Andrea Caragliu,[1] Martijn Smit,[2] Frank van Oort[3]

Who's right, Weber or Glaeser?[4]

This draft: Dec. 31, 2021

Abstract

Recent evidence about the relative strength of localization and urbanization economies is rather contradictory. Most empirical work on agglomeration economies has been devoted to the analysis of pure density mechanisms, which actually subsume two possible amenity effects. On the one hand, vast empirical evidence suggests the importance of productivity-enhancing features of spatially concentrated settlement structures, which attract profit-maximizing firms despite the costs associated to large population concentrations. On the other hand, cities also function as major market areas, concentrated in space and thereby offering within a limited area a large number of consumption possibilities.

These two strands of literature have seldom spoken and to date the relative importance of these sources of agglomeration benefits is not yet clear. In this paper we address this gap. We exploit two large data bases comprising (i.) 70 per cent of all house transactions in the Netherlands in the period 2005-2011, and (ii.) ORBIS data covering balance sheets of Dutch firms in the period 2005-2011. We also merge these two main data sources with Statistics Netherlands neighborhood data from the *Wijk- en Buurtkaart*, data on monuments from the Cultural Heritage Agency, and, lastly, LISA data of all registered firms in the Netherlands.

The paper provides two main contributions: (i.) We measure the intensity of the productivity effect of consumption and production-related advantages for the Dutch case, and (ii.) We observe whether the relative intensity of the two effects change over the observed time span, following recent theoretical predictions, that suggest the growing importance of consumption amenities as sources of agglomeration benefits (Glaeser et al., 2001; Koster et al., 2019). This empirical framework is for the first time to date explored by looking at the three main indicators of relative locational

---

[1] Politecnico di Milano, ABC Department, Piazza Leonardo da Vinci, 32, 20133 Milan (MI), Italy. andrea.caragliu@polimi.it.

[2] Utrecht University, Faculty of Geosciences, P.O. Box 80.115, 3508 TC Utrecht, the Netherlands. m.j.smit@uu.nl.

[3] Erasmus School of Economics and Institute for Housing and Urban Development Studies (IHS), Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. vanoort@ese.eur.nl.

[4] The Authors would like to thank Andrés Rodriguez-Pose; David Castells-Quintana; Edward Glaeser; Gilles Duranton; James LeSage; John Carruthers; Kieran Donaghy; Mildred Warner; Roberto Camagni; Roberta Capello; Simona Iammarino; and participants in the North American meeting of the Regional Science Association International, San Antonio, Nov. 7-10, 2018; PRIN Workshop 'Le Città Metropolitane", Milan, Dec. 3, 2018; European meeting of the Urban Economics Association, Amsterdam, May 31-June 1, 2019; Sustainable Built Environment conference, Milan, Sep. 3-6, 2019; 40th annual AISRe congress, L'Aquila, Sep. 16-18, 2019; Regional Science, Planning, and Policy Analysis Seminar Series, Cornell University, Nov. 7, 2019; Utrecht University School of Economics Seminar Series, Nov. 4, 2020; REAL Seminar on Regional Science and Economics, University of Illinois at Urbana-Champaign, USA, March 19, 2021; 13th World Congress of the Regional Science Association International, May 25-28, 2021; and the 2022 ASSA Annual Meeting, Jan. 7-9, 2022, for useful comments and suggestions. All remaining errors are our own.

advantages, i.e. house prices, firm productivity, and worker wages, thus providing evidence on all three main sources of income (rent, profits, and wages, respectively).

Our findings suggest that both consumption and production-related externalities are reflected in house prices, firm productivity, and wages. In particular, urban land rent increases with the intensity of competition, as well as with the presence of local consumption amenities (major monuments, theatres, and restaurants). Firms also tend to be more productive, and workers better paid, when located closer to sources of consumption amenities, although the evidence is less compelling.

Results are robust to a number of consistency checks, as well as to the use of Instrumental Variables, with soil composition, historical population density, and historical cinemas as the three main instruments. Instead, we find little evidence of a relative decline of the importance of production externalities with respect to consumption amenities.

Keywords: agglomeration economies, urban externalities, consumer amenity, producer amenity

JEL Classification codes: O18, R11, R12

1. Introduction

Ever since the very first city, Uruk, was founded about six thousand years ago (Liverani et al., 2006), cities have provided an inextricable bundle of amenities (and disamenities) for both consumers and firms. On the one hand, vast empirical evidence suggests the importance of productivity-enhancing features of spatially concentrated settlement structures (Duranton and Puga, 2004; Melo et al., 2009; Marshall, 1920), which attract profit-maximizing firms despite the costs associated to large population concentrations. On the other hand, cities also function as major market areas, concentrated in space and thereby offering within a limited area a large number of consumption possibilities. Consumers interested in the accessibility to diversified consumption bundles find in urban areas the widest possible freedom to choose.

These effects are intertwined and make the identification of the relative importance of each of them inherently complicated. In fact, both effects are reflected in land rents, measuring the relative price of accessibility on the urban land market. In the sociological literature, this dichotomy has been first discussed by Weber (1969). Explaining the spatial and historical evolution of different types of European cities, he argued that during the Middle Age some cities thrived mainly on a flow of monopolistically sheltered consumption, where different types of locals and non-locals played a role.[5] In contrast, others specialized in manufacturing or service activities, and in fact, these anticipated long-run economic trends by engaging in long-distance external trade which increased the size of their potential market, thus paving the way for producer cities to outperform consumer ones.

This classical dichotomy, while clearly representing an extreme form of typification,[6] has more recently been criticized in the light of the recent evolutions of the ways market institutions regulate production in present-day urban systems. In fact, the massive shift of manufacturing activities from industrialized to developing countries, coupled with the relocation of production plants from urban to non-urban areas (UNIDO, 2015), has left cities in the Western hemisphere void of their production-oriented nature. This debate has been revamped by Glaeser et al. (2001), who argue that, due to the increased mobility of firms, cities increasingly thrive as consumption centers.

While both views are based on a clear logic and are supported by sound evidence, to date no attempt has been made to test the two theories on a unique sample with the aim to provide an empirical answer to the question whether consumer or producer amenities matter the most in the formation of agglomeration externalities.

In this paper we address this gap. We exploit three main data bases comprising (i.) 70 per cent of all house transactions in the Netherlands in the period 2005-2011, (ii.) balance sheets of Dutch firms in the period 2005-2011, and (iii.) linked employer-employee data for the period 2005-2011. Using these,

[5] Weber (1969) presents Arnhem and Wiesbaden as examples for the two types of consumer cities.

[6] Most cities actually encompass both natures, and their identity stems from the prevailing function. The very nature of a city is in itself not time-invariant. Leiden, in the Netherlands, is a good example in this sense. The city hosted a buoyant textile sector booming from the late 16th century until the 18th. After the beginning of the 19th century, the city shrunk drastically, reviving as a consumer city from the beginning of the 20th century (de Vries and van der Woude, 1997, pp 280-287).

we are able to provide four main contributions: firstly,) We measure the intensity of the productivity effect of consumer and producer amenities for the Dutch case. Secondly, we observe whether the relative intensity of the two effects evolve over the observed time span, following recent theoretical predictions. Thirdly, we disentangle the location of agglomeration benefits reflected in workers' wages between their location of residence and that where they work. Finally, we are able to jointly estimate house prices, firm productivity and worker wages within the same framework.

Our results show that both consumption and production-related externalities are indeed reflected in house prices, productivity, and wages. In particular, house prices increase with the intensity of competition between firms, through the urban land rent market, as well as with the presence of local consumption amenities (major monuments, theatres, and restaurants). Firms also tend to be more productive when located closer to sources of consumption amenities, although the evidence for this connection is less compelling. Lastly, wages in places of residence are found to benefit from the presence of specialized industries, while some consumption amenities also show a positive relationship with wages.

Results are robust to a number of consistency checks, as well as to the use of instrumental variables, with soil composition, historical population density, and historical cinemas as the three main instruments (Combes et al., 2010a; Koster et al., 2019). However, we find little evidence of the predicted decline of production externalities compared to consumption amenities.

In order to reach this goal, we proceed as follows. Section 2 summarizes the literature on the debate between the consumer city and the producer city. In Section 3 we propose an empirical model and the identification strategy for the empirical analyses. Data collected are described in Section 4, while empirical results are summarized in Section 5 (hedonic price regressions), Section 6 (productivity regressions), Section 7 (wage regressions), and Section 8 (Structural Equations model). Section 9 concludes.

2 Are cities sources of consumption or production amenities?

Overwhelming evidence suggests that denser and larger urban areas make workers and firms more productive (Rosenthal and Strange, 2004; de Groot et al., 2016). Many explanations have been provided for this effect. Theoretically, the classical Hoover classification in terms of scale, localisation, and urbanisation economies (Hoover, 1937) has been rephrased in Duranton and Puga (2004), to take account of the many changes in the price-competitive systems that took place in the twentieth century while also focusing on the dynamic outcomes of agglomerated settlement structures. In Duranton and Puga (2004), the main microfoundations of agglomeration economies are identified in sharing, matching, and learning mechanisms.

Empirically, the evidence about the relative strength of the different sources of agglomeration economies is summarized in Puga (2010),[7] where three main approaches to the measurement of these externalities are identified: (i.) through the mechanism of clustering of production beyond the threshold traditionally explained by random mechanisms or comparative advantage; (ii.) through observing the spatial distribution of wages and rents; and (iii.) by means of structural spatial variations in productivity. Puga (2010) concludes that while the empirical literature seems to have a

[7] On this point, additional evidence is presented in Ellison et al. (2010); Faggio et al. (2017); Diodato et al. (2018).

good set of explanations for sharing and matching mechanisms, learning would deserve much more work to be better understood.

While overarching in many aspects, this review leaves out of the picture some interpretative approaches that could further help open the black box of agglomeration economies. One such insight is offered by the related variety concept, which is grounded in the evolutionary economic geography literature. As argued in Frenken et al. (2007) and Van Oort et al. (2015), agglomeration effects can be decomposed in knowledge-enhancing externalities, whereby variety represents a source of regional knowledge spillovers, and a pure diversification effect, whereby variety plays the role of a portfolio effect sheltering an urban region from idiosyncratic shocks. The former type of effect is measured with the notion of related variety, and acts within industries and technological paradigms, while the portfolio effect is measured by the notion of unrelated variety, and is at play across sectors.[8]

Another relevant classification of different approaches to measuring agglomerative forces is discussed in Camagni et al. (2016), where three methodological patterns are identified: (i.) a micro-industrial approach (within which most literature discussed in Puga, 2010 would fall), focusing on the economies of scale characterizing production functions at different spatial scales (from within firms to the aggregate urban level); (ii.) a geographic approach, reviving the concept of borrowed size first introduced in Alonso (1973) to explain the technological externalities stemming from the proximity of small urban areas to large urban agglomerations; and (iii.) a macro-territorial approach, dealing instead with the local elements enhancing a city's agglomerative effects through higher quality features, mostly in the form of city network embeddedness and in the capacity to attract high-level functions.

The empirics of agglomeration economies have offered a wide range of increasingly precise estimates of the strength of such forces. A relatively recent development in this sense took place with the identification of the major role that sorting processes play in the generation of agglomeration externalities. Combes et al. (2008) argue that the process of spatial segmentation of skilled workers actually accounts for a non-negligible (in some cases, roughly half) of the existing spatial wage disparities: skilled individuals sort along their skills, with the most skilled workers concentrating in dense urban areas. Recent evidence suggests that the type of additional skills learned in cities tend to stay with workers living in agglomerated settlements even after they leave the urban location where the skill has been learned (Roca and Puga, 2017).

Across this whole literature, attention is mostly paid to production-oriented types of externalities. Firms are the actors reaping the true benefits from agglomeration economies, with a focus on the firm location. However, recent evidence suggests that externalities accruing to consumers may be equally important in determining urban locational advantages. For instance, Glaeser et al. (2001) find indirect evidence of the growing relevance of consumer externalities on the basis of (i.) a faster growth of high amenity cities w.r.t. low amenity ones; (ii.) a faster growth of urban land rent w.r.t. urban wages; and (iii.) the rise of reverse commuting. Accessibility to a diverse and broad-ranging portfolio of consumer products can represent a relevant microfoundation for these stylised facts; for instance, Öner (2017) finds that an easier access to stores within municipal market boundaries

[8] Caragliu et al. (2016) provides a first empirical test of the simultaneous strength of related and unrelated variety, and specialization and diversity forces, on jobs growth in European NUTS2 regions.

increases the attractiveness of urban areas, while no such effect is found for rural municipalities. Moreover, although an extensive literature on daily urban systems exists (e.g., Poorthuis & Van Meeteren 2019), this is rarely applied to consumption or production externalities, even though it is easily imaginable a gym near the place of work can just as easily lead to consumption benefits as a gym near the place of residence can lead to knowledge spillovers.

In this niche literature, a milestone is represented by Blomquist et al. (1988), who for the first time exploit hedonic prices to identify the relative strength of spatially-differentiated amenities in determining relative housing prices and wages. On the basis of a classical spatial equilibrium based on the Rosen and Roback framework (Rosen, 1979; Roback, 1982), indices of quality of life for 253 counties are calculated, controlling for different types of amenities, related to both consumption and production activities. Following up to Blomquist et al. (1988), Gabriel and Rosenthal (2004) calculate quality of life indices for 37 US cities with separate indices for consumers and firms. Their findings suggest that in many cases cities that turn out to be attractive to firms seem not to represent potentially attractive locations for households, and vice versa. An implication of this finding is that cities tend to attract workers (thereby increasing their size) by means of increasing the quality of the business environment they offer – once again, an argument in favour of the production story.

In fact, the wage-in-kind argument based on the Rosen and Roback framework implicitly posits that wages can compensate for the loss in quality of life in places that offer poor consumption amenities. For instance, two otherwise identical workers, with similar experience and skills, could be offered different salaries if either receives a job offer in a town with a nice historical center, a dense CBD with nice shops, nice parks, or access to an aesthetically nice coastline. By the same token, an equally skilled and experienced worker could be offered a higher wage to accept a job located in a landlocked area, with cold winters and poor natural and man-made amenities (Albouy, 2008, 2016).

We are not the first to tackle the empirical assessment of the relative importance of different microfoundations of agglomeration economies. This literature has evolved through two major steps.

Earlier work departs from the seminal empirical analysis in Glaeser et al. (1992) and Henderson et al. (1995); they show the differential impact of proxies for concentration or diversity indices (respectively, Marshallian and Jacobsian sources of agglomeration economies) on different economic outcomes (productivity, employment, and GDP growth). This literature has recently been showing that along with a diversity effect and a specialization effect, empirical support is also found for the portfolio effect underlying the related variety literature (Caragliu et al., 2016).

A second advancement in the empirical work on agglomeration economies derives from Rosenthal and Strange (2001), who use proxies for the three microfoundations of agglomeration economies (knowledge spillovers, labor market pooling, and input sharing) to show that labor market pooling is the single most important factor for firms to co-agglomerate. Prior works focusing on individual factors include Audretsch and Feldman (1996) and Dumais et al. (2002). The former finds that R&D intensive activities tend to concentrate spatially more than non-R&D intensive ones. Dumais et al. (2002) propose instead an index of co-agglomeration, and decompose it in shares due to plant births, expansions, contractions, and closures.

A natural evolution of this literature is to show horse races – works showing the differential impacts of different determinants of agglomeration economies on urban productivity. This is the case of

Duranton and Puga (2010), who show that industries more exposed to idiosyncratic volatility tend to concentrate spatially more than the average.

