TI 2010-087/3
Tinbergen Institute Discussion Paper

Estimating Firms' Demand for Agglomeration

Hans R.A. Koster
Jos N. van Ommeren
Piet Rietveld

VU University Amsterdam, Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Estimating Firms' Demand for Agglomeration

Hans R.A. KOSTER, VU University Amsterdam*

Jos N. VAN OMMEREN, VU University Amsterdam

Piet RIETVELD, VU University Amsterdam

26-8-2010

Abstract

The market for commercial properties is characterised by extreme heterogeneity in demand. In this paper, we aim to gain more insight in the heterogeneity in demand for employment agglomeration and size of the rental property using a two-stage hedonic approach following Bajari and Benkard (2005). We use unique micro-data of properties' attributes as well as of firm characteristics. Given assumptions on the functional form of the production function, we identify firm-specific parameters using a nonparametric control function approach that corrects for endogeneity. The results show that agglomeration benefits are capitalised in rents: a one standard deviation increase in agglomeration leads to an increase in the annual rents of about 6 percent. It is found that larger and business services firms are willing to pay (substantially) more for agglomeration. Furthermore, for office buildings a 10 percent increase in number of employees increases the marginal willingness to pay for floor space with 8 percent, which suggests that internal returns to scale are present.

Keywords: Demand estimation; nonparametric estimation; hedonics; commercial properties.

Acknowledgements
The authors gratefully acknowledge PropertyNL for providing data. We would like to thank NICIS-KEI for financial support.

---

*Corresponding author. Department of Spatial Economics, VU University, De Boelelaan 1105 1081 HV Amsterdam, e-mail : hkoster@feweb.vu.nl

1. Introduction

Economists have generated an impressive number of studies on the demand for attributes of residential properties. In contrast, there is a lack of knowledge about the market for commercial properties, mainly because of insufficient data (Wheaton and Torto, 1994). However, commercial properties are an important input of production. For example for business services, the largest capital input component is the stock of buildings (Drennan and Kelly, 2010).

The market for commercial properties is in some aspects fundamentally different from the market for residential housing. The market is far more heterogeneous in terms of supply and demand (Adair et al., 1996). As an illustration, we observe that the coefficient of variation (CoV) for size (in square meters) for commercial properties is roughly five times higher than for residential properties.[1] The large differences in physical attributes of commercial real estate are likely caused by extreme heterogeneity in demand. For example, a large manufacturing firm will have other building preferences than a small business services firm because of different production functions and, of course, different output levels. This type of extreme heterogeneity is not observed in the market for residential properties.

In this paper, we aim to improve our understanding concerning heterogeneity in demand for size of commercial properties and for agglomeration of rental properties. The demand for agglomeration is relevant because firms that are located near others enforce agglomeration economies that are external to the firm. Agglomeration may benefit firms and so has an important role in firms' location decisions due to labour market pooling, input sharing, knowledge spillovers, and a decrease in transportation costs (Marshall, 1920; Head et al., 1995; Henderson, 2002; DeBlasio and DiAddario, 2005). Rosenthal and Strange (2004) argue that wages, high productivity employment and rents reflect the presence of agglomeration economies, although very few studies use rents as a proxy for the existence of external economies (Drennan and Kelly, 2010). Furthermore, urban economics usually focuses on the manufacturing sector when investigating the magnitude of *extra*-metropolitan agglomeration economies (Eberts and McMillen, 1999; Drennan and Kelly,

[1] This example is based on commercial property data introduced later on and a representative sample of residential properties.

2010). In this paper we are able to compare the WTP for *intra*-metropolitan agglomeration across different industries using commercial rents.

The demand for size of the rental property is of particular interest because of internal returns to scale in terms of employment (Coase, 1937; Tybout, 1993).[2] For example, it has been shown that workers are more productive in larger services firms because of higher arrival rates of customers. In the goods-producing sector, larger firms are organised more efficiently in teams, establishing higher effort standards (Idson and Oi, 1999). As employment and size of the rental property are complementary in the production function, larger firms may be willing to pay more for an additional square meter than smaller firms.

Theory suggests that the magnitude of both internal and external economies to scale is heavily dependent on firm characteristics (see e.g. Helpman et al., 2004). For example, business services firms are expected to experience more intense internal returns to workforce size than retailers. It also may be expected that there are major sectoral differences in the magnitude of agglomeration economies. It has been found for example that business services are willing to pay more for local agglomeration than manufacturing firms (Mun and Hutchinson, 1995; Dekle and Eaton, 1999). We focus on agglomeration based on employment, which is particularly important for sectors where inter-industry interactions are relevant (e.g. business services), but less so for other sectors (e.g. manufacturers).

We employ a two-stage hedonic price approach introduced by Bajari and Benkard (2005). We use micro information on property attributes and firm characteristics avoiding the use of aggregate data, which complicates the estimation procedure (see Bajari and Kahn, 2008). In the first stage, we estimate a nonparametric hedonic price function of buildings' attributes using nonparametric techniques that control for the endogeneity of agglomeration. This approach, advocated by Newey et al. (1999) and Blundell and Powell (2003), is not much applied yet in empirical literature. We combine this approach with insights of Robinson (1988). It involves the linearisation of a part of the function to be estimated, to reduce the curse of multidimensionality, common in nonparametric estimation. In the second stage, we regress firm-specific

[2] In labour and industrial economics, workforce size plays an important role. For example, labour economists have shown that larger firms pay higher wages (Lester, 1967; Masters, 1969; Mellow, 1982; Brown and Medoff, 1989).

willingness to pay (WTP) parameters for these attributes on the firms' sector and workforce size, which are arguably the most important property demand characteristics. Given assumptions on the functional form of the production function and the assumption that firms are profit maximisers, we show that the proposed procedure enables us to identify underlying production function parameters of firms. To our knowledge, this is the first study that is able to establish the relationship between firm characteristics and demand for attributes of real estate. Our approach is also more general in the sense that we do not concentrate on one specific type of building (e.g. offices) or one specific type of industry (e.g. manufacturing).

This paper proceeds as follows. In Sections 2 and 3, we outline the underlying theoretical model and describe our data. In Section 4 we pay attention to our nonparametric estimation procedure. In Section 5 we discuss the results. This is followed by a sensitivity analysis. Section 7 concludes.

2. The model

2.1 Identification of hedonic price models

Rosen (1974) developed a two-stage model where consumers maximise utility, which is a function of the property attributes and a composite good. In the first stage, the derivative of the property's hedonic price function with respect to an attribute is computed, which represents the attribute's implicit price. In the second stage, these implicit prices are regressed on the consumers' demographic characteristics. However, there exists an identification problem, because consumers choose the hedonic price and the quantity of an attribute simultaneously. As a result, consumers with a stronger preference for a certain attribute will purchase properties that contain larger amounts of this attribute (Brown and Rosen, 1982; Bartik, 1987; Epple, 1987). Bartik (1987) proposes to use data from multiple markets, assuming that unobserved preferences do not vary across markets, while the hedonic function does. This should identify hedonic demand parameters. Ekeland et al. (2004) provide another solution to this identification problem for a product with only one observable attribute: when the hedonic price function is estimated nonparametrically, data on multiple markets is not required to identify preferences. Bajari and Benkard (2005) adapt this to a multi-attribute product and propose a two-stage procedure to identify consumer demand. The first stage is estimated nonparametrically, which identifies individual-specific taste parameters.[3] In the second stage, assuming a parametric form of the utility function, the taste parameters are regressed on consumers' characteristics.

