*, R, and T after conditioning out potential supply determinants to keep only the variation coming from demand factors. The second strategy is to use homogeneous sub-samples for which factor heterogeneity is less of a concern. In this subsection, we present results using predicted values, first for K^* and R using equation (12) and then for R, K^*, and T using equation (13). We explore results for more homogeneous sub-samples in the next subsections.

As sources of exogenous demand variation that determine the price of housing P, and in turn K^* and R, we rely on the urban area to which a parcel belongs and the distance to its centre. This is in the spirit of simple monocentric urban models (Alonso, 1964, Muth, 1969, Mills, 1967). In simple versions of these models, the price of housing, land prices, and capital investment at each location are fully explained by the distance to the centre and urban area population. We also use measures of local income, a demand factor in more elaborate models of urban structure with heterogeneous residents (Duranton and Puga, 2015). More specifically, in the demand vector X of equation (12), we include urban area fixed effects, distance to the centre (allowing the effect to vary across urban areas), and three municipal socioeconomic characteristics (log mean income, its standard deviation, and the share of population with a university degree). Although we do not develop a procedure to assess the predictive power of our demand-related variables, there is little doubt that these variables strongly predict our quantities of interest. Combes *et al.* (2019) find that urban area fixed effect and (log) distance to the centre explain 63% of the variation of the price of land per square metre in France.

We worry nonetheless that the variables entering our demand vector may be correlated with the ease of building and other forms of factor heterogeneity. For instance, terrain characteristics may vary systematically with distance to the centre. To avoid this issue, we include a number of geographic characteristics as part of our vector of potential supply characteristics Y to be conditioned out in the estimation of equation (12). More specifically, in the supply vector Y, we include seven geological variables (ruggedness, and three classes of soil erodability, soil hydrogeological class, and soil dominant parent material) and three land use variables (share of built-up land, share of urbanised land, and share of agricultural land).

In addition, we want to condition out the local wage of workers in the construction industry since construction costs may vary across urban areas (Gyourko and Saiz, 2006). We cannot include urban area wages in the construction industry directly among the explanatory variables of equation (12) because they would be collinear with urban area fixed effects. Instead, we regress urban area fixed effects on wages in the construction industry and retain the estimated residual in the vector of demand factors X. This amounts to conditioning out construction wages from our predicted values for K^*, R, and T.

For our preferred set of demand-related factors used to predict K^* and R, panel (A) of table 3 reports a first series of estimations. The estimated capital elasticities of housing are close to those of table 2 but slightly higher by 0.01 to 0.02. While some of these differences are significant in a statistical sense, they are economically small. We also still observe the same pattern of higher elasticity of housing production as higher deciles of parcel size are considered.

Panel (B) of table 3 adds a quadratic term for $\log K$. The results indicate the presence of a mild log concavity for all deciles of parcel size except the first one. This is in contrast with table 2 where the results point towards modest log convexity. While there is some variation in the estimated coefficient on the quadratic term in $\log K$ across deciles of parcel size, the largest one in absolute value is 0.07 for the fifth decile of parcel size. With $\log K$ varying from 11.4 to 12.2 between the bottom and top decile of capital, this log concavity implies a capital elasticity of housing production of about 0.70 for the bottom decile of capital and 0.60 for the top decile.

Table 3: log housing production in urban areas obtained from predicted values of variables, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Predicted R and K									
$\log(K)$	0.645^a	0.647^a	0.649^a	0.652^a	0.658^a	0.666^a	0.670^a	0.676^a	0.683^a
	(0.0010)	(0.00061)	(0.00061)	(0.00061)	(0.00086)	(0.0011)	(0.0015)	(0.0020)	(0.0023)
Panel (B): Predicted R and K									
$\log(K)$	0.105	0.954^a	1.633^a	2.095^a	2.257^a	2.217^a	1.773^a	1.647^a	1.880^a
	(0.129)	(0.080)	(0.077)	(0.075)	(0.087)	(0.128)	(0.169)	(0.205)	(0.226)
$[\log(K)]^2$	0.023^a	-0.013^a	-0.042^a	-0.061^a	-0.068^a	-0.066^a	-0.047^a	-0.041^a	-0.051^a
	(0.0055)	(0.0034)	(0.0032)	(0.0032)	(0.0037)	(0.0054)	(0.0071)	(0.0086)	(0.0095)
Panel (C): Predicted R, K, and T									
$\log(K)$	0.652^a	0.664^a	0.654^a	0.653^a	0.650^a	0.645^a	0.648^a	0.663^a	0.677^a
	(0.0018)	(0.0018)	(0.0025)	(0.0025)	(0.0026)	(0.0021)	(0.0033)	(0.0026)	(0.0032)
Panel (D): Predicted R, K, and T									
$\log(K)$	0.006	0.734^a	1.080^a	2.342^a	3.304^a	3.007^a	3.365^a	2.456^a	2.250^a
	(0.407)	(0.240)	(0.379)	(0.368)	(0.411)	(0.408)	(0.464)	(0.415)	(0.385)
$[\log(K)]^2$	0.027	-0.003	-0.018	-0.071^a	-0.112^a	-0.100^a	-0.115^a	-0.076^a	-0.066^a
	(0.010)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.020)	(0.018)	(0.016)

Notes: OLS regressions with a constant in all columns. 900 observations for each regression. The R^2 is 1.00 in all specifications. Capital and parcel price are constructed using urban area fixed effects (after conditioning out construction wages), distance to the centre (urban-area specific), and income variables (log mean municipal income, log standard error of income, and share of population with a university degree). We also condition out the effect of the seven geological variables and the three land use variables listed in the text. For parcel size, observed values are used in panels (A) and (B), and values predicted from the same variables as we use to predict capital and parcel price are used in panels (C) and (D). Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. Non-parametric estimates of housing production rely on 213,786 observations (instead of 218,767 when we do not use predicted values of land prices and capital).

Before interpreting these findings further, we confirm them in a variety of ways. Panels (C) and (D) of table 3 duplicate the previous two panels but also consider that parcel size is affected by supply factors and net out their effects. The results are qualitatively similar and quantitatively close. In Appendix E, table 12 report results experimenting with the variables we include in the vectors X and Y of demand and supply factors. The similarity with table 3 shows that our results are not sensitive to the exact details of what we include and exclude to predict K^*, R, and T.

Overall, our results using predicted values for K^* and R suggest a marginally higher elasticity of housing production with respect to capital. The differences with our base results of table 2 are nonetheless too small to be economically meaningful. More importantly, predicted values for K^*

and R imply that the production function for housing is mildly log concave rather than log convex when using observed quantities. These new results imply an elasticity of substitution between capital and land slightly below 1 instead of slightly above 1. This difference makes intuitive sense in relation to the possible biases described above. For parcels of the same size, we expect parcels that are more difficult to build to require more capital. The price of these parcels will then be lower due to this and the share of capital will thus be higher. This can bias our results and generate an apparent log convexity for the production function of housing when we do not control for supply factors.

5.3 *Factor heterogeneity: Results by location*

We now examine differences across locations. We first assess differences in capital intensity for new constructions depending on how far they are from the center of their urban area. We measure distances in relative terms. While a location five kilometers from the centre is still 'central' in large urban areas, it is often 'peripheral' in small urban areas. Table 13 in Appendix F reports a moderately greater capital intensity for new constructions in more peripheral locations. Since parcels are larger in more peripheral locations, these results are consistent with the greater capital intensity for larger parcels reported in tables 2 and 3.

Next, table 4 reports results for different classes of urban areas, grouped by population size. Panel (A) regresses again the log of housing production on log capital. Unlike previous tables, in each regression we pool observations for all deciles of parcel size and include decile fixed effects. The first column considers the entire population of transactions. The estimated capital elasticity of housing is 0.65, about equal to that of the middle decile in table 2. Column 2 uses only observations from urban areas and estimates a similar elasticity. The following six columns consider urban areas of increasing sizes. For the smallest urban areas with population below 50,000 the estimated capital elasticity is 0.71. This elasticity falls to 0.58 for large urban areas with population above 500,000 and 0.54 for Paris.

These differences in capital elasticities across urban areas are confirmed by the other panels of table 4. Table 14 in Appendix F, which provides more detailed results by decile of parcel size for each size class of urban area, further corroborates these differences in capital intensity across urban areas. This appendix table also shows that the differences in capital intensity across deciles of parcel size are generally half or less the already modest differences reported in table 2. This

Table 4: log housing production, OLS by class of urban area population

City size class	Country	Urban areas	0-50	50-100	100-200	200-500	500+	Paris
Panel (A): Observed data								
log (K)	0.645[a]	0.645[a]	0.712[a]	0.698[a]	0.686[a]	0.646[a]	0.575[a]	0.539[a]
	(0.00088)	(0.00087)	(0.0020)	(0.0022)	(0.0015)	(0.0014)	(0.0017)	(0.0026)
Panel (B): Observed data								
log (K)	0.082[a]	0.086[a]	-0.605[a]	-0.419[a]	-0.135[a]	-0.317[a]	-0.443[a]	-0.288[a]
	(0.039)	(0.030)	(0.078)	(0.079)	(0.056)	(0.054)	(0.055)	(0.087)
$[\log(K)]^2$	0.024[a]	0.024[a]	0.056[a]	0.047[a]	0.035[a]	0.041[a]	0.043[a]	0.034[a]
	(0.0017)	(0.0013)	(0.0033)	(0.0033)	(0.0024)	(0.0023)	(0.0023)	(0.0036)
Panel (C): Predicted data								
log (K)	-	0.661[a]	0.727[a]	0.708[a]	0.698[a]	0.663[a]	0.582[a]	-
	-	(0.00062)	(0.0012)	(0.0013)	(0.0011)	(0.0013)	(0.0011)	-
Panel (D): Predicted data								
log (K)	-	1.618[a]	0.976[a]	1.142[a]	2.345[a]	1.943[a]	1.349[a]	-
	-	(0.078)	(0.197)	(0.269)	(0.130)	(0.185)	(0.080)	-
$[\log(K)]^2$	-	-0.040[a]	-0.011	-0.018[c]	-0.070[a]	-0.054[a]	-0.032[a]	-
	-	(0.0033)	(0.0084)	(0.011)	(0.0055)	(0.0078)	(0.0034)	-

Notes: OLS regressions with decile fixed effects in all columns. 8,100 observations for each regression. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. In panels (C) and (D), capital and parcel price are constructed using urban area fixed effects (after conditioning out construction wages), distance to the centre (urban-area specific), and income variables. Observed values of parcel size are used. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. We cannot report results for the entire country given that our construction of capital and land price relies on urban area fixed effects and distance to the centre, which are unavailable in rural areas. Similarly we cannot implement our approach when we consider Paris alone.

suggests that the greater capital intensity of larger parcels found above reflects in part the greater capital intensity of housing construction in smaller urban areas where parcels are larger.

To explain these sizeable differences in capital intensity across urban areas, we first dismiss a number of candidate explanations before turning to our preferred one. First, this heterogeneity across urban areas may be caused by a complementarity between land and capital leading to less capital investment where land is more expensive. Appendix G shows that this explanation requires an implausibly low elasticity of substitution between land and capital of about 0.2.

Next, this heterogeneity across urban areas is also unlikely to be caused by differences in construction costs. First, differences in capital intensity still occur in panels (C) and (D) of table 4 where we use predicted values of K^* and R or (K^*, R, and T) which condition out the wage of construction workers. Second, relative to small urban areas, construction wages are 14.2% higher in Paris and 5.8% higher in large urban areas. Simple calculations in Appendix G show that for these modest differences in construction wages to explain the large difference in capital elasticity between these two groups of urban areas, the elasticity of substitution between land and capital would need to be, this time, implausibly high (above three). Third, we verify in Appendix F table 15 that the results of table 2 are unaffected when we either condition out local construction wages or directly deflate them from capital investment.

Finally, we also show below that differences in land use regulations across urban areas are also unlikely to provide an explanation for the differences in capital elasticity across urban areas that we estimate.

Our preferred explanation relies instead on differing growth rates for housing rents across urban areas. When housing capital can be adjusted over time, higher future housing rents do not affect current housing investments. They only lead to greater future housing investments, after their increase. At the same time, current land prices already capitalise future housing rents. Hence, the share of capital in housing may differ across urban areas following differences in future housing rent growth.

To formalize this intuition, we propose a two-period extension of our model in Appendix H. We show that when housing capital depreciates and can be adjusted at each period, the elasticity of housing production with respect to capital is no longer equal to the share of capital in total costs. Instead, the share of capital is multiplied, first, by a discount factor which accounts for the interest rate and capital depreciation and, second, by the ratio of housing values to housing rents at the current period. Importantly, this ratio, which captures expected changes in housing rents, can be measured in the data.[20]

Appendix I provides justifications for our choices of 4% for the annual interest and 1% for the annual rate of capital depreciation. This appendix also provides details regarding our estimation of the ratio of housing values to annual housing rents for each class of urban area population. We estimate values ranging from 18.1 in small urban areas to 24.1 in Paris. We note that these ratios

[20]Appendix H also shows that the capital elasticity in this dynamic extension can also be expressed as a user-cost corrected share where capital is adjusted by the sum of the interest rate, r, and capital depreciation τ, while land values are adjusted by the difference between the interested rate and their expected rate of appreciation, g. The elasticity of housing production with respect to capital is then given by $(r+\tau)K^*/[(r+\tau)K^* + (r-g)R]$. While this expression makes intuitive sense, implementing it requires a measure of the expected growth rate of land prices, g. This quantity can only be obtained indirectly using housing rents and housing values, which takes us back to the expression we use.

Table 5: log housing production, OLS by class of urban area population with corrected cost shares

City size class	0-50	50-100	100-200	200-500	500+	Paris
Panel (A): Observed data						
log (K)	0.639[a]	0.632[a]	0.638[a]	0.632[a]	0.644[a]	0.643[a]
	(0.0018)	(0.0020)	(0.0014)	(0.0014)	(0.0019)	(0.0031)
Panel (B): Observed data						
log (K)	-0.543[a]	-0.379[a]	-0.126[b]	-0.310[a]	-0.496[a]	-0.344[a]
	(0.070)	(0.071)	(0.052)	(0.053)	(0.062)	(0.103)
$[\log(K)]^2$	0.050[a]	0.043[a]	0.032[a]	0.040[a]	0.048[a]	0.041[a]
	(0.0029)	(0.0030)	(0.0022)	(0.0022)	(0.0026)	(0.0043)
Panel (C): Predicted data						
log (K)	0.652[a]	0.641[a]	0.649[a]	0.649[a]	0.651[a]	-
	(0.0011)	(0.0012)	(0.0010)	(0.0012)	(0.0012)	-
Panel (D): Predicted data						
log (K)	0.876[a]	1.034[a]	2.178[a]	1.902[a]	1.510[a]	-
	(0.177)	(0.244)	(0.121)	(0.181)	(0.089)	-
$[\log(K)]^2$	-0.009	-0.017	-0.065[a]	-0.053[a]	-0.036[a]	-
	(0.0075)	(0.010)	(0.0051)	(0.0076)	(0.0038)	-

Notes: OLS regressions with decile fixed effects in all columns. 8,100 observations for each regression. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. This table duplicates table 5 for the same six size classes of urban areas but uses the correction described in the text with details provided in Appendix I.

are consistent with findings from prior research (Chapelle and Eyméoud, 2017).

Table 5 duplicates the estimations of table 4 for the same six size classes of urban areas. The estimated elasticities in panel (A) are now all tightly centered around 0.64 with no systematic variation across size classes of urban areas. This stability is confirmed by the other panels of the same table.

While the raw cost shares increase from 0.54 in Paris to 0.71 in small urban areas, the value-to-rent ratio varies in the opposite direction with similar magnitudes. Hence, when computing our corrected cost shares, these two differences offset each other nearly exactly. Less obviously, the two terms in our correction, the discount factor common to all urban areas of about 5% and the value-to-rent ratio of nearly 20 on average also offset each other so that, on average for the country, the estimated capital elasticity is essentially unchanged relative to the results reported above.

5.4 Factor heterogeneity: Land use regulations

In France, housing development is regulated like in many other countries. The three main regulatory instruments during our study period are (i) the zoning designation, (ii) minimum lot size and severe restrictions on parcel division, and, most importantly for our purpose, (iii) the maximum intensity of development.[21]

Starting with the zoning designation, it indicates whether a parcel can be developed and, if yes, whether this can be for residential purpose. Given that we only observe parcels with a development permit for a single-family home, this creates no further issue for us beyond the fact that we estimate the production function for single-family homes on parcels allowing this type of development. Then, restrictions on parcel divisions are the main reason why we take parcel size as given.

Turning to the maximum intensity of development that applies to a parcel, it is essentially a maximum floor-to-area ratio (FAR).[22] While the information about maximum FARs is not centrally collected by the French government, we know the actual FAR of each newly built house from our main data. We also know the FAR for each and every single-family house in France from exhaustive data about all buildings (see Appendix C for details).

We propose a measure the *absolute* FAR stringency that applies to a new construction by taking the 30th percentile of the distribution of FARs of all single-family homes in the same municipality.[23] This measure relies on the notion that if a new construction is subject to a more stringent maximum FAR, its neighbors will have been subject to a more stringent maximum FAR as well and will thus exhibit a lower realized FAR.

A limitation here is that a maximum FAR of, say, 0.8 where parcels are tiny may be more constraining than a maximum FAR of 0.2 where parcels are large. To avoid this pitfall, we can also measure the *relative* stringency of the local FARs for new constructions. To do this, we compute the centile to which any new construction belongs in its municipal distribution of FARs for all single-family homes.

[21]We ignore constraints arising from the building codes and local regulations forcing particular architectural styles or the use of specific construction materials. We think of these constraints as a form of supply heterogeneity, which we explored above.

[22]Building height is also regulated but this should not be binding for single-family homes since there is no constraint on the share of a parcel that is built-up until the last year of our study period.

[23]We choose the 30th percentile as it roughly corresponds to the median FAR percentile of newly-built houses in the distribution of all homes in the same municipality.

With these measures at hand, we then re-estimate our base results for separate quintiles of absolute and relative FAR stringency. Tables 16 and 17 in Appendix J report the results. First, we find that the capital elasticity for houses built in municipalities with a higher FAR is modestly smaller. This is likely because high FARs disproportionately occur in larger urban areas, where the capital elasticity is generally lower as documented in table 4 above.[24] We also find that the capital elasticity for newly-built houses with a higher FAR relative to their neighbours is only modestly larger. Taken together, these results suggest that FAR limits only have a minor effect on the production of housing, either because they may not bind much or because the capital used to produce more floorspace and that is used to produce better quality housing are fairly substitute.

To provide further evidence about the small effects of land use regulations on how housing is produced, we conduct two further checks. First, we take advantage of a change in land use regulations starting in 2012 after which new constructions are also subject to a building coverage ratio. Table 18 in Appendix C duplicates our baseline analysis for 2006-2011 and 2012 separately and fails to uncover any difference in capital intensity for new constructions between these two subperiods. Second, we regress log housing on log capital and deciles of parcel sizes as previously but also include indicators for quintiles of municipal FAR. For the same parcel size and capital invested, we find that the amount of housing is only about 1% higher in the top quintile of municipality FAR relative to the bottom quintile.

Obviously, these findings do not imply that land use regulations are irrelevant. They indicate instead that we fail to find evidence about an important role for land use regulations to explain the differences in capital share across locations documented above. More generally, we feel that land use regulations in France strongly limit where and whether new developments may occur but only impose modest constraints on what can be built when a parcel is constructible.