The heterogeneous nature of agglomeration economies is more specifically dealt with in Faggio et al. (2017), who propose an empirical framework to disentangle different forces driving co-agglomeration patterns across different industries. Their results show that (i.) Marshall's agglomeration forces still represent a major driver of co-agglomeration; (ii.) Knowledge spillovers and labor pooling are mostly unintended results;[9] (iii.) Differential results in terms of entry and industry age of companies analysed supports the adaptive nature of agglomeration economies; (iv.) Agglomerative forces work both for high-tech as well as for less technologically intensive industries, although high-tech industries seem to benefit more from knowledge spillovers, while low-tech sectors appear to benefit more from input sharing and labor pooling mechanisms; (v.) agglomeration effects are stronger for small firms.

The focus on the production side of the agglomeration story is complemented with a theoretical analysis of the determinants of location patterns of talented individuals in the model presented in Behrens et al. (2014). One of the main predictions of the model is that utility levels are not equalized across cities, so that talented individuals will sort along the urban hierarchy according to individual preferences. This result represents a major deviation from the standard spatial equilibrium assumption derived in Rosen and Roback, and summarized in Glaeser and Gottlieb (2009).

The empirical literature on agglomeration economies mostly focuses on the production side. A by-product of this evolution is the focus on manufacturing industries; sources of agglomeration economies are much more complex to trace among tertiary industries.[10] Theoretically, the concentration of services in cities feeds back on the hyper-concentration of variety-loving consumers in metro areas (Rivera-Batiz, 1988; and Krugman, 1991 for a treatment less focused on urban areas proper). Empirically, a notable exception is Tabuchi and Yoshoida (2000), who set for a goal similar to ours, but their work (i.) uses metropolitan-area level data, and (ii.) does not benefit from the use of specific proxies for production and consumption sources of agglomeration economies. Our work is also related to Diamond (2016), who finds that the hyper-concentration of skilled workers in productive cities can be reconciled with the rent increases observed in the same urban areas only by considering endogenous increases in amenities within higher skill cities.

Despite the vast interest in the microfoundations of agglomeration economies, the production city vs. consumer city debate still remains a puzzle. To date, very few systematic attempts have been made to analyse the relative importance of amenities related to production to those associated to consumption behaviours. This is still an open issue in the urban economics field and this paper contributes this literature by disentangling consumption and production amenities from the pure agglomerative effects generically ascribed to size and density. In what follows we will test the following hypothesis:

---

[9] For labor pooling, Faggio et al. (2017) report that labor pooling "*effects diminish drastically as coagglomeration increases. In other words, both of these sorts of interactions between industries have a larger effect when the industries collocate less frequently and interactions are more likely to be the sort of unexpected connections on which Jacobs focuses*" (p. 89).

[10] A major exception to this trend is Kanemoto et al. (1996).

H1. What is the relative impact of production and consumption externalities on urban productivity?

At the same time, the debate emerging from the seminal work of Glaeser et al. (2001) suggests that consumption-related externalities would be increasingly important for generating agglomeration economies. Consequently, our work also aims at testing a second hypothesis:

H2. Is the importance of consumption externalities growing over time?

Before discussing the details of our methodology, it is first important to delve into the proxies for net advantages we are going to use. As summarized in Combes et al. (2010a), the recent empirical literature exploits large micro data bases with information on workers and firms to trace workers' locations and activities, and infer on the impact of urban density on their wages or productivities the magnitude of agglomerative effects.

Wages and firm productivity levels have been found to only imperfectly capture agglomeration economies, in particular if the goal to identify disentangle consumption from production amenities. Albouy (2015), Camagni et al. (2016), and de Groot et al. (2015) argue that urban land rent would be a much more effective proxy. This is due to the competition for land lots in metro areas: the same location can be bid for by both residential and commercial users, with an ensuing increase in land rent irrespective of whether both segments, or just one, of the urban land market receive a positive exogenous shock to the demand side. Therefore, in our empirical analyses, urban land rent will be first used as a proxy for urban productivity.

However, the spatial equilibrium approach proposed in Rosen (1979) and Roback (1982) also suggests that both firms and consumers bid for locations richer in both consumption and production amenities. This implies that both consumers and firms will bid for the same locations, in particular when amenities are productivity-enhancing for both. In this case, "*the sign of the wage gradient is unclear while the rent gradient is positive*" (Roback, 1982, p. 1257). In other words, while rents will be univocally increasing in both consumption and production amenities, wages may not, as firms may potentially offer a wage-in-kind by hiring workers in cities rich in consumption amenities. We will address this issue by (i.) looking for the effects of consumption and production amenities on both house prices and firm productivity, with an additional research avenue on identifying the same effects on individual wages, and (ii.) controlling for local wages in house price regressions, and median house prices in wage regressions.

3. Empirical model and identification strategy

3.1 Empirical model

The theoretical problem underlying the empirical disentanglement of consumption from production related externalities can be framed in a cost-benefit approach whereby both consumers and producers maximize their utility or, respectively, profit function. In particular, this issue can be formalized within the Quality of Life literature (Gabriel and Rosenthal, 2004; Blomquist et al., 1988). Within this framework, consumers maximize a utility function

$$\bar{u} = u\left(w_j, r_j | A_j\right) \tag{1.}$$

while firms maximize a profit function

$$\bar{\pi} = \pi(w_j, r_j | A_j) \tag{2.}$$

with w indicating wage in city j relative to a reference city, and similarly r indicating rent w.r.t. a reference urban area. Lastly, A_j is a vector of attributes that describes city j.

The isoutility function is such that higher wages must be offset by higher rents, holding A_j constant. Similarly, the isoprofit function requires that higher wages (a cost in the firm's perspective) must be offset by lower rent.

This implies that firms look for more productive places (those where wages are on average higher), with lower rents; amenities, in turn, could be lower in exchange for lower wages. On the contrary, consumers will look for places with more (consumption-related) amenities, with lower rent, and higher wages.

This setting translates into the following estimable reduced form (Eq. 3):

$$\log r_{i,j,t} = \propto_{r,0,t} + \propto_{r,1} X_{i,j,t} + \gamma_{r,j,t} D_{i,j,t} + u_{r,j,t} \tag{3.}$$

In Eq. 3., X_{ijt} includes controls for apartment-specific features, while metropolitan area attributes are included as fixed effects for each city, D_{jt}.

Within this theoretical setting, firms and consumers are undifferentiated and free to move in space in response to spatial arbitrage, which, as a consequence, does not exist if not in the very short run. This is the assumption of spatial equilibrium: nicer places translate into higher rents, which thus reflect the relative attractiveness of a metropolitan area for both consumers and firms. Spatial equilibrium is a much less effective no arbitrage condition with respect to analogous equilibrium conditions assumed in financial economics: people and firms are mobile only to a minor extent, and respond to exogenous shocks to location advantages imperfectly, and typically with a lag (Glaeser and Gottlieb, 2009).

The relationship between urban amenities and land rent is inescapably endogenous. On the one hand, urban land rent capitalizes future utility levels accruing to both consumers and firms locating in a metropolitan area; in this sense, land rent discourages individuals from relocating when the net benefit is not sufficiently high for justifying the extra locational cost. On the other hand, over time city attributes change as consequences of exogenous shocks (Country-level policies, exogenous supply-side shocks), which an ensuing increased level of rent reflecting changing incentives.

3.2 Identification strategy

In the absence of quasi-natural experiments, in this paper we resort to a traditional identification strategy with the use of Instrumental Variables estimates (henceforth, IVs). Data exploited for this empirical analysis are described in detail in Section 4; for the time being, it is important to stress the relatively static nature of location decisions in the Dutch context, where data are collected for the period 2005-2011.

Within applied urban economics, IVs are by far the most frequently adopted tool for assessing causality direction (Baum-Snow and Ferreira, 2015). A new wave of empirical studies dealing with the identification of the strength of agglomeration economies starts from Ciccone and Hall (1996), who use historical density measure to instrument present-day density. The use of historical instruments satisfies the exogeneity condition (long-lags location decisions are typically explained by different factors w.r.t. present-day patterns), while at the same time satisfying the relevance condition, because of hysteresis in location patterns (Combes et al., 2010b). However, the use of historical instruments has been criticized on the grounds of the clash between the inherently qualitative narrative underlying historical studies and the need to quantify relations proper of the economics discipline (Glaeser, 2007).

More recently, a different approach for the identification of causality directions in urban economics has been proposed (Combes et al., 2010a), based on the use of geological instruments. The rationale for this class of instruments lies in the strict exogeneity of geological/soil composition to present-day productivity. It is instead conceivable that the decision to found a city in the past was based also on the likely fertility of the soil, whose main use was, until the first industrial revolution, agricultural (Debord, 2008).

Our empirical estimates aim at disentangling consumption from production-related sources of agglomeration economies. Ideally, therefore, we would like to exploit a quasi-natural experiment which over the last decade caused an exogenous shift in either of the two types of externalities. In the absence of a convincing exogenous shock to either determinant of agglomeration economies, in this paper we combine both prevailing approaches. In particular, following Combes et al. (2010a) we exploit soil composition data to extract modal soil typologies at the urban level, while, following Ciccone and Hall (1996), we also exploit information on historical population density.

In both cases, it can be argued that a positive correlation exists between soil composition and production and consumption amenities – we would expect a traditional location advantage to cause co-agglomeration of firms and historical buildings, but no link should exist between soil composition and present-day house prices or firm productivity. By the same token, population densities in remote periods of time are expected to be correlated to present-day consumption and production amenities, but no positive link should exist between the former and present-day house prices and firm productivity. These assumptions will be statistically tested with first-stage regressions and the standard battery of IV tests.

4. Data description

4.1 Housing prices

We combine data on house transactions from the *Nederlandse Vereniging van Makelaars en Taxateurs in onroerende goederen* (NVM) with both microdata and publicly available spatial data from Statistics Netherlands. Moreover, we use data on monuments from the Cultural Heritage Agency (*Rijksdienst* *voor het Cultureel Erfgoed*), published soil data from de Vries et al. (1999), and the LISA dataset of all registered firms in the Netherlands. We briefly discuss each of these in turn.

The Dutch association of real estate agents NVM keeps a database of transactions, with details of the house, and the asking price. This database, which covers about 70 per cent of all transactions for the period 1985-2011, has been used extensively in hedonic pricing analysis (Koster & Rouwendal, 2017). We use data for the years 2005, 2008 and 2011, and remove outliers.[11]

4.2 Firm productivity

ORBIS gathers data from their yearly reports on many companies around Europe. Although their database is continually updated, we used the so-called Historical ORBIS (Bureau Van Dijk, 2019), where past information is kept. Balance sheet data are available for 5,593,363 firms in the Netherlands, of which 774,532 are also georeferenced – many of the others are holdings without a physical presence. We select data for 2005, 2008, and 2011. In order to obtain productivity from the ORBIS data, we use the Levinsohn & Petrin method, which we will further discuss in §6 below.

4.3 Wages

The Netherlands offers excellent linked employer-employee data. The System of Social Statistical Datasets (SSB; Bakker et al., 2014) of Statistics Netherlands gives access to both survey and census data on employment, where the surveys offer the additional advantage of good educational variables and a reliable registration of the municipality of work[12]. We select data for 2005, 2008, and 2011, using jobs as of the 31st of December of every year, and harmonize municipalities to 2015. Jobs with fewer than 20 working days as well as hourly wages that are over 50× or under 1/10 of the average are dropped. This leaves ca. 8.3 million jobs.

4.4 Agglomeration externalities

There is a large literature studying agglomeration externalities in the wake of Glaeser et al. (1992). Some of the debate in that literature revolves around the operationalization of proxies for specialization, competition and diversity, which remain the three key aspects to be studied (Caragliu et al. 2016). Following overviews like Beaudry and Schiffauerova (2009, particularly Tables 4 and 5), and meta-analytical results showing how in fact different proxies do not necessarily lead to different results (De Groot et al. 2016), we chose the following operationalizations:

---

[11] In particular, we follow Koster and Rouwendal (2017) by dropping transactions over €1 million or under €25,000, with a square meter price over €5,000 or below €500, or with more than 250 m^2 or fewer than 25 m^2. This drops less than 1% of the dataset. Descriptives for the full and the resulting dataset can be found in Appendix A2.

[12] The main files on the municipality of work (gem_sp) are accurate for single-location firms and for employees that have participated in the Labour Force surveys. For all others, firm plants are allocated the employees that live closest by. This creates a small degree of imprecision in our analysis, since for workers that graduated in the last few years, educational data has fortunately been added from registry sources, and thus they are included in our analysis without perfect knowledge of their location of employment.

- For specialization, we take Krugman's Specialization Index (Krugman 1991). It is "*the standard index among the specialization measures*" (Palan 2010), operationalized as the sum over all industries of the absolute differences between the regional industry structure and the national reference.
- For competition, we take the number of firms per employee, which is a dominant measure in the literature (De Groot et al. 2016).
- For diversity, we take Shannon's entropy, a measure quite similar to the ubiquitous Hirschman-Herfindal index (over a third of the total, according to Beaudry and Schiffauerova 2009), but instead of taking the square of each industry share and then summing these, Shannon's entropy multiplies the share by its natural logarithm, and then sums. Across all specifications, results are tested for the use of either of these indicators.
- For general urbanization effects, i.e. those not related to the presence or even variety of sectors in a region, we use log population density. We also control for the size of local laor markets by including the city's total employment.

4.5 Consumption externalities

We base our urban amenities on the Statistics Netherlands neighbourhood dataset ("*CBS Wijk- en Buurtkaart*"), in particular the years 2006, 2008, and 2011.[13] We choose the following variables:

Variable	Indicator
Access to secondary education	no. of schools within 5 km
Stages	no. of concert and theatre stages within 10 km, not counting festivals
Museums	no. of museums within 10 kms
Library	distance to nearest
Green spaces	distance to nearest
Cinemas	distance to nearest
Restaurants	no. of establishments within 1 km
Cafes	no. of establishments within 1 km

Table 1. Variables and indicators from the Statistics Netherlands neighbourhood dataset

Secondly, we compute the amount of monuments in close geographic proximity (2 kms) based on data from the Cultural Heritage Agency. This includes major monuments, such as the Rijksmuseum or the Palace on Dam Square in Amsterdam, but also a large amount of individual houses with high monumental quality. In Amsterdam, the *Grachtengordel* neighbourhood alone counts 1,830 such monuments, but the post-WWI Nieuw-West area counts only 6, including two windmills and two churches. The number is an approximation of the monumental quality of a neighbourhood.

Thirdly, we compute the distance to the beach, a more recreational attraction that is however an important feature of Dutch cities like The Hague and Haarlem. We exclude the Frisian Wadden

[13] Neighborhood data is not gathered for every year, and the variables included are not the same in every round. We assume 2006 equals 2005. For three rather dynamic variables where 2005 (2006) is unavailable, viz. cafés, restaurants, cinemas, we linearly extrapolate numbers based on 2011-2008. For three rather stable variables where 2011 is missing, viz. museums, libraries, and green space, we assume the numbers from 2008 still apply.

islands, since they require ferry rides to be reached, and include the linear distance from each house to the nearest coastline west of 4°45′. Since the whole North Sea coast has a continuous beachline, with only very short exceptions in between, this is a good approximation of distance to the nearest beach.

Fourthly, we include Central Business Districts. Instead of resorting to administrative units for these, we calculate the density of retail establishments (NACE code 47) within a 2 km radius around each observed house. Next, density hotspots are identified by means of Local Indicators of Spatial Association (Anselin, 1995). Details on the methodology can be found in Appendix A3.

4.6 Instruments

Soil data

Data on soil composition of the Netherlands are available in raster format from Vries et a. (2003). Figure 1 shows the spatial distribution of soil typologies, which have been classified according to their subgroup (with a grand total of 203 soil composition classes[14]).