We employ a similar procedure as proposed by Bajari and Benkard (2005) and, importantly, show that it can be applied to the market of commercial properties. Our approach requires some assumptions on the firms' production function. We distinguish between building inputs (e.g. the size of the rental property) and non-building inputs (e.g. number of workers) in the production function. Using a bid rent methodology and assuming a competitive market with profit-maximising firms, we show that the proposed procedure enables one to identify firm-specific production function parameters related to building inputs.[4]

2.2 A model for the commercial property market

More formally, firms are assumed to maximise a profit function subject to a production function (see also Palmquist, 1988; Bollinger et al., 1998). Let π_i denote the profit of firm i, p is the price of output y_i, c denotes the price vector of non-building inputs x_{mi}, where $m = 1, \dots, M$ and $i = 1, \dots, I$. $\theta_{ij}(z_j)$ equals the bid rent for a property j, which depends on a vector of building inputs z_{kj}, where $k = 1, \dots, K$ and $j = 1, \dots, J$. The production function ϕ_i is a continuous function of building and non-building inputs. Hence, for each building j, firm behaviour is characterised by maximising profits with respect to non-building inputs x_i:

$$\max \pi_i = p y_i - c x_i - \theta_{ij}(z_j). \quad \text{s.t. } y_i = \phi_i(x_i, z_j). \tag{1}$$

Given prices p and c, and building inputs z_j, this maximisation problem solves for non-building input $x_i^* = x(p, c, z_j)$. We assume a perfect competitive market, so π_i equals zero. Then:

$$p\,\phi_i(x_i^*, z_j) - c x_i^* - \theta_{ij}(z_j) = 0. \tag{2}$$

So, equation (2) defines the bid rent of firm i for property j. In equilibrium, each property is rented to the firm with the highest bid rent, so equilibrium rent R_j is defined by $R_j = \max_i\{\theta_{ij}\}$. To obtain the partial

[3] The estimates are nonparametric in the sense that there is not a distribution imposed on the individual-specific taste parameters.
[4] The bid rent methodology is standard in the field of urban economics; see for example Fujita et al. (2001).

derivative of the equilibrium rent R_j with respect to variable z_{jk}, we take first derivatives and, using the envelope theorem, we arrive at the following condition:

$$\frac{\partial R_j}{\partial z_{kj}} = \frac{\partial \theta_{ij}}{\partial z_{kj}} = p\,\frac{\partial \phi_i}{\partial z_{kj}}. \tag{3}$$

This condition is intuitive, as $p\phi_i$ denotes revenue. So, it simply states that the marginal effect of a building input on the rent is equal to its marginal effect on the revenue function (the production function multiplied with the output price). To identify parameters of the production function, we normalise price p to 1.

As we have only one observation per firm, we cannot learn a firm's production function without making additional assumptions on the functional form. To identify firm-specific production parameters, we assume a semi-logarithmic production function:[5]

$$y_i = \phi_i(x_i^*, z_j) = f(x_{mi}^*) + \sum_k \alpha_{ki} \log(z_{kj}), \tag{4}$$

where the production is an arbitrary function $f(\cdot)$ of non-building inputs, and α_{ki} are structural parameters of the production function related to building inputs. We will allow that $\alpha_{ki} = \alpha_{ki}(x_{mi}^*)$, so these parameters may be a function of optimally chosen non-building inputs (e.g. workforce size).[6] Applying the functional form assumption (4) yields:

$$\alpha_{ki} = z_{kj}\frac{\partial R_j}{\partial z_{kj}}. \tag{5}$$

Given an estimate of $\partial R_j / \partial z_{kj}$ one may derive the firm-specific structural parameters α_{ki} of the production function.

---

[5] For convenience, we also assume continuous attributes z in this exposition.

[6] Note that our assumptions on the functional form of the production function are not so much restrictive. In case that α_{ki} is a constant, then the production function is additive in x_{mi}^* and $\log(z_{kj})$. However, it also allows for the possibility that $f(x_{mi}^*) = 0$ and it may be that x_{mi}^* and $\log(z_{kj})$ interact via α_{ki}. Without making any assumptions on the functional form of the production firm, we still are able to identify firm-specific *marginal* willingness to pay, which is equal to $\partial R_j / \partial z_{kj}$, but not any *structural* parameters of the production function.

3. Data

3.1 Datasets

We make use of two datasets. The first consists of transactions of commercial properties provided by real estate agents between 2001 and 2007 in the NUTS3-region Zuid-Holland, located in the west of the Netherlands. This region includes Rotterdam, the second largest city of the Netherlands, and The Hague, where the national government is located, but also cities such as Leiden and Gouda. This region covers about 20 percent of national economic activity. The dataset contains information about annual rent (or purchase price) and the buildings' attributes, such as address, size (gross floor area in square meters), type of the building (e.g. shop, office), number of parking spaces, year of construction and last renovation. It also contains information about the sector in which the firm, that occupies the building, operates.[7]

We select rent transactions, which cover about 80 percent of the observations, and exclude observations which do not provide information about the building's size or rent, or which refer to properties smaller than 100 square meter, larger than 7,500 square meter or with yearly rents above 275,000 euro.[8] When firms rent a part of a building, the rent and building attributes of the transaction refer to the rental property (the rented part of the building). Our second dataset contains information about characteristics of *all* establishments in the Zuid-Holland region in 2005. This information comes from administrative sources and is very reliable, as Dutch firms are obliged by law to provide this information. We have information on the establishment's exact location, sector and number of employees. In the remainder of this paper, we label establishments as firms and usually refer to rents as prices.

For our two-stage estimation procedure, discussed later on, one has to match property and firm data. We match the information on firms with the information on property transactions using address and sectoral information. To minimise the probability that the firm is a holding company that does not occupy the

[7] Sector information is not always reliable and is missing in 15 percent of the cases, probably because real estate agents do not consider this variable as essential, as it is not used by them for commercial purposes. This information will be helpful to match the two datasets, but is not used otherwise.

[8] We have obtained data from PropertyNL, a publisher that collects data on transactions that usually refer to properties that exceed 100m^2. The chosen maximum rent (275,000) is about two standard deviations above the mean rent. We have compared the medians of the attributes before and after the selections. They are similar except, of course, for size (because we select properties larger than 100 square meters). So, the selected sample appears to be a representative selection of the whole sample.

building, we select firms that have at least two employees, because single-employee firms seldom occupy a building larger than 100 square meter. When there is only one firm at an address, the match is based on addresses. When there are several firms located at one address, we match using sector information as well. When there is no exact match (30 percent of the data), we weight observations of firms using the proportion of firms located at one address.[9] Because we use weights (based on individual observations), rather than taking averages, the second stage estimates are consistent (and more efficient than without weighting). Note that we match firm data from 2005 with building transaction data from 2001-2007. So, we implicitly assume that, when a firm moves out of a property, there is a high positive correlation between the firm's characteristics and those of the new firm that will occupy the property. However, if the latter does not hold, the matching procedure just implies that for some observations we randomly match firms and properties. This does not affect the consistency of the second stage estimates. Property observations without a match are used in the first stage but not in the second stage.[10] In the first stage we estimate a nonparametric hedonic price function based on 3,595 observations. In the second stage, we use 1,366 property observations that are matched to 2,431 observations of firms.

3.2 Description of data

We are interested in the willingness to pay for size of the rental property and employment agglomeration, controlling for other building and environmental attributes. We use the number of parking spaces, distance to the nearest railway station, dummy indicators whether a property is within 1200 meter of a railway station, within 150 meter of highways, railways and rivers, year of last renovation/construction, and whether the property already exists, or has to be constructed. For descriptives of our data we refer to Appendix A. On average, the yearly rent is about € 70,000, the transacted size is 635 square meters and the average rent per square meter is about € 150. However, there is a large difference between the average size of different building types (industrial buildings, offices and shops). Industrial buildings are on average 1,000

[9] For example, if there are two firms on the same address, then we use two weights of 0.5 each.

[10] Alternatively, in the first stage one may use only the property observations that are matched to firms' observations. The results based on this sample are similar. We prefer the results presented in the current paper as the number of observations is larger, resulting in more reliable estimates.

square meters, offices 500 square meters and shops 300 square meters. For each property we compute a so-called employment agglomeration potential, following scholars such as Lucas (2001) and Lucas and Rossi-Hansberg (2002), who argue that workers are more productive when they are employed in the vicinity of other workers. Agglomeration is measured by a weighted average of the number of jobs located in the neighbourhood of the property using an exponential distance decay function, which is continuous over space.[11] Formally:

$$\mathcal{A}(\tilde{\ell}) = \delta \int e^{-\delta d(\ell,\tilde{\ell})} n(\ell) d\ell, \quad \forall \ell \tag{6}$$

where $\mathcal{A}(\tilde{\ell})$ denotes agglomeration of location $\tilde{\ell}$, $n(\ell)$ is the number of jobs of location ℓ, $d(\ell,\tilde{\ell})$ denotes the distance in kilometres between ℓ and $\tilde{\ell}$, and δ is a decay coefficient. We set δ is equal to one, implying that most of the weight of this agglomeration potential is within a few kilometres from the buildings' location.[12] We present a map of the pattern of agglomeration in Appendix A. It clearly reveals the locations of Rotterdam and The Hague, the main employment centres where almost 45 percent of the commercial properties are transacted.