5.5 *Housing heterogeneity*

Aside from factor heterogeneity, we also worry about housing heterogeneity since houses in the data are built for specific buyers. These buyers may have idiosyncratic preferences, which could affect construction costs. Because the information we have about construction costs is for one of three levels of completion, we can compare results across these levels of completion (fully finished

[24] This result also runs contrary to the prediction of the extension of our model proposed in Appendix A where higher FARs should lead to a higher capital elasticity in the Cobb-Douglas case as capital can be deployed with fewer constraints.

units, ready-to-decorate units, and units with only a bare-bone structure). Any unobserved heterogeneity associated with the customisation of houses should have a greater effect on fully finished units than on bare-bone structures. Table 19 in Appendix K duplicates panel (A) of table 2 but splits observations by level of completion. Unsurprisingly, we find a slightly higher capital elasticity for houses at a more advanced degree of completion. For the median parcel, the capital elasticity is 0.66 for fully finished units, 0.64 for ready-to-decorate units, and 0.61 for units with only a bare-bone structure. The corresponding elasticity in table 2 is 0.64. For all levels of completion, we find again a modestly increasing capital elasticity as we consider higher parcel size deciles as in table 2.

While we cannot track the heterogeneity of houses directly, it may be reflected in the heterogeneity of buyers. We can split the sample of transactions by buyers' occupation: executives, intermediate occupations, and clerical and blue-collar workers. We report results for these three groups in table 20 in Appendix K. The differences between occupational categories are small. For the median parcel, the capital elasticity is 0.63 for executives, 0.64 for intermediate occupations, and 0.65 for clerical and blue-collar workers with the same general pattern of modestly increasing elasticities as we consider larger parcels.

6. Functional forms and comparisons with alternative approaches

So far, we have non-parametrically estimated the production of housing as a function of capital for a given parcel before using simple regressions to assess the shape of this non-parametric function. While the production function of housing can be described by a Cobb Douglas function with a coefficient on capital of about 0.65, a more detailed look suggests some mild log convexity when using our base approach and, perhaps more reasonably, modest log concavity when we rely on predicted values for parcel prices and capital. In this section, we asses a variety of functional forms for the production function of housing. This is useful to characterise our results further and to compare with alternative approaches.

6.1 Assessing functional forms and recovering their parameters

We proceed as follows. First, we consider a specific parametric form for the production function and recover the underlying parameters by fitting the theoretical cost shares to their empirical counterparts. Next, we use the estimated parameters to compute the value of the parametric production function at each point of our grid. Then, we duplicate our estimation of the capital elasticity for each decile of parcel size. Finally, we compare the results we obtain using pre-imposed functional forms to our earlier, non-parametric estimations results.[25] Let us develop these four steps in turn.

For the exposition to remain concrete, consider a CES production function. The production of housing is given by $H = A\left(\alpha K^{(\sigma-1)/\sigma} + (1-\alpha)T^{(\sigma-1)/\sigma}\right)^{\sigma/(\sigma-1)}$ where σ is the elasticity of substitution between land and capital and A is a productivity shifter. Using equation (4) and the partial derivative of the CES production function with respect to K, we obtain the following cost share:

$$\frac{K^*}{K^* + R(K^*,T)} = \frac{\alpha(K^*)^{1-1/\sigma}}{\alpha(K^*)^{1-1/\sigma} + (1-\alpha)T^{1-1/\sigma}}\,. \tag{14}$$

From values on a 300×300 grid, we can estimate α and σ using equation (14) by minimising the sum of the squared distances between empirical costs shares and those predicted by a CES production function.[26] We then compute the "CES productions" of housing at the points of our grid using the estimated values for α and σ and perform the same regressions as in table 2. We also repeat the same exercise using values of parcel prices and capital predicted by demand factors as in table 3. Beyond the CES, we also assess Cobb-Douglas and second- and third-order translog production functions.

Tables 21 and 22 in Appendix L report results for these four functional forms. These two tables duplicate the base results of table 2 and those of table 3 for predicted values for K^* and R, respectively. For Cobb-Douglas production functions, we recover a capital elasticity of 0.634, close to the cross-decile mean of the coefficients estimated in table 2. Using predicted values, the capital elasticity is 0.652 and again close to the corresponding mean in table 3. Obviously, the Cobb-Douglas form is unable to replicate any of the modest differences in capital elasticity we observe across deciles of parcel size in tables 2 and 3. This functional form also fails to replicate the slight log convexity estimated in table 2 and the slight log concavity in table 3.

[25] An alternative to what we do here would be to use standard specification tests to isolate the 'best' functional form. This would be problematic in our case. First, standard specification tests provide an arbitrary resolution to the tradeoff between goodness of fit and the number of parameters. Our non-parametric smoothing yields a very high R^2, even with the simplest Cobb-Douglas specification, making this tradeoff one sided. Second, these tests would only assess the fit with respect to K and not with respect to T.

[26] While for expositional reasons we report results for only 9 deciles of parcel size and estimate regressions for 900 capital points for each decile, we prefer to use a grid that treats both factors symmetrically to estimate functional forms. To take into account the distribution of observations in the data, we also weight observations with the kernel weights used for determining the land price on the grid. Finally, taking a finer grid does not affect the point estimates, but makes the computation of standard errors highly time-consuming

For the CES case, the estimated parameter values for the production function are $\alpha = 0.601$ and $\sigma = 1.028$ when using observed values of K and R and $\alpha = 0.767$ and $\sigma = 0.900$ when using predicted values for K and R. These values of the elasticity of substitution σ close to one confirm that the Cobb Douglas form provides a good approximation. CES functions provide a better fit since they are able to duplicate finer features of our non-parametric results such as the tendency of the capital elasticity to be larger for higher deciles of parcel size. The CES functions we estimate also display the log convexity and concavity estimated in tables 2 and 3, albeit attenuated. With more parameter, and thus more flexibility to fit the data, the second- and third-order translog production functions can match the patterns of non-parametric results reported in tables 2 and 3 even more closely. Despite this, we can reject that the coefficients we estimate when regressing $\log H$ and on $\log K$ with a translog are statistically equal to those we estimate after fitting the data non-parametrically. This said, despite these statistical differences, the coefficients are economically very close.

We draw three conclusions from this analysis. First, we confirm that a Cobb-Douglas specification provides a good first-order description of the data. Second, and consistent with this, we find that the estimation of a CES production function for housing implies an elasticity of substitution between land and capital inputs close to one, either 0.90 or 1.03 depending on whether we predict capital and parcel prices with demand factors. Overall, the third-order translog offers the closest approximation to our non-parametric results but the gain from this more flexible functional form relative to the Cobb-Douglas case remains small. Third, none of the functional forms we consider is able to match the results of our non-parametric estimation exactly. Put differently, these results suggest that it is better to use a non-parametric approach and then provide a functional form approximation than impose a functional form directly into the estimation.

6.2 *Comparisons with alternative methodologies*

We now compare our results to those obtained from alternative approaches. Given past literature, we focus on the estimation of CES production functions and use the CES approximation of our non-parametric approach as benchmark. The first alternative is the traditional regression of $\log(K/T)$ on $\log(R/T)$, which comes out of a rewriting of equation (14). We also duplicate the approach of Ahlfeldt and McMillen (2020). They follow Epple *et al.* (2010) to obtain an estimate of $\log(K/T)$, which they then regress on $\log(R/T)$ as in the traditional regression. Finally, closer to the spirit of

Table 6: Comparison across alternative approaches

	(1) Traditional	(2) Traditional smoothed	(3) EGS AM	(4) EGS AM smoothed	(5) Cost share	(6) Cost share smoothed	(7) Our approach
Panel (A): Observed data							
σ	0.490	0.994	0.745	1.043	1.046	1.102	1.028
	(0.0015)	(0.0027)	(0.013)	(0.030)	(0.0028)	(0.0038)	(0.0048)
α	0.997	0.631	0.915	0.573	0.600	0.509	0.601
	(0.0003)	(0.0033)	(0.0096)	(0.0341)	(0.0033)	(0.0042)	(0.0057)
Panel (B): Predicted data							
σ	0.644	0.933	0.816	0.952	0.963	0.962	0.900
	(0.0024)	(0.0033)	(0.0025)	(0.0032)	(0.0032)	(0.0032)	(0.0046)
α	0.971	0.729	0.860	0.707	0.704	0.695	0.767
	(0.0008)	(0.0039)	(0.0023)	(0.0038)	(0.0037)	(0.0038)	(0.0052)

Notes: 218,767 observations in columns (1)-(6) of both panels and 90,000 in column (7). Observed parcel prices and capital investment are used for regressions in panel (A). Predicted parcel prices and capital investment as per table 3 are used for regressions in panel (B). The coefficients of panels (A) and (B) are obtained from OLS regressions in columns (1)-(4) (including for the Ahlfeldt-McMillen estimates, AM), OLS regressions weighted by the sum of kernel weights in column (7), and non-linear least squares regressions in columns (5) and (6). A non-linear transformation of the estimated coefficients is applied to recover coefficient α in column (1)-(4) and (7). Bootstrapped standard errors in parentheses. ^{a}: significant at 1% level; ^{b}: significant at 5% level; ^{c}: significant at 10% level.

our main approach, we can estimate equation (14) directly using non-linear least squares without computing the quantity of housing. An important feature of this alternative is that the variable which may be the most affected by measurement error, R, is now part of the dependent variable and no longer an explanatory variable like in the traditional regression. Appendix M provides further details about these approaches. Because of our concern regarding the measurement of parcel prices, R, we use both raw and smoothed values for parcel prices. We also provide results when we predict K and R as above.

In table 6, we report the estimated coefficients for the elasticity of substitution, σ, and the 'share', α for the three alternatives we have just described and for the CES approximation of our approach. Panel (A) reveals stark differences between two groups of estimates. In the first group, the traditional regression with raw data for parcel prices in column (1) estimates an elasticity of substitution of 0.490. The approach of Ahlfeldt and McMillen (2020), which also uses R as explanatory variable, estimates again an elasticity of substitution well below one with un-smoothed data in column (3). By contrast, the other estimations, which either use smoothed data or have land prices in the dependent variable, estimate an elasticity of substitution close to one. These results are consistent with measurement error on R being of first-order importance. Further support for this interpretation can be found in panel B of table 6, which duplicates the results of panel A using predicted values of R and K. Interestingly, the elasticities of substitution for the traditional regression and for the approach of Ahlfeldt and McMillen (2020) with predicted data are higher than when using the raw data as we expect the use predicted data to correct for measurement error. For the other approaches including ours, we estimate an elasticity of substitution slightly below one.[27]

While the results of table 6 strongly suggest that how we handle measurement error for parcel prices is crucial to explain the differences in the estimated elasticity of substitution across approaches, we keep in mind that smoothing R is not innocuous. Smoothing reduces (classical) measurement error but may also bias our estimates of σ, possibly towards one since over-smoothing reduces the underlying concavity or convexity of the data. To explore this issue, Appendix M provides a characterisation of the bias introduced by smoothing. Monte-Carlo simulations in the same appendix show that the systematic error associated with smoothing is small, even when the elasticity of substitution is far from one. This result is consistent with the stability of our base results with respect to the smoothing bandwidth. We also use simulations calibrated to the variations we observe in the data to confirm the sensitivity of the traditional regression to measurement error. Finally, these simulations confirm the robustness to measurement error of the approaches that smooth parcel prices.

7. Full identification?

Our approach identifies how the production of housing varies with housing capital given parcel size. Full identification, including how the production of housing varies with parcel size, would obviously be desirable. This section provides four results. First, we prove the impossibility of full identification in absence of restriction on the returns to scale of the production function. Second, we show that, if we impose constant returns to scale in production, full identification is possible. Third, the data we use strongly reject a key corollary of constant returns: the linearity of parcel

[27] Importantly, in columns (1)-(4) of panel (B) we do not perform a two-stage least-square estimation with some instruments to identify a causal effect. This is because $R \equiv PH - K^*$ and thus any determinant of R must also be a direct determinant of K^*, making the exclusion restriction impossible to satisfy. Estimating the 'true' σ in the traditional regression is not about estimating the causal effect of R on K, it is about estimating the simultaneous relationship between K, R, and T in absence of interference from unobserved factor heterogeneity.

prices in parcel size. Fourth, while the capital elasticity we estimate when imposing constant returns is statistically different from the capital elasticity we estimate under partial identification, the difference between the two is economically modest.

In Appendix N, we show that the elasticity of housing production with respect to parcel size is equal to the share of land in the total value of the house multiplied by the elasticity of parcel price with respect to parcel size, $\frac{\partial \log H(K^*,T)}{\partial \log T} = \frac{R(x,T)}{K^*+R(x,T)} \frac{\partial \log R(x,T)}{\partial \log T}$. This is the counterpart for land of equation (4), which states that the elasticity of housing production with respect to capital is equal to the share of capital in the value of the house. To recover the quantity of housing, the expression for the elasticity of housing with respect to parcel size should be integrated over parcels of increasing size T for a given (unobserved) set of characteristics x and a given K^*. This is impossible since K^* generally depends on T.

The elasticity of housing production with respect to parcel size can nonetheless be integrated when the elasticity of parcel price with respect to parcel size is equal to one. In turn, parcel price can only increase proportionately with parcel size given x when the production function is constant returns to scale as we demonstrate in Appendix O. In this specific case, we show how the production function for housing can be recovered by integrating the relevant cost shares over the variation of capital, the variation of parcel size, and the joint variation of capital and parcel size.

To assess whether the price of parcels per unit of land is constant at a given location, we regress the log of the price of parcels per unit of land on log parcel size and other parcel characteristics. The results are reported table 7. Column 1 regresses the log price of parcels per square metre on the log of their size. Strikingly, the coefficient is about minus one. Adding parcel controls, urban area indicators, distance to the centre (with a coefficient specific to each urban area), and many municipal controls in columns 2 to 5 lowers the magnitude of the coefficient on log parcel size. Nonetheless, even with a full set of controls, the coefficient on parcel size remains large in magnitude at about -0.66.[28] Adding a quadratic term on log parcel size in column 6 provides evidence of some log concavity indicating that the marginal price of land declines faster for larger parcels. Columns 7 and 8 use kernel-smoothed land price data instead of the actual transaction price data used in columns 1 to 6. The results in these last two columns essentially confirm the

[28]For a very similar specification using US land price data, Albouy and Ehrlich (2013) estimate a coefficient of -0.61.

Table 7: Explaining the price of land per square metre

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Observed values						Smoothed values	
Log parcel size	-0.991[a]	-0.971[a]	-0.759[a]	-0.644[a]	-0.656[a]	0.922[a]	-1.084[a]	-0.254[a]
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.024)	(0.001)	(0.063)
Log parcel size squared						-0.114[a]		-0.060[a]
						(0.002)		(0.005)
Parcel controls	No	Yes	Yes	Yes	Yes	Yes	No	No
Urban area indicator	No	No	Yes	Yes	Yes	Yes	No	No
Distance to the centre	No	No	No	Yes	Yes	Yes	No	No
Municipal controls	No	No	No	No	Yes	Yes	No	No
R^2	0.43	0.45	0.74	0.79	0.80	0.81	0.81	0.81

Notes: OLS regressions with year effects in all columns with 213,786 observations in columns (1)-(6) and 90,000 observations in columns (7)-(8). [a]: significant at 1% level; [b]: significant at 5% level; [a]: significant at 10% level. Parcel controls include indicator variables for whether the parcel is serviced and three types of intermediaries through whom the parcel may have been bought. Municipal controls include log area, log mean income of the year, log standard error of income of the year, share of municipal land that is urbanised (covered) in 2006, share of municipal land for agriculture, ruggedness, soil erodability, soil hydrogeological class, dominant parent material for two main classes of (lighter) soils.

result that unit land prices strongly decline with parcel size. We take these results as a strong rejection of constant parcel prices per unit of land and thus a rejection of constant returns in the production of housing.

For our last exercise, we proceed in the spirit of section 6 and re-estimate housing production under the added restriction of constant returns to scale. We then regress $\log H$ computed under constant returns on $\log K$. If imposing constant returns to scale was appropriate, we should find results similar to those obtained above under partial identification when this assumption is not imposed. The results are reported in table 8.

For the first decile of parcel sizes, the results from panels (A) and (B) of table 8 are similar to those of table 2. For subsequent deciles, the capital elasticity falls from 0.61 to 0.55 in table 8 while this elasticity increases from 0.64 to 0.66 in table 2. We observe a similar pattern of divergence, albeit in the opposite direction, in the last two panels of table 8 relative to the two corresponding panels of table 3 when using predicted quantities.

We interpret this divergence between the results obtained with a constant-return assumption and those obtained without it as evidence that imposing constant returns to scale may be warranted when considering small parcels but becomes increasingly less appropriate when we

Table 8: log housing production with constant returns to scale, OLS by size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Observed data									
$\log(K)$	0.624^a	0.606^a	0.590^a	0.577^a	0.568^a	0.561^a	0.556^a	0.553^a	0.549^a
	(0.00093)	(0.0011)	(0.0013)	(0.0016)	(0.0019)	(0.0022)	(0.0024)	(0.0025)	(0.0022)
Panel (B): Observed data									
$\log(K)$	0.113^a	-0.650^a	-1.279^a	-1.935^a	-2.629^a	-3.295^a	-3.932^a	-4.482^a	-4.927^a
	(0.038)	(0.075)	(0.107)	(0.129)	(0.153)	(0.169)	(0.187)	(0.211)	(0.246)
$[\log(K)]^2$	0.022^a	0.053^a	0.079^a	0.106^a	0.135^a	0.162^a	0.189^a	0.212^a	0.231^a
	(0.0016)	(0.0032)	(0.0045)	(0.0055)	(0.0065)	(0.0072)	(0.0080)	(0.0090)	(0.010)
Panel (C): Predicted data									
$\log(K)$	0.645^a	0.640^a	0.652^a	0.670^a	0.690^a	0.708^a	0.723^a	0.733^a	0.744^a
	(0.0010)	(0.0027)	(0.0035)	(0.0039)	(0.0042)	(0.0046)	(0.0052)	(0.0059)	(0.0066)
Panel (D): Predicted data									
$\log(K)$	0.105	1.028^c	1.461^b	1.002	-0.095	-1.592^c	-3.416^a	-5.302^a	-7.137^a
	(0.137)	(0.565)	(0.720)	(0.787)	(0.847)	(0.923)	(0.991)	(1.051)	(1.109)
$[\log(K)]^2$	0.023^a	-0.016	-0.034	-0.014	0.033	0.097^b	0.175^a	0.255^a	0.333^a
	(0.0059)	(0.024)	(0.031)	(0.033)	(0.036)	(0.039)	(0.042)	(0.044)	(0.047)

Notes: OLS regressions with a constant in all columns. 300 observations for each regression. In panels (C) and (D), capital and parcel price are predicted from demand-related factors as in table 3. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%.

consider larger parcels. This interpretation is consistent with the results of table 7 showing that the price of land per square metre declines with parcel size. In turn, parcels are probably best viewed as exogenous because of their indivisibility rather than the product of a maximising choice by house builders. This said, our rejection of constant returns to scale is like our rejection of the Cobb-Douglas functional form. Although, we can formally reject that houses are produced under constant returns, this remains a reasonable first-order approximation.

8. Conclusions

We develop a novel approach to estimate the production function of housing. We rely on the notion that, although heterogeneous in many dimensions, houses all provide units of housing. The price of a house is then the product of the price of housing per unit (which varies across locations) and the number of units of housing provided by this house. To separate these two quantities, we assume that housing is competitively provided. Then, the first-order condition for house builders determines the *marginal value product* of capital. Using the zero-profit condition, we can eliminate the price of housing per unit from the first-order condition and isolate the *marginal product* of capital when building a house. For parcels of a given size, we can sum this marginal product across houses in different locations that have optimally received different levels of capital and recover the production of housing associated with each level of capital. Although our approach could potentially be applied to other production function estimations, we believe that using it for housing is particularly appropriate because we can rely on the large spatial variations of land prices, a fundamentally important input in our context.