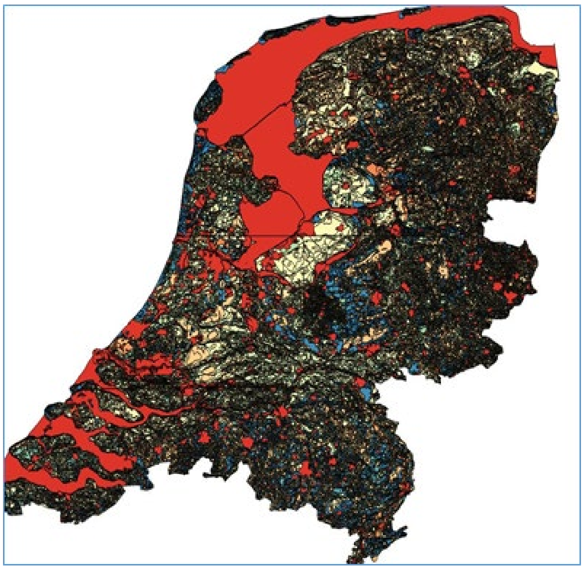


Figure 1. Soil composition map for the Netherlands

Source of raw data: De Vries et al. (2003), Authors' elaboration

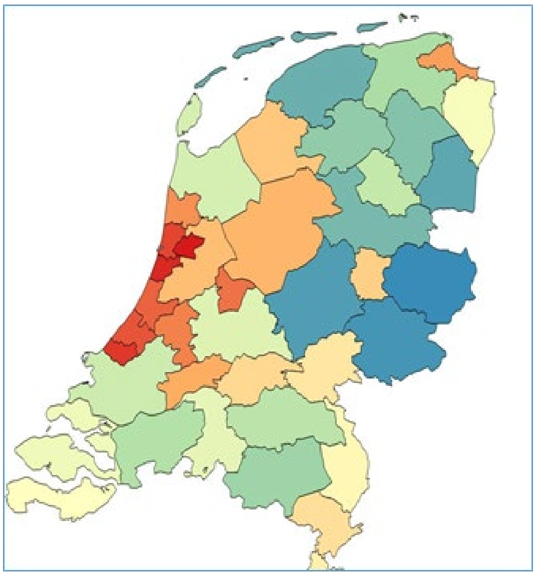


Figure 2. Prevailing type of soil for each Dutch COROP region

Source: Authors' elaboration, using for the regional borders a shapefile by Imergis, Jan-Willem van Aalst, www.imergis.nl, which in turn is based on information of Kadaster

It can be argued that soil composition has in the past represented a considerable locational advantage, causing the decision to found a city – technically often born out of the need to trade agricultural goods – to *also* be based on the productivity advantages available in places rich in soil apt for agricultural purposes (Dickinson, 1951). While urban soil also performs several of the functions also undertaken by purely rural soil (Pouyat et al., 2015), the importance of urban soil for non-artificial activities has over the past decades decreased with respect to typically urban functions,

[14] Appendix A1 shows a full list of the classes included.

with an increase relative importance of first manufacturing, then commercial activities (Krugman, 1993). These arguments make soil composition a potentially very good instrument: the importance of soil composition for early city founding decisions makes it for the relevance of the instrument, while the relative smaller importance of soils for present-day economic activities in cities makes a strong case for instrument validity.

The raster map shown in Figure 1 has been then interpolated with the administrative unit level (COROP region, corresponding to the European NUTS3 level). For each COROP region the modal value of soil composition has been calculated. Color-wise, this corresponds to the shift from Figure 1 to Figure 2, where for each COROP region the prevailing color is represented.

Historical population density

Data on historical population density are instead retrieved from the Dutch Statistical Institute's historical archive, and in particular from the Atlas "*Bevolking en bevolkingsdichtheid der gemeenten van Nederland 1926*". Data used include average population density for Dutch provinces (NUTS2 regions in the present-day EUROSTAT classification) for the years 1830, 1920, and 1926. The levels of population density in the years 1830, 1920, and 1926 are shown in Figures 3.1, 3.2, and 3.3, respectively.

Historical (long) lags of population density is a classical way to deal with identification issues in empirical urban economics (Combes et al., 2010a). The associated identification storyline goes as follows. While on the one hand it is trivial to think of a hysteretic behavior of urban population (instrument relevance), it can hardly be argued that nowadays location decisions are made on account of the century-long time lag of urban populations here considered.

A possible future research avenue could possibly include historical population density data at the municipal level, which nevertheless requires specific treatment because of the gerrymandering processes and boundary changes that took place over the past century.

Time lags of cinemas

An additional source of identification comes from the analysis of remote time lags of cinema distribution in the Netherlands, also used in Koster et al. (2019). Data are collected by VU University's Spinlab.[15] The map of 1910 cinemas is shown in Figure 4 below.

The idea behind this strategy is that early movie theatres would be located in large population concentrations (McKernan, 2007), so as to draw large crowds to this inexpensive entertainment. Also, buildings where cinemas would be located are very often the same typology where presently retail trade establishments are concentrated (Koster et al., 2019), thus suggesting a positive correlation with present-day sources of agglomeration externalities. We may also expect, nevertheless, no link to exist between a century-long lag in the spatial distribution of cinemas and today's house prices, productivity, and wages.

[15] Historical of historical cinemas are available at the project web site http://www.cinemacontext.nl/, designed by Karel Dibbets, currently edited by Julia Noordegraaf.

Figure 3. Historical population density in the Netherlands, 1830-1926

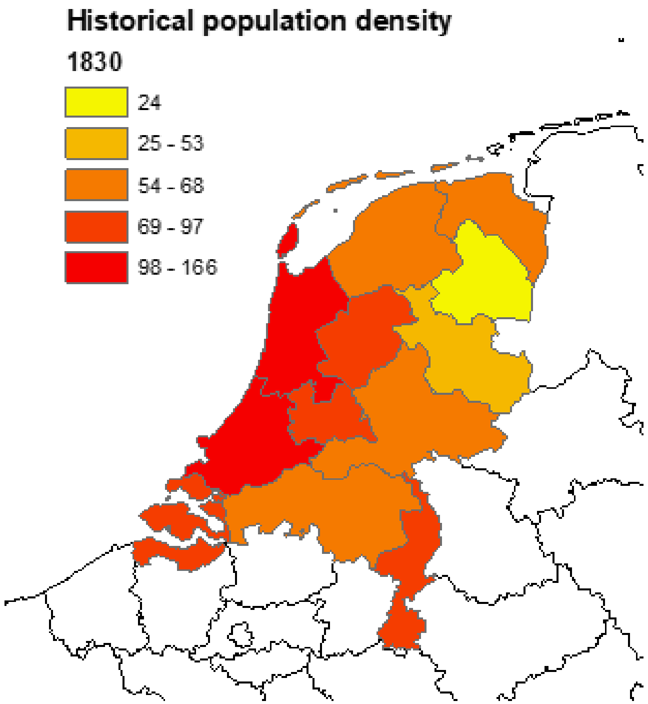


Figure 3.1. Historical population density in the Netherlands, 1830

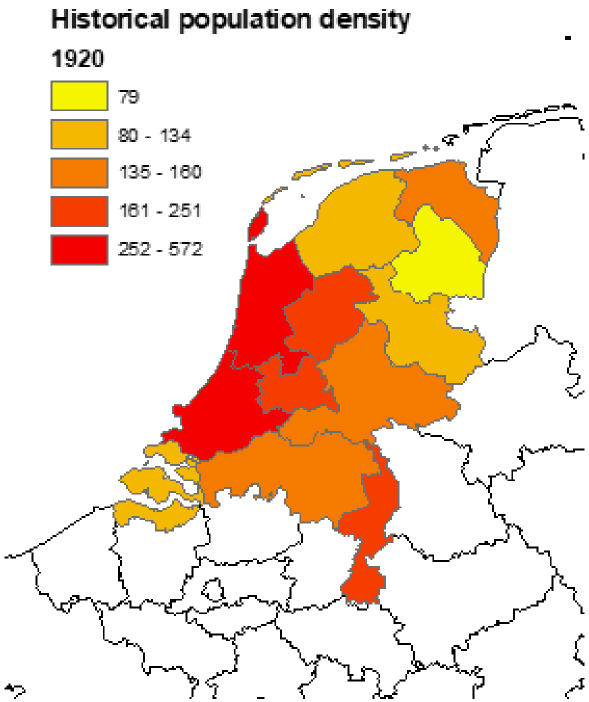


Figure 3.2. Historical population density in the Netherlands, 1920

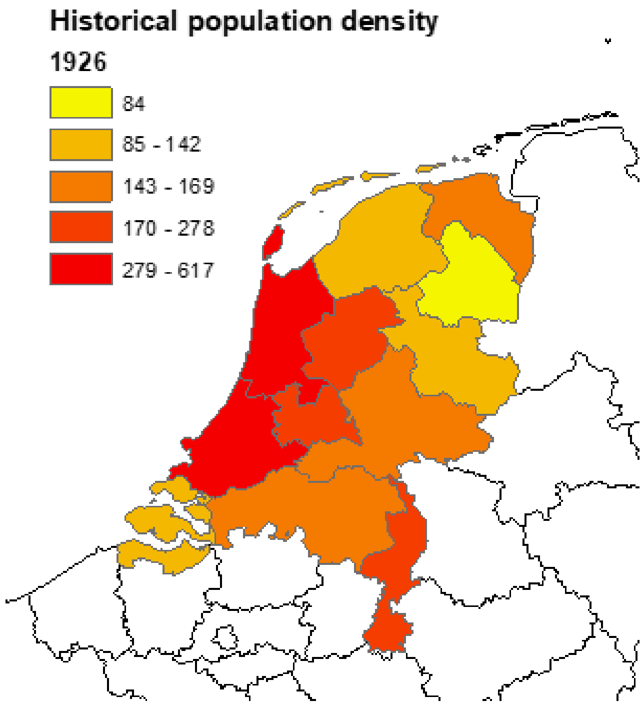


Figure 3.3. Historical population density in the Netherlands, 1926

Source: Bevolking en bevolkingsdichtheid der gemeenten van Nederland 1926, Authors' elaboration

Figure 4. Spatial distribution of cinemas in 1910

Source: VU's Spinlab

5. Consumption vs. production amenities: empirical results (urban land rent determinants)

5.1 Baseline regressions

Table 2 presents the baseline results for answering H1 and H2 on the basis of house prices regressions. The unit of observation is the single house transaction (roughly 70% of all house transactions in the Netherlands) for the years 2005, 2008, and 2011. Results are organized as follows.

Column 1 presents a baseline model with a constant term only. In Column 2 we also include a population density term, while also controlling for the location of each house within the metro area; this is controlled for with the inclusion of a distance in kms. from the Central Business District of each city.[16]In Column 3 we add production amenities, i.e. standard employment and diversity indices.

[16] Density is calculated at the municipality level.

Dep. Variable: Housing price (Control variables relating to the individual house have been omitted)						
	(1) Base model	(2) Urbanization	(3) Production	(4) Consumption 1	(5) Consumption 2	(6) Consumption 2
Distance to CBD		-0.00296***	-0.00262***	-0.00239***	-0.00243***	-0.00231***
		(-5.92)	(-5.65)	(-5.65)	(-6.93)	(-6.78)
Log population density		0.129***	0.141***	0.132***	0.111***	0.109***
		(5.49)	(6.54)	(5.94)	(4.99)	(4.81)
Total employment			-0.845***	-0.806***	-0.644***	-0.411***
			(-6.80)	(-6.44)	(-4.59)	(-3.03)
Specialization: Krugman Specialization Index			-0.0943	-0.105	-0.406***	-0.401***
			(-0.47)	(-0.58)	(-2.69)	(-2.73)
Competition: Firms per employee			58.45***	55.78***	42.27***	29.86***
			(7.14)	(6.92)	(4.83)	(3.61)
Diversity: Shannon Diversity Index			0.168	0.170	-0.0397	0.0470
			(1.05)	(1.10)	(-0.32)	(0.38)
Count of monuments in close geographic proximity				0.0244***	0.0187***	0.0150***
				(4.80)	(3.97)	(3.64)
Distance from the nearest coastline					-0.00115***	-0.00121***
					(-5.46)	(-5.80)
Secondary education (no. of schools within 5 km)					-0.000998	-0.00501***
					(-0.43)	(-3.14)
Stages (no. of concert and theatre stages within 10 km, not counting festivals)					0.0134***	0.00877***
					(3.89)	(3.35)

Museums (no. of museums within 10 km)					-0.0110***	-0.00749**
					(-2.64)	(-2.23)
Library (distance to nearest)					-0.00533**	-0.00388
					(-2.02)	(-1.39)
Green spaces (distance to nearest)					-0.0178*	-0.0150**
					(-1.96)	(-1.99)
Cinemas (distance to the nearest)						0.0200**
						(2.56)
Restaurants (no. of establishments within 1 km)						0.00499***
						(5.68)
Cafes (no. of establishments within 1 km)						-0.00478***
						(-5.83)
Constant	10.94***	10.22***	9.393***	9.259***	10.59***	10.23***
	(225.76)	(73.47)	(12.20)	(12.17)	(16.91)	(16.43)
Observations	291,190	291,190	291,190	291,190	277,791	277791
Adjusted R^2	0.618	0.693	0.712	0.716	0.732	0.747

Table 2. Impact of production and consumption externalities on house prices: baseline OLS estimates

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Variables relating to the individual house omitted.

The employment index captures labor pooling effects, in that larger urban labor markets enhance matching of workers' skills and labor demand; the latter are instead aimed at measuring Jacobsian externalities. In Column 4 we also include a first measure of consumption amenities. In particular, we interpolate the geo-referenced micro data base on house transactions with the shape file with the stock of Dutch monuments described by the Cultural Heritage Agency; the variable included in this estimate is the number of monuments located within a 2 kms. radius from each house object of transaction.

Columns 5 and 6 provide additional evidence on the role of artificial sources of consumption-based sources of agglomeration externalities. Among the amenities included, restaurants (Glaeser et al., 2001; Glaeser and Gottlieb, 2006), green spaces (Bartik and Smith, 1987), and cinemas (Dalmazzo and De Blasio, 2011) are found to be positively associated with urban land rent, after controlling for production and other consumption externalities.

Results show a remarkable degree of consistency across different specifications. The absolute population density indicator remains consistently positive and significant irrespective of the inclusion of consumption and production measures, although the estimated parameter declines by roughly 20 per cent with the inclusion of relevant controls.

We also find robust evidence of a positive role of labor pooling effects; larger urban labor markets are associated with a higher urban land rent, suggesting that firms and individuals are ceteris paribus more eager to bid a higher price for locations with this production-related externality.

Columns 4-6 provide a first answer to H1. Measuring consumption externalities with the number of monuments close to each house transaction, as well as with cultural heritage and pure retail sources of consumption possibilities, we find robust evidence that this amenity causes a mark-up in urban land rent, with an estimated impact roughly equal to 23% of the total density effect (1-0.109, Column 6/0.1141, Column 3).

The estimated density parameter also deserves specific discussion. While its estimate increases in modulus as we include production sources of agglomeration externalities, the inclusion of consumption related sources of agglomeration economies increasingly shrinks the estimated parameter (Figure 5). This result needs further test from our additional regressions (Section 6), with the aim to verify whether (i.) production and consumption-related sources of agglomeration effects go in the same direction, and (ii.) whether other measures of benefits (productivity and wages) reveal similar results.