As argued by Bayer and Timmins (2007), location decisions alone are insufficient in distinguishing the potential of local spillovers from those of local locational advantages. As a result, any positive effect of agglomeration is likely to be overstated (Ellison and Glaeser, 1999; Bayer and Timmins, 2007). We therefore need an instrument which is correlated with agglomeration but uncorrelated with any unobserved locational advantage. We use population density of *municipalities in 1830* as an instrument for agglomeration. Note that municipalities in 1830 were much smaller and do not overlap with the current ones. Zuid-Holland, the region which our data refer to, consisted in 1830 of 267 municipalities, whereas nowadays it consists of only 77 municipalities. The instrument's validity rests on the assumption that population density in 1830 is unrelated to current locational advantages (and therefore profit of firms), but has a causal effect on the current agglomeration pattern (see also Ciccone and Hall, 1996; Rice et al., 2006;

[11] We focus on urbanisation economies, which instead of sector-specific localisation economies, relate to positive externalities of diversity (Henderson, 1986; Glaeser *et al.,* 1992).
[12] In the sensitivity analysis, we will also employ other values of δ.

Combes et al., 2008). This instrument is strong as population density is strongly autocorrelated and (current) population and employment densities are positively correlated (McMillen and McDonald, 1998).[13]

4. Estimation procedure

4.1 First stage estimation procedure

We estimate the implicit prices faced by firm i that occupies property j, consisting of k attributes. We suppose that rent R_j is some nonparametric function $\psi(\cdot)$ of employment agglomeration, size of the rental property and a number of control variables. We interact size with building type because the unobserved quality of an additional square meter is highly correlated with building type.[14] For example, office space refers to a much higher quality building, in particularly internally, whereas shops and industrial buildings are usually bare. To control for unobserved heterogeneity we add transaction year dummies t and municipality dummies s to our specification, where $t = 1, \dots, T$ and $s = 1, \dots, S$. To reduce the number of nonparametric parameters in the function to be estimated, we assume that the latter dummies are linearly related to the rent. This reduces the curse of multidimensionality, a limitation of nonparametric applications (Yatchew, 2003). We then have the following rent function:

$$R_j = \psi\left(aggl_j, size\ office_j, size\ shop_j, size\ indb_j, controls_j\right) + \sum_{\forall s} \varphi_s\ municipality\ dummy_s \\ + \sum_{\forall t} \gamma_t\ year\ dummy_{tj} + \xi_j, \qquad (7)$$

where ξ_j denotes the property-specific error term. We employ an estimation approach proposed by Fan and Gijbels (1996). $\psi(\cdot)$ is then estimated using locally weighted regression. Local methods have a lower asymptotic bias than the Nadaraya-Watson estimator and a lower asymptotic variance than the Gasser-Müller estimator, but more importantly, have been shown to generate more plausible estimates of implicit prices (Bajari and Kahn, 2005; 2008). For each observation $\tilde{j}$ we run a locally weighted regression,

[13] Following Ciccone and Hall (1996), we also examined an instrument that captures the distance to the nearest station in 1900. Railway stations were an important factor enforcing agglomeration of firms and people in 1900. However, it appears that this instrument is weak, so we do not use it in the present analysis.

[14] Note that some sectors occupy one type of building. For example, the government occupies offices in 95 percent of the cases. Equivalently, different types of buildings are predominantly occupied by one sector: about 70 percent of industrial buildings are rented by manufacturing, logistic or wholesale firms, 50 percent of offices are rented by business or other services firms and about 70 percent of the shops are rented by retailers, restaurants or bars.

where $1 \leq \tilde{j} \leq J$. So, more informally, local linear regressions assign greater importance to observations with attributes that are similar to $\tilde{j}$. A kernel is employed which is a product of standard normal distributions based on differences between attributes that are nonlinearly related to the rent. We then assume that the hedonic price function *locally* satisfies:

$$R_j = \beta_{0j} + \beta_{1j} \log(agglomeration_j) + \beta_{2j} \log(size\ office_j) + \beta_{3j} \log(size\ shop_j) + \beta_{4j} \log(size\ indb_j) + \sum_{k=5}^{K} \beta_{kj}\, controls_j + \sum_{\forall s} \varphi_s\, municipality\ dummies_s + \sum_{\forall t} \gamma_t\, year\ dummies_{tj} + \xi_j, \qquad (8)$$

where $\beta_{kj}, \varphi_s, \gamma_t$ are coefficients to be estimated.

There are three main issues in our set-up which complicate the estimation procedure. We therefore adopt a two-step estimation procedure which overcomes these issues. The first issue is that we have to determine the bandwidth of our kernel. A lower bandwidth leads to a lower mean-squared error, but to higher variance of the estimator. A larger bandwidth may create a larger bias when the underlying function is nonlinear (Fan and Gijbels, 1996). In the current context, it is less relevant to reduce the variance of the local estimator, as the estimated coefficients will be used as the dependent variable in the second stage. According to theory, a small bandwidth and therefore a high variance of the local estimator only results in additional *random* 'measurement' error of the dependent variable in the second stage (Yatchew, 2003). This does *not* create any inconsistency in the second-stage coefficients. In contrast, any bias in the estimates of the first stage due to oversmoothing (too high values of the bandwidth) will induce a bias *towards zero* in the second-stage coefficients because the 'measurement' error is not independent of the explanatory variables. Using these considerations, we employ a bandwidth of 2. Lower bandwidths lead to nearly singular weight matrices for a number of observations, which in turn lead to unreliable estimates.[15]

Second, as already noted, to reduce the curse of dimensionality we linearise some part of the hedonic price function. We employ the framework of Robinson (1988), who proposes a procedure that leads to $\sqrt{n}$-

[15] Conventional bandwidth selection methods, such as minimizing the Akaike Information Criterion, Generalised Cross Validation (see Hurvich et al., 1998) and the T-value of Rice (1984) lead to oversmoothing and less variation in the coefficients, whereas Silverman's rule of thumb and the Zheng rule lead to undersmoothing and unreliable estimates (see Silverman, 1986, Bishop and Timmins, 2008). Our bandwidth is more or less in between. In the sensitivity analysis, we will show that our results are rather robust to changes in the bandwidth.

consistent estimates for the linearly related variables.[16] Robinson (1988) demonstrates that the coefficients may be estimated at parametric rates of convergence, which make them rather efficient.

A third issue is the endogeneity of agglomeration. To account for this, we employ a control function approach (see Blundell and Powell, 2003; Yatchew, 2003). This approach treats endogeneity as an omitted variable problem, comparable to Heckman's correction for selectivity bias, through the introduction of an appropriately estimated control function (Heckman, 1979). An important restrictive assumption of the control function approach is that the endogenous variable should be continuously distributed, which is fulfilled in our application. Given the use of local linear models, this approach is preferred to two main alternative approaches to correct for endogeneity such as IV and plugging in fitted values (Blundell and Powell, 2003).[17] The procedure to apply the control function is to first regress the endogenous independent variable on all independent variables and instruments. The predicted errors of this step are used as a nonparametric control function which is additive to the dependent variable in the second step. This solves the inconsistency of standard nonparametric estimation (Newey et al., 1999; Pinkse and Ng, 2007). We employ a series approximation of the nonparametric function of the first stage errors. The main advantage is that one can apply the semiparametric estimation procedure of Robinson (1988). Otherwise, very computational intensive procedures such as backfitting or additive separable nonparametric least squares have to be employed (see Yatchew, 2003, pp. 102-103).

We refer to Appendix B for a detailed description of the full estimation procedure of the first stage.

[16] To correct for unobserved spatial heterogeneity, Bajari and Kahn (2008) first run a standard linear regression with zip code fixed effects. The zip code fixed effects are then subtracted from the rent. This procedure does only lead to $\sqrt{n}$ estimates if the linear model is the true model, so we slightly improve on this procedure.

[17] Note that in linear models, the control function approach, instrumental variables, and plugging in fitted values in the second stage will lead to the same results. This is not the case in nonlinear and nonparametric models. Furthermore, Newey and Powell (2003) propose a nonparametric two-step least squares estimator (NP2SLS), which is applicable to series approximation (such as polynomials) but not to local linear methods, as we use in our work. As is well known, the intuitive approach to plug fitted values of the first stage into the second stage leads to inconsistent estimates of nonlinear and nonparametric parameters (Ameniya, 1974; Angrist and Pischke, 2009).