Our main result is that the production function of housing is reasonably well approximated by a Cobb-Douglas production function under constant returns. This said, we can nonetheless show that this is not exactly true. Our preferred results indicate a mild amount of log concavity in the production function of housing and an elasticity of housing production with respect to capital increasing with parcel size, which is consistent with an elasticity of substitution between land and capital slightly below one. We also statistically reject that the production function for housing exhibits constant returns. Nonetheless, the capital elasticity we estimate when imposing constant returns is close to our unrestricted estimates.

There are three challenges that future work will need to deal with. First, we implicitly assume that housing is perfectly divisible (unlike parcels). We do not expect households who purchase a new house to get exactly the quantity of housing they wanted. In turn, the willingness to pay of a household for a unit of housing may decline as the house they consider deviates from their preferred choice. Exploring the implications of the indivisibility of housing in our framework is a natural next step. Second, we assume that houses only vary in the amount of housing units that they provide. While we show that this may be a reasonable assumption for new houses in a given city or for buyers that belong to the same occupational group, it will be important in future work to consider richer forms of heterogeneity in the demand for housing. Finally, with richer data about the houses being built, it will be interesting to decompose the quantity of housing we estimate into a house's observable characteristics.

References

Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes. 2007. Econometric tools for analyzing market outcomes. In James J. Heckman and Edward E. Leamer (eds.) *Handbook of Econometrics*, volume 6A. Amsterdam: North-Holland, 4271–4276.

Ahlfeldt, Gabriel M. and Daniel P. McMillen. 2020. New estimates of the elasticity of substitution between land and capital. Processed, London School of Economics.

Albouy, David and Gabriel Ehrlich. 2013. The distribution of urban land values: Evidence from market transactions. Processed, University of Illinois.

Albouy, David and Gabriel Ehrlich. 2018. Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics* 107(0):101–120.

Albouy, David, Gabriel Ehrlich, and Minchul Shin. 2018. Metropolitan land values. *Review of Economics and Statistics* 100(3):454–466.

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

Chapelle, Guillaume and Jean-Benoît Eyméoud. 2017. The housing supply elasticity and the cost of agglomeration. Processed, Science Po.

Chatterjee, Satyajit and Burcu Eyigungor. 2015. Quantitative analysis of the US housing and mortgage markets and the foreclosure crisis. *Review of Economic Dynamics* 18(2):165–184.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2019. The costs of agglomeration: House and land prices in French cities. *Review of Economic Studies* 86(4):1556–1589.

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

Commissariat Général au Développement Durable. 2012. *RéférenceS: Economie du Logement*. Paris: Ministère de l'Ecologie, du Développement Durable, des Transports et du Logement.

Davis, Morris, Jonas D.M. Fisher, and Toni M. Whited. 2014. Macroeconomic implications of agglomeration. *Econometrica* 82(2):731–764.

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

De Loecker, Jan. 2011. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica* 79(5):1407–1451.

De Loecker, Jan and Frederic Warzynski. 2012. Markups and firm-level export status. *American Economic Review* 106(6):2437–2471.

Duranton, Gilles and Diego Puga. 2015. Urban land use. In Gilles Duranton, J. Vernon Henderson, and William C. Strange (eds.) *Handbook of Regional and Urban Economics*, volume 5A. Amsterdam: North-Holland, 467–560.

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.

Gandhi, Amit, Salvador Navarro, and David Rivers. 2020. On the identification of gross output production functions. *Journal of Political Economy* 128(forthcoming).

Gyourko, Joe and Jacob Krimmel. 2020. The impact of local residential land use restrictions on land values across and within single family housing markets. Processed, Wharton School of the University of Pennsylvania.

Gyourko, Joseph. 2009. Housing supply. *Annual Review of Economics* 1(1):295–318.

Gyourko, Joseph and Albert Saiz. 2006. Construction costs and the supply of housing structure. *Journal of Regional Science* 46(4):661–680.

Head, Allen and Huw Lloyd-Ellis. 2012. Housing liquidity, mobility, and the labour market. *Review of Economic Studies* 79(4):1559–1589.

Hsieh, Chang-Tai and Enrico Moretti. 2019. Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics* 11(2):1–39.

Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov. 2011. Winners and losers in housing markets. *Journal of Money, Credit and Banking* 43(2-3):255–296.

Klein, Lawrence R. 1953. *A Textbook of Econometrics*. Evanston (IL): Row, Peterson and Co.

Klette, Tor Jakob and Zvi Griliches. 1996. The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics* 11(4):343–361.

Larson, William and Anthony Yezer. 2015. The energy implications of city size and density. *Journal of Urban Economics* 90(1):35–49.

Mauro, Léa. 2013. Le patrimoine économique national en 2011. *INSEE Première* 0(1431):1–4.

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

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

Muth, Richard F. 1975. Numerical solution of urban residential land-use models. *Journal of Urban Economics* 4(2):307–332.

Olley, G. Steven and Ariel Pakes. 1996. The dynamics of productivity in the telecommunication equipment industry. *Econometrica* 64(6):1263–1297.

Poterba, James M. 1984. Tax subsidies to owner-occupied housing: An asset-market approach. *Quarterly Journal of Economics* 99(4):729–752.

Rupert, Peter and Etienne Wasmer. 2012. Housing and the labor market: Time to move and aggregate unemployment. *Journal of Monetary Economics* 59(1):24–36.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.

Solow, Robert M. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics* 39(3):312–320.

Syverson, Chad. 2011. What determines productivity? *Journal of Economic Literature* 49(2):326–365.

Thorsnes, Paul. 1997. Consistent estimates of the elasticity of substitution between land and non-land inputs in the production of housing. *Journal of Urban Economics* 42(1):98–108.

Yoshida, Jiro. 2016. Structure depreciation and the production of real estate services. Processed, Smeal College of Business, Penn State University.

Appendix A. Extending the base model to two types of housing capital

We now extend our base model to two types of capital. The first type of capital produces raw floorspace while the second produces housing quality. The production function now contains three arguments instead of two in the base model: $H(K_1,K_2,T)$ where T is the land area of the parcel. We distinguish here between floorspace and quality but finer distinctions could obviously be envisioned (flooring, finishes, structure quality, etc) since the reasoning would be the same.

The builder, who develops a parcel of exogenously given size T and characteristics x purchased at the endogenously determined price R, seeks to maximise profits now given by $\pi = P(x)H(K_1,K_2,T) - K_1 - K_2 - R$ with respect to both K_1 and K_2. The first-order conditions for profit maximisation are:

$$P(x)\frac{\partial H(K_1,K_2,T)}{\partial K_1} = 1 \qquad \text{and} \qquad P(x)\frac{\partial H(K_1,K_2,T)}{\partial K_2} = 1\,. \tag{A1}$$

As $H(K_1,K_2,T)$ is increasing and strictly concave in K_1 and K_2, the profit-maximising capital investments K_1^* and K_2^* are unique. Free entry still dissipates the profits from construction into the price of land: $R = P(x)H(K_1^*,K_2^*,T) - K_1^* - K_2^* \equiv R(K_1^*,K_2^*,T)$. This condition can be used to eliminate the price of housing from the first-order conditions. Writing the resulting expressions as elasticities, we obtain:

$$\frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_1^*} = \frac{K_1^*}{K_1^* + K_2^* + R(K_1^*,K_2^*,T)} \quad \text{and} \quad \frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_2^*} = \frac{K_2^*}{K_1^* + K_2^* + R(K_1^*,K_2^*,T)}. \tag{A2}$$

These two expressions duplicate equation (4) in the main text for each type of capital. We cannot use these two expressions empirically since we do not observe K_1 and K_2 separately.

To obtain the elasticity of housing with respect to total capital K^* $(= K_1^* + K_2^*)$, we need K_1^* and K_2^* to be functions of K^* (and T). From the first-order conditions (A1) and the restrictions imposed to the production function, both K_1^* and K_2^* are functions of P and T, increasing in both arguments. Their sum, $K^* = K_1^* + K_2^*$ is thus also an increasing function of P and T. Hence, this sum can be inverted so that P is written as function of K^* and T. We can use this property to rewrite K_1^* and K_2^* as functions of K^* and T using equation (A1). We can then replace the two resulting expressions for K_1^* and K_2^* into the production function of housing and write it as: $H\left(K_1^*(K^*,T),K_2^*(K^*,T),T\right)$. With a slight abuse of notations to keep the algebra readable, we then derive the housing production function:

$$\begin{aligned}
\frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K^*} &= \frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_1^*}\frac{\partial \log K_1^*}{\partial \log K^*} + \frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_2^*}\frac{\partial \log K_2^*}{\partial \log K^*} \\
&= \frac{K^*}{K_1^*}\frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_1^*}\frac{\partial K_1^*}{\partial K^*} + \frac{K^*}{K_2^*}\frac{\partial \log H(K_1^*,K_2^*,T)}{\partial \log K_2^*}\frac{\partial K_2^*}{\partial K^*} \\
&= \frac{K^*}{K^* + R(K_1^*,K_2^*,T)}\left(\frac{\partial K_1^*}{\partial K^*} + \frac{\partial K_2^*}{\partial K^*}\right) \\
&= \frac{K^*}{K^* + R(K^*,T)} \qquad \text{(A3)}
\end{aligned}$$

where the third equality uses equation (A1) and the last one relies on $K^* = K_1^* + K_2^*$. Equation (A3) is equivalent to equation (4) and we can integrate it in the same manner to recover the quantity of housing. Hence, equation (A3) shows that, although the composition of capital may be highly heterogeneous, our approach still recovers $H(K,T)$ for a given T.

Importantly, we recover the quantity of housing from the variations of K^* and R only along the optimal path for K_1^* and K_2^* for a given T. To be concrete, we can determine how the quantity of housing varies with the overall housing investment but cannot recover how floorspace and quality map into housing units.

We now introduce a regulatory constraint taking the form of a cap on floorspace: $K_1 \leq \overline{K}_1$. Because our estimation is for T given, this constraint readily generalises to $K_1 \leq \overline{K}_1(T)$. While a simple cap on floorspace is not realistic, recall that maximum floor-to-area ratio constraints are the main tools used to regulate the intensity of land use in France.

The first-order conditions in equation (A1) imply that K_1 and K_2 are strictly increasing in $P(x)$ given T. Hence, there is always a high enough level of housing price such that the cap on floorspace is binding and $K_1 = \overline{K}_1$. The other derivations remain however the same as above. When the constraint on K_1 is not binding, equation (A3) applies. When it is binding, the capital elasticity is:

$$\frac{\partial \log H(\overline{K}_1,K_2^*,T)}{\partial \log K^*} = \frac{\partial H(\overline{K}_1,K_2^*,T)}{\partial K_2^*}\frac{K^*}{H(\overline{K}_1,K_2^*,T)} = \frac{K^*}{K^* + R(K^*,T,\overline{K}_1)}, \qquad \text{(A4)}$$

where the first equality arises because $\partial K^* = \partial K_2^*$ when the constraint $K_1 = \overline{K}_1$ is binding and the second equality uses equation (A3). This rewriting highlights that we cannot integrate equation (A4) like in the base case because $\overline{K}_1$ may depend on x and thus vary jointly with K^*. It also shows that the housing elasticity with respect to capital is estimated given the constraint.

In the Cobb-Douglas case where H is proportional to $K_1^{\beta_1} K_2^{\beta_2}$, it is easy to show using equations (A3) and (A4) that the capital elasticity is equal to $\beta_1 + \beta_2$ when the constraint is not binding and

$\beta_2(\overline{K}_1 + K_2^*)/K_2^*$ when the constraint is binding. Considering a gradual increase in P from 0, we first observe a capital elasticity of $\beta_1 + \beta_2$ followed by a gradual decrease to β_2 as the constraint becomes more binding.

Appendix B. Comparison with Epple *et al.* (2010)

In this appendix, we provide a detailed comparison between our approach and that of Epple *et al.* (2010), hereafter, EGS.

A. Applying our approach with the data of EGS

We first show that our approach can be used with data similar to those of EGS. The main apparent difference is that EGS observe the value of the house whereas we observe the investment made to build the house. More specifically, EGS have information on (V,R,T) where $V \equiv PH$ is the house value, instead of (K,R,T) in our case. It is nonetheless possible to implement our approach with the data of EGS. To do this, note that the first-order condition for profit maximisation (1) used in the main text can be rewritten as:

$$\frac{1}{H(K,T)} \frac{\partial H(K,T)}{\partial K} = \frac{1}{V}\,, \tag{B1}$$

after dividing both sides by $V = P\,H$ and omitting the argument x for brevity. Although the value of capital, K^*, is not directly observed with the data of EGS, the zero-profit condition readily yields $K^* = V - R$. From equation (1), we have $P = P(K^*,T)$. Inserting this into the expression for house values, we get $V = P(K^*,T)H(K^*,T) \equiv V(K^*,T)$ and we end up with the differential equation:

$$\frac{1}{H(K^*,T)} \frac{\partial H(K^*,T)}{\partial K} = \frac{1}{V(K^*,T)}\,. \tag{B2}$$

This differential equation can be integrated over K^* to recover the housing production function (up to a multiplicative function of T).

B. The approach of EGS *in our setting*

A corollary of the previous result is that we can also apply the approach of EGS to estimate the housing production function with the data at hand for France.

To compare the two approaches further, it is insightful to re-derive the approach of EGS in the spirit of our paper. Note first that EGS make two additional assumptions relative to our approach.

First, they assume that the housing production function is constant returns to scale. Second, the value of parcels is linear in their size. We show below that it is possible to adapt their approach and avoid making these two assumptions. This leads to the partial identification of the housing production function as in our case.

The crux of the approach of EGS is to make K^*, which is unobserved in their case, disappear from the first-order condition by substituting its expression as a function of the house price V and parcel size T to recover a differential equation for the supply function of housing $S(P,T)$, which links housing production to house prices for a given land area. Their differential equation contains a function that can be estimated using the data at hand. Once the supply of housing is recovered, house prices, $P = P(K^*,T)$, can be computed as a function of the optimal structure and land area from the zero-profit condition. Finally, inserting this expression for house prices into the supply function for housing yields the housing production function.

More formally, note first that we have $K^* = K^*(V,T)$ from (B1). The zero-profit condition then implies that $R = V - K^*(V,T) \equiv R(V,T)$. It is possible to recover non-parametrically the function $R(V,T)$ since (R,V,T) is observed.

The first-order condition for profit maximisation implies that we have $K^* = K^*(P,T)$. Using this equation and following the derivation of EGS, the first-order condition can be rewritten to obtain a differential equation for the supply function of housing:

$$P\,\frac{\partial H(K,T)}{\partial K} = 1 \iff P\left(\frac{\partial K^*(P,T)}{\partial P}\right)^{-1}\frac{\partial H\left(K^*(P,T),T\right)}{\partial P} = 1\,, \tag{B3}$$

$$\iff P\frac{\partial S(P,T)}{\partial P} = \frac{\partial K^*(P,T)}{\partial P}\,, \tag{B4}$$

$$\iff P\frac{\partial S(P,T)}{\partial P} = \frac{\partial\left(V - R(V,T)\right)}{\partial P}\,. \tag{B5}$$

As $V = PS(P,T)$, we have:

$$\frac{\partial V}{\partial P} = S(P,T) + P\frac{\partial S(P,T)}{\partial P} \tag{B6}$$

Substituting this expression into equation (B5) yields:

$$S(P,T) = \frac{\partial R\left(PS(P,T),T\right)}{\partial P}\,. \tag{B7}$$

This is the equation used by EGS to estimate the supply function for housing for a given land area. Note that this expression could alternatively be obtained directly from Hotelling's lemma applied to the short-run profit given by $PH - K\,(= R)$.

Using the fact that $\frac{\partial R}{\partial P} = \frac{\partial V}{\partial P}\frac{\partial R}{\partial V}$, expression (B6), and dropping the arguments of S for readability, equation (B7) can be developed to obtain the following differential equation:

$$S = \frac{\partial R\,(PS,T)}{\partial V}\left(S + P\frac{\partial S}{\partial P}\right) \tag{B8}$$

Since the function R can be recovered from the data through the zero-profit condition, so can its partial derivative with respect to V. The resulting differential equation can then be solved to recover S as a function of P for a given land area T. Note that the differential equation (B8) is made intricate by the presence of S in the function R, which makes it implicit only, contrary to our approach.

Once the supply of housing is recovered, the optimal amount of capital corresponding to price P can be obtained using the zero-profit condition:

$$K^*\,(P,T) = PS(P,T) - R(PS(P,T),T)\,. \tag{B9}$$

This function can be inverted to obtain $P = P\,(K^*,T)$ and the variations of the production function of housing with respect to K (holding T fixed) can be recovered using the fact that $S\,(P\,(K^*,T)\,,T) = H\,(K^*,T)$. Note that, as in our case, the differential equation can be solved up to a function of T and the production function of housing is only partially identified. Under the constant return-to-scale assumption made by EGS, there is full identification since there is only one differential equation to solve regardless of the value of T. This single differential equation is simply equation (B8) where T is set to one.

Appendix C. Additional data

Urban areas. We use the 1999 delineation of urban areas from the French statistical institute (INSEE).

Wages. We construct measures of wages for blue collar workers in the construction industry for all French urban areas from the French labour force administrative records (DADS - Déclarations Annuelles des Données Sociales).

Education. We construct measures of the share of population with a college or university degree for all French municipalities from the French census for 2006. We consider all higher education degrees that sanction two years of study or more after high school.

Income. Mean household income and its standard deviation by municipality can be constructed using information from each cadastral section (about 100 housing units on average) contained in the FILOCOM repository. This repository is managed by the *Direction Générale des Finances Publiques* of the French Ministry of Finance. It contains a record of all housing units and their occupants which they match to income tax records.

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

Land use. We use information from the 2006 *Corine Land Cover* dataset to compute the share of agricultural and impervious land in each municipality. We compute the fraction of land that is built up in each municipality using information from *BD Topo* (version 2.1) from the French National Geographical Institute. This data set is originally produced using satellite imagery combined with the French land registry. It reports information for nearly all buildings in the country including their footprint, height, and use with an accuracy of one metre.

We also use the *Fichiers Fonciers du CEREMA* 3.0 for all dwellings to compute observed floor-to-area ratios in all municipalities. We select all residential houses (maisons) with three stories or less. We then allocate each of these houses to the parcel it sits on. Then, for each of the resulting parcels we compute the floor-to-area ratio by summing the built-up area of all dwellings on the parcels and divide by the area of the parcel. For each municipality, we keep the entire distribution of floor-to-area ratios for all parcels with a single-family home.

Values, rents, and mortgage rates for properties. Monthly rent data in euros per m^2 are from the *Clameur* consortium for 2012 (and published as "2013" rent data). The data are for 2,932 municipalities with population above 2,000 inhabitants in mainland France. The nearly 400,000 individual new leases or lease renewals that underlie the data are collected from members of the *Clameur* consortium, including financial institutions, large property management firms, associations of small property managers and real estate brokers, etc. A municipality is included only if 30 or more leases are observed. Direct rentals, which represent about 30% of the market are absent.

Property prices (which we refer to as 'values' in the main text to avoid any confusion with rental prices) are municipal indices constructed from the 2012 census of all transactions of non-new dwellings conducted by regional notary associations. The log of the price per square metre is regressed on indicator variables for the construction period (before 1850, 1850-1913, 1914-1947, 1948-1959, 1960-1980, 1981-1991, after 1991) and for the quarter of the transaction. The municipal index is the average of the residuals for each municipality after adding the regression constant (see Combes *et al.*, 2019, for further details). We obtain an index for 26,972 municipalities in mainland France.

Average annual rates for mortgages are from l'Observatoire Crédit Logement / CSA, a consortium of the main French banks (which provide a joint-mortgage guarantee for a subset of properties akin to that of Fanny Mae in the US) and CSA, a market study firm with a long-running survey of housing finance in France.

Appendix D. Supplementary results for section 5.1: Technical checks

Table 9 duplicates table 2 using different smoothing bandwidths equal to a half, a quarter, and a tenth of the rule-of-thumb bandwidth we use in table 2 and other estimations. The results show that even strong under-smoothing barely affects the results.