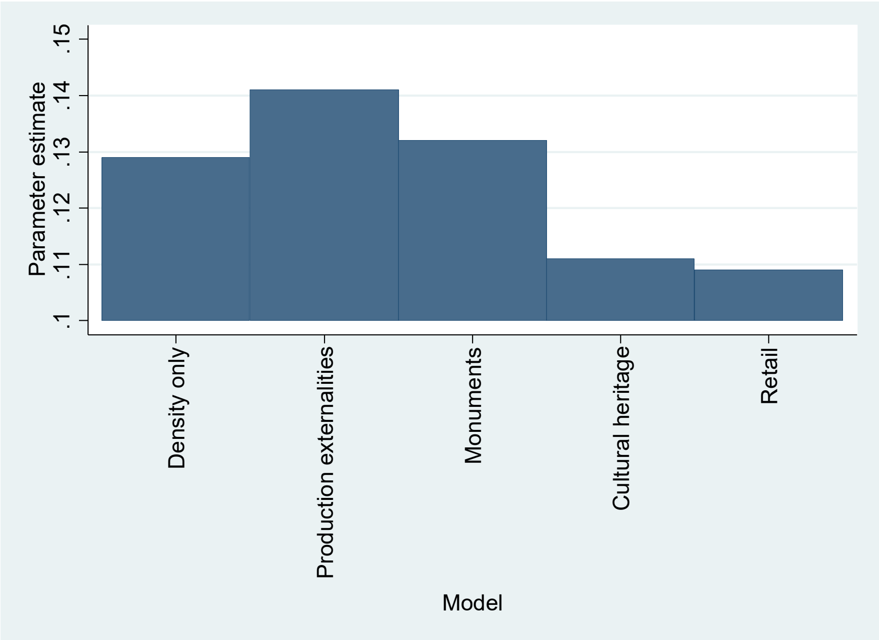
5.1 Robustness checks

In this second subsection we deal with two major issues. We first provide a reply to H2, i.e. whether evidence supports the consumer city argument, assuming a growing importance of consumption possibilities as sources of agglomeration economies. We then verify whether these findings hold after dealing with (possible) endogeneity and reverse causality.

Results of both types of controls are presented in Table 3, which is organized as follows. Column 1 shows results of estimating a model where a variable capturing consumption externalities (count of monuments in close geographical proximity) is interacted with time dummies. The base year is 2005;

therefore, we read the two interactions as the additional effect that monuments have on house prices in 2008 and 2011, respectively.

Figure 5. OLS estimates of the density effect on house prices



Source: Authors' elaboration

Column 2 does the same with an indicator of production externalities (firms per employee). Column 3 then looks at whether these two last indicators exert a joint effect, by including the interaction between the count of proximate monuments and firms per employee. Lastly, Columns 4 through 6 show 2SLS estimates whereby first monuments, then firms per employee, then both are instrumented with soil composition and historical population density.

Results provide weak support for the dynamic version of the consumer city argument, i.e. that cities would be increasingly a place for consumption, rather than of production. Column 1 shows a positive and significant estimate only for the 2008 mark-up on the monuments effect: this means that in 2008 the effect of touristic amenities was stronger than in 2005. In 2011, no significant difference is detected. The flipside of this result is that (Column 2) competition mechanisms matter less in 2011 than in 2005, as suggested by the negative and significant interaction between the 2011 period dummy with the firms per employee vector.

Column 3 suggests that the simultaneous presence of production and consumption sources of agglomeration economies may be beneficial to a city's productivity. To better visualize this point we also represent a marginal effects graph (Figure 6), where we plot marginal effects of consumption amenities (monuments in close proximity, sorted by quintiles) for the five quintiles of the production amenity (firms per employee) captured in Table 3, Column 3.

Figure 6 highlights that while the impact of consumption amenities is positive and significant throughout the interval of definition of the production amenities indicator, higher levels of competition among firms also correspond to stronger consumption amenities impacts.

Dep. Variable: Housing prices						
	(1) Consumption over time	(2) Production over time	(3) Combined (interaction)	(4) Consumption IV	(5) Production IV	(6) General IV
Distance to CBD	-0.00224***	-0.00224***	-0.00226***	-0.0393	-0.00231***	-0.00129***
	(-6.61)	(-6.60)	(-6.66)	(-1.48)	(-88.24)	(-10.98)
Log population density	0.105***	0.108***	0.107***	1.337	0.109***	0.0703***
	(4.62)	(4.78)	(4.68)	(1.52)	(101.75)	(16.67)
Total employment	-0.415***	-0.553***	-0.448***	-0.082	-0.411***	8.10e-08*
	(-2.98)	(-3.54)	(-3.20)	(-1.47)	(-46.25)	(1.92)
Specialization: Krugman Specialization Index	-0.404***	-0.466***	-0.422***	-2.214*	-0.401***	-0.256***
	(-2.72)	(-2.98)	(-2.82)	(-1.68)	(-43.31)	(-15.66)
Competition: Firms per employee	29.72***	42.81***	38.15***	541.2	29.86***	-2.858
	(3.45)	(3.85)	(3.74)	(1.48)	(53.64)	(-1.02)
Diversity: Shannon Diversity Index	0.0533	0.000452	0.0198	-2.287	0.0470***	0.172***
	(0.41)	(0.00)	(0.16)	(-1.36)	(5.68)	(11.48)
Count of monuments in close geographic proximity	0.0113***	0.0154***	0.0201***	-5.212	0.0150***	0.173***
	(2.99)	(3.43)	(4.26)	(-1.39)	(37.62)	(10.35)
Count of monuments in close geographic proximity (additional effect 2008 over 2005	0.0206**					
	(2.23)					
Count of monuments in close geographic proximity (additional effect 2011 over 2005	0.0126					
	(1.39)					
Distance from the nearest coastline	-0.00119***	-0.00121***	-0.00121***	-0.0158	-0.00121***	-0.000762***

	(-5.76)	(-5.80)	(-5.81)	(-1.51)	(-89.27)	(-15.27)
Secondary education (no. of schools within 5 km)	-0.00473***	-0.00434**	-0.00496***	0.00215	-0.00501***	-0.00614***
	(-2.95)	(-2.59)	(-3.10)	(0.36)	(-40.41)	(-34.76)
Stages (no. of concert and theatre stages within 10 km, not counting festivals)	0.00834***	0.00803***	0.00912***	0.0709	0.00877***	0.00769***
	(3.18)	(2.99)	(3.54)	(1.59)	(64.78)	(33.19)
Museums (no. of museums within 10 km)	-0.00735**	-0.00692**	-0.00757**	-0.0227**	-0.00749***	-0.00792***
	(-2.19)	(-2.04)	(-2.26)	(-1.98)	(-52.59)	(-43.35)
Library (distance to nearest)	-0.00384	-0.00366	-0.00376	-0.256	-0.00388***	0.00351***
	(-1.40)	(-1.32)	(-1.34)	(-1.42)	(-10.55)	(3.82)
Green spaces (distance to nearest)	-0.0154**	-0.0148*	-0.0147*	-0.0462	-0.0150***	-0.0143***
	(-2.07)	(-1.94)	(-1.95)	(-1.38)	(-15.01)	(-11.35)
Cinemas (distance to the nearest)	0.0197**	0.0183**	0.0200**	0.310	0.0200***	0.0146***
	(2.57)	(2.28)	(2.54)	(1.49)	(46.60)	(15.68)
Restaurants (no. of establishments within 1 km)	0.00498***	0.00502***	0.00499***	0.0186*	0.00499***	0.00464***
	(5.70)	(5.83)	(5.70)	(1.89)	(106.11)	(65.17)
Cafes (no. of establishments within 1 km)	-0.00470***	-0.00484***	-0.00479***	0.0371	-0.00478***	-0.00611***
	(-5.65)	(-5.83)	(-5.86)	(1.23)	(-81.19)	(-39.17)
Firms per employee (additional effect 2008 over 2005)		-0.290				
		(-0.11)				
Firms per employee (additional effect 2011 over 2005)		-9.502***				
		(-3.12)				

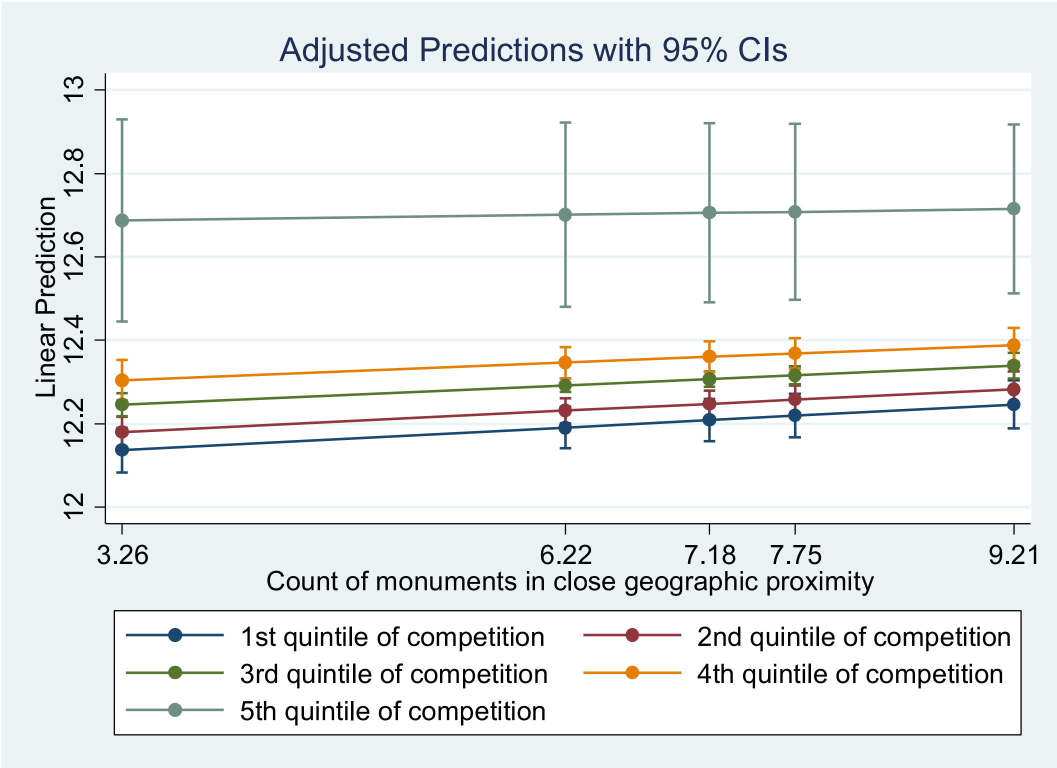
Interaction production-consumption amenities			-0.877^{***}			
			(-2.66)			
Constant	10.10^{***}	10.44^{***}	10.33^{***}	53.30^{*}	10.23^{***}	8.663^{***}
	(15.23)	(16.18)	(16.39)	(1.73)	(248.99)	(53.68)
Observations	277791	277791	277791	277791	277791	277791
Adjusted R^2	0.747	0.748	0.747	-157.075	0.747	0.601

Table 3. Impact of production and consumption externalities on house prices: robustness checks and 2SLS

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Control variables relating to the individual house have been omitted*

Lastly, Table 3 deals with identification issues. We first instrument the count of monuments (Column 4), then firms per employee (Column 5), and both (Column 6), with modal soil composition at the NUTS3 level, and historical (1826-1920-1926) population density at the NUTS2 level. While all variables of interest retain sign and (mostly) significance levels across all three IV specifications, in the last column with the most general (preferred) specification we obtain a point estimate for the consumption amenity that is slightly larger in modulus w.r.t. the OLS equivalent, while the parameter estimate associated to firms per employee becomes not significant.

Figure 6. Interaction between consumption and production externalities



Source: Authors' elaboration

Across all specifications, the null hypotheses of all standard IV tests (underidentification test/Anderson LM statistic; Weak identification test/ Cragg-Donald Wald F statistic; and the overidentification test of all instruments/Sargan statistic) are rejected at all standard significance level.

All in all, results suggest that both sources of agglomeration economies matter for engendering productivity advantages for cities. Additional evidence will now be sought with firm productivity data.

6. Consumption vs. production amenities: empirical results (productivity determinants)

This section discusses in detail an additional piece of information related to the role of consumption and production amenities in determining productivity advantages. To measure the impact of agglomeration economies on firm productivity, we use data on firm-level productivity as measured by the Levinsohn and Petrin's TFP estimator (Levinsohn and Petrin, 2003). This estimator addresses the problem of simultaneity in estimating firm-level productivity, and finds a way out by instrumenting the stock of capital with the costs for intermediate inputs. We first regress the log of firm level value added on the log of employees and the log of fixed assets as a proxy for the stock of capital. The latter is instrumented with the costs of goods sold. The predicted value added is then used as dependent variable in this set of regressions.

Formally, we test the following reduced form model:

$$\log prod_{i,j,t} = \propto_{r,0,t} + \propto_{r,1} X_{i,j,t} + \gamma_{r,j,t} D_{i,j,t} + u_{r,j,t} \quad (4.)$$

6.1 Baseline OLS estimates

Eq. (4.) is first tested by means of heteroscedasticity-robust standard errors. Baseline OLS results are shown in Table 4.

Table 4 is organized as follows. Regressors are added one per column, starting from population density, which subsumes both production and consumption-related agglomeration amenities. We then start including consumption amenities, viz. distance from the closest monument, distance from the nearest coastline, cafés, childcare establishments, secondary education institutions, cinemas, and libraries.

We then add production-related controls: firms per employee as a measure of competition, and the Hirschmann-Herfindahl Index and Krugman specialization index in order to capture MAR or Jacobs externalities.

Results remain substantially consistent across different specifications. Density always remains positively and significantly associated to firm productivity. Its parameter estimate (illustrated in Figure 7 for the different model specifications in Table 4) increases in modulus as we include further controls. This suggests that pure density mechanisms actually subsume productivity-enhancing characteristics that also correlate with it. In particular, sorting out from the density parameters facilities that are often densely located in cities (childcare establishments, schools, cafés, and restaurants) our estimates imply a stronger pure density mechanism that truly reflects pure urban externalities.

In terms of magnitude, comparing density estimates in the baseline specification (1.3, Column 1) to the most general specification (4.9, Column 10 implies a 3.7 increase in the size of parameter estimates, which suggests a rather substantial role of consumption amenities in the explanation of the true city size effect.