4.2 Second stage estimation procedure

In the second stage, we identify the coefficients of the production function (defined by (4)) and are able to identify the structural WTP of a firm: i.e. these parameters hold at any point away from its observed building/location choice, by assuming the following functional form:

$$y_i = \phi_i(x_i^*, z_j) = \alpha_{1i} \log(aggl_j) + \alpha_{2j} \log(size\ office_j) + \alpha_{3j} \log(size\ shop_j) + \alpha_{4j} \log(size\ indb_j) + \sum_{k=5}^{K} \alpha_{kj}\, controls_j + f(x_{mi}^*). \quad (9)$$

Given (9) we are able to recover the structural parameters α_{ki}, because (5) and (8) imply that $\hat{\alpha}_{ki} = \hat{\beta}_{kj}$. The equality of these two types of coefficients is due to functional form assumptions. Given other functional form assumptions, the relationship between these two types of coefficients will be more complicated. For example, when one locally assumes a double-log hedonic price function, it can be shown that $\hat{\alpha}_{ki} = \hat{\beta}_{kj} R_j$.

So, in essence the estimation procedure allows for an estimation of firm's i structural willingness to pay for attribute k. For the attribute agglomeration we report the willingness to pay for a standard deviation increase from the mean: $\hat{\alpha}_{aggl,i}^* = \hat{\alpha}_{aggl,i}(\log(32726.14) - \log(17618.86))$.[18] For the attribute size, we report the willingness to pay for a one unit increase from the mean (in square meter). This is computed as $\hat{\alpha}_{sizeoffice,i}^* = \hat{\alpha}_{sizeoffice,i}(\log(503.41) - \log(502.41))$, $\hat{\alpha}_{sizeshop,i}^* = \hat{\alpha}_{sizeshop,i}(\log(315.09) - \log(314.09))$ and $\hat{\alpha}_{sizeindb,i}^* = \hat{\alpha}_{sizeindb,i}(\log(1013.45) - \log(1012.45))$. The estimation procedure allows us to examine whether $\hat{\alpha}_{ki}^*$ varies systematically with characteristics of firms, in particular, whether they depend on the firms' employment and the type of sector. These sector dummies capture many characteristics of the firm, not explicitly included, such as the input of capital. Then:

$$\hat{\alpha}_{ki}^* = \hat{\alpha}_{ki}^*(x_{mi}^*) = \theta_1 + \theta_2 \log(employment\ size_i) + \sum_r \theta_r sector_{ri} + \epsilon_i. \quad (10)$$

where θ are parameters to be estimated and ϵ_i denotes an error term.[19]

---

[18] Note that the mean of the agglomeration potential is 17618.86 and the standard deviation 15107.28, see Table A1 in Appendix A.

[19] It is assumed that employment size and industry are exogenous. One may however argue that the size of the workforce of a firm is endogenous. For example, firms that for unobserved reasons prefer larger more prestigious buildings may hire more workers. However, these endogenous changes in workforce size should be minimal compared to the large differences in workforce size. For agglomeration, one may argue that firms which have for unobserved reasons a strong preference for agglomeration may hire fewer workers, because wages tend to be higher in agglomerated areas (Wheaton and Lewis, 2001). However, we include municipality fixed effects that control for this.

5. Results and discussion

5.1 First stage results

We first pay attention to the mean values of the willingness to pay. In Table 1 we compare the estimation results of a nonparametric regression with the control function approach, as discussed above, with three other approaches: a standard OLS, instrumental variables regression and a nonparametric regression without a control function. In Figures 1-4 the distributions of the coefficients of agglomeration and size of the rental property are presented. They indicate for example that heterogeneity in the WTP for agglomeration is relatively large compared to the WTP for a square meter.

Table 1: Estimates of Marginal WTP for continuous attributes

	Parametric regression						**Nonparametric regression**			
	Ordinary Least Squares			Instrumental Variables			Without Control Function		With Control Function	
	Coeff.	Std.Error		Coeff.	Std.Error		Mean Pref.	CoV	Mean Pref.	CoV
Agglomeration	3689.43	(823.86)	***	3708.02	(1264.85)	***	3261.32	0.67	2765.62	0.81
Size in m², Office	128.06	(3.73)	***	128.06	(3.73)	***	120.86	0.08	120.58	0.08
Size in m², Shops	126.89	(5.81)	***	126.89	(5.86)	***	116.83	0.11	117.00	0.11
Size in m², Ind. Building	44.69	(1.61)	***	44.69	(1.61)	***	42.67	0.15	42.67	0.15
Number of observations	3595			3595			3595		3595	
R-squared	0.6998									

*NOTES: The dependent variable is the yearly rent. Standard errors (parametric regressions) are between parentheses and are clustered on municipality level. Coefficients are significant at *0.10, **0.05 and ***0.01 levels. We test in the parametric model whether the instrument (population density) is strong. It appears that the instrument is strong and has an F-value of 958.95.*

Figure 1: Distribution of WTP for agglomeration

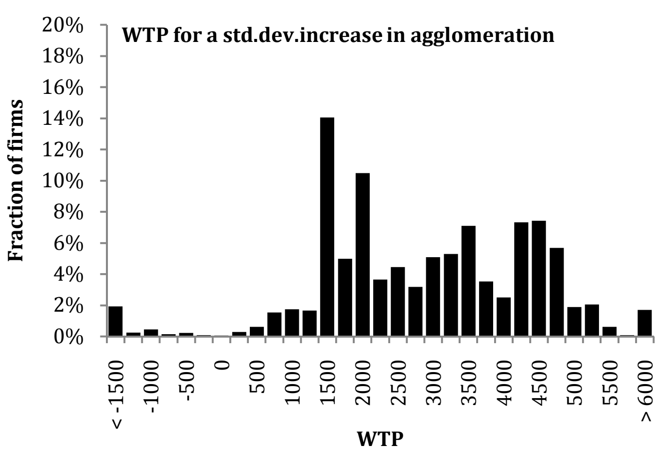


Figure 2: Distribution of WTP for office size

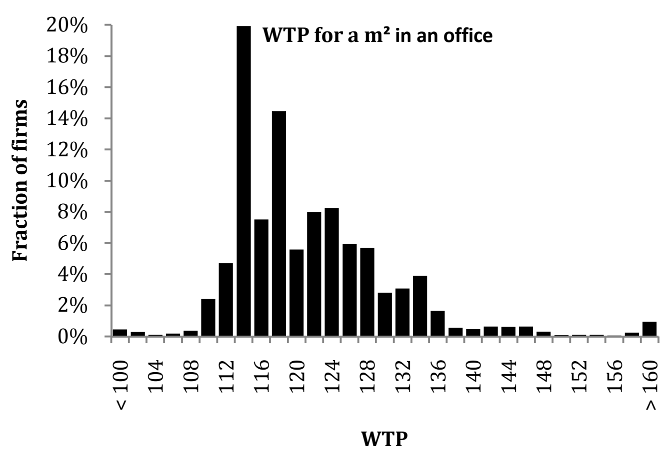


Figure 3: Distribution of WTP for shop size

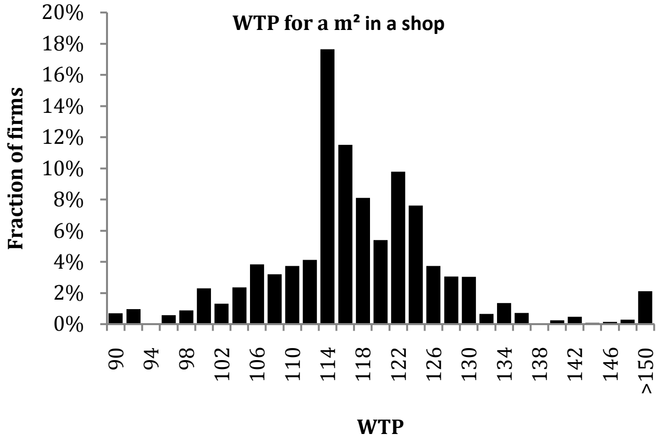
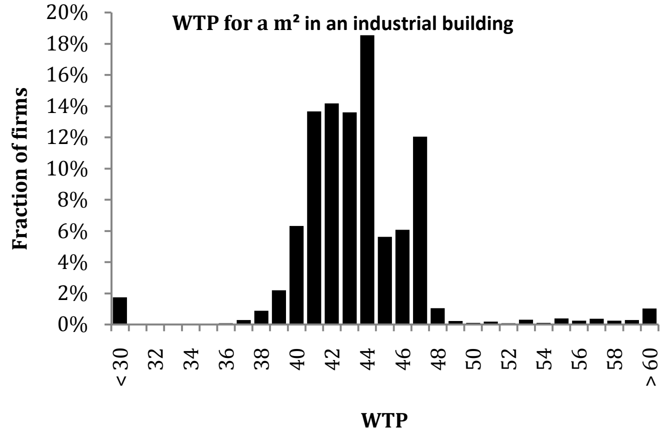


Figure 4: Distribution of WTP for size in an industrial bldng



5.1.1 Agglomeration

In Table 1 we see that the four different estimation procedures indicate that the average marginal WTP for a standard deviation increase in agglomeration ranges from € 2765 to € 3708. The coefficients of the parametric regressions are similar to the mean preferences of the nonparametric regressions. This increases our confidence in the estimation procedure, as Bayer et al. (2007) argue that standard hedonic price regressions should reflect mean preferences for attributes that vary continuously over space. Melo et al. (2009) suggest that instrumenting agglomeration will not lead to a substantial change in estimates, which is confirmed by our IV-estimates. Nevertheless, as is suggested by Bayer and Timmins (2007), non-instrumented estimates of agglomeration are usually overstated: indeed, the estimates of the mean preferences using a nonparametric control function approach are about 15 percent lower. The coefficient of variation reveals that there is much more heterogeneity in the demand for agglomeration than in the demand for size of the rental property. This makes sense because some industries experience more agglomeration economies than other industries, and are therefore willing to pay more for agglomeration.