Table 10 considers a broader support for K between the 3rd and 97th percentiles of all land values in the data instead of between the 10th and 90th percentile of parcel values in the distribution in table 2. This leads us to consider 94% of all land values instead of 75% in our base estimation. While we lose some precision in the estimates when introduce a quadratic term in panel (B), the results of panel (A) are similar to those of table 2.

Panels (A) and (B) of table 11 also duplicate table 2 but directly smooth the cost share $K^*/(K^* + R)$ instead of smoothing R prior to computing the cost share. Panels (C) and (D) of table 11 duplicate panels (A) and (B) of table 3 but smooth the cost share directly again. The results of the first two panels are similar to our base results but with mildly higher estimates of the capital elasticity by about 3 to 4 percentage points. The results of the last two panels are virtually undistinguishable from the corresponding results of table 3.

Appendix E. Supplementary results for section 5.2: Predicted values of K, R, and T

Table 12 report results experimenting with the set of demand-related factors. Panels (A) and (B) only include the urban area of a parcel to predict its price and capital investment. The results are qualitatively the same as those of the two panels of table 3. Despite the bluntness of this rudimentary exercise, we note a greater dispersion of the capital elasticity in panel (A) and more log

Table 9: log housing production with different smoothing bandwidth, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Bandwidth $= 0.5\times$ rule-of-thumb bandwidth									
$\log(K)$	0.618^a	0.639^a	0.637^a	0.638^a	0.646^a	0.651^a	0.652^a	0.662^a	0.664^a
	(0.0011)	(0.00087)	(0.0011)	(0.0012)	(0.0014)	(0.0021)	(0.0022)	(0.0035)	(0.0038)
Panel (B): Bandwidth $= 0.5\times$ rule-of-thumb bandwidth									
$\log(K)$	0.229^a	0.019	-0.107^b	0.089	-0.028	0.122	0.355^a	0.141	0.440^b
	(0.058)	(0.043)	(0.044)	(0.057)	(0.083)	(0.105)	(0.124)	(0.170)	(0.183)
$[\log(K)]^2$	0.016^a	0.026^a	0.031^a	0.023^a	0.028^a	0.022^a	0.013^b	0.022^a	0.009
	(0.0024)	(0.0018)	(0.0019)	(0.0024)	(0.0035)	(0.0044)	(0.0053)	(0.0071)	(0.0078)
Panel (C): Bandwidth $= 0.25\times$ rule-of-thumb bandwidth									
$\log(K)$	0.616^a	0.638^a	0.635^a	0.640^a	0.650^a	0.653^a	0.650^a	0.663^a	0.666^a
	(0.0015)	(0.0011)	(0.0013)	(0.0020)	(0.0023)	(0.0026)	(0.0030)	(0.0046)	(0.0047)
Panel (D): Bandwidth $= 0.25\times$ rule-of-thumb bandwidth									
$\log(K)$	0.258^a	-0.011	-0.076	0.184^b	-0.085	0.182	0.342^b	-0.100	0.548^b
	(0.085)	(0.049)	(0.063)	(0.088)	(0.113)	(0.152)	(0.159)	(0.233)	(0.230)
$[\log(K)]^2$	0.015^a	0.027^a	0.030^a	0.019^a	0.031^a	0.020^a	0.013^c	0.032^a	0.005
	(0.0036)	(0.0021)	(0.0027)	(0.0037)	(0.0048)	(0.0064)	(0.0067)	(0.010)	(0.010)
Panel (E): Bandwidth $= 0.1\times$ rule-of-thumb bandwidth									
$\log(K)$	0.621^a	0.635^a	0.638^a	0.649^a	0.653^a	0.658^a	0.650^a	0.664^a	0.670^a
	(0.0024)	(0.0017)	(0.0032)	(0.0030)	(0.0038)	(0.0039)	(0.0064)	(0.0064)	(0.0071)
Panel (F): Bandwidth $= 0.1\times$ rule-of-thumb bandwidth									
$\log(K)$	0.251^c	-0.027	-0.094	0.295	-0.101	0.147	0.183	-0.228	1.046^a
	(0.140)	(0.068)	(0.084)	(0.194)	(0.134)	(0.177)	(0.188)	(0.307)	(0.322)
$[\log(K)]^2$	0.016^a	0.028^a	0.031^a	0.015^c	0.032^a	0.022^a	0.020^b	0.038^a	-0.016
	(0.0059)	(0.0029)	(0.0036)	(0.0082)	(0.0057)	(0.0075)	(0.0080)	(0.013)	(0.014)

Notes: OLS regressions with a constant in all columns. 900 observations for each regression. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%.

concavity, especially for the lower deciles of parcel size in panel (B). Panels (C) and (D) rely again on the urban area of a parcel to predict its price and capital investment but condition out local wages in the construction industry from the estimated urban area fixed effects. With this specification, the differences in capital elasticity across parcel size deciles are minimal. Depending on the deciles, the production function of housing is either marginally log concave or marginally log convex. Panels (E) and (F) include urban area fixed effects, distance to the centre (with an effect specific to each urban area), income, and land-use variables among the demand determinants but do not condition out construction wages. Finally, panels (G) and (H) additionally condition out construction wages

Table 10: log housing production, OLS with expanded support for R

Decile	1	2	3	4	5	6	7	8	9
Panel (A)									
$\log(K)$	0.633^a	0.647^a	0.650^a	0.647^a	0.648^a	0.654^a	0.654^a	0.662^a	0.663^a
	(0.015)	(0.016)	(0.017)	(0.017)	(0.016)	(0.015)	(0.015)	(0.016)	(0.016)
Panel (B)									
$\log(K)$	0.344	0.373	0.332	0.391	0.506	0.606	0.678	0.619	0.698
	(0.484)	(0.531)	(0.567)	(0.578)	(0.649)	(0.719)	(0.817)	(0.787)	(0.823)
$[\log(K)]^2$	0.012	0.011	0.013	0.010	0.006	0.002	-0.001	0.002	-0.001
	(0.021)	(0.023)	(0.024)	(0.025)	(0.028)	(0.030)	(0.035)	(0.034)	(0.035)

Notes: OLS regressions with a constant in all columns. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Table 11: log housing production with smoothed cost shares, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Observed data									
$\log(K)$	0.648^a	0.658^a	0.661^a	0.663^a	0.670^a	0.679^a	0.684^a	0.691^a	0.695^a
	(0.00077)	(0.00072)	(0.00071)	(0.00079)	(0.00080)	(0.0011)	(0.0011)	(0.0014)	(0.0016)
Panel (B): Observed data									
$\log(K)$	0.126^a	0.033	-0.021	0.085^a	0.118^a	0.198^a	0.333^a	0.243^a	0.321^a
	(0.038)	(0.025)	(0.025)	(0.031)	(0.036)	(0.046)	(0.068)	(0.083)	(0.082)
$[\log(K)]^2$	0.022^a	0.026^a	0.029^a	0.024^a	0.023^a	0.020^a	0.015^a	0.019^a	0.016^a
	(0.0016)	(0.0010)	(0.0011)	(0.0013)	(0.0015)	(0.0020)	(0.0029)	(0.0035)	(0.0035)
Panel (C): Predicted data									
$\log(K)$	0.653^a	0.654^a	0.655^a	0.658^a	0.663^a	0.671^a	0.675^a	0.680^a	0.686^a
	(0.0011)	(0.00065)	(0.00056)	(0.00069)	(0.00071)	(0.00092)	(0.0011)	(0.0014)	(0.0021)
Panel (D): Predicted data									
$\log(K)$	0.359^a	1.026^a	1.538^a	1.969^a	2.183^a	2.270^a	1.947^a	1.884^a	2.168^a
	(0.111)	(0.069)	(0.071)	(0.070)	(0.081)	(0.117)	(0.147)	(0.168)	(0.193)
$[\log(K)]^2$	0.012^a	-0.016^a	-0.037^a	-0.055^a	-0.064^a	-0.068^a	-0.054^a	-0.051^a	-0.063^a
	(0.0047)	(0.0029)	(0.0030)	(0.0030)	(0.0034)	(0.0049)	(0.0062)	(0.0071)	(0.0081)

Notes: OLS regressions with a constant in all columns. In panels (C) and (D), cost shares are predicted directly from the same demand-related factors used in table 3 to predict investment and parcel price separately. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

from urban area fixed effects and predict parcel size with demand-related factors.

Table 12: log housing production in urban areas obtained from predicted values, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Urban area fixed effects only									
$\log(K)$	0.608[a]	0.634[a]	0.651[a]	0.661[a]	0.669[a]	0.681[a]	0.689[a]	0.695[a]	0.701[a]
	(0.0022)	(0.0018)	(0.0014)	(0.0011)	(0.0014)	(0.0019)	(0.0022)	(0.0023)	(0.0028)
Panel (B): Urban area fixed effects only									
$\log(K)$	5.741[a]	5.389[a]	3.508[a]	2.682[a]	2.904[a]	2.957[a]	2.927[a]	2.921[a]	3.207[a]
	(0.331)	(0.239)	(0.236)	(0.169)	(0.224)	(0.285)	(0.332)	(0.362)	(0.439)
$[\log(K)]^2$	-0.217[a]	-0.201[a]	-0.121[a]	-0.085[a]	-0.094[a]	-0.096[a]	-0.095[a]	-0.094[a]	-0.106[a]
	(0.014)	(0.010)	(0.010)	(0.0072)	(0.0095)	(0.012)	(0.014)	(0.015)	(0.019)
Panel (C): Urban area fixed effects net of construction wages									
$\log(K)$	0.651[a]	0.640[a]	0.634[a]	0.632[a]	0.634[a]	0.637[a]	0.639[a]	0.642[a]	0.650[a]
	(0.0034)	(0.00092)	(0.00086)	(0.0010)	(0.0013)	(0.0020)	(0.0027)	(0.0032)	(0.0054)
Panel (D): Urban area fixed effects net of construction wages									
$\log(K)$	-1.132[b]	-0.576[b]	0.706[a]	1.260[a]	1.438[a]	1.005[a]	-0.122	-0.616	0.521
	(0.541)	(0.225)	(0.166)	(0.149)	(0.220)	(0.350)	(0.501)	(0.676)	(1.115)
$[\log(K)]^2$	0.075[a]	0.051[a]	-0.003	-0.027[a]	-0.034[a]	-0.016	0.032	0.053[c]	0.005
	(0.0230)	(0.0095)	(0.0070)	(0.0063)	(0.0091)	(0.015)	(0.021)	(0.028)	(0.047)
Panel (E): Urban area fixed effects, distance effects, income, and land use									
$\log(K)$	0.615[a]	0.633[a]	0.643[a]	0.651[a]	0.660[a]	0.672[a]	0.678[a]	0.687[a]	0.697[a]
	(0.0011)	(0.00074)	(0.00074)	(0.00089)	(0.0012)	(0.0013)	(0.0016)	(0.0017)	(0.0021)
Panel (F): Urban area fixed effects, distance effects, income, and land use									
$\log(K)$	3.795[a]	3.532[a]	3.508[a]	3.708[a]	3.944[a]	4.091[a]	4.191[a]	4.117[a]	4.034[a]
	(0.087)	(0.071)	(0.064)	(0.067)	(0.085)	(0.093)	(0.108)	(0.126)	(0.159)
$[\log(K)]^2$	-0.134[a]	-0.122[a]	-0.121[a]	-0.129[a]	-0.139[a]	-0.144[a]	-0.148[a]	-0.145[a]	-0.141[a]
	(0.0037)	(0.0030)	(0.0027)	(0.0029)	(0.0036)	(0.0039)	(0.0046)	(0.0054)	(0.0067)
Panel (G): ——— net of construction wages with predicted T									
$\log(K)$	0.638[a]	0.648[a]	0.653[a]	0.649[a]	0.653[a]	0.657[a]	0.656[a]	0.674[a]	0.685[a]
	(0.0020)	(0.0014)	(0.0018)	(0.0020)	(0.0015)	(0.0023)	(0.0029)	(0.0029)	(0.0030)
Panel (H): ——— net of construction wages with predicted T									
$\log(K)$	-0.767[a]	1.192[a]	2.857[a]	3.295[a]	3.185[a]	2.510[a]	2.876[a]	2.236[a]	1.551[a]
	(0.291)	(0.240)	(0.277)	(0.290)	(0.238)	(0.260)	(0.317)	(0.369)	(0.403)
$[\log(K)]^2$	0.059[a]	-0.023[b]	-0.093[a]	-0.112[a]	-0.107[a]	-0.078[a]	-0.094[a]	-0.066[a]	-0.037[b]
	(0.012)	(0.010)	(0.012)	(0.012)	(0.010)	(0.011)	(0.013)	(0.016)	(0.017)

Notes: OLS regressions with a constant in all columns. In all panels (E)-(H), distance to the centre is urban-area specific; income variables are log mean municipal income, log standard error of income, and share of population with a university degree; geology variables are ruggedness, soil erodability, soil hydrogeological class, dominant parent material for two main classes of (lighter) soils; land use variables are three land use variables share of built-up land, share of urbanised land, and share of agricultural land. 900 observations for each regression. The R^2 is 1.00 in all specifications. Robust standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Appendix F. Supplementary results for section 5.3: Differences across locations

Table 13 reports results for different bands of distance to the centre. For each new construction, we compute the distance between the centroid of its municipality and the centroid of its urban area. We then normalise this distance by the maximum centroid-to-centroid distance in the urban area. Finally, we partition new constructions by their quintiles in the distribution of relative distances.

Table 14 complements table 4 with separate regressions for each decile of parcel size and confirms its results.

In table 15, we assess the effects of our measure of construction costs on the main results of table 2. In the first two panel, we used residualised values of K after conditioning out local construction wages (measured for each urban area). In the last two panels, we use a blunter approach and directly deflate capital investment by construction wages.

Table 13: log housing production, OLS by centiles of distance to the centre

Centiles	All urban areas	0-20	20-40	40-60	60-80	80-100
Panel (A): Observed data						
$\log(K)$	0.645^a	0.640^a	0.616^a	0.640^a	0.664^a	0.684^a
	(0.00087)	(0.0024)	(0.0015)	(0.0017)	(0.0017)	(0.0029)
Panel (B): Observed data						
$\log(K)$	0.086^a	-0.106^a	0.106^a	-0.116^a	-0.517^a	-0.509^a
	(0.030)	(0.076)	(0.051)	(0.066)	(0.064)	(0.099)
$[\log(K)]^2$	0.024^a	0.031^a	0.021^a	0.032^a	0.050^a	0.051^a
	(0.0013)	(0.0032)	(0.0022)	(0.0028)	(0.0027)	(0.0042)
Panel (C): Predicted data						
$\log(K)$	0.661^a	0.644^a	0.634^a	0.659^a	0.673^a	0.692^a
	(0.00062)	(0.0017)	(0.0016)	(0.0014)	(0.0070)	(0.0022)
Panel (D): Predicted data						
$\log(K)$	1.618^a	1.071^a	0.791^a	0.660^a	-0.203	1.270^c
	(0.078)	(0.203)	(0.181)	(0.175)	(1.771)	(0.694)
$[\log(K)]^2$	-0.040^a	-0.018^b	-0.007	-0.00018	0.037	-0.025
	(0.0033)	(0.0085)	(0.0076)	(0.0074)	(0.075)	(0.029)

Notes: OLS regressions with parcel size decile fixed effects in all columns. Centiles of distances are computed relative to the maximum distance from the centre in each urban area. In panels (C) and (D), K and R are predicted as in panel A of table 3. Bootstrapped standard errors in parentheses. 8,100 observations for each regression. The R^2 is 1.00 in all specifications. a, b, c: significant at 1%, 5%, 10%.

Table 14: log housing production by size class of urban areas, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Urban areas with 0 to 50,000 inhabitants, observed values									
log (K)	0.707[a]	0.705[a]	0.702[a]	0.703[a]	0.709[a]	0.715[a]	0.718[a]	0.722[a]	0.729[a]
	(0.0018)	(0.0019)	(0.0021)	(0.0023)	(0.0025)	(0.0031)	(0.0039)	(0.0046)	(0.0052)
Panel (B): Urban areas with 50,000 to 100,000 inhabitants, observed values									
log (K)	0.694[a]	0.688[a]	0.683[a]	0.685[a]	0.694[a]	0.707[a]	0.714[a]	0.710[a]	0.706[a]
	(0.0020)	(0.0018)	(0.0020)	(0.0022)	(0.0025)	(0.0027)	(0.0034)	(0.0053)	(0.0053)
Panel (C): Urban areas with 100,000 to 200,000 inhabitants, observed values									
log (K)	0.684[a]	0.681[a]	0.678[a]	0.677[a]	0.678[a]	0.683[a]	0.692[a]	0.700[a]	0.705[a]
	(0.00150)	(0.0014)	(0.0014)	(0.0017)	(0.0022)	(0.0024)	(0.0026)	(0.0025)	(0.0027)
Panel (D): Urban areas with 200,000 to 500,000 inhabitants, observed values									
log (K)	0.651[a]	0.650[a]	0.648[a]	0.644[a]	0.638[a]	0.636[a]	0.639[a]	0.651[a]	0.657[a]
	(0.0015)	(0.0013)	(0.0013)	(0.0016)	(0.0019)	(0.0022)	(0.0028)	(0.0026)	(0.0031)
Panel (E): Urban areas with more than 500,000 inhabitants (except Paris), observed values									
log (K)	0.621[a]	0.604[a]	0.586[a]	0.571[a]	0.567[a]	0.566[a]	0.559[a]	0.554[a]	0.547[a]
	(0.0016)	(0.0015)	(0.0016)	(0.0018)	(0.0020)	(0.0026)	(0.0039)	(0.0042)	(0.0047)
Panel (F): Paris, observed values									
log (K)	0.516[a]	0.521[a]	0.529[a]	0.537[a]	0.546[a]	0.552[a]	0.552[a]	0.551[a]	0.548[a]
	(0.0025)	(0.0024)	(0.0025)	(0.0029)	(0.0032)	(0.0034)	(0.0041)	(0.0042)	(0.0062)

Notes: OLS regressions with a constant in all columns. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Table 15: log housing production taking out construction costs, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Observed data after conditioning out constructions wages									
$\log(K)$	0.626^a	0.639^a	0.640^a	0.638^a	0.641^a	0.648^a	0.652^a	0.657^a	0.659^a
	(0.0011)	(0.00077)	(0.00083)	(0.0011)	(0.0012)	(0.0014)	(0.0017)	(0.0023)	(0.0027)
Panel (B): Observed data after conditioning out constructions wages									
$\log(K)$	-0.410^a	-0.476^a	-0.552^a	-0.410^a	-0.373^a	-0.274^a	-0.186^b	-0.254^b	-0.201
	(0.041)	(0.035)	(0.037)	(0.043)	(0.056)	(0.077)	(0.081)	(0.101)	(0.135)
$[\log(K)]^2$	0.044^a	0.047^a	0.050^a	0.044^a	0.043^a	0.039^a	0.035^a	0.038^a	0.036^a
	(0.0018)	(0.0013)	(0.0015)	(0.0018)	(0.0024)	(0.0033)	(0.0034)	(0.0043)	(0.0057)
Panel (C): Observed data after deflating by constructions wages									
$\log(K)$	0.625^a	0.639^a	0.640^a	0.638^a	0.642^a	0.649^a	0.653^a	0.658^a	0.660^a
	(0.0010)	(0.00080)	(0.00088)	(0.00097)	(0.0012)	(0.0017)	(0.0019)	(0.0022)	(0.0030)
Panel (D): Observed data after deflating by constructions wages									
$\log(K)$	-0.185^a	-0.290^a	-0.371^a	-0.241^a	-0.201^a	-0.101	0.011	-0.050	-0.005
	(0.044)	(0.032)	(0.035)	(0.046)	(0.064)	(0.066)	(0.092)	(0.097)	(0.127)
$[\log(K)]^2$	0.034^a	0.039^a	0.043^a	0.037^a	0.035^a	0.032^a	0.027^a	0.030^a	0.028^a
	(0.0019)	(0.0013)	(0.0015)	(0.0019)	(0.0027)	(0.0028)	(0.0039)	(0.0041)	(0.0054)

Notes: OLS regressions with a constant in all columns. In panels (A) and (B), in a first step log capital investment is regressed on log constructions wages to derive a predicted value for K. In panels (C) and (D), capital investment is directly deflated by construction wages (beyond the year effects that we also use to make R comparable across years). Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. a, b, c: significant at 1%, 5%, 10%.