Another interesting result is related to the positive and significant impact of being remote from the coast. This finding is apparently counterintuitive, given the location on the coast on in near proximity of cities such as Amsterdam, Utrecht, and The Hague. Still, the Dutch industrial structure is mostly concentrated realize many large firms in the Netherlands are located closer to the South-Eastern area, where the headquarters of major MNCs, such as Phillips, are also located. In fact, holding density constant, the key urban centers in the periphery of the Country (therefore, far from

Dep. Variable: Levinsohn and Petrin's TFP										
	(1) Density	(2) Monumen ts	(3) Coastlin e	(4) Cafés	(5) Restaurant s	(6) Childcare	(7) Education	(8) Theatres	(9) Libraries	(10) Production
Population density	1.326***	1.326***	1.402***	1.480***	3.254***	3.254***	3.230***	3.182***	3.256***	4.876***
	(23.18)	(22.92)	(23.28)	(24.28)	(27.99)	(27.99)	(28.02)	(27.85)	(27.93)	(30.05)
Distance from the closest monument		0.951	5.271	-35.89***	-121.3***	-121.3***	-123.8***	-80.09***	-32.24	-7.984
		(0.15)	(0.79)	(-5.27)	(-6.06)	(-6.06)	(-6.18)	(-3.78)	(-1.49)	(-0.38)
Distance from the nearest coastline			2.996***	3.923***	8.041***	8.041***	8.325***	7.587***	7.379***	5.096***
			(18.16)	(21.33)	(19.29)	(19.29)	(19.53)	(18.81)	(18.44)	(11.97)
Cafes (no. of establishments within 1 km)				-6.489***	-36.78***	-36.78***	-38.91***	-33.10***	-34.36***	-37.18***
				(-15.21)	(-27.36)	(-27.36)	(-26.85)	(-24.66)	(-21.97)	(-23.27)
Childcare (no. of daycare establishments within 3 km)					1.955*	1.955*	-4.516***	-28.78***	-35.22***	-41.40***
					(2.39)	(2.39)	(-4.73)	(-15.39)	(-17.86)	(-18.42)

Secondary education (no. of schools within 5 km)	25.27***	10.38**	-41.85***	-71.83***
	(7.21)	(3.02)	(-9.93)	(-15.20)
Cinema (distance to nearest)		312.1***	718.8***	1059.0***
		(19.36)	(26.49)	(28.81)
Library (distance to nearest)			-102.3***	-115.9***
			(-9.53)	(-10.38)
				-423627.7***
Specialization: Krugman specialization index				-4815.6***
				(-15.97)
Competition: Firms per employee				
				(-26.16)
Diversity: Hirschmann-Herfindahl Index				294803.0***
				(32.89)

Constant	-449.9***	775.5	-514.4***	-949.5	11189.6	11189.6	11085.6	11665.3	126.2	-287171.8***
	(-23.08)	(1.51)	(-18.57)	(-0.00)	(0.00)	(0.00)	(.)	(.)	(1.61)	(-32.65)
Observations	774,532	774,532	774,532	687,518	235,007	235,007	235,007	235,007	216,781	216,781
Adjusted R^2	0.063	0.063	0.063	0.083	0.207	0.207	0.207	0.207	0.214	0.221

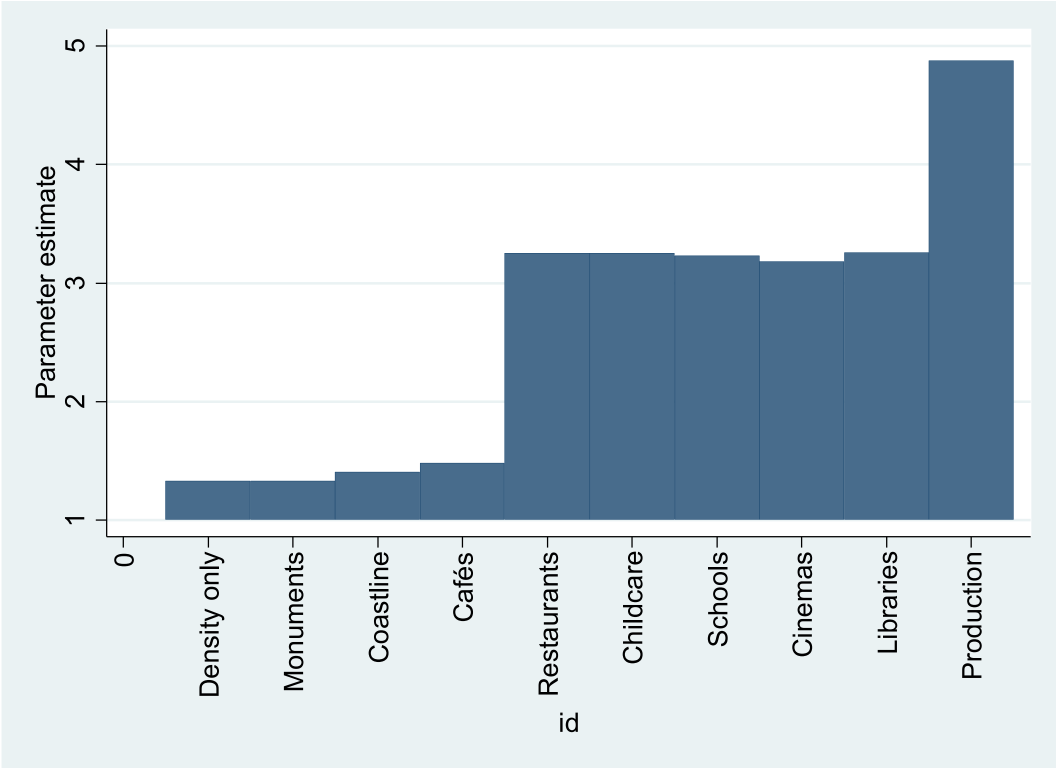
Table 4. Impact of production and consumption externalities firm productivity: baseline OLS estimates

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

-

the cost) tend to be characterized by stronger agglomeration economies than their size-equivalents in the Randstad, where many dense urban areas act as suburbs to larger nearby cities.

Figure 7. OLS estimates of the density effect on firm productivity



Source: Authors' elaboration

Results for production-related agglomeration externalities are also robust. We consistently find that firms tend to be more productive in places with (i.) lower competition (as measured by firms per employee), (ii.) more industrially concentrated (as measured by the Hirschmann-Herfindahl Index) and more diversified (as captured by the Krugman Specialization Index).

6.2 Baseline OLS estimates

Table 5 presents a first round of robustness checks also aiming at further supporting our reply to H2, i.e. whether cities are increasingly becoming places for consumption, rather than production hotspots. To answer (again) this question we interact consumption and production agglomeration externalities with year dummies; interaction terms can be interpreted as additional marginal impacts of each variable on top of the baseline (2005) correlation.

Results hint quite univocally at a growing importance of consumption amenities also for firm productivity, with the only exception of libraries, whose impacts steadily and uniformly declines over time. This last result is in line with the recent evolution of knowledge consumption devices, increasingly dematerialized with a growing percentage of regular readers using e-readers and mobile phones for their everyday reading activities (Miranda et al., 2011).

Results on production externalities are slightly more complex. While a univocally decreasing negative impact of competition is detected, we also find that the impact of concentration (HHI) and

Dep. Variable: Levinsohn and Petrin's TFP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Cafés	Childcare	Cinemas	Libraries	Competition	HHI	KSI
Population density	1.739***	1.735***	1.740***	1.740***	1.762***	1.660***	1.691***
	(23.48)	(23.54)	(23.48)	(23.48)	(23.57)	(24.02)	(23.69)
2008 dummy	728.8***	685.7***	807.1***	944.8***	734.5***	-43884.3***	1587.5***
	(82.19)	(63.54)	(88.10)	(75.90)	(55.83)	(-14.48)	(27.69)
2011 dummy	1784.5***	1345.4***	1359.6***	2107.0***	1006.1***	273898.6***	-3088.1***
	(51.23)	(44.53)	(67.03)	(47.71)	(29.85)	(15.43)	(-11.69)
Cafes (no. of establishments within 1 km)	-11.47***	-9.977***	-9.223***	-10.60***	-10.48***	-11.43***	-10.94***
	(-16.75)	(-10.69)	(-9.87)	(-10.54)	(-10.42)	(-10.90)	(-10.70)
Cafes (no. of establishments within 1 km) in 2008	10.41***						
	(19.01)						
Cafes (no. of establishments within 1 km) in 2011	-12.63***						
	(-10.54)						
Secondary education (no. of schools within 5 km)	9.884***	-19.74***	2.115	8.168***	5.542**	12.63***	8.714***
	(5.53)	(-12.08)	(1.25)	(4.58)	(3.22)	(6.62)	(4.84)

Cinema (distance to nearest)	57.30***	73.63***	-4.017	66.04***	81.17***	53.93***	64.56***
	(8.08)	(9.40)	(-0.74)	(8.98)	(10.75)	(7.86)	(8.88)
Library (distance to nearest)	-99.21***	-100.5***	-97.41***	14.55***	-101.7***	-93.22***	-91.21***
	(-30.38)	(-30.48)	(-30.02)	(4.10)	(-30.67)	(-28.95)	(-28.38)
Firms per employee	-70915.0***	-70683.1***	-70354.1***	-70230.1***	-144847.8***	-61294.2***	-54077.0***
	(-23.72)	(-23.66)	(-23.34)	(-23.49)	(-32.31)	(-23.05)	(-21.91)
Hirschmann-Herfindahl Index	102424.4***	101147.4***	105918.9***	103378.0***	110965.7***	179480.0***	89441.3***
	(38.45)	(38.74)	(38.08)	(38.47)	(39.24)	(23.26)	(40.40)
Krugman Specialization Index	-367.2***	-362.1***	-213.2*	-369.5***	-112.5	121.2	-3923.6***
	(-3.88)	(-3.82)	(-2.29)	(-3.87)	(-1.25)	(1.32)	(-13.07)
Secondary education (no. of schools within 5 km) in 2008		17.03***					
		(15.00)					
Secondary education (no. of schools within 5 km) in 2011		53.55***					
		(11.84)					
Cinemas in 2008			3.530				
			(1.03)				

Cinemas in 2011	289.8***				
	(16.36)				
Libraries in 2008		-92.71***			
		(-18.15)			
Libraries in 2011		-246.1***			
		(-19.69)			
Specialization: Krugman Specialization Index in 2008					-2091.7***
					(-13.70)
Specialization: Krugman Specialization Index in 2011					12422.0***
Competition: Firms per employee in 2008			24863.0***		
			(11.60)		
Competition: Firms per employee in 2011			155454.9***		
			(23.25)		
Diversity: Hirschmann-Herfindahl Index in 2008				45451.8***	

						(14.75)	
Diversity: Hirschmann-Herfindahl Index in 2011						-276953.0***	
						(-15.35)	
							(16.71)
Constant	-101152.8***	-99725.6***	-104541.3***	-102278.8***	-109367.6***	-177076.3***	-87074.0***
	(-38.12)	(-38.45)	(-37.82)	(-38.19)	(-38.94)	(-23.21)	(-40.39)
Observations	657,774	657,774	657,774	657,774	6577,74	657,774	657,774
Adjusted R^2	0.018	0.018	0.018	0.018	0.018	0.020	0.021

Table 5. Impact of production and consumption externalities on firm productivity: time trends in production and consumption amenities

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

sectoral diversity (KSI) is maximum in 2008 to decline again in 2011. Results could also be affected by the timing of the data observed for our estimates. Measuring GDP in chain linked volumes (base year 2010), Dutch GDP reached its pre-crisis zenith in 2008, a level that was reached again only in 2015. 2011 is therefore still an overall contraction year for the Country as a whole.

6.3 2SLS estimates-

Table 6 presents 2SLS estimates of the impact of production and consumption externalities on firm productivity. For the sake of symmetry, we exploit the same battery of instruments described in Section 4.4. The exclusionary restriction here is that while prevailing soil composition and historical population density matter for present-day production and consumption amenities, the same link should not be valid with present-day firm productivity.

Table 6 is organized as follows. Column 1 presents a baseline OLS specification controlling for density only. Column 2 presents the 2SLS version of the same specification. In Columns 3 and 4 we also control for consumption and production amenities, respectively, while instrumenting again density alone; and, lastly, Column 5 presents the most general specification, based on Instrumental Variables estimates of density, all consumption, and all production amenities.

Throughout the different specifications, we obtain the following main messages:

- Density remains constantly positively and significantly associated with firm productivity, safe for the reduced form specification where it is the only variable being instrumented (and controlled for). This result confirms again that compact urban forms (probably by enhancing face-to-face contacts, and facilitating knowledge generation and accumulation) are more prone to engender productivity advantages, also controlling for other characteristics that come with dense urban environments;
- Both after instrumenting individual sources of agglomeration economies, or when all are instrumented, most the main messages found with baseline OLS techniques still hold. In particular, firms become more productive in (i.) less competitive environments and (ii.) in ore industrially diverse environments. Instead, on the basis of 2SLS estimates we find a negative and significant association between the Hirschmann-Herfindahl index, capturing the intensity of concentration in the urban area, and firm productivity.
- As for consumption amenities, the results are instead quite different than in the baseline OLS case, thus suggesting that the bias stemming from incorrectly treating our covariates as exogenous is larger than the estimated parameter ($E\left(\beta - \hat{\beta}\right) = \beta + bias$. In particular, while estimated parameters hold the same sign and significance for childcare establishments (firms further way from this amenity become less productive), secondary schools (with a negative impact on firm productivity), cinemas (with a positive impact) and libraries (with a negative association to firm productivity), opposite findings are in particular identified for touristic amenities (distance from the coastline being associated with a negative and significant impact on firm productivity) and proximity to monuments, which seemed to have no impact when production amenities would be taken into account into a general specification estimated by means of standard robust OLS (Table 4).

Dependent variable: Levinsohn and Petrin's TFP					
	(1)	(2)	(3)	(4)	(5)
Population density	1.326***	-0.114	2.849***	14.09***	9.027***
	(23.18)	(-1.30)	(14.89)	(17.56)	(40.20)
Firms per employee			-107490.4***	-894430.9***	-631113.2***
			(-12.01)	(-21.41)	(-39.72)
Hirschmann-Herfindahl Index			138064.4***	1151955.3***	-156463.8***
			(8.20)	(15.52)	(-7.25)
Krugman Specialization Index			-5293.5***	-24208.4***	-61452.5***
			(-12.94)	(-13.32)	(-27.59)
Distance from the closest monument				575.3***	6333.2***
				(8.60)	(18.88)
Distance from the nearest coastline				7.280***	-14.61***
				(7.03)	(-13.78)
Cafes (no. of establishments within 1 km)				-67.53***	31.08***
				(-16.24)	(5.46)
Childcare (no. of daycare establishments within 3 km)				17.12**	-75.69***
				(2.65)	(-12.03)

Secondary education (no. of schools within 5 km)				-109.5***	-197.7***
				(-10.01)	(-15.00)
Cinema (distance to nearest)				1537.6***	1562.5***
				(32.85)	(29.31)
Library (distance to nearest)				-152.1***	-146.3***
				(-6.08)	(-5.46)
Museums (no. of museums within 10 km)				-177.2***	58.89***
				(-16.17)	(6.18)
Constant	-449.9***	53.26	-135214.7***	-1133509.2***	162279.9***
	(-23.08)	(0.01)	(-7.93)	(-15.31)	(6.45)
Observations	774,532	774,532	774,532	205,263	205,263
Adjusted R^2	0.063	0.053	0.061	0.171	0.070
Estimation method	OLS	2SLS	2SLS	2SLS	2SLS
Variables instrumented	-	Density	Density	Density	Density, all production amenities, all consumption

Table 6. Impact of production and consumption externalities on firm productivity: 2SLS estimates

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Across all specifications, the null hypotheses of all standard IV tests (underidentification test/Anderson LM statistic; Weak identification test/ Cragg-Donald Wald F statistic; and the overidentification test of all instruments/ Sargan statistic) are rejected at any standard significance level.

7. Consumption vs. production amenities: empirical results (wage determinants)

In this section we look for signs of consumption and production externalities in workers' wages. To this aim we exploit information on the universe of Dutch workers for the same period analysed for prior analyses (2005, 2008, and 2011).

This entails a two-step procedure. In the first stage we estimate the determinants of individual wages regressed against individual traits and locational and time fixed effects, in a classical Mincerian framework (Mincer, 1974; Groot et al., 2014). The estimated wage regression is as follows:

$$\log wage_{i,j,t} = \alpha_{i,j,t} + \beta * X_{i,j,t} + \gamma * \mu_{j,t} + u_t \quad (5.)$$

where indices i, j, and t stand for individual i, area j, and time t, respectively. Matrix X contains individual-specific traits; we control for gender, migrant status, part time work, age, and education, and we also control for the year and industry observed, with t = *[2005, 2008, 2011]*. Finally, we calculate the residuals for each combination of municipality of residence and of work, assuming employees access both in their daily systems. These form the basis for our second stage. Results of the first stage are available upon request.