Figure 1 presents the distribution of $\hat{\alpha}^*_{aggl,i}$. The willingness to pay for one standard deviation increase in employment agglomeration is almost always positive and less than € 6000. Only 3 percent of the WTP-parameters for agglomeration are negative.[20] To get an idea of the magnitude of the agglomeration effect, consider a firm on a location with an average agglomeration potential of about 17,000. When an additional

[20] An example of a negative agglomeration effect is the increased competitiveness when firms cluster together.

firm with 250 employees locates at 100 meter distance of the firm, the firm's average willingness to pay for this increase in agglomeration is € 57 on a yearly basis, about 0.1 percent of the yearly rent.[21] When the additional firm is ten times larger (e.g. the size of a headquarter of a multinational), the WTP is € 540, about 1 percent of the yearly rent. Firms located in non-urban areas, say a location with an agglomeration potential of 2500, value additional agglomeration more. The WTP for an additional firm with 250 employees located at 100 meter distance is then about € 387. These findings are consistent with the notion that agglomeration benefits are capitalised into higher prices.

5.1.2 Size of the rental property

We find that the annual marginal willingness to pay for an additional square meter from the mean is between € 35 and € 170 with an average of € 140, see Figures 2, 3 and 4. Industrial buildings are on average much larger, the WTP for a one meter increase from the mean is much lower than for example an additional meter in a shop, because of our semi-log specification. When we consider the WTP for a one meter increase of a building of the same size, let's say an average building, the WTP for offices, shops and industrial buildings are respectively € 100, € 63 and € 72. We now see that office space indeed implies higher quality, because the WTP is much higher than for example shop space. Figures 2, 3 and 4 reveal that the willingness to pay for size is almost always positive, a feature that is not imposed.

5.2 Second stage results

5.2.1 Agglomeration

The results of the second stage estimates are presented in Table 2. We have defined agglomeration in such a way (see (1)) that we identify the willingness to pay for employment agglomeration within a couple of kilometres of each firm. We first note that there is a significant positive effect of firm size on the WTP for agglomeration: a 10 percent increase in workforce size leads to an increase in the WTP for agglomeration of about € 950, about 2 percent of the annual rent.

---

[21] Using (5) and (6): $\hat{\alpha}_{ki}\left[\log\left(17618.86+(250e^{-0.1})\right)-\log(17618.86)\right]$.

Table 2: Estimates of Marginal WTP for attributes

	Agglomeration			Size in m² Office			Size in m² Shop			Size in m² Industrial Building		
Employment (log)	95.23	(43.84)	**	1.22	(0.19)	***	-0.30	(0.29)		0.38	(0.11)	***
Transport	31.87	(274.86)		4.37	(1.33)	***				0.71	(0.69)	
Wholesale	23.54	(213.16)		1.17	(0.53)	**	-4.63	(0.89)	***	0.65	(0.24)	***
Retail	374.42	(153.31)	**	0.22	(0.54)		3.76	(1.03)	***	-1.82	(0.34)	***
Hotel & Recreation	-198.95	(298.90)		0.36	(1.07)		1.33	(2.08)				
Business Services	1047.82	(153.07)	***	5.95	(0.53)	***	-2.30	(0.99)	**	-2.53	(0.38)	***
Other Services	788.56	(173.27)	***	2.95	(0.78)	***				-2.75	(0.52)	***
Government	746.01	(686.58)		14.83	(2.64)	***						
Education	1297.49	(263.74)	***	3.39	(1.72)	**						
Healthcare	911.74	(240.59)	***	7.77	(1.44)	***				-0.54	(0.60)	
Other	616.44	(302.96)	**	5.54	(2.60)	**						
Constant	1992.43	(170.15)	***	114.35	(0.61)	***	119.05	(1.04)	***	42.87	(0.33)	***
Number of observations	2431			2431			1787			2255		
R-squared	0.0568			0.1551			0.0633			0.0648		

*NOTES: Robust standard errors are between parentheses. Coefficients are significant at *0.10, **0.05 and ***0.01 levels. The manufacturing sector is the omitted category. The WTP for agglomeration refers to a standard deviation increase from the mean. We exclude observations of sectors which almost always occupy one specific type of building.*

So, when the workforce size increases, the WTP for agglomeration increases less than proportional. This suggests that for larger firms agglomeration economies are relatively less important, probably because a large firm is more adept at exploiting knowledge created in the own establishment (Audretsch, 1998).

It appears that there is a substantial difference in the WTP for agglomeration between the Services and Manufacturing sector (the reference category in Table 2), which is also found by Mun and Hutchinson (1995) and Dekle and Eaton (1999). Business services are willing to pay € 1047 more for a one standard deviation increase in agglomeration than manufacturers. This effect also holds for Other Services. Retailers are willing to pay more for agglomeration than manufacturers, probably because there may exist localisation economies by allowing customers to go a shop nearby (Eberts and McMillen, 1999). We also establish that the educational sector is willing to pay € 1300 more for agglomeration than the manufacturing sector. In our sample, the educational sector refers predominantly to private firms. [22] For these firms customers are usually workers from other companies that take off-the-job trainings. Proximity to other companies may therefore beneficial, because it implies increased accessibility to potential customers. Another explanation

[22] Public firms, such as universities and schools for higher vocational training, generally do not rent buildings, but purchase properties, which are not included in our sample.

may be that proximity to other companies may lead to more intense ties with labour markets, which can be beneficial for students in their search for jobs.

5.2.2 Size of the rental property

In Table 2, we see that there is a substantial employment effect on the WTP for an additional square meter in an office. Larger firms are willing to pay more for additional office space than smaller firms: 10 percent increase in workforce size leads to an increase in WTP for a square meter of € 12.20. Consider a firm with 30 employees and a slightly smaller firm with 25 employees. For an average office building, the larger firm is then willing to pay € 0.22 more for an additional square meter than the smaller firm, in line with the idea that there are internal returns to workforce size which makes it profitable to operate larger establishments (Coase, 1937; Helpman and Krugman, 1985; Tybout, 1993). This result also holds for additional floor space in industrial buildings, but then the effect is somewhat smaller.

There are also some notable sector-specific differences in the WTP for size of the rental property. Interestingly, the government sector is willing to pay more for office space than the private sector. For example, it is willing to pay € 14.83 more per square meter than a manufacturer. An explanation is that (local) governments are less flexible in their location choice. For example, the local governments of Rotterdam have to locate in the relatively expensive municipality of Rotterdam and cannot locate elsewhere forcing them to outbid other firms on these locations. Another explanation is that the government does not strive for profit maximisation or cost minimisation, so a less efficient policy regarding the use of buildings is adopted and, it may happen that civil servants receive larger offices than employees in the private sector. Retailers are willing to pay € 8.39 more for shop space than wholesalers, probably because retailers usually locate in shops, while wholesalers also locate in industrial buildings.

Retailers also may have a higher turnover per square meter than wholesalers. The manufacturing, transport and wholesale sectors are willing to pay more for an additional square meter in an industrial building than any other sector. Overall, the average willingness to pay for an additional meter of a specific type of building strongly depends on sector.

We also include interactions of sector and employment size, presented in Appendix C, to verify whether the effects of an increased workforce vary over different sectors. These results indicate that business services experience internal returns to scale in offices: a one percent increase in workforce size increases the WTP for size with € 1.87. Manufacturers and wholesalers that occupy offices also experience internal returns to scale. In industrial buildings, only wholesalers experience statistically significant internal returns to scale.

6. Sensitivity Analysis

In this section we will demonstrate that our results are robust to weighting in the second stage, changes in bandwidth, excluding extreme values in the second stage, exclusion of fixed effects, other values for the decay coefficient δ, and the functional form of the hedonic price function. We already showed in Table 1 that the results for agglomeration are not very sensitive to its instrumentation.