Appendix G. Dismissing two explanations for differences in capital elasticity across urban areas

We first show that differences in capital elasticities across size classes of urban areas are unlikely to be the result of a complementarity between land and capital. To do this, we can use expression (14) derived in section 6 for a constant-elasticity-of-substitution approximation of the housing production function. Following expression (14), the capital elasticity for a parcel of size T in city c is equal to:

$$\frac{\partial \log H(K_c^*,T_c)}{\partial \log K_c^*} = \frac{\alpha (K_c^*)^{1-1/\sigma}}{\alpha (K_c^*)^{1-1/\sigma} + (1-\alpha) T_c^{1-1/\sigma}} = \frac{1}{1 + \frac{1-\alpha}{\alpha}\left(\frac{K_c^*}{T_c}\right)^{1/\sigma - 1}}\,, \tag{G1}$$

where σ is the elasticity of substitution and α is the share parameter in the CES production function.

We consider Paris for which the capital elasticity estimated in table 4 is equal to 0.539 and small urban areas (with population below 50,000) for which the capital elasticity estimated in the same table is equal to 0.712. We also consider parcels of 1,000 m^2. This is close to mean parcel size across all urban areas, as reported in table 1. To avoid computing construction costs from too few parcels, we compute the average construction cost for parcels of areas between 920 and 1100 m^2. There are 1,405 of them with average parcel size of 1,002 m^2 in Paris and 5,065 of them with average parcel size of 1,001 m^2 in small urban areas. Mean construction costs are 137,225 euros in small urban areas and 164,432 euros in Paris. For Paris, equation (G1) implies:

$$\frac{1}{1 + \frac{1-\alpha}{\alpha}(164.1)^{1/\sigma - 1}} = 0.539\,. \tag{G2}$$

For small urban areas, the same expression implies:

$$\frac{1}{1 + \frac{1-\alpha}{\alpha}(137.1)^{1/\sigma - 1}} = 0.712\,. \tag{G3}$$

Simple algebra shows that $\sigma = 0.194$ would be needed to satisfy these two equations. This value of σ is only a fraction of what we estimate below and is inconsistent with the stability of the capital elasticity we estimate within each class of urban areas in table 4. Comparing small urban areas with large urban areas with population above 500,000 instead of Paris leads to an even smaller value of 0.125 for σ.

The differences in capital elasticity we estimate across urban areas are also unlikely to be caused by differences in construction costs. As mentioned in the main text, construction wages are 14.2% higher in Paris than in small urban areas with a population less than 50,000. Construction wages are also 5.8% higher in large urban areas (with population above 500,000, excluding Paris) than in small urban areas. These figures are arguably an upper bound for the difference in construction costs since price differences for materials are expected to be less.

As above, we can use a constant-elasticity-of-substitution approximation of the production function. With $H_c = A\left(\alpha K_c^{(\sigma-1)/\sigma} + (1-\alpha)T_c^{(\sigma-1)/\sigma}\right)^{\sigma/(\sigma-1)}$ and the cost of capital equal to r_c in city c, profit maximization by house builders implies:

$$\frac{\partial H(K_c,T_c)}{\partial K_c} = \frac{H(K_c,T_c)^{1/\sigma}}{K_c^{1/\sigma}} = \frac{r_c}{P_c}, \tag{G4}$$

Then using this last expression, it is easy to show that:

$$\frac{\partial \log H(K_c,T_c)}{\partial \log K_c} = \frac{H(K_c,T_c)^{1/\sigma-1}}{K_c^{1/\sigma-1}} = \left(\frac{P_c}{r_c}\right)^{\sigma-1}, \tag{G5}$$

For a higher cost of capital in larger cities to explain a lower share of capital in construction costs as we observe in the data, we need $\sigma > 1$. Recall that table 4 reports that the cost share is 32.1% higher in small urban areas than in Paris (0.712 vs. 0.539) and 23.8% in small urban areas than in large urban areas (0.712 vs. 0.575). Then, even if we ignore the higher cost of housing in Paris and large urban areas relative to small urban areas, we need $\sigma = 3.10$ for a 14.2% difference in cost between Paris and small urban areas to explain a 32.1% difference in capital elasticity. We also need $\sigma = 4.79$ for a 5.8% difference in cost between large and small urban areas to explain a 23.8% difference in capital elasticity.

Appendix H. A two-period extension of our base model

We develop a two-period extension of the model in the main text. The argument readily generalises to many periods or to an infinite time horizon but the calculations below are simpler and more transparent with two periods, 1 and 2. With multiple periods, we need to distinguish between the asset value of housing and its rental value. We note the rental value of housing per unit P_1 in period 1 and P_2 in period 2.

If housing investment can only occur at the beginning of period 1, house builders seek to maximise $\pi = (P_1 + \delta P_2)\, H(K_1,T) - K_1$ where $\delta < 1$ is the actualisation factor from period 2 to period 1. We can define $\delta \equiv \frac{1}{1+r}$ where r is now the interest rate for the first period. With $P \equiv P_1 + \delta P_2$, this model is equivalent to the static model in the main text. That is, higher future housing prices, once appropriately discounted, are no different than higher current prices. Then, the elasticity of output with respect to capital is still equal to the share of capital in construction costs as per equation (4). Hence, with this framework the large differences in capital shares across urban areas reported in table 4 remain difficult to reconcile with the 'more stable' results of table 2.

We now consider that housing investment can be made in both period 1 and period 2. To distinguish between stocks and flows, we note I_1, the investment in housing made in period 1 and I_2, the investment made in period 2. We have $K_1 = I_1$ and $K_2 = (1-\tau)\,K_1 + I_2$ where $0 \leq \tau \leq 1$ is the depreciation of capital between period 1 and period 2.

Profit in period 1 is given by:

$$\pi_1 = P_1\,H(K_1,T) - I_1 - R_1 + \delta V_2 = P_1\,H(K_1,T) - K_1 - R_1 + \delta V_2\,, \tag{H1}$$

where V_2 is the resale value of the house at the beginning of period 2 before any second-period investment in housing. This resale value can be written as $V_2 = R_2 + (1-\tau)K_1$, which sums the remaining housing capital and the endogenously determined value of the land in period 2.

The profit of the builder rebuilding a house in period 2 is given by:

$$\pi_2 = P_2\,H(K_2,T) - I_2 - V_2 = P_2\,H(K_2,T) - I_2 - (1-\tau)K_1 - R_2 = P_2\,H(K_2,T) - K_2 - R_2\,, \tag{H2}$$

where we used the expressions above giving K_2 and V_2 to rewrite the profit in period 2 as a function of K_2 to obtain the last equality.

Maximising profits in period 2 with respect to K_2 implies that the optimal level of capital in the second period K_2^* is given implicitly by:

$$P_2\frac{\partial H(K_2^*)}{\partial K_2^*} = 1\,. \tag{H3}$$

Then, optimal investment is immediately obtained from $I_2^* = K_2^* - (1-\tau)K_1^*$ where K_1^* is derived below. Zero profit for builders in period 2, $\pi_2 = 0$, also implies:

$$R_2 = P_2\,H(K_2^*) - I_2^* - (1-\tau)K_1\,. \tag{H4}$$

Turning to period 1, we can use $V_2 = R_2 + (1-\tau)K_1$ and equation (H4) to rewrite period-1 profit in equation (H1) as:

$$\pi_1 = P_1\,H(K_1,T) - K_1 - R_1 + \delta(P_2\,H(K_2^*,T) - I_2^*)\,. \tag{H5}$$

The first-order condition for profit maximisation with respect to K_1 implies that optimal period-1 capital, K_1^*, is given by:

$$P_1\frac{\partial H(K_1^*,T)}{\partial K_1^*} = 1 - \delta\left(P_2\frac{\partial H(K_2^*)}{\partial K_2^*}\frac{\partial K_2^*}{\partial K_1^*} - \frac{\partial I_2^*}{\partial K_1^*}\right)\,. \tag{H6}$$

We can use expression (H3) (the envelop theorem) and $I_2^* = K_2^* - (1-\tau)K_1^*$ to rewrite the first-order condition (H6) as:

$$P_1 \frac{\partial H(K_1^*)}{\partial K_1^*} = 1 - \delta(1-\tau)\,. \tag{H7}$$

This expression corresponds to the optimal investment equation (1) in the main text for a two-period setting.

Then, it is useful to rewrite equation (H7) in elasticity form:

$$\frac{\partial \log H(K_1^*)}{\partial \log K_1^*} = \frac{[1-\delta(1-\tau)]\,K_1^*}{P_1 H(K_1^*)} = [1-\delta(1-\tau)]\,\frac{V_1}{P_1 H(K_1^*)}\frac{K_1^*}{K_1^* + R_1}\,, \tag{H8}$$

where the last equality is obtained using $V_1 = K_1^* + R_1$. Recall that with optimal investment and zero profit for the builder the value of a house is equal to the value of its land plus the optimal capital invested. Equation (H8) differs from equation (4) in the main text because it multiplies the cost share $\frac{K_1^*}{K_1^*+R_1}$ on the right-hand side by $1-\delta(1-\tau)$, a factor which accounts for the discount rate and the depreciation of capital, and by $\frac{V_1}{P_1 H(K_1^*)}$, the ratio of the value of the house by its rental value for the period.

While we use this expression in our empirical exercise in section 5.3, further insight can be gained from replacing $P_1 H(K_1^*)$ using equation (H5) after setting π_1 to zero. After simplifications, equation (H8) becomes:

$$\frac{\partial \log H(K_1^*)}{\partial \log K_1^*} = \frac{(1-\delta(1-\tau))K_1^*}{(1-\delta(1-\tau))K_1^* + (R_1 - \delta R_2)}\,. \tag{H9}$$

Denoting the growth rate of value of parcels $g \equiv R_2/R_1 - 1$ and expressing the discount rate as function of the interest rate as defined above, $\delta \equiv \frac{1}{1+r}$, equation (H9) can finally be rewritten as

$$\frac{\partial \log H(K_1^*)}{\partial \log K_1^*} = \frac{(r+\tau)K_1^*}{(r+\tau)K_1^* + (r-g)R_1}\,, \tag{H10}$$

where K_1^* is implicitly defined by equation (H7) and depends on P_1 but not on P_2. This expression for the elasticity of housing production with respect to capital is no longer equal to the cost share $\frac{K^*}{K^*+R}$ like in equation (4) of the main text as already noted. Instead, it is equal to the ratio of the cost of capital for the first period $(r+\tau)K_1^*$ divided by the sum of the cost of capital and the cost of land $(r-g)R_1$. Each factor is now adjusted by its user cost. This user cost is equal to the sum of the interest rate and the rate of depreciation for capital. It is equal to the interest rate minus the rate of appreciation for land. This correction is in the spirit of the user cost correction first proposed by Poterba (1984).

Appendix I. Supplementary results for section 5.3: Corrected cost shares

To compute the corrected cost shares described by equation (H8), we need empirical values for τ, the rate of depreciation of housing capital, $\delta = \frac{1}{1+r}$, the discount factor, and for each of the size classes of cities used in table 4, $V_1 / [P_1 H(K_1^*)]$, the ratio of housing value to housing rent.

Starting with the rate of depreciation of housing capital, we take an annual value of 1% for the entire country. In the French national accounts, housing depreciation can be computed as the difference between investment in housing and the increase in housing stocks. According to Commissariat Général au Développement Durable (2012), this difference in 2009 was about 15 bn euros, which corresponds to slightly less than 1% of GDP or just below 0.6% of the value of the stock. This is arguably a lower bound as much housing maintenance falls under home production and is not accounted for in national accounts.

For the discount rate, we compute it using $r = 4\%$ which corresponds to the average annual rate for mortgages in France during our study period according to Observatoire Crédit Logement / CSA.

To compute the ratio of property values to annual rents for each class of city size we proceed as follows. We use monthly rent and property price data for 2012 as described in Appendix C. Three caveats are worth keeping in mind: (i) rent data are for an average of all observed transactions by the data provider, (ii) they only cover municipalities with a population above 2,000, and (iii) property values are for a reference property computed from all transactions. We consider only the 1,938 municipalities in an urban area for which we observe a new construction during our period, rent data, and property price data. Our sample covers 85% of municipalities with a population above 5,000. Because some urban areas are much larger than others, we regress property rents and values on the inverse hyperbolic sine of the distance between the centroid of a property's municipality and the centroid of the urban area (which corresponds to the centroid of the main municipality) allowing for a different coefficient for each urban area. See Combes *et al.* (2019) for further discussion. We use the results of this estimation to compute a distance-corrected rent and property value for each municipality in the data. We then take the ratio of property value to annualized rents for each municipality. Finally, for each size class of urban area, we take the median ratio among all represented municipalities rather than the mean to avoid giving too much weight to a few outlier municipalities.

Appendix J. Supplementary results for section 5.4: Land use regulations

Table 16 re-estimates our base results with decile indicators for each quintiles of absolute and relative FAR stringency using either smoothed observed data or predicted data for R and K.

Table 17 duplicates table 16 but computes local floor-to-area ratios using only single-family homes built after 2000 instead of the entire stock.

Table 18 estimates our bases results separately for 2006 to 2011 and for 2012 because the planning regime changes in 2012.

Appendix K. Supplementary results for section 5.5: Housing heterogeneity

Table 19 report results for different levels of completion.

Table 20 reports result by occupational groups of buyers.

Table 16: log housing production, OLS land use regulations

Centiles	All urban areas	0-20	20-40	40-60	60-80	80-100
Panel (A): Floor-to-area ratio, observed data						
$\log(K)$	0.645^a	0.679^a	0.655^a	0.634^a	0.621^a	0.619^a
	(0.00088)	(0.0019)	(0.0020)	(0.0016)	(0.0018)	(0.0020)
Panel (B): Floor-to-area ratio, observed data						
$\log(K)$	0.068^a	-0.457^a	-0.107^a	-0.203^a	-0.184^a	-0.173^a
	(0.031)	(0.080)	(0.067)	(0.071)	(0.071)	(0.061)
$[\log(K)]^2$	0.024^a	0.048^a	0.032^a	0.035^a	0.034^a	0.033^a
	(0.0013)	(0.0034)	(0.0029)	(0.0030)	(0.0030)	(0.0026)
Panel (C): Floor-to-area ratio, predicted data						
$\log(K)$	0.661^a	0.698^a	0.673^a	0.649^a	0.637^a	0.612^a
	(0.00014)	(0.00006)	(0.00012)	(0.00006)	(0.00007)	(0.00013)
Panel (D): Floor-to-area ratio, predicted data						
$\log(K)$	1.362^a	1.112^a	1.629^a	1.138^a	0.852^a	0.767^a
	(0.027)	(0.013)	(0.022)	(0.010)	(0.013)	(0.025)
$[\log(K)]^2$	-0.030^a	-0.018^a	-0.040^a	-0.021^a	-0.009^a	-0.007^a
	(0.0011)	(0.00053)	(0.00092)	(0.00044)	(0.00055)	(0.00103)
Panel (E): Relative floor-to-area ratio, observed data						
$\log(K)$	0.645^a	0.612^a	0.642^a	0.649^a	0.653^a	0.662^a
	(0.00085)	(0.0027)	(0.0018)	(0.0017)	(0.0015)	(0.0012)
Panel (F): Relative floor-to-area ratio, observed data						
$\log(K)$	0.068^a	-0.254^a	0.216^a	0.315^a	0.158^a	-0.007
	(0.037)	(0.083)	(0.070)	(0.060)	(0.056)	(0.055)
$[\log(K)]^2$	0.024^a	0.036^a	0.018^a	0.014^a	0.021^a	0.028^a
	(0.0016)	(0.0034)	(0.0030)	(0.0026)	(0.0024)	(0.0023)
Panel (G): Relative floor-to-area ratio, predicted data						
$\log(K)$	0.661^a	0.622^a	0.658^a	0.666^a	0.671^a	0.680^a
	(0.00055)	(0.0040)	(0.0013)	(0.0010)	(0.00092)	(0.00097)
Panel (H): Relative floor-to-area ratio, predicted data						
$\log(K)$	1.362^a	0.876^a	0.811^a	0.537^a	0.345^a	0.599^a
	(0.084)	(0.469)	(0.119)	(0.099)	(0.093)	(0.110)
$[\log(K)]^2$	-0.030^a	-0.011^a	-0.006^a	0.005^a	0.014^a	0.003^a
	(0.0036)	(0.020)	(0.0050)	(0.0042)	(0.0040)	(0.0047)

Notes: OLS regressions with parcel size decile fixed effects in all columns. In panels (A)-(D), centiles of floor-to-area ratio are computed using the 30th percentile of property level floor-to-area ratio of all existing houses in each municipality. In panels (E)-(H), centiles of relative floor-to-area ratio are computed by measuring for each new construction the centile of their floor-to-area ratio in their municipal distribution before dividing them into five quantiles from least binding to most binding. In panels (B), (D), (F) and (H), K and R are predicted as in panel A of table 3. Bootstrapped standard errors in parentheses. 8,100 observations for each regression. The R^2 is 1.00 in all specifications. a, b, c: significant at 1%, 5%, 10%.

