Department of Economics and Statistics
Middlesex University Business School

Discussion Paper

No 145

September 2011

Is there a limit to agglomeration? Evidence from productivity of Dutch firms

Marian Rizov, Arie Oskam and Paul Walsh

Is there a limit to agglomeration? Evidence from productivity of Dutch firms

Marian Rizov[1,2], Arie Oskam[2] and Paul Walsh[3] *
[1] Middlesex University Business School and SPU Nitra
[2] Wageningen University
[3] University College Dublin

Abstract

We compute aggregate productivity of three categories of regions, classified by level of urbanization and density of economic activity in the Netherlands, from firm-specific total factor productivity (TFP) measures. TFP measures are estimated by a semi-parametric algorithm, within 2-digit industries, covering agriculture, manufacturing, construction, trade and services, using AMADEUS data over the period 1997-2006. We analyse the productivity differentials across urbanization categories by decomposing them into industry productivity effect and industry composition effect. Our analysis indicates that there is non-linear net effect of agglomeration on productivity growth but in levels agglomeration is associated with higher productivity.

Key words: Agglomeration, factor prices, total factor productivity, structural estimation, The Netherlands

JEL classification: D24, R11, R30

Addresses:
M. Rizov
Department of Economics and Statistics
Middlesex University Business School
London NW4 4BT
E-mail: m.rizov@mdx.ac.uk

* We thank Eric Bartelsman for ideas and useful discussion at an early stage of this research and Geerte Cotteleer for data and discussion. The paper was presented and benefited from discussions at the 2011 ERSA congress in Barcelona. The financial support of the Mansholt Graduate School of Social Sciences is acknowledged. The usual disclaimer applies.

Is there a limit to agglomeration? Evidence from productivity of Dutch firms

1 Introduction

Agglomeration and thus, the geographic concentration of economic activity in urbanized regions can result in a snowball effect, where firms tend to agglomerate to benefit from higher diversity and specialization in production processes. There are also benefits to firms from co-locating in close proximity to other firms in the same industry (Marshall, 1920; Henderson, 1974; 2003). Both urbanization and localization economies can be considered centripetal (agglomeration) forces leading to concentration of economic activity. Theoretical models (e.g., Glaeser et al., 1992; Ciccone and Hall, 1996) and empirical studies (e.g., Carlino and Voith, 1992; Ciccone, 2002; Combes et al., 2009; Graham, 2009; Combes et al., 2010) show that agglomeration associated with high density of economic activity positively affects productivity.

Agglomeration can affect productivity in several ways. If technologies have constant returns themselves, but the transportation of products from one stage of production to the next involves costs that rise with distance, then the technology for the production of all goods within a particular geographical area will have increasing returns - the ratio of output to input will rise with density. If there are positive externalities associated with the physical proximity of production, then density will contribute to productivity for this reason as well. A third source of density effects is the higher degree of beneficial specialization possible in areas of dense economic activity.

A second branch of the literature on agglomeration hypothesises economies of scale internal to firms (e.g., Fujita, 1988; Hanson, 1996; Davis and Weinstein, 2008). Models with internal increasing returns build on theories of the firm and its market and commonly employ the formalisation of monopolistic competition suggested by Spence (1976) and Dixit and

Stiglitz (1977) to demonstrate that non-transportable intermediate inputs produced with increasing returns imply agglomeration. In a related model, Krugman (1991) demonstrates that agglomeration will result even when transportation costs are small, if most workers are mobile. The essence of all these models is that when local markets are more active, a larger number of producers of the differentiated intermediate inputs break even and the production of final goods is more efficient when a greater variety of intermediate inputs is available.

However, Henderson (1974) building on work by Mills (1967) demonstrates that, in an equilibrium, disamenities from agglomeration and high density may offset the productivity advantages thus acting as centrifugal forces.[1] For example, these include increased costs resulting from higher wages driven by competition among firms for skilled labour, higher rents due to increased demand for housing and commercial land, and negative externalities such as congestion. Recent studies (e.g., Rappaport, 2008; Broersma and Oosterhaven, 2009) confirm that there are limits to agglomeration and point to a negative effect of congestion (crowdedness) on productivity growth.[2] Furthermore, evidence suggests that increases in estimated productivity are insufficient to sustain the high levels of crowdedness in heavily urbanized areas (Rappaport, 2008).