In the second stage we explain the residuals for each commuting combination of municipalities. We include commuting inside a municipality, but remove all combinations based on fewer than 200 workers. We then explain the residuals using externalities for the municipality of residence (Res) and of work (Work). Results of this set of regressions are shown in Table 7.

Table 7 is organized as follows. Column 1 shows results of the OLS specification. In Columns 2 through 5, instead, results for 2SLS estimates are displayed. Across all estimates, excluded instruments are the time lags of population, modal soil composition, and 1910 distribution of cinemas. Columns 2 through 4 show results of estimating with Instrumental Variables the main production-related externalities, while Column 5 shows results of using the same instruments for the two main consumption amenities in the data, i.e. cinemas and restaurants in the place of work.

After netting out all control variables, pure density effects range between 1 and 2.5 per cent of total agglomerative effects, in line with the recent evidence summarized in Combes and Gobillon (2015). This estimate becomes roughly twice as large (nearly two per cent) w.r.t. the baseline OLS specification, as usual in the IV literature. Similar magnitudes are found when instrumenting competition, diversity, and consumption externalities.

Columns 3 and 4 show results of using our historical and geological instruments for assessing the relevance of competition and diversity. In the former case, the parameter estimate for the competition measure for the place of work remains positively and significantly associated with local

	(1) OLS estimates	(2) Instrumenting density	(3) Instrumenting competition	(4) Instrumenting diversity	(5) Instrumenting consumption amenities
Res.: av. distance to supermarket	-0.000165	-0.000461	-0.000116	-0.00209	0.00279
	(-0.07)	(-0.20)	(-0.05)	(-0.87)	(0.51)
Res.: av. distance to cafe	0.0150***	0.0157***	0.0154***	0.0192***	0.0138***
	(9.25)	(9.31)	(7.52)	(9.93)	(7.24)
Res.: av. distance to restaurant	-0.00583**	-0.00407	-0.00665**	-0.00242	0.00331
	(-2.40)	(-1.54)	(-2.13)	(-0.84)	(0.48)
Res.: av. distance to train station	-0.00000305	0.0000384	0.0000290	-0.000169	0.000888*
	(-0.01)	(0.19)	(0.14)	(-0.81)	(1.83)
Res.: av. distance to cinema	-0.0000928	-0.0000749	-0.000148	0.0000269	-0.00216***
	(-0.50)	(-0.40)	(-0.77)	(0.14)	(-2.75)
Work: av. distance to supermarket	-0.000500	-0.00122	-0.000976	0.000260	0.00366
	(-0.18)	(-0.47)	(-0.36)	(0.10)	(0.25)
Work: av. distance to cafe	-0.00441**	-0.00243	-0.00233	-0.00335*	-0.00454*
	(-2.22)	(-1.24)	(-1.05)	(-1.66)	(-1.77)
Work: av. distance to restaurant	-0.0131***	-0.0103***	-0.0130***	-0.0125***	-0.0533**
	(-4.41)	(-3.27)	(-4.06)	(-3.96)	(-2.43)
Work: av. distance to train	-0.000248	-0.000114	-0.000114	0.000120	-0.00242***

station					
	(-1.07)	(-0.52)	(-0.51)	(0.50)	(-2.76)
Work: av. distance to cinema	-0.000652***	-0.000734***	-0.000803***	-0.000959***	0.00493**
	(-3.26)	(-3.66)	(-3.94)	(-4.48)	(2.53)
Work: competition	2.866***	2.267***	1.822**	0.711	3.231***
	(10.24)	(7.93)	(2.45)	(1.40)	(9.02)
Work: diversity	1.530***	2.335***	2.496***	8.441***	0.664
	(4.11)	(5.51)	(4.19)	(5.37)	(1.34)
Res.: competition	0.957***	0.910***	-0.641	1.559***	1.146***
	(3.50)	(4.09)	(-0.53)	(4.10)	(4.56)
Res.: diversity	-0.373	-0.419	0.135	-3.828**	-0.201
	(-1.13)	(-1.31)	(0.22)	(-2.50)	(-0.57)
Work: log pop. density	0.0100***	0.0198***	0.0184***	0.0250***	0.00238
	(5.34)	(6.15)	(6.49)	(7.42)	(0.75)
Res.: distance to coastline	-0.000361***	-0.000359***	-0.000386***	-0.000266***	-0.000384***
	(-6.35)	(-9.16)	(-7.63)	(-4.80)	(-8.16)
Work: distance to coastline	-0.000269***	-0.000221***	-0.000239***	-0.000336***	-0.000243***
	(-4.63)	(-5.11)	(-4.82)	(-6.65)	(-5.23)
year==2008	0				
	(.)				
year==2011	-0.0232***	-0.0227***	-0.0211***	-0.0218***	-0.0226***
	(-17.39)	(-16.43)	(-14.18)	(-15.04)	(-15.09)
Constant	-1.011***	-1.828***	-2.503***	-4.509***	-0.275

	(-3.21)	(-4.51)	(-5.01)	(-5.34)	(-0.61)
Observations	4930	4930	4930	4930	4930
Adjusted R^2	0.485	0.482	0.474	0.447	0.392

Table 7. Impact of production and consumption externalities on wage premia: OLS and 2SLS estimates

Instruments: lagged population density (1830, 1920, 1936); modal soil composition; distance to closest cinema in 1910. Variables for specialization, competition and diversity are defined as in the previous tables.
t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$

wages, while the same does not apply for the impact of the competition measure for the place of residence. A similar effect takes place for the diversity measure, positively and significantly associated with local wages for the place of work in both the OLS and IV specification, while for the place of residence this parameter estimate switches sign, suggesting that local wages in places of residence (mostly characterized by large concentrations of dwellers) thrive on the specialization of areas in a few industries, arguably related to the nature of such places.

Lastly, in Column 5 we deal with potential endogeneity in consumption amenities. While this should not represent an issue for geographical amenities, exogenous by nature (e.g.: access to the natural amenity represented by the coastline), the same does not hold true for man-made amenities, such as restaurants, cafes, etc.

Results in this case show two main facts:

- Restaurants still represent a source of substantial productivity advantages as measured by local wages. In line with the IV literature, the estimated parameter maintains sign and significance, while the magnitude increases as much as fourfold (from about 1.25 per cent as an average estimate across Columns 1-4 to 5 per cent in Column 5).
- Access to cinemas appears instead to flip sign; under the IV specification we find a positive and significant association between proximity to cinemas and local wages (in the place of work).

Across all estimates, the null hypotheses of both the underidentification and weak identification tests are rejected at all conventional levels.

8. Structural Equations Model results

The big question whether and how wages, rent, and productivity would actually be "*as tangled as well-tossed spaghetti*" (Putnam, 2000) remains open. In order to identify causality links among these three mechanisms of spatial price equilibrium, this section discusses results of pooling the three prior sets of estimates within a Structural Equations Model (henceforth, SEM; Wooldridge, 2010) whereby each agglomeration indictor (i.e. wages, productivity, and rent) is regressed against consumption *and* production amenities while also controlling for mutual dependence.

SEM allows to model the mutual interactions among agglomeration benefit proxies so far independently analysed. In fact, unconvincing evidence has been provided on whether the wage-in-kind argument prevails, i.e. whether firms can offer lower wages based on complementing the wage offered with consumption amenities. In this view, while all else being equal workers are paid their marginal product, they still can be paid less if a firm can bargain on the in-kind amenities offered (such as the possibility for the worker to live in a safe and nice environment).

Identification requires two conditions to hold for the system of equations to be solvable:

- The *order condition*, positing that that for each equation the number of excluded exogenous variables is at least as large as included endogenous variables;
- The *rank condition*, both necessary and sufficient, requests that Π_{i0} is full column rank.

For this specific SEM application, we start with a neutral set of assumptions on which amenities matter for which type of places (i.e. places of residence or work). This implies including the same set of controls across all three specifications. To identify our estimates, thus, we resort to the same set of instruments used in prior individual reduced form equations, thereby controlling for historical population density, historical cinemas distribution, and soil composition.

Because geographies for the three indicators are different in the first place, we homogenize geographies by calculating municipality-level residuals for each reduced form specification. The latter can therefore be interpreted as the premium capacity of each urban area to benefit from internal characteristics.

Table 8 shows baseline OLS regressions for the three outputs (wages, productivity, and rent) in odd columns, and their equivalent obtained through SEM estimates in even ones. Parameter estimates can thus be compared, although estimates' precision improves in even columns, due to SEM controlling for mutual endogeneity.

Results show a positive impact of several consumption amenities (accessibility to transportation networks, restaurants, and natural amenities (here captured by proximity to the coastline). At the same time, substantial evidence is found for industrial diversity, while less compelling evidence is found for competition.

Moving to the mutual relations among productivity, wages, and rent, the path analysis graph in Figure 8 suggests that while positive and statistically significant evidence is found for the association between housing prices and wages on the one hand, and housing prices and productivity on the other hand, a mutual negative influence exists between wages and productivity, after controlling for local amenities. This last result resonates with arguments typical of the wage-in-kind assumption, positing that firms can bid lower wages once in-kind amenities are offered along with monetary remuneration.

9. Conclusions

Cities offer an inextricable bundle of positive and negative externalities, enhancing both consumption and production activities for consumers and firms alike. Most extant literature focuses on the latter, and in particular deals with the theoretical and empirical microfoundations of agglomeration economies, seen through the lens of workers and firms. However, the nature of cities is quickly shifting towards being hotspots of consumption; metro areas offer natural and artificial amenities, with a consequent higher quality of life. This is expected, all else being equal, to translate into higher willingness to pay for urban locations, but also higher firm productivity.

We contribute to the measurement of the intensity of the microfoundations of agglomeration economies from the consumption side, linking to the literature dealing with quality of life indices. Our work encompasses for the first time a comprehensive analysis of wages, productivity, and rent.

Model	(1.)	(2.)	(3.)	(4.)	(5.)	(6.)
Dep. variable	*Municipality wages*		*Municipality productivity*		*Municipality rent*	
Dist. from closest supermarket (res.)	0.00128	-0.00478	-411.2	-1310	0.0285***	0.0247**
	(0.45)	(-1.21)	(-0.85)	(-1.06)	(3.45)	(2.45)
Dist. from closest café (res.)	0.0106***	0.00925***	670.3**	1835.3**	0.0127**	-0.00812
	(4.68)	(3.18)	(2.17)	(2.10)	(1.98)	(-0.90)
Dist. from closest restaurant (res.)	0.00517	0.0240***	396.6	6348.1**	-0.103***	-0.0980***
	(1.50)	(3.58)	(0.68)	(2.08)	(-10.02)	(-8.15)
Distance to train station (res.)	0.000521*	0.000697*	-118.4***	104.7	-0.00287***	-0.00248**
	(1.76)	(1.85)	(-2.83)	(0.92)	(-3.59)	(-2.56)
Distance from closest cinema (res.)	-0.0000947	-0.000177	-0.405	-30.48	0.000345	0.00043
	(-0.38)	(-0.58)	(-0.01)	(-0.40)	(0.43)	(0.52)
Dist. from closest supermarket (work)	-0.0204***	-0.0197***	193.9	-2982.2**	0.012	0.0396***
	(-5.21)	(-4.46)	(0.24)	(-2.03)	(1.24)	(2.79)
Dist. from closest café (work)	0.00221	-0.00704	292.3	-2880.7*	0.0620***	0.0485***
	(0.81)	(-1.48)	(0.76)	(-1.65)	(7.96)	(5.20)
Dist. from closest restaurant (work)	0.00275	0.00248	-2476.8**	-1482.6	-0.0443***	-0.0272*
	(0.55)	(0.41)	(-2.52)	(-0.81)	(-3.45)	(-1.73)
Distance to train station (work)	-0.0000773	0.000256	58.65	24.94	-0.000381	-0.000572
	(-0.21)	(0.62)	(0.93)	(0.23)	(-0.43)	(-0.52)
Distance from closest cinema (work)	-0.000779***	-0.00049	152.7**	-28.95	0.0012	0.00125
	(-2.58)	(-1.36)	(2.56)	(-0.30)	(1.39)	(1.33)
Competition (work)	2.411***	1.558**	-362950.8***	-107751.8	-0.085	-1.365
	(4.38)	(2.28)	(-3.70)	(-0.28)	(-0.07)	(-0.81)
Diversity (work)	1.493*	1.63	353545.0**	556348.8	10.78***	4.780*
	(1.70)	(1.20)	(2.34)	(1.03)	(5.04)	(1.72)
Competition (res.)	0.361	-2.228**	54756.1	-600091.9	15.40***	12.25***
	(0.69)	(-2.39)	(0.79)	(-1.42)	(14.42)	(7.37)
Diversity (res.)	-1.731**	-2.783***	-468855.8***	-741204.5***	1.196	6.739***
	(-2.33)	(-3.47)	(-3.09)	(-3.36)	(0.64)	(2.74)

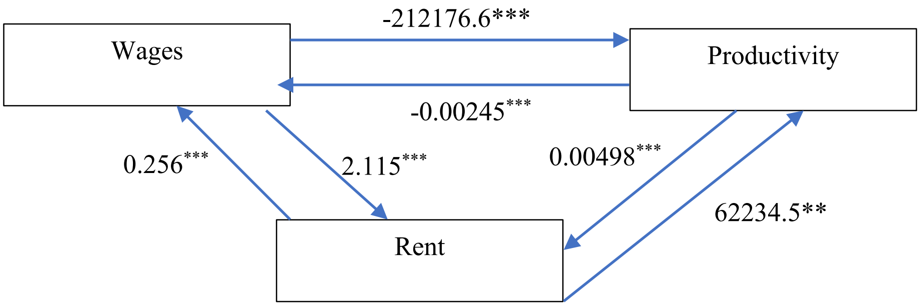
Log population density	-0.00286	-0.0272^{***}	261.3	-3578.1	0.111^{***}	0.0989^{***}
	(-0.77)	(-3.33)	(0.38)	(-1.15)	(11.37)	(8.36)
Distance from the closest coastline (res.)	-0.000124^{*}	0.000199^{*}	9.134	57.85	-0.00158^{***}	-0.00123^{***}
	(-1.82)	(1.87)	(1.40)	(1.23)	(-9.18)	(-5.83)
Distance from the closest coastline (work)	-0.000287^{***}	-0.000456^{***}	-46.21^{***}	-83.88^{**}	0.000674^{***}	0.00130^{***}
	(-4.02)	(-5.04)	(-5.64)	(-2.48)	(3.80)	(5.17)
Municipality productivity	0.000000259	-0.00000245^{***}	-	-	-0.000000902^{**}	0.00000498^{***}
	(1.64)	(-2.87)			(-2.05)	(3.00)
Municipality rent	0.105^{***}	0.256^{***}	-3040.8^{**}	62234.5^{**}	-	-
	(10.94)	(5.49)	(-2.09)	(2.27)		
Municipality wages	-	-	6357.2	-212176.6^{***}	0.763^{***}	2.115^{***}
			(1.63)	(-4.10)	(10.67)	(5.40)
Year=2008	-	0.0169^{***}	-	3043.7^{**}	-	-0.0234^{**}
		(6.36)		(2.31)		(-2.13)
Year=2011	-0.0207^{***}	-	363.3	-	-0.00621	-
	(-10.20)		(1.05)		(-1.09)	
Constant term	0.425	1.448	113392.8	265069.9	-12.63^{***}	-12.33^{***}
	(0.71)	(1.20)	(1.34)	(0.42)	(-7.47)	(-6.10)
Instruments	-	1830, 1920, and 1926 population density	-	Soil composition statistics†	-	Distance from 1910 cinemas in work and residence areas
Number of obs.	1,723	1,723	1,723	1,723	1,723	1,723
Adj. R^2	0.468	0.281	0.078	-2.15	0.659	0.528

Table 8: Structural Equations Model estimates

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† *Statistics include the median and standard deviation of type of soil class, letter, and code.*

Figure 8. Structural Equations Model path among Wages, Productivity, and Rent



Source: Authors' elaboration

The role of consumption amenities in driving agglomeration-related savings had been mostly neglected in prior studies. Our results suggest that consumption amenities actually play a rather relevant role in generating agglomeration economies. They explain roughly 22 per cent of the total density effect in a hedonic price regression; the estimated city size effect increases 3.7 as we move from a baseline to a general specification in firm productivity regressions; and cafés and restaurants combined together contribute roughly to 1.5 per cent of wage effects for every additional kilometer saved. All three results hint at a rather substantial role of consumption amenities in the explanation of the correct city size effect.