Table 17: log housing production, OLS land use regulations post 2000 data

Centiles	All urban areas	0-20	20-40	40-60	60-80	80-100
Panel (A): Floor-to-area ratio, observed data after 2000						
$\log(K)$	0.644^a	0.692^a	0.655^a	0.635^a	0.619^a	0.592^a
	(0.00084)	(0.0020)	(0.0019)	(0.0016)	(0.0016)	(0.0016)
Panel (B): Floor-to-area ratio, observed data after 2000						
$\log(K)$	0.081^a	-0.640^a	-0.869^a	-0.779^a	-0.771^a	-0.771^a
	(0.030)	(0.071)	(0.073)	(0.054)	(0.061)	(0.043)
$[\log(K)]^2$	0.024^a	0.056^a	0.064^a	0.060^a	0.058^a	0.057^a
	(0.0013)	(0.0030)	(0.0031)	(0.0023)	(0.0026)	(0.0018)
Panel (C): Floor-to-area ratio, predicted data after 2000						
$\log(K)$	0.659^a	0.717^a	0.655^a	0.646^a	0.627^a	0.586^a
	(0.00050)	(0.0014)	(0.0042)	(0.0070)	(0.0039)	(0.0012)
Panel (D): Floor-to-area ratio, predicted data after 2000						
$\log(K)$	1.622^a	1.727^a	-0.477^a	1.364^a	0.792^a	-0.529^a
	(0.077)	(0.206)	(0.664)	(1.646)	(0.723)	(0.151)
$[\log(K)]^2$	-0.041^a	-0.043^a	0.048^a	-0.030^a	-0.007^a	0.047^a
	(0.0032)	(0.0087)	(0.028)	(0.069)	(0.030)	(0.0063)
Panel (E): Relative floor-to-area ratio, observed data after 2000						
$\log(K)$	0.644^a	0.611^a	0.633^a	0.646^a	0.653^a	0.674^a
	(0.00070)	(0.0022)	(0.0018)	(0.0017)	(0.0014)	(0.0013)
Panel (F): Relative floor-to-area ratio, observed data after 2000						
$\log(K)$	0.081^a	-0.133^a	0.364^a	0.214^a	-0.013	-0.160^a
	(0.036)	(0.076)	(0.072)	(0.060)	(0.054)	(0.049)
$[\log(K)]^2$	0.024^a	0.031^a	0.011^a	0.018^a	0.028^a	0.035^a
	(0.0015)	(0.0032)	(0.0030)	(0.0025)	(0.0023)	(0.0021)
Panel (G): Relative floor-to-area ratio, predicted data after 2000						
$\log(K)$	0.659^a	0.614^a	0.643^a	0.660^a	0.675^a	0.699^a
	(0.00055)	(0.0018)	(0.0013)	(0.0010)	(0.0011)	(0.0011)
Panel (H): Relative floor-to-area ratio, predicted data after 2000						
$\log(K)$	1.622^a	0.543^a	0.646^a	0.561^a	0.456^a	0.728^a
	(0.078)	(0.128)	(0.119)	(0.080)	(0.079)	(0.078)
$[\log(K)]^2$	-0.041^a	0.003^a	0.001	0.004^a	0.009^a	-0.001
	(0.0033)	(0.0054)	(0.0050)	(0.0034)	(0.0033)	(0.0033)

Notes: OLS regressions with parcel size decile fixed effects in all columns. In panels (A)-(D), centiles of floor-to-area ratio are computed using the 30th percentile of property level floor-to-area ratio in each municipality for single-family homes built after 2000. In panels (E)-(H), centiles of relative floor-to-area ratio are computed by measuring for each new construction the centile of their floor-to-area ratio in their municipal distribution before dividing them into five quantiles from least binding to most binding. In panels (B), (D), (F) and (H), K and R are predicted as in panel A of table 3. Bootstrapped standard errors in parentheses. 8,100 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Table 18: 2006-2011 vs. 2012, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Observed data 2006-2011									
$\log(K)$	0.625[a]	0.639[a]	0.641[a]	0.639[a]	0.644[a]	0.651[a]	0.653[a]	0.660[a]	0.659[a]
	(0.0014)	(0.0013)	(0.0014)	(0.0015)	(0.0017)	(0.0017)	(0.0022)	(0.0027)	(0.0029)
Panel (B): Observed data 2006-2011									
$\log(K)$	0.142[a]	-0.032	-0.124[a]	-0.024	0.017	0.177[a]	0.339[a]	0.263[b]	0.155
	(0.042)	(0.031)	(0.037)	(0.050)	(0.058)	(0.080)	(0.105)	(0.127)	(0.138)
$[\log(K)]^2$	0.020[a]	0.028[a]	0.032[a]	0.028[a]	0.026[a]	0.020[a]	0.013[a]	0.017[a]	0.021[a]
	(0.0018)	(0.0013)	(0.0016)	(0.0021)	(0.0024)	(0.0037)	(0.0044)	(0.0054)	(0.0058)
Panel (C): Observed data 2012									
$\log(K)$	0.635[a]	0.638[a]	0.641[a]	0.643[a]	0.648[a]	0.652[a]	0.657[a]	0.664[a]	0.672[a]
	(0.0015)	(0.0013)	(0.0014)	(0.0015)	(0.0018)	(0.0023)	(0.0029)	(0.0035)	(0.0039)
Panel (D): Observed data 2012									
$\log(K)$	-0.158[b]	-0.095	-0.074	-0.044	-0.062	-0.189	-0.259[c]	-0.211	-0.116
	(0.079)	(0.065)	(0.061)	(0.068)	(0.083)	(0.115)	(0.143)	(0.147)	(0.178)
$[\log(K)]^2$	0.033[a]	0.031[a]	0.030[a]	0.029[a]	0.030[a]	0.035[a]	0.039[a]	0.037[a]	0.033[a]
	(0.0034)	(0.0028)	(0.0026)	(0.0029)	(0.0035)	(0.0048)	(0.0060)	(0.0062)	(0.0075)

Notes: OLS regressions with a constant in all columns. In panels (A) and (B), data from 2006 to 2011. In panels (C) and (D), data from 2012. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Table 19: log housing production in urban areas at various degrees of completion, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Fully finished units, observed data									
log (K)	0.645[a]	0.654[a]	0.656[a]	0.653[a]	0.657[a]	0.662[a]	0.667[a]	0.669[a]	0.661[a]
	(0.0017)	(0.0015)	(0.0016)	(0.0018)	(0.0020)	(0.0024)	(0.0028)	(0.0033)	(0.0051)
Panel (B): Fully finished units, predicted data									
log (K)	0.658[a]	0.660[a]	0.663[a]	0.666[a]	0.671[a]	0.677[a]	0.682[a]	0.688[a]	0.691[a]
	(0.0017)	(0.0011)	(0.0011)	(0.0013)	(0.0016)	(0.0020)	(0.0026)	(0.0032)	(0.0039)
Panel (C): Ready to decorate, observed data									
log (K)	0.620[a]	0.633[a]	0.636[a]	0.636[a]	0.639[a]	0.647[a]	0.652[a]	0.654[a]	0.661[a]
	(0.0011)	(0.0010)	(0.0011)	(0.0011)	(0.0012)	(0.0015)	(0.0020)	(0.0022)	(0.0035)
Panel (D): Ready to decorate, predicted data									
log (K)	0.644[a]	0.646[a]	0.647[a]	0.649[a]	0.654[a]	0.661[a]	0.667[a]	0.669[a]	0.677[a]
	(0.0013)	(0.00081)	(0.00074)	(0.00071)	(0.00094)	(0.00132)	(0.00156)	(0.0019)	(0.0024)
Panel (E): Structure completed, observed data									
log (K)	0.593[a]	0.603[a]	0.608[a]	0.607[a]	0.607[a]	0.611[a]	0.615[a]	0.614[a]	0.614[a]
	(0.0030)	(0.0028)	(0.0031)	(0.0040)	(0.0043)	(0.0045)	(0.0053)	(0.0068)	(0.0090)
Panel (F): Structure completed, predicted data									
log (K)	0.619[a]	0.615[a]	0.611[a]	0.608[a]	0.606[a]	0.612[a]	0.622[a]	0.632[a]	0.636[a]
	(0.0047)	(0.0032)	(0.0025)	(0.0026)	(0.0035)	(0.0043)	(0.0052)	(0.0071)	(0.0081)

Notes: OLS regressions with a constant in all columns. In panels (B), (D), and (F), K and R are predicted as in panel A of table 3. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. *a*, *b*, *c*: significant at 1%, 5%, 10%.

Table 20: log housing production in urban areas across owners' occupations, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Executives, observed data									
$\log(K)$	0.612[a]	0.624[a]	0.627[a]	0.625[a]	0.628[a]	0.628[a]	0.628[a]	0.631[a]	0.620[a]
	(0.0020)	(0.0017)	(0.0017)	(0.0021)	(0.0026)	(0.0026)	(0.0032)	(0.0036)	(0.0062)
Panel (B): Executives, predicted data									
$\log(K)$	0.633[a]	0.632[a]	0.634[a]	0.638[a]	0.644[a]	0.650[a]	0.652[a]	0.656[a]	0.655[a]
	(0.0019)	(0.0012)	(0.0012)	(0.0013)	(0.0016)	(0.0023)	(0.0032)	(0.0040)	(0.0050)
Panel (C): Intermediate occupations, observed data									
$\log(K)$	0.627[a]	0.636[a]	0.639[a]	0.639[a]	0.642[a]	0.645[a]	0.653[a]	0.657[a]	0.650[a]
	(0.0033)	(0.0032)	(0.0035)	(0.0036)	(0.0040)	(0.0047)	(0.0050)	(0.0060)	(0.0087)
Panel (D): Intermediate occupations, predicted data									
$\log(K)$	0.652[a]	0.650[a]	0.647[a]	0.645[a]	0.647[a]	0.651[a]	0.658[a]	0.663[a]	0.666[a]
	(0.0024)	(0.0017)	(0.0015)	(0.0018)	(0.0022)	(0.0027)	(0.0035)	(0.0039)	(0.0048)
Panel (E): Clerical and blue-collar workers, observed data									
$\log(K)$	0.641[a]	0.645[a]	0.647[a]	0.649[a]	0.654[a]	0.658[a]	0.663[a]	0.669[a]	0.674[a]
	(0.0014)	(0.0011)	(0.0010)	(0.0012)	(0.0014)	(0.0016)	(0.0019)	(0.0020)	(0.0025)
Panel (F): Clerical and blue-collar workers, predicted data									
$\log(K)$	0.663[a]	0.657[a]	0.654[a]	0.653[a]	0.654[a]	0.656[a]	0.659[a]	0.660[a]	0.660[a]
	(0.0020)	(0.0012)	(0.00090)	(0.00090)	(0.0012)	(0.0017)	(0.0023)	(0.0025)	(0.0029)

Notes: OLS regressions with a constant in all columns. In panels (B), (D), and (F), K and R are predicted as in panel (A) of table 3. Bootstrapped standard errors in parentheses. 900 observations for each regression. The R^2 is 1.00 in all specifications. a, b, c: significant at 1%, 5%, 10%.

Appendix L. Supplementary results for section 6

Table 21 duplicates table 2 after fitting the data to a Cobb-Douglas function in panels (A) and (B), a CES function in panels (C) and (D), a second-order translog function in panels (E) and (F), and a third order translog in panels (G) and (H).

Table 22 duplicates table 3 in the same way using predicted values for housing capital and parcel price.

Appendix M. More on measurement error and smoothing

A. Details about the regressions of table 6

We consider seven sets of estimators for our parameters of interest:

1. Traditional regression. Following much of the literature, we estimate a linear specification of the log of housing capital per square metre of land, $\log(K/T)$, on the log parcel value per square metre, $\log(R/T)$. As shown below, the estimated constant $\widehat{\alpha}$ of the regression is an estimator of $a = \sigma \log[\alpha/(1-\alpha)]$ and the estimated coefficient of the explanatory variable is an estimator of the elasticity of substitution $\widehat{\sigma}$. An estimator of parameter α can then be recovered using the formula: $\widehat{\alpha} = \exp(\widehat{a}/\widehat{\sigma}) / [1 + \exp(\widehat{a}/\widehat{\sigma})]$.

2. Traditional approach with smoothing. We estimate a linear regression of $\log(K/T)$ on a smoothed version of $\log(R/T)$ that is obtained by replacing R with its kernel estimator, a bivariate normal kernel with rule-of-thumb bandwidth as used in our base approach. As previously, an estimator of α can be derived from the coefficients of the regression.

3. EGS. As described in Appendix B, Epple *et al.* (2010) show that with competitive house builders facing a concave, constant returns-to-scale production function, there is a relationship between R/T, the parcel price per square metre, and $V/T = PH/T$, the housing value per square metre of land: $R/T = f(V/T)$. Thanks to the zero profit condition, we also have: $K/T = V/T - f(V/T)$. We can use these expressions to recover predictors of parcel price per square metre $\widehat{R/T}$ and housing capital per square metre of land $\widehat{K/T}$ after reconstructing the house value as $V = K + R$ and approximating the function $f(\cdot)$ with a polynomial expansion, which we choose to be of order 10. Following Ahlfeldt and McMillen (2020), we then estimate a linear specification of

Table 21: log housing production fitting specific functional forms, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Cobb-Douglas									
$\log(K)$	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]
	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)
Panel (B): Cobb-Douglas									
$\log(K)$	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]	0.634[a]
	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)	(0.00070)
$[\log(K)]^2$	-1.4e-7	1.4e-7	1.1e-7	-0.18e-7	-0.57e-7	-1.2e-7	0.93e-7	1.9e-7	1.3e-7
	(1.5e-7)	(1.5e-7)	(1.6e-7)	(1.4e-7)	(1.7e-7)	(1.7e-7)	(1.5e-7)	(1.6e-7)	(1.7e-7)
Panel (C): CES									
$\log(K)$	0.638[a]	0.636[a]	0.634[a]	0.633[a]	0.632[a]	0.631[a]	0.630[a]	0.630[a]	0.629[a]
	(0.0010)	(0.00080)	(0.00071)	(0.00070)	(0.00075)	(0.00081)	(0.00089)	(0.0010)	(0.0011)
Panel (D): CES									
$\log(K)$	0.563[a]	0.561[a]	0.559[a]	0.557[a]	0.556[a]	0.555[a]	0.554[a]	0.554[a]	0.553[a]
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)
$[\log(K)]^2$	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]	0.0032[a]
	(0.00053)	(0.00053)	(0.00053)	(0.00053)	(0.00053)	(0.00054)	(0.00054)	(0.00054)	(0.00054)
Panel (E): Second-order translog									
$\log(K)$	0.628[a]	0.633[a]	0.637[a]	0.641[a]	0.643[a]	0.646[a]	0.648[a]	0.650[a]	0.651[a]
	(0.0010)	(0.00079)	(0.00072)	(0.00075)	(0.00084)	(0.00094)	(0.0010)	(0.0011)	(0.0012)
Panel (F): Second-order translog									
$\log(K)$	-0.149[a]	-0.144[a]	-0.140[a]	-0.136[a]	-0.134[a]	-0.131[a]	-0.129[a]	-0.127[a]	-0.126[a]
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
$[\log(K)]^2$	0.033[a]	0.033[a]	0.033[a]	0.033[a]	0.033[a]	0.033[a]	0.033[a]	0.033[a]	0.033[a]
	(0.00082)	(0.00082)	(0.00082)	(0.00082)	(0.00082)	(0.00082)	(0.00082)	(0.00082)	(0.00082)
Panel (G): Third-order translog									
$\log(K)$	0.633[a]	0.635[a]	0.638[a]	0.641[a]	0.645[a]	0.648[a]	0.652[a]	0.655[a]	0.658[a]
	(0.00116)	(0.00082)	(0.00081)	(0.00078)	(0.00082)	(0.00099)	(0.0013)	(0.0016)	(0.0020)
Panel (H): Third-order translog									
$\log(K)$	-0.048	-0.032	-0.017	-0.005	0.006	0.015	0.024	0.033	0.040
	(0.037)	(0.024)	(0.021)	(0.024)	(0.030)	(0.035)	(0.041)	(0.046)	(0.051)
$[\log(K)]^2$	0.029[a]	0.028[a]	0.028[a]	0.027[a]	0.027[a]	0.027[a]	0.026[a]	0.026[a]	0.026[a]
	(0.0016)	(0.0010)	(0.00087)	(0.0010)	(0.0012)	(0.0015)	(0.0017)	(0.0020)	(0.0022)

Notes: OLS regressions with a constant in all columns. 900 observations for each regression. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. For the second-order translog, there is a single coefficient for all deciles of parcel size for the term in log K squared by definition.

Table 22: log housing production fitting specific functional forms and using predicted values, OLS by parcel size decile

Decile	1	2	3	4	5	6	7	8	9
Panel (A): Cobb-Douglas									
$\log(K)$	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]
	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)
Panel (B): Cobb-Douglas									
$\log(K)$	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]	0.652[a]
	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)	(0.00044)
$[\log(K)]^2$	3.8e-7	-3.4e-7	2.8e-7	-3.2e-7	7.6e-7	1.9e-7	1.5e-7	5.4e-7	1.8e-7
	(5.8e-7)	(6.0e-7)	(6.0e-7)	(6.3e-7)	(5.6e-7)	(5.7e-7)	(5.8e-7)	(5.6e-7)	(5.0e-7)
Panel (C): CES									
$\log(K)$	0.635[a]	0.644[a]	0.650[a]	0.655[a]	0.659[a]	0.663[a]	0.666[a]	0.669[a]	0.671[a]
	(0.00091)	(0.00056)	(0.00044)	(0.00050)	(0.00062)	(0.00076)	(0.00089)	(0.0010)	(0.0011)
Panel (D): CES									
$\log(K)$	0.938[a]	0.944[a]	0.948[a]	0.951[a]	0.953[a]	0.955[a]	0.957[a]	0.959[a]	0.960[a]
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
$[\log(K)]^2$	-0.013[a]	-0.013[a]	-0.013[a]	-0.012[a]	-0.012[a]	-0.012[a]	-0.012[a]	-0.012[a]	-0.012[a]
	(0.00066)	(0.00065)	(0.00064)	(0.00063)	(0.00062)	(0.00062)	(0.00061)	(0.00060)	(0.00060)
Panel (E): Second-order translog									
$\log(K)$	0.634[a]	0.643[a]	0.649[a]	0.654[a]	0.659[a]	0.662[a]	0.666[a]	0.668[a]	0.671[a]
	(0.00095)	(0.00060)	(0.00047)	(0.00052)	(0.00065)	(0.00079)	(0.00093)	(0.0011)	(0.0012)
Panel (F): Second-order translog									
$\log(K)$	1.593[a]	1.602[a]	1.608[a]	1.613[a]	1.618[a]	1.621[a]	1.624[a]	1.627[a]	1.630[a]
	(0.054)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
$[\log(K)]^2$	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]	-0.041[a]
	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Panel (G): Third-order translog									
$\log(K)$	0.646[a]	0.645[a]	0.649[a]	0.654[a]	0.659[a]	0.665[a]	0.672[a]	0.678[a]	0.684[a]
	(0.00140)	(0.00056)	(0.00054)	(0.00061)	(0.00067)	(0.00080)	(0.0010)	(0.0013)	(0.0014)
Panel (H): Third-order translog									
$\log(K)$	0.579[a]	1.126[a]	1.538[a]	1.868[a]	2.144[a]	2.382[a]	2.591[a]	2.777[a]	2.945[a]
	(0.130)	(0.077)	(0.054)	(0.060)	(0.080)	(0.102)	(0.122)	(0.141)	(0.159)
$[\log(K)]^2$	0.0028	-0.020[a]	-0.038[a]	-0.051[a]	-0.063[a]	-0.073[a]	-0.081[a]	-0.089[a]	-0.096[a]
	(0.0055)	(0.0033)	(0.0023)	(0.0025)	(0.0034)	(0.0043)	(0.0052)	(0.0060)	(0.0067)

Notes: OLS regressions with a constant in all columns. Capital and parcel price are predicted from demand-related factors as in table 3. Observed values of parcel size are used. 900 observations for each regression. The R^2 is 1.00 in all specifications. Bootstrapped standard errors in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. For the second-order translog, there is a single coefficient for all deciles of parcel size for the term in log K squared by definition.

$\log\left(\widehat{K/T}\right)$ on $\log\left(\widehat{R/T}\right)$ as with the traditional regression. Using the same transformation as in 1. and 2., we end up with an estimated share parameter and an estimated elasticity of substitution for the corresponding CES production function.

4. *EGS with smoothing*. We follow the same approach as in 3. but use a smoothed version of R when reconstructing the value of houses.

5. *Cost share*. We obtain estimates of α and σ by minimising the difference between the cost shares computed from the data and the theoretical expression obtained for a CES production function in equation (14) using non-linear least squares.

6. *Cost share with smoothing*. This duplicates approach 5. except that we use the smoothed version of R when computing the cost shares from the data.

7. *Our approach*. It is similar to the cost share approach with smoothed parcel prices except that, instead of considering differences between the cost shares computed from the data and the theoretical cost shares for every observation, we consider values on a 300×300 grid, and we weight each point on the grid with the sum of kernel weights (since the predictor of parcel price is more accurate when there are more observations in the neighborhood). This leads to a regression with 90,000 observations.

Table 6 in the text reports the results of these estimations. We also assess the robustness of these estimation techniques in section B. below with Monte-Carlo simulations.