First, recall that we used weighting in the second stage for 30 percent of the building observations, as it is unknown for these observations which specific firm occupies a certain building. When we exclude these observations, it appears that the standard errors are somewhat larger but the results are (almost) identical.

Second, Bajari and Kahn (2005) argue that the choice of bandwidth of the kernel is important as it determines the smoothness of the function to be estimated. Our reference bandwidth is two. We also tried bandwidths of 1.5 and 2.5. In Table 3 we observe that the mean preferences are about the same. The coefficient of variation is somewhat larger for lower bandwidths, implying larger effects in the second stage. Indeed, the coefficients are respectively about 50 percent larger and 35 percent smaller, when bandwidths of 1.5 and 2.5 are employed. Nevertheless, the signs of the coefficients remain the same.

Third, we checked whether excluding WTP-estimates that are more than three standard deviations away from the mean of the WTP-parameters affect our results. The mean preferences remain unaffected (Table 3). It appears that the second stage coefficients are very similar, although the exclusion of extreme values leads to estimates that are closer to zero in the second stage (about 10 percent, see Table C2 in Appendix C).

Table 3: Summary of sensitivity analysis

	Nonparametric regression									
	Bandwidth=1.5		Bandwidth=2.5		Excluding extreme values		No Fixed Effects		Double log hedonic	
	Mean Pref.	CoV	Mean Pref.	CoV	Mean. Pref.	CoV	Mean. Pref.	CoV	Mean. Pref.	CoV
Agglomeration	2826.34	1.63	2859.89	0.30	2805.82	0.52	4203.5	0.54	1209.61	1.08
Size in m^2, Office	117.23	0.15	122.23	0.05	119.99	0.06	120.70	0.08	109.83	0.71
Size in m^2, Shops	109.23	0.28	120.28	0.07	116.00	0.08	116.97	0.11	94.91	0.73
Size in m^2, Ind. Building	40.66	0.28	43.59	0.05	43.13	0.08	42.87	0.16	19.41	0.74
Number of observations	3595		3595		3595		3595		3595	

Fourth, we have examined whether excluding municipality fixed effects generate other results. It appears that the results for size of the rental property are identical. However, the average willingness to pay for agglomeration is substantially higher (about 50 percent), indicating that municipality specific, but unobserved factors, are correlated with agglomeration. Examples come into mind are the presence of a large harbour (Rotterdam), the presence of national government buildings (The Hague), and universities (Delft, Leiden).

Fifth, the value of the agglomeration potential is dependent on a decay coefficient δ. As already noted, most of the weight of this agglomeration potential is then within a few kilometres from the property's location. The higher the value of δ, the more localised this potential is. The mean WTP for agglomeration appears to be insensitive with regard to δ (see Table C3 in Appendix C). The second stage results also hardly change. For example, it is still found that for business services are willing to pay more for agglomeration than manufacturers, but the effect is somewhat smaller, about € 400. One important exception is that for large values of δ, the employment effect changes sign: one percent increase in workforce size leads to a decrease in the WTP for agglomeration of € 177. Apparently, for very local agglomeration smaller firms are more efficient in capturing agglomeration economies.

Finally, in (8) we assumed a local semi-log hedonic price function. We have examined whether a double-log hedonic price function leads to substantially different results. The mean preference of agglomeration is € 1200, which is about the half of our semi-log estimates (see Table 3). Also the mean preferences of size of the rental property are somewhat lower. When we regress elasticities of agglomeration and size of the rental property on workforce size and sector, the main results discussed above still hold. So, the approach introduced by Bajari and Benkard (2005) is quite robust to the specification of the hedonic price function. As elasticities do not directly reveal parameters of the production function, we prefer estimates produced by the semi-log hedonic price function.

7. Conclusions

The market for commercial property is characterised by extreme heterogeneity in demand and therefore in supply of properties. In the current paper, we estimated the firms' demand for agglomeration and size of the rental property. On the one hand, firms may want to be located near others, leading to returns external to the firm. On the other hand, occupying larger properties may lead to increasing returns to scale in terms of workforce size. We employ a two-stage estimation method proposed by Bajari and Benkard (2005) and applied by Bajari and Kahn (2005). Given assumptions on the production function, we identify firm-specific parameters related to building inputs. Nonparametric methods that control for the endogeneity of agglomeration are employed, proposed by Newey et al. (1999) and Blundell and Powell (2003).

We showed that the agglomeration potential is positively related to rents, revealing the presence of external economies. Our results also indicate that there are substantial differences in firms' willingness to pay for agglomeration. Manufacturing firms are for example willing to pay substantially less for agglomeration than business services firms, retailers and firms in the educational sector. We were also able to show that larger firms are willing to pay more for space, suggesting internal returns to scale. A 10 percent increase in number of employees increases the marginal willingness to pay for a square meter of offices with about 8 percent per year. These returns are especially strong in the business services and manufacturing sector.

The results contribute to our understanding of spatial patterns (for example: labour-extensive manufacturing firms are often clustered on remote industrial sites; business services firms have more employees and are located in city centres near other employment). However, more research is needed to corroborate these results and provide more evidence on the magnitude and nature of differences in firms' preferences.

References

Adair, A., Hutchison, N., MacGregor, B., McGreal, S., Nanthakumaran, N. (1996). An Analysis of Valuation Variation in the UK Commercial Property Market. *Journal of Property Valuation and Investment 14(5): 34-47.*

Ameniya, T. (1974). The Nonlinear Two-Stage Least Squares Estimator. *Journal of Econometrics 2: 105-110.*

Angrist, J.D., Pischke, J.-D. (2009). Mostly Harmless Econometrics: An Empiricists Companion. *Princeton: Princeton University Press.*

Audretsch, B., (1998). Agglomeration and the Creation of Innovative Activity. *Oxford Review of Economic Policy 14: 18-29.*

Bajari, P., Benkard, C.M. (2005). Demand Estimation with Heterogeneous consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy 113(6): 1239-1276.*

Bajari, P., Kahn, M.E., (2005). Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business & Economic Statistics 23(1): 20-35.*

Bajari, P., Kahn, M.E., (2008). Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl. *In: A Baranzini* et al. *(eds). Hedonic Methods in Housing Markets.*

Bartik, T.J. (1987). The Estimation of Demand Parameters in Hedonic Price Models. *Journal of Political Economy 95(1): 81-88.*

Bayer, P., Ferreira, F., McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighbourhoods. *Journal of Political Economy 115(4): 588-638.*

Bayer, P., Timmins, C. (2007). Estimating Equilibrium Models of Sorting Across Locations. *The Economic Journal 117: 353-374.*

Bishop, K., Timmins, C. (2008). Simple, Consistent Estimation of the Marginal Willingness to Pay Function: Recovering Rosen's Second Stage Without Instrumental Variables. *Working Paper.*

Blundell, R., Powell, J.L. (2003). Endogeneity in Nonparametric and Semiparametric Regression Models. *In: Dewatripont, M., Hansen, L.P., Turnovsky, S.J. (eds). Advances in Economics and Econometrics: Theory and Applications. Cambridge: Cambridge University Press.*

Bontemps, C., Simioni, M., Surry, Y. (2008). Semiparametric Hedonic Price Models: Assessing the Effects of Agricultural Nonpoint Source Pollution. *Journal of Applied Econometrics 23: 825-842.*

Brown, C., Medoff, J. (1989). The Employer Size-Wage Effect. *Journal of Political Economy 97(5): 1027-1059.*

Brown, J.N., Rosen, H.S. (1982). On the Estimation of Structural Hedonic Price Models. *Econometrica 50(3): 765-768.*

Bollinger, C.R., Ihlandfeldt, K.R., Bowes, D.R. (1998). Spatial Variation in Office Rents within the Atlanta Region. *Urban Studies 35(7): 1097-1118.*

Coase, R.H. (1937). On the Nature of the Firm. *Economica 4(16): 386-405.*

Combes, P.P., Duranton, G., Gobillon, L. (2008). Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics 63(2): 723-742.*

Ciccone, A., Hall, R.E. (1996). Productivity and the Density of Economic Activity. *American Economic Review 86(1): 54-70.*

DeBlasio, G., DiAddario, S.L. (2005). Do Workers Benefit from Agglomeration. *Journal of Regional Science 45(4): 797-827.*

Dekle, R, Eaton, J. (1999). Agglomeration and Land Rents: Evidence from the Prefectures. *Journal of Urban Economics 46: 200-214.*