These findings are robust to a number of consistency checks, and to the use of geological instruments, historical population density, and historical spatial distribution of consumption amenities for instrumenting consumption- and production-related sources of agglomeration economies.

These results are particularly relevant for policy reasons. The prior focus of the agglomeration economies literature fails to account for the changing nature of cities, which are becoming increasingly attractive more for the consumption possibilities they offer, rather than for the productivity mark-up they grant to firms, especially if we focus on manufacturing firms, where tasks are more routinary, and the price mark-up of urban locations is thereby less easy to justify. This effect appears already quite substantial, and appears as potentially increasing in relevance over time. While we find only weak evidence of the increasing role of consumption amenities as drivers of agglomeration economies, this point deserves further research and will shape the policy debate for the years to come.

While urban land rent reflects bids from both consumers and firms, wages and productivity narrow down the scope of agglomeration economies to the production side of the economy. Thus, the fact that in both cases, wages and productivity reflect a positive role of consumption amenities, supports the wage-in-kind argument typical of the Rosen-Roback framework.

While the types of amenities consumers look for when relocating in cities not necessarily mirrors the needs of firms (Gabriel and Rosenthal, 2004), the comparison of our findings allows to more precisely pinpoint the extent of different sources of agglomerative effects on different markets.

References

Albouy, D. (2008). "Are big cities bad places to live? Estimating quality of life across metropolitan areas", NBER Working Paper 14472.

Albouy, D. (2016). "What are cities worth? Land rents, local productivity, and the total value of amenities", *Review of Economics and Statistics*, 98 (3): 477–487.

Alonso, W. (1973). "Urban zero population growth", *Daedalus*, 102: 191–206.

Audretsch, D. B., and Feldman, M. P. (1996). "R&D spillovers and the geography of innovation and production", *The American Economic Review*, 86 (3): 630–640.

Bartik, T. J., and Smith, V. K. (1987). "Urban amenities and public policy", in E. S. Mills (ed), "*Handbook of regional and urban economics"*, Vol. 2, Amsterdam (North Holland): Elsevier, pp. 1207–1254.

Baum-Snow, N., and Ferreira, F. (2015). "Causal inference in Urban and Regional Economics", in G. Duranton, J. Vernon Henderson and W. C. Strange (eds), "*Handbook of Regional and Urban Economics*", 5th Vol., Amsterdam (NL): Elsevier, pp. 3–68.

Beaudry, C., and Schiffauerova, A. (2009). "Who's right, Marshall or Jacobs? The localization versus urbanization debate", *Research Policy*, 38 (2): 318–337.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). "Productive cities: Sorting, selection, and agglomeration", *Journal of Political Economy*, 122 (3): 507–553.

Blomquist, G. C., Berger, M. C., and Hoehn, J. P. (1988). "New estimates of quality of life in urban areas", *The American Economic Review*, 78 (1): 89-107.

Bureau van Dijk, (2019). "*Orbis - Detailed global private company information*", retrieved online on Oct. 18, 2019 at the URL https://www.bvdinfo.com/en-apac/our-products/company-information/international-products/orbis

Burger, M. J., Meijers, E.J., Hoogerbrugge, M.M., and Masip Tresserra, J. (2015). "Borrowed size, agglomeration shadows and cultural amenities in North-West Europe", *European Planning Studies*, 23 (6): 1090–1109.

Camagni, R., Capello, R., and Caragliu, A. (2016). "Static vs. dynamic agglomeration economies. Spatial context and structural evolution behind urban growth", *Papers in Regional Science*, 95 (1): 133–158.

Caragliu, A., de Dominicis, L., and de Groot, H. L. (2016). "Both Marshall and Jacobs were right!", *Economic Geography*, 92 (1): 87–111.

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

Combes, P. P., Duranton, G., Gobillon, L., and Roux, S. (2010a). "Estimating agglomeration economies with history, geology, and worker effects", in E. L. Glaeser (ed), "*Agglomeration economics*", Chicago (IL): University of Chicago Press, pp. 15–66.

Combes, P. P., Duranton, G., and Gobillon, L. (2010b). "The identification of agglomeration economies", *Journal of Economic Geography*, 11 (2): 253–266.

Combes, P. P., and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics* (Vol. 5, pp. 247-348). Elsevier.

Debord, G. (2008). "Introduction to a critique of urban geography", in H. Bauder and S. Engel-Di Mauro (eds), "*Critical geographies: A collection of readings*", Kelowna (BC): Praxis (e) Press, pp. 23–27.

Dalmazzo, A., and de Blasio, G. (2011). "Amenities and skill-biased agglomeration effects: Some results on Italian cities", *Papers in Regional Science*, 90 (3): 503–527.

de Groot, H. L.F., Marlet, G., Teulings, C., & Vermeulen, W. (2015). "Cities and the urban land premium", Cheltenham (UK): Edward Elgar.

de Groot, H. L. F., Poot, J., and Smit, M. J. (2016). "Which agglomeration externalities matter most and why?", *Journal of Economic Surveys*, 30 (4): 756–782.

De Vries, J., and van der Woude, A. (1997). "*The first modern economy: Success, failure, and perseverance of the Dutch economy, 1500–1815*", Cambridge (MA): Cambridge University Press.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3), 479-524.

Dickinson, R. E. (1951). "*The West European city: a geographical interpretation*", London (UK) : Routledge & Paul.

Diodato, D., Neffke, F., and O'Clery, N. (2018). "Why do industries coagglomerate? How Marshallian externalities differ by industry and have evolved over time", *Journal of Urban Economics*, 106, 1-26.

Dumais, G., Ellison, G., and Glaeser, E. L. (2002). "Geographic concentration as a dynamic process", *Review of Economics and Statistics*, 84 (2): 193–204.

Duranton, G., and Puga, D. (2004). "Micro-Foundations of Urban Agglomeration Economies", in J. Vernon Henderson and Jacques-François Thisse (eds), "*Handbook of Regional and Urban Economics*", Vol. 4: pp. 2063-2117. Amsterdam (NL): Elsevier.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, 100(3), 1195-1213.

Faggio, G., Silva, O., and Strange, W. C. (2017). "Heterogeneous agglomeration", *Review of Economics and Statistics*, 99 (1): 80–94.

Frenken, K., Van Oort, F., and Verburg, T. (2007). "Related variety, unrelated variety and regional economic growth", *Regional Studies*, 41 (5): 685–697.

Gabriel, S. A., and Rosenthal, S. S. (2004). "Quality of the business environment versus quality of life: do firms and households like the same cities?", *The Review of Economics and Statistics*, 86 (1): 438–444.

Glaeser, E. L. (2007). "The economics approach to cities", National Bureau of Economic Research WP. w13696, retrieved online on Jul. 10, 2018 at the URL http://www.nber.org/papers/w13696.

Glaeser, E. L., and Gottlieb, J. D. (2006). "Urban resurgence and the consumer city", *Urban Studies*, 43 (8): 1275–1299.

Glaeser, E. L., and Gottlieb, J. D. (2009). "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States", *Journal of Economic Literature*, 47 (4): 983–1028.

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

Glaeser, E. L., Kolko, J. and Saiz, A. (2001). "Consumer city", *Journal of Economic Geography*, 1 (1): 27-50.

Groot, S. P. T., de Groot, H. L. F., & Smit, M. J. (2014). "Regional wage differences in the Netherlands: Micro-evidence on agglo7 meration externalities", *Journal of Regional Science*, 54 (3): 503–523.

Henderson, V., Kuncoro, A., and Turner, M. (1995). "Industrial development in cities", *Journal of Political Economy*, 103 (5): 1067–1090.

Hoover, E. M. (1937). "*Location theory and the shoe and leather industry*", Cambridge (MA): Harvard University Press.

Kanemoto, Y., Ohkawara, T., and Suzuki, T. (1996). "Agglomeration economies and a test for optimal city sizes in Japan", *Journal of the Japanese and International Economies*, 10 (4): 379–398.

Koster, H. R. A., and Jan Rouwendal. "Historic Amenities and Housing Externalities: Evidence from the Netherlands", *The Economic Journal* 127, 605: F396–420.

Koster, H. R. A., Pasidis, I., and van Ommeren, J. (2019). Shopping externalities and retail concentration: Evidence from Dutch shopping streets. *Journal of Urban Economics*, 114, 103194.

Krugman, P. (1991). "Increasing returns and economic geography", *Journal of Political Economy*, 99(3), 483-499.

Krugman, P. (1993). "First nature, second nature, and metropolitan location", *Journal of Regional Science*, 33(2), 129-144.

Levinsohn, J., and Petrin, A. (2003). "Estimating production functions using inputs to control for unobservables", *The Review of Economic Studies*, 70 (2): 317–341.

Liverani, M., Zainab, B., and Van de Mieroop, M. (2006). "*Uruk: The first city*", London (UK): Equinox Publishing.

Marshall, A. (1920). "*Principles of Economics*", London (UK): Macmillan.

Melo, P. C., Graham, D. J., and Noland, R. B. (2009). "A meta-analysis of estimates of urban agglomeration economies", *Regional Science and Urban Economics*, 39 (3): 332–342.

Miranda, T., Williams-Rossi, D., Johnson, K. A., and McKenzie, N. (2011). "Reluctant readers in middle school: Successful engagement with text using the e-reader", *International Journal of Applied Science and Technology*, 1 (6): 81–91.

Öner, Ö. (2017). "Retail city: the relationship between place attractiveness and accessibility to shops", *Spatial Economic Analysis*, 12 (1): 72–91.

Overman, H. G., and Puga, D. (2010). "Labor pooling as a source of agglomeration: An empirical investigation", in E. L. Glaeser (ed), "*Agglomeration economics*", Chicago (IL): University of Chicago Press, pp. 133–150.

Palan, N. (2010). "Measurement of specialization the choice of indices", FIW working paper No. 62. Retrieved online on May 15, 2019 at the URL https://www.econstor.eu/bitstream/10419/121065/1/N_062.pdf

Poorthuis, A., & van Meeteren, M. (2019). Containment and Connectivity in Dutch Urban Systems: A Network-Analytical Operationalisation of the Three-Systems Model. *Tijdschrift Voor Economische En Sociale Geografie*, Early View

Pouyat, R. V., Yesilonis, I. D., Dombos, M., Szlavecz, K., Setälä, H., Cilliers, S., and Yarwood, S. (2015). "A global comparison of surface soil characteristics across five cities: a test of the urban ecosystem convergence hypothesis", *Soil Science*, 180(4/5), 136-145.

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

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York (NY): Simon and Schuster.

Rivera-Batiz, F. L. (1988). "Increasing returns, monopolistic competition, and agglomeration economies in consumption and production", *Regional Science and Urban Economics*, 18 (1): 125–153.

Roback, J. (1982). "Wages, rents, and the quality of life", *Journal of Political Economy*, 90 (6): 1257–1278.

Roca, J. D. L., and Puga, D. (2017). "Learning by working in big cities", *The Review of Economic Studies*, 84 (1): 106-142.

Rosen, S. (1979). "Wage-based indexes of urban quality of life", in P. Mieszkowski and M. Straszheim (eds), "*Current Issues in Urban Economics*", Baltimore (MD): Johns Hopkins University Press.

Rosenthal, S. S., and Strange, W. C. (2001). "The determinants of agglomeration", *Journal of Urban Economics*, 50 (2): 191–229.

Rosenthal, S.S., and Strange, W. (2004): "Evidence on the Nature and Sources of Agglomeration Economies", in J. V. Henderson and J.-F. Thisse (eds), "*Handbook of Regional and Urban Economics, Vol.4*", Amsterdam: North-Holland, pp. 2119–2171.

Tabuchi, T., and Yoshida, A. (2000). "Separating urban agglomeration economies in consumption and production", *Journal of Urban Economics*, 48 (1): 70–84.

United Nations Industrial Development Organization (2015). "*Industrial Development Report 2016. The Role of Technology and Innovation in Inclusive and Sustainable Industrial Development*", Vienna (AT): UNIDO.

van Oort, F., de Geus, S., and Dogaru, T. (2015). "Related Variety and Regional Economic Growth in a Cross-Section of European Urban Regions", *European Planning Studies*, 23 (6): 1110–1127.

Vries, F. de, de Groot, W.J.M. , Hoogland, T., and Denneboom, J. (2003). "*De bodemkaart van Nederland digital*", retrieved online on Jul. 13, 2018 at the URL http://www2.alterra.wur.nl/Webdocs/PDFFiles/Alterrarapporten/AlterraRapport811.pdf

Vries, Jan de, and Ad van der Woude. *The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy*, 1500-1815. Cambridge University Press, 1997.

Weber, M. (1969). "The nature of the city", in R. Sennet (ed), "*Classical essays on the culture of cities*", Englewood Cliffs (NJ): Prentice Hall.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambirdge (MA): MIT press.