B. Estimating the traditional regression with heterogeneous coefficients

To capture the effects of factor heterogeneity, we extend the standard CES production function to heterogeneous coefficients:

$$H_i(K_i,T_i) = A_i\left[\alpha_i K_i^{\frac{\sigma_i-1}{\sigma_i}} + (1-\alpha_i)T_i^{\frac{\sigma_i-1}{\sigma_i}}\right]^{\frac{\sigma_i}{\sigma_i-1}}, \tag{M1}$$

for any parcel i. Following a derivation analogous to that leading to equation (14), we obtain:

$$\frac{K_i}{K_i+R_i} = \frac{\alpha_i K_i^{\frac{\sigma_i-1}{\sigma_i}}}{\alpha_i K_i^{\frac{\sigma_i-1}{\sigma_i}} + (1-\alpha_i)T_i^{\frac{\sigma_i-1}{\sigma_i}}}, \tag{M2}$$

where K_i is the profit-maximising capital investment (which we do not star to ease notations). Simple algebra shows that this last equation is equivalent to:

$$\log\left(\frac{K_i}{T_i}\right) = a_i + \sigma_i \log\left(\frac{R_i}{T_i}\right), \tag{M3}$$

where $a_i = \sigma_i \log\left(\frac{\alpha_i}{1-\alpha_i}\right)$ and $R_i = R\left(K_i, T_i, a_i, \sigma_i\right)$.

The observed price of parcels may contain some measurement error: $\log \widetilde{R}_i = \log R_i + \varepsilon_i$ where R_i is the 'true' price of parcel i and ε_i is an independent, identically distributed, and centred error term orthogonal to all other quantities. To simplify the notations further, we note $X_i^0 \equiv \log\left(\frac{R_i}{T_i}\right)$ the true value of log parcel price per square metre, $X_i \equiv \log\left(\frac{\widetilde{R}_i}{T_i}\right)$ its observed value, and $Y_i \equiv \log\left(\frac{K_i}{T_i}\right)$ the log of housing investment per square metre of land. In practice, the production function is not exactly CES as argued in section 6. Hence, there is a mis-specification term ψ_i that we add to the equilibrium relationship between capital per square metre and parcel price per square metre in equation (M3). The specification brought to the data is given by:

$$Y_i = a_i + \sigma_i X_i + \psi_i - \sigma_i \varepsilon_i \,. \tag{M4}$$

The literature usually estimates a version of this equation with homogenous coefficients a and σ. We assess what the presence of heterogeneous coefficients in the data generating process implies for the estimator of the main coefficient of interest, σ, the elasticity of substitution (a similar argumentation holds for a). We can rewrite equation (M4) as:

$$Y_i = a + \sigma_i X_i + \left[\psi_i - \sigma_i \varepsilon_i + a_i - a + \left(\sigma_i - \sigma\right) X_i\right] = a + \sigma X_i + \eta_i \,, \tag{M5}$$

with $\eta_i \equiv \psi_i - \sigma \varepsilon_i + a_i - a + \left(\sigma_i - \sigma\right) X_i$. Denote $X_\bullet$ and $Y_\bullet$ the sample means of, respectively, X_i and Y_i. According to the Frisch-Waugh theorem, the OLS estimator of σ is given by:

$$\begin{aligned} \widehat{\sigma} &= \left[\frac{1}{N}\sum_i (X_i - X_\bullet)'(X_i - X_\bullet)\right]^{-1} \frac{1}{N}\sum_i (X_i - X_\bullet)'(Y_i - Y_\bullet) \\ &\longrightarrow V(X_i)^{-1}\mathrm{cov}(X_i, Y_i)\,, \end{aligned} \tag{M6}$$

where we have:

$$\begin{aligned} \mathrm{cov}(X_i, Y_i) &= \mathrm{cov}\left(X_i, a_i + \sigma_i X_i^0 + \psi_i\right) \\ &= \mathrm{cov}\left(X_i^0 + \varepsilon_i, a_i + \sigma_i X_i^0\right) \\ &= \mathrm{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right) \end{aligned} \tag{M7}$$

and the limit of the estimator $\widehat{\sigma}$ can be decomposed as:

$$\widehat{\sigma} \longrightarrow E\left(\sigma_i\right) + \mathit{Unobs.\,Bias} + \mathit{M.\,Error\,Bias}\,. \tag{M8}$$

where *Unobs. Bias* and *M. Error Bias* capture the divergence from the average elasticity of substitution due to the heterogeneity of production function parameters and the existence of measurement errors, respectively. We have:

$$M.\,Error\;Bias = \left[V(X_i)^{-1} - V(X_i^0)^{-1}\right] \mathrm{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right)$$
$$= -\frac{V(\varepsilon_i)}{V(X_i^0)\left[V(X_i^0) + V(\varepsilon_i)\right]} \mathrm{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right) . \quad \text{(M9)}$$

When there is heterogeneity in the parameters, this term is hard to sign. When, production function parameters are constant ($a_i = a$ and $\sigma_i = \sigma$), we get:

$$M.\,Error\;Bias = -\sigma \frac{V(\varepsilon_i)}{V(X_0) + V(\varepsilon_i)} < 0\,. \quad \text{(M10)}$$

The measurement error bias is negative and depends on the respective importance of the variances of measurement error and log parcel price per square metre, consistently with standard results in the literature. We also have:

$$Unobs.\,Bias = V\left(X_i^0\right)^{-1} \left[\mathrm{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right) - E(\sigma_i) V\left(X_i^0\right)\right] . \quad \text{(M11)}$$

This bias is different from zero because the log parcel price per square metre depends on a_i and σ_i. If these two terms were independent, the bias would be zero, as we can check that: $\mathrm{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right) = \mathrm{cov}\left(X_i^0, \sigma_i X_i^0\right) = E(\sigma_i)V(X_i^0)$. When these two terms are not independent, the object to which $\widehat{\sigma}$ converges has no clear interpretation. Even in the absence of measurement errors, this object is the sum of the average elasticity of substitution and a deviation due to the correlation between individual-specific coefficients and the log of parcel price per square metre at the equilibrium.

C. Estimating the traditional regression with smoothing

We now assess the effect of replacing $\widetilde{R}_i$ on the right-hand side of specification (M5) with a smoothed version of it. Consider for simplicity that the smoothing amounts to averaging parcel prices over the M constructions most similar to i (including i), denoted Θ_i, in the sense that for any $j \in \Theta_i$, we have K_j close to K_i and T_j close to T_i. The intuitions are similar when using a kernel for smoothing instead of considering the closest neighbors. In this case, we have:

$$X_j = X_j^0 + \varepsilon_j = X_i^0 + \zeta_{i,j} + \varepsilon_j\,, \quad \text{(M12)}$$

with $\zeta_{i,j} \equiv X_j^0 - X_i^0$. We now consider $X_i^\bullet = \frac{1}{M}\sum_{j \in \Theta_i} X_j$ instead of X_i in the estimated specification:

$$X_i^\bullet = X_i^0 + \zeta_i^\bullet + \varepsilon_i^\bullet\,, \quad \text{(M13)}$$

where $\zeta_i^\bullet = \frac{1}{M}\sum_{j\in\Theta_i}\zeta_{i,j}$ and $\varepsilon_i^\bullet = \frac{1}{M}\sum_{j\in\Theta_i}\varepsilon_j$. The specification can then be written as:

$$\begin{aligned} Y_i &= a_i + \sigma_i X_i^0 + \psi_i \\ &= a + \sigma X_i^0 + \left[\psi_i + a_i - a + (\sigma_i - \sigma) X_i^0\right] \\ &= a + \sigma X_i^\bullet + \left[\psi_i + a_i - a + (\sigma_i - \sigma) X_i^0 + \sigma\left(X_i^0 - X_i^\bullet\right)\right] \\ &= a + \sigma X_i^\bullet + \widetilde{\eta}_i\,, \end{aligned} \tag{M14}$$

where the residual is now:

$$\widetilde{\eta}_i = \psi_i + a_i - a + (\sigma_i - \sigma) X_i^0 - \sigma\left(\varepsilon_i^\bullet + \zeta_i^\bullet\right)\,. \tag{M15}$$

There are two differences relative to the situation without smoothing. We consider an average of measurement errors in the neighborhood of i instead of the measurement error of i and there is an additional term corresponding to the mis-specification introduced by the use of the average parcel price per square metre in the neighbourhood of i instead of parcel price per square metre of i. We now have:

$$\begin{aligned} \widehat{\sigma} &= \left[\frac{1}{N}\sum_i \left(X_i^\bullet - X_\bullet^\bullet\right)'\left(X_i^\bullet - X_\bullet^\bullet\right)\right]^{-1}\frac{1}{N}\sum_i \left(X_i^\bullet - X_\bullet^\bullet\right)\left(Y_i - Y_\bullet\right) \\ &\longrightarrow V\left(X_i^\bullet\right)^{-1}\operatorname{cov}\left(X_i^\bullet, Y_i\right)\,, \end{aligned} \tag{M16}$$

where $X_\bullet^\bullet$ is the sample average of $X_i^\bullet$. The covariance term verifies:

$$\begin{aligned} \operatorname{cov}\left(X_i^\bullet, Y_i\right) &= \operatorname{cov}\left(X_i^\bullet, a_i + \sigma_i X_i^0 + \psi_i\right) \\ &= \operatorname{cov}\left(X_i^0 + \zeta_i^\bullet + \varepsilon_i^\bullet, a_i + \sigma_i X_i^0\right) \\ &= \operatorname{cov}\left(X_i^0, a_i + \sigma_i X_i^0\right) + \operatorname{cov}\left(\zeta_i^\bullet, a_i + \sigma_i X_i^0\right)\,. \end{aligned} \tag{M17}$$

This covariance involves the same term as without smoothing and an additional one coming from the approximation of X_i^0 with a local average. The limit of the estimator $\widehat{\sigma}$ can then be decomposed in the following way:

$$\widehat{\sigma} \longrightarrow E\left(\sigma_i\right) + Unobs.\ Bias\ +\ M.\ Error\ Bias\ +\ Smooth.\ Bias\,. \tag{M18}$$

Compared to the case without smoothing, the bias due to measurement errors is modified but the one due to the heterogeneity in production function parameters remains the same. There is also an additional bias *Smooth. Bias* that comes from the smoothing of parcel prices per square metre. We now have:

$$
\begin{aligned}
\textit{M. Error Bias} &= \left[V\left(X_i^0+\varepsilon_i^\bullet\right)^{-1} - V\left(X_i^0\right)^{-1}\right]\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right) \\
&= -\frac{V\left(\varepsilon_i^\bullet\right)}{V\left(X_i^0\right)\left[V\left(X_i^0\right)+V\left(\varepsilon_i^\bullet\right)\right]}\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right) \\
&= -\frac{1}{M}\frac{V\left(\varepsilon_i\right)}{V\left(X_i^0\right)\left[V\left(X_i^0\right)+\frac{1}{M}V(\varepsilon_i)\right]}\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right) \\
&= -\left[\frac{V\left(X_i^0\right)+V(\varepsilon_i)}{MV\left(X_i^0\right)+V(\varepsilon_i)}\right]\frac{V(\varepsilon_i)\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right)}{V\left(X_i^0\right)\left[V\left(X_i^0\right)+V(\varepsilon_i)\right]}\,. \qquad \text{(M19)}
\end{aligned}
$$

Since the first right-hand side term in brackets is lower than 1, the bias due to measurement errors is smaller in absolute term than without smoothing and decreasing with the number of neighbors used in the approximation. In fact, we can also see that when $V(\varepsilon_i) \ll V\left(X_i^0\right)$ this bias is $1/M$ lower than that without smoothing. The bias due to heterogeneity in production function parameters is the same as before and is still given by:

$$
\textit{Unobs. Bias} = V\left(X_i^0\right)^{-1}\left[\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right) - E(\sigma_i)V\left(X_i^0\right)\right]\,. \qquad \text{(M20)}
$$

Finally, the bias coming from smoothing is given by:

$$
\begin{aligned}
\textit{Smooth. Bias} &= \left[V\left(X_i^0+\zeta_i^\bullet+\varepsilon_i^\bullet\right)^{-1} - V\left(X_i^0+\varepsilon_i^\bullet\right)^{-1}\right]\operatorname{cov}\left(X_i^0, a_i+\sigma_i X_i^0\right) \\
&\quad + V\left(X_i^0+\zeta_i^\bullet+\varepsilon_i^\bullet\right)^{-1}\operatorname{cov}\left(\zeta_i^\bullet, a_i+\sigma_i X_i^0\right)\,. \qquad \text{(M21)}
\end{aligned}
$$

It is the sum of two terms. The first one comes from the fact that smoothing affects the variance of the explanatory variable and the second one arises from the disruption of the covariance between the outcome and explanatory variable. We note that the first term is influenced by the variance of measurement errors. In fact, is it not possible to provide a linear decomposition that additively separates strictly all the effects.

It can be seen from our decomposition that smoothing with a large number of parcels (M large) decreases the bias due to measurements error but introduces a specification bias, as it usually does.

D. Monte-Carlo simulations with homogeneous coefficients

We now conduct Monte-Carlo simulations to compare the performances of the various approaches analysed in section 6. More specifically, we assess their robustness to measurement error and the importance of the specification bias introduced by smoothing.

Consider that the data generating process is such that the production function is, for now, CES with constant coefficients. We can generate a parcel price for every observation in the data from the amount of housing capital and the area of the parcel, provided that values for α and σ are available. From equation (M3), the price of parcels must verify:

$$R = K\left(\frac{1-\alpha}{\alpha}\right)\left(\frac{K}{T}\right)^{\frac{\sigma-1}{\sigma}} . \qquad \text{(M22)}$$

In practice, we need values of α and σ that make empirical sense. The first set of values we use is from the cost share estimation in table 6, panel A, and column 5: $\alpha = 0.60$ and $\sigma = 1.05$. These values are close to those obtained from the CES approximation of our base approach. We then consider a range of alternative values for σ between 0.6 and 2. To retain empirically meaningful values of α, we impose set values of σ to our cost share non-linear regression to estimate the corresponding values for α. Cobb-Douglas is a special case here obtained with $\sigma = 1$ and $\alpha = 0.65$ above. In the end, we consider:

Table 23: Values of σ and α for Monte-Carlo simulations

σ	1.05	0.6	0.8	1	1.2	1.5	2
α	0.60	0.98	0.87	0.65	0.45	0.26	0.13

We also introduce some measurement error. The simulated price of parcel i verifies: $\log \widetilde{R} = \log R + \sqrt{\delta}.\sigma_V.u$ where $\log R$ is obtained from equation (M22) and $\sqrt{\delta}.\sigma_V.u$ is the noise we introduce. The chosen value for parameter σ_V is the standard deviation of the difference between the observed and the smoothed prices of parcels (where smoothing is as per our base approach). In essence, we interpret the difference between the observed value and the prediction from smoothing as the measurement error in the data. This variation represents 91.6% the standard deviation of the logarithm of observed land prices. The term u is drawn in a centred normal law with unit variance. In results not reported below, we also experimented with a uniform law, also with unit variance (where we draw d uniformly from [0,1] and apply the transformation $u = \sqrt{12}.(d - .5)$ since the variance of a uniform law [0,1] is $1/12$). The results are very similar. Finally, δ is a scale parameter that determines the importance of the variance of measurement errors as a fraction of σ_V^2, since we have $V\left(\sqrt{\delta}.\sigma_V.u\right) = \delta\sigma_V^2$. For this parameter, we consider five values: 0, 0.25, 0.5, 1, and 2.

E. Results for homogeneous coefficients

We first consider the case with no heterogeneity in the parameters of the CES production function. The estimated parameters are reported in table 24. They show that the traditional approach behaves well as long as there is no measurement error (panel A). As soon as measurement error is introduced in the data generating process (panels B to E), the estimated parameters are biased and the biases increase with the variance of the measurement error. Unsurprisingly the bias is towards zero for the elasticity of substitution which is directly estimated and more complex for the share parameter α which is inferred from the constant and the elasticity of substitution. Interestingly, the traditional approach with smoothed parcel prices behaves well even with heavy measurement error.

The patterns are similar for the adaptation of the approach of Epple *et al.* (2010) by Ahlfeldt and McMillen (2020) to estimate a CES production function. We observe the same difference between estimates using observed parcel prices and those using smoothed parcel prices, with albeit smaller biases for the estimates using observed parcel prices. We also note that the cost share approach yields estimated parameters that are only mildly biased, even when we use observed parcel prices which are infused with a lot of measurement error. Recall however that, with this approach, it is the dependent variable rather than an explanatory variable that is affected by the measurement error. The two other approaches, cost share with smoothed land prices and our approach, which involves a similar smoothing, behave very well. In absence of measurement error, the mis-specification bias introduced by smoothing is minimal while the bias caused by measurement errors remains small.

F. Monte-Carlo simulations with heterogeneous coefficients

In another set of simulations, we now consider that the parameters of the CES production function are heterogeneous. More specifically, we have α_i and σ_i drawn from uniform distribution such that they can at most be 10% above or below the values of parameters reported in table 23. That is, we have $\alpha_i = (1 + u_\alpha)\alpha$ and $\sigma_i = (1 + u_\sigma)\sigma$ where u_α and u_σ are drawn independently and uniformly from $[-0.1, 0.1]$. Note that we do not consider heterogeneous coefficients when $\alpha = 0.98$ (and $\sigma = 0.6$) since consistency requires the share parameter α_i to be less than one.

The results reported in table 25 are very close to those reported in table 24 with homogeneous coefficients. This implies that our estimations are only barely affected by the introduction of some heterogeneity in the parameters of the CES production function.

In conclusion, we find that all the approaches we consider behave well when there is no issue of measurement error. However, if one believes that there is measurement error on parcel prices, the traditional approach and EGS without smoothing should be avoided while the other approaches should be favoured.

Appendix N. Full identification in absence of restrictions on the returns to scale

According to first-order condition in equation (1), profit-maximising housing capital is a function of both parcel size T and the location characteristics x: $K^* = K^*(x,T)$. Then, from the zero-profit condition (2), we deduce that the price of parcel depends on the same two arguments: $R \equiv R(x,T)$. To remain consistent with our assumption of exogenous parcels, we consider builders bidding for parcels of different sizes and anticipating the amount of capital they will use, K^*. To know how much a profit-maximizing developer is willing to pay for a marginal increase of parcel size, we derive the zero-profit condition with respect to T:

$$\begin{aligned}\frac{\partial R(x,T)}{\partial T} &= P(x)\frac{\partial H(K^*,T)}{\partial K^*}\frac{\partial K^*(x,T)}{\partial T} + P(x)\frac{\partial H(K^*,T)}{\partial T} - \frac{\partial K^*(x,T)}{\partial T} \\ &= P(x)\frac{\partial H(K^*,T)}{\partial T}, \end{aligned} \tag{N1}$$

where the simplification arises from the envelop's theorem and the builder's first-order condition for profit maximisation with respect to K. Although we formally derive the zero profit condition, we note that expression (N1) is equivalent to a first-order condition with respect to T.[29]

We can then use the zero-profit condition to eliminate $P(x)$ from equation (N1) and obtain:

$$\frac{1}{H(K^*,T)}\frac{\partial H(K^*,T)}{\partial T} = \frac{1}{K^* + R(x,T)}\frac{\partial R(x,T)}{\partial T}, \tag{N2}$$

or equivalently:

$$\frac{\partial \log H(K^*,T)}{\partial \log T} = \frac{R(x,T)}{K^* + R(x,T)}\frac{\partial \log R(x,T)}{\partial \log T}. \tag{N3}$$

Importantly, these two expressions are conditional on x. Because we do not know x, we cannot integrate them over T to recover $H(K^*,T)$. Even for a given location with a fixed x, integrating

[29]This avoids having to deal explicitly with the unit price of land and the fact that optimal parcel size is zero under decreasing returns. See the next appendix for more.

$\partial \log R(x,T)$ would require K^* to be constant, which is not the case. Put slightly differently, we cannot integrate the equation (N3) over T because there is no one-to-one mapping between K^* and R as x differs across new houses.

Since we observe K and T in the data but not x, we could try to use R written as a function of K^* and T (as in the zero-profit condition 2 in the main text):

$$\frac{\partial \log R\left(x,T\right)}{\partial \log T} = \frac{\partial \log R\left(K^*,T\right)}{\partial \log T} + \frac{\partial \log R\left(K^*,T\right)}{\partial \log K}\frac{\partial \log K^*}{\partial \log T} \tag{N4}$$

Again, this expression cannot be integrated over T since K^* is also a function of T.