In this paper we study the net impact of agglomeration (accounting both for agglomeration and congestion) on total factor productivity (TFP) using Dutch firm level data for the 1997-2006 period. The Netherlands is particularly suitable for studying agglomeration-congestion effects given the fact that the country is one of the most urbanized and densely populated in the world but it clearly exhibits diversity in the degree of

[1] Alonso-Villar (2008) using features of Forslid and Ottaviano's (2003) framework analytically shows that when considering the effects of congestion costs, the dispersion of economic activity is possible not only at high, but also at low transport costs which suggests limits to agglomeration.

[2] In somewhat different but related context Saito and Gopinath (2009) and Combes et al. (2009) study the impact of firm self-selection and agglomeration on regional or city productivity. The first paper finds that firm's self-selection outweighs the contribution of agglomeration economies in increasing a region's productivity level in Chile while the second paper finds the opposite for the case of French cities.

urbanization. Three main categories of regions can be distinguished, according to their level of urbanization and population density, characterised by increasing aglomeration and congestion effects. Analysing the effects Broersma and Oosterhaven (2009) find non-linear net impact of agglomeration on labour productivity growth in the Dutch regions. We extend their analysis by taking a production function approach as in Henderson (2003) and related studies to directly account for the net efect of agglomeration on productivity. We contribute to the literature by applying an advanced TFP estimation technique following modelling ideas in Olley and Pakes (1996) and Ackerberg et al. (2007).[3] We explicitly model unobservable productivity utilising unique disaggregated land price and wage data and incorporate directly the effects of these and other location characteristics into the structural estimation algorithm. Our results add robust empirical evidence to the small but growing literature on the limits of agglomeration. In line with Mitra's (1999), Rappaport's (2008) and Broersma and Oosterhaven's (2009) results we find a non-linear relationship between net agglomeration and productivity growth.

The paper is organised as follows. In Section 2 we characterize the three urbanization categories used in the analysis and introduce a simple economic geography model to motivate the link between agglomeration, land prices and productivity. Next, in Section 3, we describe our econometric framework and develope the model of unobservable productivity. In Section 4 we describe the AMADEUS data used in our empirical analysis and report results from estimating production functions. In Section 5 we analyse aggregate productivity in levels and changes by the means of decompositions. We also analyse samples of manufacturing and

[3] Previous studies attempting to link agglomeration and productivity apply a two-stage analysis. In the first stage authors estimate firm productivity, and in a second stage they proceed to link productivity to agglomeration measures. In our view testing for a relationship between agglomeration and (unobservable) productivity, *ex post*, is admitting that there is omitted information that should have been used while estimating the production function in the first instance.

service industries separately as agglomeration effect are likely to differ across industies. Section 6 concludes.

2 Agglomeration effects in the Dutch regions and theoretical considerations

The territory of the Netherlands is subdivided into 40 COROP (Coördinatie Commissie Regionaal Onderzoeks Programma) regions, based on functional regionalization principles, which form the NUTS3 (Nomenclature of Units for Territorial Statistics) level EU classification. For the analysis of regional differentiation, a functional typology based on degree of urbanization and population density is used by the CBS (Het Centraal Bureau voor de Statistiek) and other government departments. According to the typology the 40 COROP regions are divided into three categories: less urbanized, urbanized and highly urbanized. Given that the meaning of the concept of rural economy is largely a misnomer in the Netherlands, the typology based on degree of urbanization is quite appropriate for the analysis of the socio-economic developments in the Dutch regions.[4]

A comparative analysis of main characteristics of the three urbanization categories, for the 2002-2003 period, summarised in Table 1 reveals that employment growth, usually associated with productivity growth, in all three categories was positive as the growth rates were the highest in the less urbanized regions.[5] Population growth in less urbanized regions also exceeded that of the highly urbanized regions. However, economic growth in the less urbanized regions was of about 1 percent annually which was lower compared to the growth

[4] There is also a separate geographical classification of the Randstad urban zone (Amsterdam, Rotterdam, The Hague and Utrecht) which has resulted in fourty subregions: large and medium-sized cities and designated growth towns are treated as separate units, whereas for each of the functional urban regions in the Randstad all other municipalities are aggregated into one subregion.

[5] We acknowledge that the analysis at aggregate urbanization categories may mask differences at more disaggregated level such as municipalities (gemeenten). It is important to point out, however, that Terluin et al. (2005) who focussed on a number of selected municipalities did not find any substantial differences in socio-economic indicators from the national average, employment growth being the exception.

in the other two urbanization categories. Age distribution was quite similar in all categories. Unemployment rates in the less urbanized regions were slightly higher compared to other urbanization categories while disposable income per capita was below that in the urbanized and highly urbanized regions.