Appendices

Appendix A1: soil composition classes

Subgroup code	Frequency	Percent	Cum.
AAK	7	0.01	0.01
AAP	25	0.04	0.05
ABk	53	0.08	0.13
ABl	5	0.01	0.13
ABv	161	0.24	0.37
ABz	23	0.03	0.4
AEk	3	0.00	0.41
AEm	23	0.03	0.44
AEp	42	0.06	0.5
AFk	13	0.02	0.52
AFz	11	0.02	0.54
AHa	10	0.01	0.55
AHc	26	0.04	0.59
AHk	32	0.05	0.64
AHl	65	0.10	0.74
AHs	9	0.01	0.75
AHt	2	0.00	0.75
AHv	1	0.00	0.75
AHz	4	0.01	0.76
AK	26	0.04	0.8
ALu	26	0.04	0.84
AM	10	0.01	0.85
AMm	9	0.01	0.86
AO	121	0.18	1.04
AP	212	0.31	1.35
AQ	4	0.01	1.36
AR	1	0.00	1.36
AS	48	0.07	1.43
AVk	13	0.02	1.45
AVo	211	0.31	1.76
AWg	30	0.04	1.81
AWo	12	0.02	1.82
AWv	2	0.00	1.83
AZ	1	0.00	1.83
AZW	85	0.13	1.95
BKd	38	0.06	2.01
BKh	14	0.02	2.03
BLb	192	0.28	2.31
BLd	167	0.25	2.56
BLh	8	0.01	2.57
BLn	13	0.02	2.59
BZd	8	0.01	2.6
EK	97	0.14	2.75
EL	2	0.00	2.75
EZ	33	0.05	2.8
EZg	194	0.29	3.08
FG	64	0.09	3.18
FK	12	0.02	3.2
Hd	1,185	1.75	4.94
Hn	10,948	16.14	21.08
KK	25	0.04	21.11
KM	2	0.00	21.12
KRd	180	0.27	21.38
KRn	355	0.52	21.91
KS	15	0.02	21.93
KT	43	0.06	21.99
KX	216	0.32	22.31
Ld	121	0.18	22.49
Ldd	81	0.12	22.61
Ldh	88	0.13	22.74
Lh	12	0.02	22.75
Ln	21	0.03	22.79
Lnd	40	0.06	22.84
Lnh	1	0.00	22.85
MA	7	0.01	22.86
MK	4	0.01	22.86
MOb	184	0.27	23.13
MOo	124	0.18	23.32
MZk	3	0.00	23.32
MZz	22	0.03	23.35
Mn	6,269	9.24	32.59
Mo	402	0.59	33.18
Mv	617	0.91	34.09
ROb	6	0.01	34.1
Rd	880	1.30	35.4
Rn	2,335	3.44	38.84
Ro	54	0.08	38.92
Rv	218	0.32	39.24
Sn	121	0.18	39.42
U0102	25	0.04	39.46
U01W	38	0.06	39.51
U02O	15	0.02	39.54
U04T	11	0.02	39.55
U06T	29	0.04	39.59
U0708	14	0.02	39.62
U07W	35	0.05	39.67
U09T	31	0.05	39.71
U10T	4	0.01	39.72

U11O	62	0.09	39.81	U5455	60	0.09	44.77
U11W	15	0.02	39.83	U5657	30	0.04	44.81
U12O	3	0.00	39.84	U5758	101	0.15	44.96
U12OR	93	0.14	39.97	U58O	10	0.01	44.98
U12W	22	0.03	40.01	U5960	16	0.02	45
U1318	29	0.04	40.05	U6162	5	0.01	45.01
U1419	132	0.19	40.24	V	946	1.39	46.4
U15T	51	0.08	40.32	Vo	37	0.05	46.46
U1617	126	0.19	40.5	Vz	1	0.00	46.46
U16T	4	0.01	40.51	Wg	22	0.03	46.49
U17OR	78	0.11	40.62	Wo	215	0.32	46.81
U1920	209	0.31	40.93	Y	563	0.83	47.64
U2021	47	0.07	41	Zb	469	0.69	48.33
U21O	26	0.04	41.04	Zd	759	1.12	49.45
U22T	151	0.22	41.26	Zn	2,034	3.00	52.45
U2425	58	0.09	41.35	aV	1,644	2.42	54.87
U25O	21	0.03	41.38	bEZ	1,023	1.51	56.38
U2627	46	0.07	41.45	bRn	21	0.03	56.41
U2632	32	0.05	41.49	cHd	126	0.19	56.59
U27OR	105	0.15	41.65	cHn	3,144	4.63	61.23
U2829	28	0.04	41.69	cY	296	0.44	61.66
U28W	21	0.03	41.72	cZd	48	0.07	61.73
U30T	63	0.09	41.81	gMn	1,123	1.66	63.39
U31O	118	0.17	41.99	hEV	8	0.01	63.4
U31W	61	0.09	42.08	hV	962	1.42	64.82
U32OR	240	0.35	42.43	iV	1,908	2.81	67.63
U33T	45	0.07	42.5	iWp	645	0.95	68.58
U3435	83	0.12	42.62	iWz	213	0.31	68.89
U36O	21	0.03	42.65	kMn	376	0.55	69.45
U37O	176	0.26	42.91	kV	650	0.96	70.41
U37W	139	0.20	43.12	kVz	1	0.00	70.41
U38O	3	0.00	43.12	kWp	198	0.29	70.7
U38W	54	0.08	43.2	kWz	119	0.18	70.88
U39T	23	0.03	43.23	pKRn	32	0.05	70.92
U40T	28	0.04	43.27	pLn	34	0.05	70.97
U41T	24	0.04	43.31	pMn	631	0.93	71.9
U4248	100	0.15	43.46	pMo	220	0.32	72.23
U43O	49	0.07	43.53	pMv	75	0.11	72.34
U43W	45	0.07	43.6	pRn	107	0.16	72.5
U44O	37	0.05	43.65	pRv	32	0.05	72.54
U44W	59	0.09	43.74	pV	512	0.75	73.3
U4546	74	0.11	43.85	pZg	2,861	4.22	77.51
U45WR	32	0.05	43.89	pZn	1,705	2.51	80.03
U4849	205	0.30	44.2	tZd	16	0.02	80.05
U49O	11	0.02	44.21	uWz	3	0.00	80.05
U5051	43	0.06	44.28	vWp	692	1.02	81.07
U50WR	60	0.09	44.36	vWz	751	1.11	82.18
U51O	27	0.04	44.4	zEZ	4,166	6.14	88.32
U52O	20	0.03	44.43	zV	873	1.29	89.61
U52W	49	0.07	44.51	zWp	974	1.44	91.04
U52WR	8	0.01	44.52	zWz	352	0.52	91.56
U5354	111	0.16	44.68	aGR	89	0.13	91.69

bAF	159	0.23	91.93
cOP	496	0.73	92.66
dEG	5	0.01	92.67
eVE	23	0.03	92.7
fTE	999	1.47	94.17
gMO	240	0.35	94.53
gWA	2,070	3.05	97.58
hBE	1,360	2.00	99.58
hDI	247	0.36	99.95
iBO	28	0.04	99.99
MY	9	0.01	100

Table A1. Cumulative distribution function of soil composition subgroups

Source of raw data: De Vries (2003), authors' elaboration

Appendix A2: descriptives for NVM dataset

Province	2000	2005	2010	2015
Groningen	4,283	5,222	4,201	5,443
Friesland	3,976	5,161	3,584	5,157
Drenthe	4,201	5,013	3,394	4,803
Overijssel	6,918	8,535	5,712	8,575
Flevoland	3,221	3,529	2,025	3,408
Gelderland	14,066	17,339	11,868	17,587
Utrecht	10,63	13,125	9,323	14,281
Noord-Holland	20,523	24,781	19,064	30,287
Zuid-Holland	23,543	30,989	21,293	28,828
Zeeland	1,489	2,057	1,375	2,083
Noord-Brabant	14,418	18,623	12,217	19,405
Limburg	2,586	3,026	2,547	4,626

Table A2: Number of housing transactions in the dataset by province, for four years – the dataset as a whole covers all years from 2000 to 2015, inclusive.

Construction period	Observations	Mean price	Mean house size (m²)	Mean lot size (m²)
After 2001	93,583	€ 326.638	138	1001
1991-2000	217,088	€ 260.316	117	1167
1981-1990	147,87	€ 247.953	119	652
1971-1980	140,677	€ 216.921	107	2804
1960-1970	286,627	€ 192.298	107	1911
1945-1959	305,881	€ 216.254	125	2425
1931-1944	259,119	€ 216.170	116	317
1906-1930	279,383	€ 267.403	130	1110
Before 1906	137,746	€ 290.961	130	1981
Total sample	1867974	€ 238.936	120	1495
s.d. for total sample		*€ 151.663*	*55*	*459211*

Table A3. Number of housing transactions in the dataset by construction period, with average price, house size, and lot size.

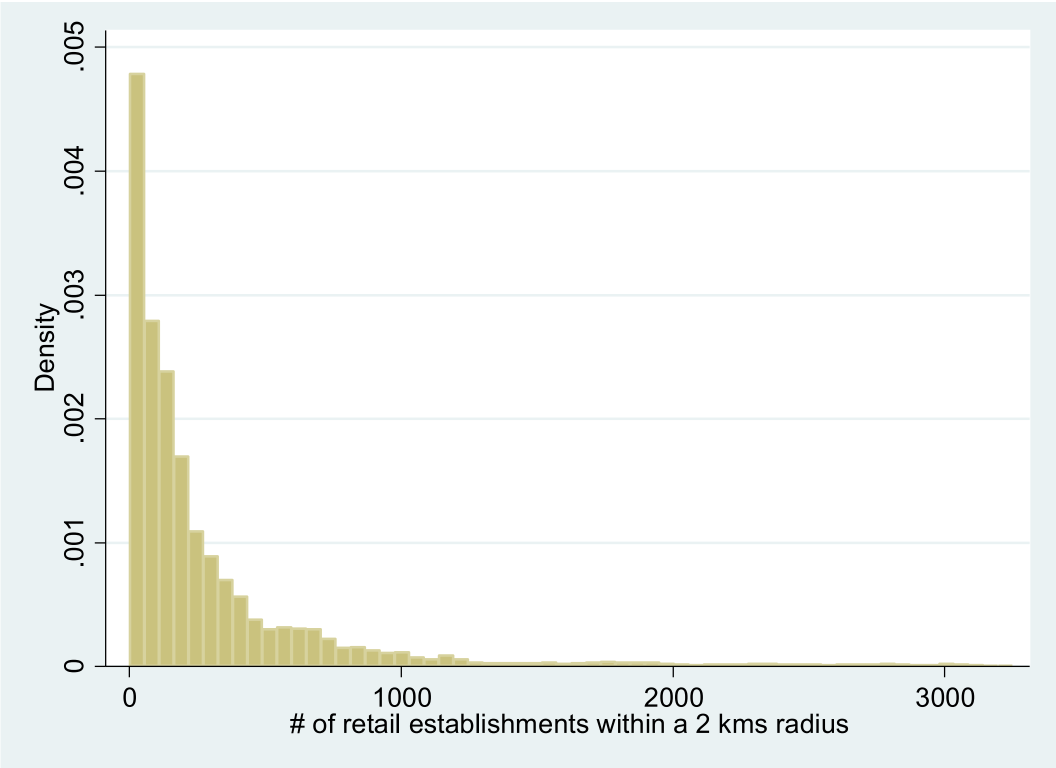
Maintenance	non-monuments	monuments
Good	88%	84%
Moderate	10%	12%
Poor	1%	3%
total number	**1,854,522**	**13,452**

Table A4. Maintenance state of houses in the dataset, cross-tabulated with the monument status of the house.

Appendix A3: Identifying Central Business Districts from LISA clusters

For each set of coordinates, we calculated the number of retail establishments, to possibly take into account multiple stores, malls, etc. Next, or each set of coordinates we then calculated the number of retail establishments within a 2 kms radius. Figure A1 shows the distribution density histogram for this count.

Figure A1. Count of retail establishments within a 2 kms radius



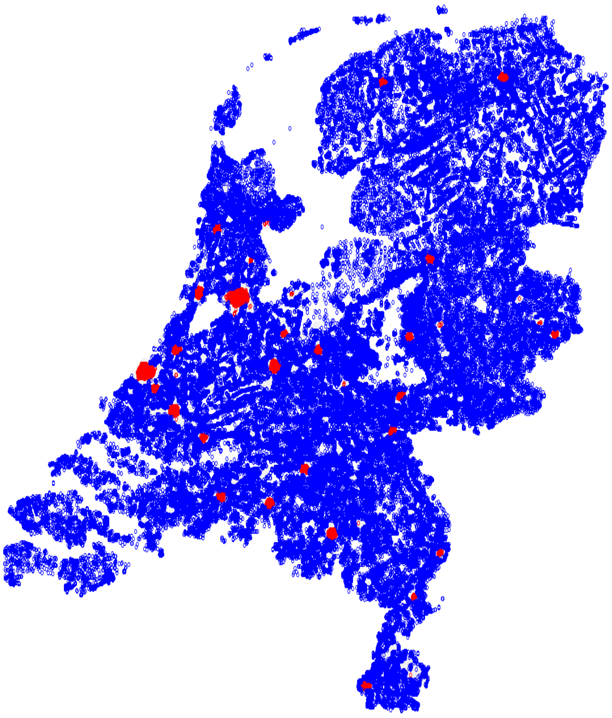
Source: Authors' elaboration

We converted the data set as a csv and shape file and produced LISA maps to endogenize the definition of CBDs in the data.

Figure A2 shows the queen contiguity matrix-based LISA map (areas share corners, not borders). Of this map clusters have been saved, thereby including high density-high density clusters that represent our notion of CBD. Across all maps, blue areas indicate clusters of low density close to low density. Red dots indicate the opposite case: high density next to high density.

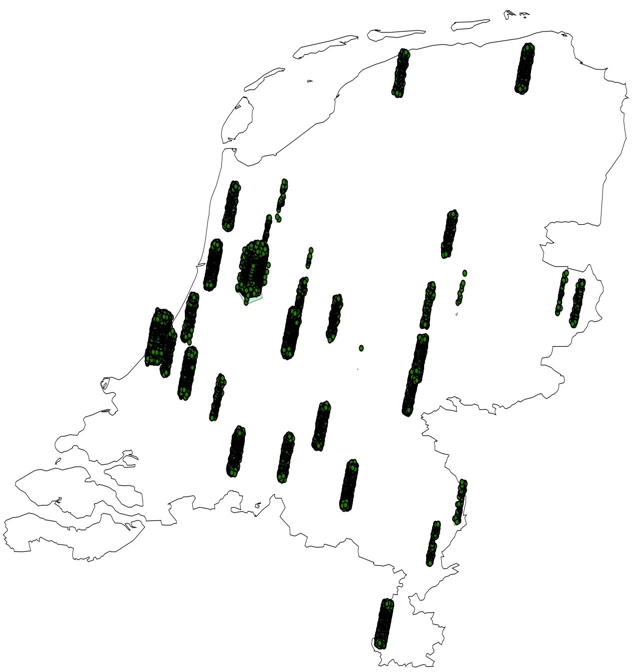
The latter have been then again used as shape file to identify clusters of high-high density (Figure A3). Finally, to create the shapes of the CBDs for which distance within each housing unit is going to be calculated we built convex hulls around each cluster of observations. The resulting shapes are shown in Figure A4.

Figure A2. LISA cluster map 1: queen contiguity



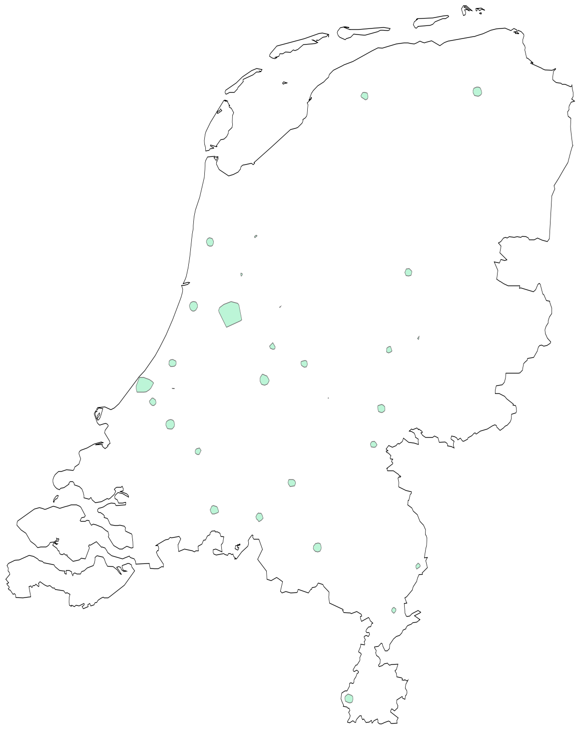
Source: Authors' elaboration

Figure A3. High-high clusters of retail establishments density according to the queen contiguity criterion



Source: Authors' elaboration

Figure A4. Central Business Districts as identified on the basis of LISA clusters and queen contiguity matrix



Source: Authors' elaboration

Appendix A4: Levinsohn and Petrin's productivity estimates

Table A5 shows the results of the first-stage estimates of the determinants of firm-level productivity determinants. As dictated in Levinsohn and Petrin (2003), capital stock is instrumented with the costs of intermediate inputs. Results in particular suggest the rejection of the null hypothesis of constant returns to scale in production.

Dependent variable: firm value added		
	Parameter	*Standard error*
Log capital stock	0.19***	(7.94))
Log employees	0.26*	(1.88)
Wald test of constant returns to scale: χ^2 = 29.32*** (p = 0.0000).		

Table A5. Levinsohn and Petrin's estimates of firm level productivity determinants

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$