Drennan, M.P., Kelly, H.F. (2010). Measuring Urban Agglomeration Economies with Office Rents. *Journal of Economic Geography: 1-27*

Eberts, R., McMillen, D. (1999). Agglomeration Economies and Urban Public Infrastructure. *In: Cheshire, P., Mills, E. (eds). Handbook of Regional and Urban Economics 3, Applied Urban Economics, pp. 1455-1495. Amsterdam: North-Holland.*

Ekeland, I., Heckman, J.J., Nesheim, L. (2004). Identification and Estimation of Hedonic Models. *Journal of Political Economy 112(1): S60-S109.*

Ellison, G., Glaeser, E.L. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review 89(2): 311-316.*

Epple, D. (1987). Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy 95(1): 59-80*

Fan, J., Gijbels I. (1996). Local Polynomial Modeling and Its Applications. *In: Monographs and Statistics and Applied Probability 66. Chapman and Hall, England.*

Fujita, M., Krugman, P., Venables, A.J. (2001). The Spatial Economy; Cities, Regions and International Trade. *MIT Press, Cambridge MA.*

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy 100(6): 1126-1152.*

Head, K., Ries, J., Swenson, D. (1995). Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States. *Journal of International Economics 38: 223-247.*

Heckman, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47: 931-959.*

Henderson, J.V. (1986). Efficiency of Resource Usage and City Size. *Journal of Urban Economics 35: 47-70.*

Henderson, J.V. (2002). Marshall's Scale Economies. *Journal of Urban Economics 53: 1-28.*

Helpman, E., Krugman, P. (1985). Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. *MIT Press, Cambridge MA.*

Helpman, E., Melitz, M.J., Yeaple, S.R. (2004). Export versus FDI with Heterogenous Firms. *American Economics Review 94(1): 300-316.*

Hurvich, C.M., Simonoff, J.S., Tsai, C.L. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society B 60(2): 271-293.*

Idson, T.L., Oi, W.Y. (1999). Workers are More Productive in Larger Firms. *American Economic Review 89(2): 104-108.*

Lester, R.A. (1967). Pay Differentials by Size of Establishment. *Industrial Relations 7: 57-67.*

Lucas, R.E., Jr. (2001). Externalities and cities. *Review of Economic Dynamics 4: 245-274.*

Lucas, R.E., Jr., Rossi-Hansberg, E. (2002). On the Internal Structure of Cities. *Econometrica 70(4): 1445-1476.*

Marshall, A. (1920). Principles of Economics. *MacMillan, London.*

Masters, S.H. (1969). An Inter-Industry Analysis of Wages and Plant Size. *Review of Economics and Statistics 51: 341-345.*

McMillen, D.P., McDonald, J.F. (1998). Suburban Employment and Employment Density in Metropolitan Chicago. *Journal of Urban Economics 43: 157-180.*

Mellow, W. (1982). Employer Size and Wages. *Review of Economics and Statistics 64: 495-501.*

Melo, P.C., Graham, D.J., Noland, R.B. (2009). A Meta-Analysis of Estimates of Urban Agglomeration Economies. *Regional Science and Urban Economics 39: 332-342.*

Mun, S.E., Hutchinson, B.G. (1995). Empirical Analysis of Office Rent and Agglomeration Economies: A Case Study of Toronto. *Journal of Regional Science 35(3): 437-455.*

Newey, W.K., Powell, J.L. (2003). Instrumental Variable Estimation of Nonparametric Models. *Econometrica 71(5): 1565-1578.*

Newey, W.K., Powell, J.L., Vella, F. (1999). Nonparametric Estimation of Triangular Simultaneous Equations Models. *Econometrica 67(3): 565-603.*

Pinkse, J., Ng, S. (2007). Nonparametric Two-Step Regression Estimation When Regressors and Errors are Dependent. *Unpublished Manuscript.*

Pagan, A., Ullah, A. (1999). Non-parametric Econometrics. *Cambridge University Press.*

Palmquist, R.B. (1988). Land as a Differentiated Factor of Production: a Hedonic Model and Its Implications for Welfare Measurement. *Land Economics 65(1): 23-28.*

Rice, J. (1984). Bandwidth Choice for Nonparametric Regression. *Annals of Statistics 12: 1215-1230.*

Rice, P., Venables, A.J., Patacchini, E. (2006). Spatial Determinants of Productivity: Analysis for Regions of Great Britain. *Regional Science and Urban Economics 36: 727-752.*

Robinson, P.M. (1988). Root-N-Consistent Semi-Parametric Regression. *Econometrica 57: 1403-1430.*

Rosen, H.S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy 82: 34-55.*

Rosenthal, S.S., Strange, W.C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. *In: Henderson, J.V., Thisse, J. (eds.). Handbook of Urban and Regional Economics, 4.*

Silverman, B. (1986). Density Estimation for Statistics and Data Analysis. *In: Monographs on Statistics and Applied Probability 26. Chapman and Hall, London.*

Tybout, J.R. (1993). Internal Returns to Scale as a Source of Comparative Advantage. *American Economic Review 83(2): 440-444.*

Yatchew, A. (2003). Semiparametric Regression for the Applied Econometrician. *Cambridge: Cambridge University Press.*

Wheaton, W.C., Lewis, M.J. (2002). Urban Wages and Labour Market Agglomeration. *Journal of Urban Economics 51: 542-562.*

Wheaton, W.C., Torto, R.G. (1994). Office Rent Indices and Their Behavior over Time. *Journal of Urban Economics 35(2): 121-139.*

Appendix A: Descriptives

Table A1: Descriptive statistics of the first stage variables

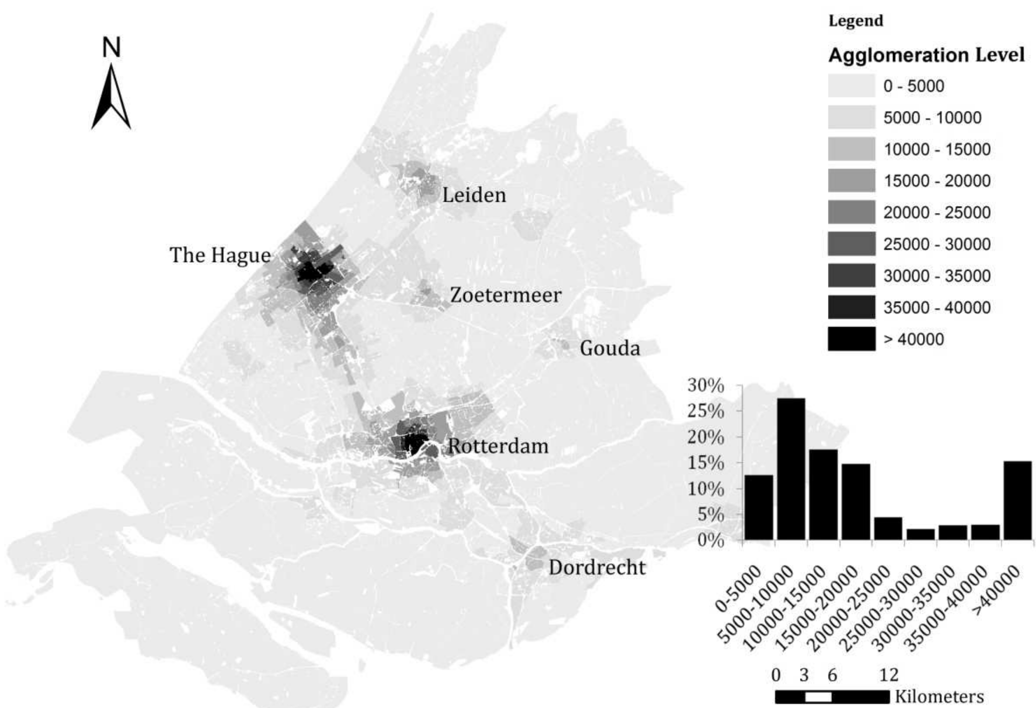
Variable	Mean	Std. Deviation
Dependent variable		
Price per year in €	68635.10	48508.51
Variables of interest		
Size in m^2	637.15	713.44
Size in m^2, Office	502.41	353.8422
Size in m^2, Shops	314.09	220.6823
Size in m^2, Ind. Building	1012.45	714.4944
Agglomeration ($\delta = 1$)	17618.86	15107.28
Agglomeration ($\delta = 0.5$)	22171.86	14542.90
Agglomeration ($\delta = 2.5$)	17615.62	15095.58
Agglomeration ($\delta = 5$)	12035.95	12609.22
Agglomeration ($\delta = 10$)	8472.62	9377.56
Control variables		
Office (dummy)	0.43	0.49
Shop (dummy)	0.23	0.42
Industrial building (dummy)	0.35	0.48
Parking spaces [a]	28.96	63.67
Data on parking spaces missing (dummy)	0.64	0.48
Distance to station (within 1200m)	0.30	0.39
Station < 1200m (dummy)	0.43	0.50
Highway <150m (dummy)	0.16	0.36
Rail <150m (dummy)	0.09	0.29
River <150m (dummy)	0.03	0.17
Building status in process (dummy)	0.02	0.16
Construction < 1900	0.01	0.12
Renovation ≤ 1971 dummy [b]	0.05	0.22
Renovation 1971-1980 (dummy) [b]	0.03	0.16
Renovation 1981-1990 (dummy) [b]	0.73	0.44
Renovation 1991-2000 (dummy) [b]	0.09	0.28
Renovation ≥2001 (dummy) [b]	0.10	0.03
Instrument		
Population 1830 per km^2	1555.10	3313.88