Appendix O. Full identification under constant returns to scale

In this appendix, we consider again the full identification of the housing production function (up to a constant). We now assume developers choose directly parcel size and that the price of parcels is linear in their size: $R = \widetilde{R}(x)T$, where $\widetilde{R}$ is the unit land price. This is consistent with the notion that, if parcels are divisible, there should be no arbitrage possibility between parcels of different sizes in equilibrium.

Builders' profit at location x is $\pi = P(x)H(K,T) - K - \widetilde{R}\left(x\right)T$, which is now maximised over both K and T. Aside from the first-order condition for profit maximisation with respect to capital (1), there is also one for land:

$$P(x)\frac{\partial H(K,T)}{\partial T} = \widetilde{R}(x). \tag{O1}$$

Plugging the two first-order conditions into the zero-profit condition and simplifying by $P(x)$ leads to:

$$H(K,T) = K\,\frac{\partial H\left(K,T\right)}{\partial K} + T\,\frac{\partial H\left(K,T\right)}{\partial T}\,. \tag{O2}$$

This is Euler's condition that characterises homogeneous functions of degree 1. For it to be verified, $H(K,T)$ must be constant returns to scale.

Then, the first-order condition with respect to K, equation (1), still shows that the housing price can be rewritten as a function of K^* and T only, and, as before, the free-entry condition then implies that it is also the case for the total land price, $R\left(K,T\right)$. We can substitute away $P\left(x\right)$ from the free-entry condition by using now the first-order condition for T given by equation (O1). Recalling that $\widetilde{R}\left(x\right) = R\left(K,T\right)/T$, this leads to:

$$\frac{\partial H\left(K,T\right)}{\partial T} = \frac{\widetilde{R}(x)}{P(x)} = \frac{H\left(K,T\right)}{K + R\left(K,T\right)}\frac{R\left(K,T\right)}{T}\;, \tag{O3}$$

which is equivalent to:

$$\frac{\partial \log H\left(K,T\right)}{\partial \log T} = \frac{R\left(K,T\right)}{K + R\left(K,T\right)}\,. \tag{o4}$$

We obtain an expression that mirrors equation (4) for the elasticity of housing production with respect to capital. To derive the production function, we substitute expression (5) into (o4) and obtain:

$$\frac{R\left(K,T\right)}{K + R\left(K,T\right)} = \frac{\partial \log Z\left(T\right)}{\partial \log T} - \int\limits_{K} \frac{KT\frac{\partial R(K,T)}{\partial T}}{\left[K + R\left(K,T\right)\right]^2} \mathrm{d} \log K\,. \tag{o5}$$

Integrating this equation with respect to $\log T$ yields:

$$\log Z\left(T\right) = C + \int\limits_{T} \frac{R\left(K,T\right)}{K + R\left(K,T\right)} \mathrm{d} \log T + \int\limits_{T}\int\limits_{K} \frac{KT\frac{\partial R(K,T)}{\partial T}}{\left[K + R\left(K,T\right)\right]^2} \mathrm{d} \log K \mathrm{d} \log T\,, \tag{o6}$$

where C is a constant. Substituting equation (o6) into (5), we get:

$$\begin{aligned} \log H\left(K,T\right) \quad = \quad & C + \int\limits_{T} \frac{R\left(K,T\right)}{K + R\left(K,T\right)} \mathrm{d} \log T + \int\limits_{K} \frac{K}{K + R\left(K,T\right)} \mathrm{d} \log K \\ & + \int\limits_{T}\int\limits_{K} \frac{K\,T\,\frac{\partial R(K,T)}{\partial T}}{\left[K + R(K,T)\right]^2} \mathrm{d} \log K \mathrm{d} \log T \end{aligned} \tag{o7}$$

Note that this expression is consistent with a Cobb-Douglas production function since, for that function, the second and third right-hand side terms are integrals of constant cost shares and collapse into $\log T$ and $\log K$. Moreover, we have $R\left(K,T\right) = \frac{1-\alpha}{\alpha}\,K$ (where α is the share of capital), which implies that $\frac{\partial R(K,T)}{\partial T} = 0$ and the third right-hand side term of (o7) is then zero.

Table 24: Monte-Carlo simulations with homogeneous parameters

Production Function σ	α		Estimated values Traditional	Traditional smoothed	EGS AM	EGS AM smoothed	Cost share	Cost share smoothed	Our approach
Panel (A): $\delta = 0$									
1.05	0.6	$\hat{\sigma}$	1.05	1.06	1.05	1.06	1.05	1.06	1.06
-	-	$\hat{\alpha}$	0.60	0.59	0.60	0.59	0.60	0.59	0.59
0.6	0.98	$\hat{\sigma}$	0.60	0.62	0.58	0.60	0.60	0.62	0.61
-	-	$\hat{\alpha}$	0.98	0.97	0.98	0.98	0.98	0.97	0.98
0.8	0.87	$\hat{\sigma}$	0.80	0.82	0.79	0.80	0.80	0.82	0.81
-	-	$\hat{\alpha}$	0.87	0.85	0.88	0.87	0.87	0.85	0.86
1	0.65	$\hat{\sigma}$	1.00	1.01	1.00	1.01	1.00	1.01	1.01
-	-	$\hat{\alpha}$	0.65	0.63	0.65	0.63	0.65	0.63	0.64
1.2	0.45	$\hat{\sigma}$	1.20	1.20	1.20	1.20	1.20	1.20	1.20
-	-	$\hat{\alpha}$	0.45	0.45	0.45	0.45	0.45	0.44	0.45
1.5	0.26	$\hat{\sigma}$	1.50	1.47	1.51	1.48	1.50	1.49	1.48
-	-	$\hat{\alpha}$	0.26	0.27	0.26	0.27	0.26	0.27	0.27
2	0.13	$\hat{\sigma}$	2.00	1.90	2.01	1.91	2.00	1.94	1.94
-	-	$\hat{\alpha}$	0.13	0.15	0.13	0.14	0.13	0.14	0.14
Panel (B): $\delta = 0.25$									
1.05	0.6	$\hat{\sigma}$	0.83	1.06	0.97	1.05	1.05	1.06	1.06
-	-	$\hat{\alpha}$	0.84	0.58	0.69	0.58	0.60	0.58	0.58
0.6	0.98	$\hat{\sigma}$	0.55	0.62	0.57	0.60	0.60	0.62	0.61
-	-	$\hat{\alpha}$	0.99	0.97	0.99	0.98	0.98	0.97	0.97
0.8	0.87	$\hat{\sigma}$	0.69	0.82	0.74	0.80	0.80	0.82	0.81
-	-	$\hat{\alpha}$	0.95	0.85	0.92	0.87	0.87	0.84	0.85
1	0.65	$\hat{\sigma}$	0.81	1.01	0.93	1.01	1.00	1.01	1.01
-	-	$\hat{\alpha}$	0.86	0.62	0.73	0.63	0.65	0.62	0.63
1.2	0.45	$\hat{\sigma}$	0.89	1.20	1.09	1.20	1.19	1.20	1.20
-	-	$\hat{\alpha}$	0.78	0.44	0.56	0.44	0.45	0.43	0.44
1.5	0.26	$\hat{\sigma}$	0.98	1.47	1.31	1.48	1.48	1.48	1.48
-	-	$\hat{\alpha}$	0.68	0.26	0.36	0.26	0.26	0.26	0.26
2	0.13	$\hat{\sigma}$	1.03	1.89	1.64	1.91	1.96	1.93	1.94
-	-	$\hat{\alpha}$	0.62	0.14	0.20	0.14	0.13	0.13	0.13
Panel (C): $\delta = 0.5$									
1.05	0.6	$\hat{\sigma}$	0.69	1.06	0.89	1.05	1.05	1.06	1.06
-	-	$\hat{\alpha}$	0.95	0.56	0.77	0.57	0.59	0.56	0.56
0.6	0.98	$\hat{\sigma}$	0.51	0.62	0.56	0.60	0.61	0.63	0.62
-	-	$\hat{\alpha}$	1.00	0.97	0.99	0.98	0.98	0.97	0.97
0.8	0.87	$\hat{\sigma}$	0.61	0.82	0.73	0.81	0.81	0.82	0.82
-	-	$\hat{\alpha}$	0.98	0.84	0.92	0.85	0.86	0.84	0.84
1	0.65	$\hat{\sigma}$	0.68	1.01	0.85	1.00	1.00	1.02	1.01
-	-	$\hat{\alpha}$	0.95	0.61	0.81	0.63	0.64	0.61	0.61
1.2	0.45	$\hat{\sigma}$	0.71	1.20	0.99	1.20	1.19	1.21	1.21
-	-	$\hat{\alpha}$	0.94	0.42	0.66	0.42	0.45	0.42	0.42
1.5	0.26	$\hat{\sigma}$	0.73	1.47	1.16	1.49	1.48	1.49	1.49
-	-	$\hat{\alpha}$	0.93	0.25	0.47	0.25	0.27	0.25	0.25
2	0.13	$\hat{\sigma}$	0.69	1.90	1.39	1.92	1.93	1.94	1.96
-	-	$\hat{\alpha}$	0.95	0.13	0.30	0.13	0.14	0.13	0.13

Monte-Carlo simulations with homogeneous parameters (continued)

Production Function σ	α		Traditional	Traditional smoothed	EGS AM	EGS AM smoothed	Cost share	Cost share smoothed	Our approach
Panel (D): $\delta = 1$									
1.05	0.6	$\hat{\sigma}$	0.51	1.06	0.79	1.04	1.05	1.06	1.06
-	-	$\hat{\alpha}$	1.00	0.55	0.87	0.56	0.59	0.54	0.55
0.6	0.98	$\hat{\sigma}$	0.45	0.62	0.53	0.60	0.62	0.62	0.62
-	-	$\hat{\alpha}$	1.00	0.97	0.99	0.97	0.97	0.97	0.97
0.8	0.87	$\hat{\sigma}$	0.50	0.82	0.67	0.81	0.81	0.82	0.82
-	-	$\hat{\alpha}$	1.00	0.83	0.96	0.84	0.85	0.83	0.83
1	0.65	$\hat{\sigma}$	0.51	1.01	0.76	1.03	1.00	1.01	1.01
-	-	$\hat{\alpha}$	1.00	0.60	0.90	0.57	0.64	0.59	0.59
1.2	0.45	$\hat{\sigma}$	0.51	1.20	0.85	1.19	1.18	1.20	1.20
-	-	$\hat{\alpha}$	1.00	0.41	0.82	0.41	0.45	0.40	0.40
1.5	0.26	$\hat{\sigma}$	0.48	1.46	0.95	1.47	1.45	1.48	1.49
-	-	$\hat{\alpha}$	1.00	0.24	0.70	0.24	0.28	0.23	0.23
2	0.13	$\hat{\sigma}$	0.42	1.88	1.06	1.91	1.86	1.92	1.95
-	-	$\hat{\alpha}$	1.00	0.13	0.57	0.12	0.15	0.12	0.12
Panel (E): $\delta = 2$									
1.05	0.6	$\hat{\sigma}$	0.34	1.06	0.64	1.07	1.05	1.07	1.07
-	-	$\hat{\alpha}$	1.00	0.50	0.97	0.48	0.58	0.49	0.49
0.6	0.98	$\hat{\sigma}$	0.36	0.62	0.48	0.60	0.64	0.62	0.62
-	-	$\hat{\alpha}$	1.00	0.96	1.00	0.97	0.97	0.96	0.96
0.8	0.87	$\hat{\sigma}$	0.36	0.82	0.56	0.80	0.82	0.82	0.82
-	-	$\hat{\alpha}$	1.00	0.80	0.99	0.83	0.84	0.79	0.80
1	0.65	$\hat{\sigma}$	0.35	1.01	0.62	1.01	1.00	1.02	1.02
-	-	$\hat{\alpha}$	1.00	0.55	0.97	0.55	0.63	0.54	0.54
1.2	0.45	$\hat{\sigma}$	0.32	1.20	0.66	1.20	1.17	1.21	1.22
-	-	$\hat{\alpha}$	1.00	0.36	0.96	0.36	0.45	0.35	0.35
1.5	0.26	$\hat{\sigma}$	0.28	1.46	0.70	1.46	1.41	1.49	1.51
-	-	$\hat{\alpha}$	1.00	0.21	0.94	0.21	0.28	0.20	0.19
2	0.13	$\hat{\sigma}$	0.23	1.87	0.74	1.93	1.78	1.93	1.98
-	-	$\hat{\alpha}$	1.00	0.11	0.91	0.10	0.16	0.10	0.10

Table 25: Monte-Carlo simulations with heterogeneous parameters

Production Function			Estimated values						
σ	α		Traditional	Traditional smoothed	EGS AM	EGS AM smoothed	Cost share	Cost share smoothed	Our approach
Panel (A): $\delta = 0$									
1.05	0.6	$\hat{\sigma}$	1.10	1.11	1.05	1.06	1.10	1.11	1.06
-	-	$\hat{\alpha}$	0.55	0.54	0.60	0.59	0.55	0.54	0.59
0.8	0.87	$\hat{\sigma}$	0.87	0.89	0.79	0.80	0.87	0.89	0.81
-	-	$\hat{\alpha}$	0.88	0.86	0.88	0.87	0.88	0.86	0.86
1	0.65	$\hat{\sigma}$	1.06	1.07	1.00	1.01	1.06	1.08	1.01
-	-	$\hat{\alpha}$	0.64	0.63	0.65	0.63	0.64	0.63	0.64
1.2	0.45	$\hat{\sigma}$	1.27	1.27	1.20	1.20	1.27	1.27	1.20
-	-	$\hat{\alpha}$	0.44	0.44	0.45	0.45	0.44	0.44	0.45
1.5	0.26	$\hat{\sigma}$	1.45	1.43	1.51	1.48	1.45	1.43	1.48
-	-	$\hat{\alpha}$	0.24	0.25	0.26	0.27	0.24	0.25	0.27
2	0.13	$\hat{\sigma}$	2.09	1.97	2.01	1.91	2.09	2.02	1.94
-	-	$\hat{\alpha}$	0.13	0.15	0.13	0.14	0.13	0.14	0.14
Panel (B): $\delta = 0.25$									
1.05	0.6	$\hat{\sigma}$	0.85	1.11	0.97	1.05	1.10	1.11	1.06
-	-	$\hat{\alpha}$	0.82	0.53	0.69	0.58	0.55	0.53	0.58
0.8	0.87	$\hat{\sigma}$	0.74	0.89	0.74	0.80	0.87	0.89	0.81
-	-	$\hat{\alpha}$	0.95	0.86	0.92	0.87	0.87	0.86	0.85
1	0.65	$\hat{\sigma}$	0.84	1.07	0.93	1.01	1.06	1.07	1.01
-	-	$\hat{\alpha}$	0.87	0.62	0.73	0.63	0.64	0.62	0.63
1.2	0.45	$\hat{\sigma}$	0.92	1.27	1.09	1.20	1.27	1.27	1.20
-	-	$\hat{\alpha}$	0.78	0.43	0.56	0.44	0.44	0.43	0.44
1.5	0.26	$\hat{\sigma}$	0.97	1.42	1.31	1.48	1.43	1.43	1.48
-	-	$\hat{\alpha}$	0.64	0.24	0.36	0.26	0.25	0.24	0.26
2	0.13	$\hat{\sigma}$	1.03	1.97	1.64	1.91	2.05	2.02	1.94
-	-	$\hat{\alpha}$	0.65	0.14	0.20	0.14	0.13	0.13	0.13
Panel (C): $\delta = 0.5$									
1.05	0.6	$\hat{\sigma}$	0.70	1.10	0.89	1.05	1.09	1.11	1.06
-	-	$\hat{\alpha}$	0.94	0.52	0.77	0.57	0.55	0.52	0.56
0.8	0.87	$\hat{\sigma}$	0.64	0.89	0.73	0.81	0.88	0.89	0.82
-	-	$\hat{\alpha}$	0.98	0.85	0.92	0.85	0.87	0.85	0.84
1	0.65	$\hat{\sigma}$	0.69	1.07	0.85	1.00	1.06	1.07	1.01
-	-	$\hat{\alpha}$	0.96	0.62	0.81	0.63	0.64	0.61	0.61
1.2	0.45	$\hat{\sigma}$	0.72	1.26	0.99	1.20	1.26	1.27	1.21
-	-	$\hat{\alpha}$	0.94	0.42	0.66	0.42	0.44	0.42	0.42
1.5	0.26	$\hat{\sigma}$	0.73	1.42	1.16	1.49	1.42	1.43	1.49
-	-	$\hat{\alpha}$	0.91	0.23	0.47	0.25	0.25	0.23	0.25
2	0.13	$\hat{\sigma}$	0.68	1.97	1.39	1.92	2.02	2.03	1.96
-	-	$\hat{\alpha}$	0.96	0.14	0.30	0.13	0.14	0.13	0.13

Monte-Carlo simulations with heterogeneous parameters (continued)

Production Function σ	α		Traditional	Traditional smoothed	EGS AM	EGS AM smoothed	Cost share	Cost share smoothed	Our approach
Panel (D): $\delta = 1$									
1.05	0.6	$\hat{\sigma}$	0.51	1.10	0.79	1.04	1.09	1.11	1.06
-	-	$\hat{\alpha}$	1.00	0.50	0.87	0.56	0.55	0.49	0.55
0.8	0.87	$\hat{\sigma}$	0.51	0.89	0.67	0.81	0.88	0.89	0.82
-	-	$\hat{\alpha}$	1.00	0.84	0.96	0.84	0.86	0.84	0.83
1	0.65	$\hat{\sigma}$	0.51	1.07	0.76	1.03	1.06	1.08	1.01
-	-	$\hat{\alpha}$	1.00	0.59	0.90	0.57	0.63	0.58	0.59
1.2	0.45	$\hat{\sigma}$	0.50	1.26	0.85	1.19	1.24	1.27	1.20
-	-	$\hat{\alpha}$	1.00	0.40	0.82	0.41	0.44	0.39	0.40
1.5	0.26	$\hat{\sigma}$	0.49	1.41	0.95	1.47	1.39	1.42	1.49
-	-	$\hat{\alpha}$	1.00	0.22	0.70	0.24	0.26	0.22	0.23
2	0.13	$\hat{\sigma}$	0.41	1.94	1.06	1.91	1.93	2.00	1.95
-	-	$\hat{\alpha}$	1.00	0.13	0.57	0.12	0.15	0.12	0.12
Panel (E): $\delta = 2$									
1.05	0.6	$\hat{\sigma}$	0.33	1.10	0.64	1.07	1.09	1.11	1.07
-	-	$\hat{\alpha}$	1.00	0.45	0.97	0.48	0.54	0.44	0.49
0.8	0.87	$\hat{\sigma}$	0.36	0.88	0.56	0.80	0.88	0.89	0.82
-	-	$\hat{\alpha}$	1.00	0.82	0.99	0.83	0.85	0.81	0.80
1	0.65	$\hat{\sigma}$	0.34	1.06	0.62	1.01	1.05	1.07	1.02
-	-	$\hat{\alpha}$	1.00	0.56	0.97	0.55	0.63	0.55	0.54
1.2	0.45	$\hat{\sigma}$	0.31	1.25	0.66	1.20	1.23	1.26	1.22
-	-	$\hat{\alpha}$	1.00	0.37	0.96	0.36	0.45	0.36	0.35
1.5	0.26	$\hat{\sigma}$	0.29	1.40	0.70	1.46	1.36	1.42	1.51
-	-	$\hat{\alpha}$	1.00	0.19	0.94	0.21	0.27	0.19	0.19
2	0.13	$\hat{\sigma}$	0.23	1.93	0.74	1.93	1.83	2.00	1.98
-	-	$\hat{\alpha}$	1.00	0.11	0.91	0.10	0.16	0.10	0.10