[a] *We have data on parking spaces for about 30 percent of the observations.*
[b] *We have information on construction year and year of last renovation and report the most recent year. When there is no information about construction and renovation year we impute the average (1986).*

Table A2: Descriptive statistics of the second stage variables

Variable	**Mean**	**Std. Deviation**
Employment (number of workers)	24.41	84.55
Manufacturing	0.12	0.32
Transport	0.06	0.23
Wholesale	0.17	0.38
Retail	0.14	0.35
Hotel, Catering & Recreation	0.03	0.17
Business Services	0.31	0.46
Consumer and Other Services	0.10	0.30
Government	0.01	0.10
Education	0.01	0.09
Healthcare	0.05	0.21
Other	0.01	0.11

Figure A1: Map of agglomeration levels in Zuid-Holland



Appendix B. Semiparametric estimation procedure

Table B1: Semi parametric estimation procedure of the first stage

STEP 1	
	Let A be a matrix consisting of all independent variables and instruments and Z a vector of all values of $\log(agglomeration_j)$. Let D be a matrix consisting of all municipality and year dummies.
Linear part	We regress the log of agglomeration and all municipality and year dummies D_u on independent variables (building size, controls) and instruments (population density 1830) nonparametrically (using local linear methods), where $u = 1, \dots, S + T$. We generate residuals $\tilde{Z} \equiv Z - \hat{Z}(A)$ and $\widetilde{D}_u \equiv D_u - \widehat{D}_u(A)$, $\forall u$ (see also Bontemps et al., 2008). Perform OLS on these residuals: $\tilde{Z} = \widetilde{D}_u'\tau + W$. Under regularity conditions, this procedure yields a $\sqrt{n}$-consistent and asymptotically normal estimator for τ (Robinson, 1988).
Nonlinear part	Let $W \equiv Z - \widetilde{D}_u'\hat{\tau}$. We regress this residual on A nonparametrically: $W = W(A) + \varrho$. The vector ϱ consists of first stage errors: $\hat{\varrho} = W - \widehat{W}(A)$ (Blundell and Powell, 2003).
STEP 2	
	Let B be a matrix consisting of all independent variables: log agglomeration, size, and other controls. R is a vector consisting of the rents. We want to estimate: $\mathrm{R} = f_{MI_x}(B) + D_u'\rho + m(\varrho) + \varsigma$. We assume that $m(\varrho) = \sum_i \upsilon_i \varrho^i$ and $i = 1, \dots, 5$.
Linear part	We regress the rent R, all location dummies D_u and all orders of ϱ on B nonparametrically. We generate residuals $\widetilde{\mathrm{R}} \equiv \mathrm{R} - \widehat{\mathrm{R}}(\mathrm{B})$, $\widetilde{D}_u \equiv D_u - \widehat{D}_u(B)$, $\forall u$, and $\tilde{\varrho}^i \equiv \varrho^i - \hat{\varrho}^i(B)$, $\forall i$. Then, OLS is performed on these residuals: $\widetilde{\mathrm{R}} = \widetilde{D}_u'\rho + \sum_i \tilde{\varrho}^{i\prime}\upsilon_i + S$.
Nonlinear part	Let $S \equiv \mathrm{R} - \widetilde{D}_u'\hat{\rho} - \widehat{m}(\varrho)$, where $\widehat{m}(\varrho) = \sum_i \hat{\upsilon}_i \varrho^i$ We regress this residual on B nonparametrically: $S = R(B) + \varepsilon$.

Appendix C. Other results

Table C1: Second stage estimates of interactions workforce size and sectoral dummies

	Size in m² Office			Size in m² Shop		Size in m² Industrial Building		
Empl*Manufacturing	1.33	(0.36)	***			0.22	(0.14)	
Empl*Transport	-1.17	(1.07)				0.06	(0.55)	
Empl*Wholesale	1.12	(0.45)	**	0.06	(0.65)	0.60	(0.22)	***
Empl*Retail	0.34	(0.42)		0.27	(0.68)	0.77	(0.47)	
Empl*Hotel & Recreation	1.48	(1.09)		-1.76	(2.21)			
Empl*Business Services	1.87	(0.36)	***	-0.61	(0.52)	0.23	(0.20)	
Empl*Other Services	0.96	(0.40)	**	-0.28	(0.40)	0.46	(0.34)	
Empl*Government	2.79	(2.39)						
Empl*Education	-0.20	(2.08)						
Empl*Healthcare	1.76	(1.52)				0.38	(0.47)	
Empl*Other	6.29	(3.06)	**					
Sectoral dummies (10)	Yes			Yes		Yes		
Number of observations	2431			1787		2255		
R-squared	0.1672			0.0644		0.0662		

NOTES: See Table 2. Employment is in logs

Table C2: Second stage estimates for agglomeration and size under exclusion of extreme valued observations

	Agglomeration			Size in m² Office			Size in m² Shop			Size in m² Industrial Building		
Employment (log)	76.19	(31.51)	**	1.08	(0.17)	***	-0.57	(0.2)	***	0.27	(0.06)	***
Transport	71.74	(164.22)		3.07	(0.93)	***				0.61	(0.29)	**
Wholesale	-139.39	(113.60)		0.68	(0.49)		-4.72	(0.75)	***	0.50	(0.20)	**
Retail	186.17	(111.10)	*	-0.04	(0.52)		2.18	(0.77)	***	-1.46	(0.22)	***
Hotel & Recreation	-98.67	(192.52)		0.26	(1.07)		-0.11	(1.49)				
Business Services	956.45	(110.87)	***	5.44	(0.50)	***	-3.97	(0.72)	***	-1.29	(0.20)	***
Other Services	591.81	(143.52)	***	2.38	(0.64)	***				-2.00	(0.23)	***
Government	605.05	(680.12)		12.88	(2.11)	***						
Education	1099.1	(240.84)	***	3.41	(1.69)	**						
Healthcare	758.74	(219.36)	***	6.52	(1.00)	***				-0.52	(0.60)	
Other	419.98	(286.84)		3.19	(1.27)	**						
Constant	2242.43	(120.85)	***	114.72	(0.56)	***	119.51	(0.87)	***	43.16	(0.23)	***
Number of observations	2431			2431			1787			2255		
R-squared	0.1039			0.1735			0.1442			0.1032		

NOTES: See Table 2.

Table C3: WTP for agglomeration employing different values of δ.

	Nonparametric regression									
	$\delta = 0.5$		$\delta = 1$		$\delta = 2.5$		$\delta = 5$		$\delta = 10$	
	Mean Pref.	CoV	Mean Pref.	CoV	Mean. Pref.	CoV	Mean. Pref.	CoV	Mean. Pref.	CoV
Agglomeration	3033.7	0.71	2765.62	0.81	3159.95	0.69	3493.16	0.63	3053.79	0.82
Number of observations	3595		3595		3595		3595		3595	

NOTE: We report the WTP for a standard deviation increase from the mean: $\hat{\alpha}^*_{aggl\delta=0.5,i} = \hat{\alpha}_{aggl\delta=0.5,i}(\log(36714.76) - \log(22171.86))$, $\hat{\alpha}^*_{aggl\delta=2.5,i} = \hat{\alpha}_{aggl\delta=2.5,i}(\log(32711.21) - \log(17615.62))$, $\hat{\alpha}^*_{aggl\delta=5,i} = \hat{\alpha}_{aggl\delta=5,i}(\log(24645.18) - \log(12035.95))$, $\hat{\alpha}^*_{aggl\delta=10,i} = \hat{\alpha}_{aggl\delta=10,i}(\log(17850.18) - \log(8472.62))$.