- Table 1 about here -

The comparative analysis based on summary statistics shows that socio-economic differences across the three urbanization categories were relatively small and employment growth, population density, and land prices seem to be the main characteristics of difference.[6] Therefore, next, we focus on the relashinship between population density and land prices reflecting the strength of agglomeration and find a nonlinear relationship which is depicted graphycally in Figure 1. This is an important first evidence for presence of congestion and other negative externalities from agglomeration in the Netherlands. The finding is consistent with results of Broersma and Oosterhaven (2009) who find negative impact of extreme agglomeration on labour productivity growth in the Dutch regions.

- Figure 1 about here -

Next, to understand better agglomeration-congestion effects in the Netherland we employ a simple economic geography model that casts light on above facts. The model is based on neoclassical trade theory and assumes equality of output prices for each industry across all regions. Individual firms in each industry have constant returns to scale and make zero profits, so the equality of output price to unit cost holds for all regions. Furthermore, the weak form of factor price invariance with respect to endowments implies that the number of industries operating in each region should be at least as great as the number of inputs with region-specific prices (Leamer, 1984). The solution to the system of equations for price

[6] The data source for land prices is the Cadastral Land Sales Database (CLSD) that contains information on land transactions, transaction prices, and the location of each parcel sold in the Netherlands. From the CLSD we obtained the transaction prices at 4-digit postcode level per hectare in 2003.

equality to unit cost within industries leads to the result that relative factor productivities are exactly equal to relative factor prices across regions (e.g., Rice et al., 2006). This result places the non-linear relationship between land prices and population density presented in Figure 1 in the context of productivity differences across the Dutch regions.

Another important implication of the model result is that although the spatial variation in factor prices is determined entirely from the production side of the economy the model is consistent with perfect mobility of some factors such as labour across regions - an important feature of the Dutch labour market. If there is perfect labour mobility, then any spatial differences in wages and in other considerations (such as amenity or disamenity) of agglomeration will be fully shifted into the prices of immobile factors in each area - land and housing (Voith, 1991; Adsera, 2000; Drennan and Kelly, 2011).[7] Variation in the degree to which factor mobility is possible is entirely consistent with the model as well. If labour is immobile then we would still observe the same wage differences, although land prices would not necessarily have adjusted to give real wage equalisation across regions. In the case of the Netherlands it is justified to assume a high degree of labour mobility and thus that land prices almost fully internalise spatial differences in agglomeration and productivity.

Even though, in general, the model offers no predictions about the structure of production in each region it is consistent with the different degrees of factor mobility and hence different factor stocks in each region.[8] Furthermore, the assumptions of the model that productivity levels are region specific, but not specific either to industries or factors, give the

[7] In competitive markets labor is paid the value of its marginal product. However, even if labor markets are not perfectly competitive, higher wages in dense urban areas can be seen as evidence of higher productivity. For workers, higher wages may be offset by larger commuting and housing costs. However, higher wages and land rents in areas with high density of economic activity would lead firms to relocate elsewhere unless there were some significant productive advantages (Roback, 1982; Combes et al., 2010; Puga, 2010).

[8] The model is consistent with the predictions of alternative theories with regard to regional and urban production structures. Models of regional specialisation include the hierarchical view of central place theory, and models of urban specialisation (e.g., Henderson, 1974; Drennan and Kelly, 2011).

benchmark case. Relaxing them could add more detail but would not change the main conclusion. For example, spatial productivity differences may be greater for some factors or for some industries than others, in which case the model would also provide an insight into the regional specialisation. We do not pursue this further theoretically but instead in our empirical analysis we, first, estimate firm productivity by 2-digit industry samples controlling for location. Next, we calculate aggregate regional productivity by urbanization category and then decompose it into industry productivity and industry composition effects.

3 Estimation algorithm: Location and firm productivity

3.1 Behabioural framework

To theoretically derive a productivity measure we rely on a behavioral framework which builds on models of industry dynamics by Ericson and Pakes (1995) and Hopenhayn (1992). Alongside econometric modeling ideas in Ackerberg et al. (2007), the framework underlines our estimation strategy and helps us specify timing and relational assumptions for the firm decisions in a manner similar to Olley and Pakes (1996).

As in Olley and Pakes (1996) we specify a log-linear production function,

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + \omega_{jt} + \eta_{jt}, \tag{1}$$

where the log of value added of firm, j at time, t, y_{jt} is modelled as a function of the logs of the firm's state variables at t, namely age, a_{jt}, capital, k_{jt}, and labour, l_{jt}. The error structure comprises a stochastic component, η_{jt}, with zero expected mean, and a component that represents unobserved productivity, ω_{jt}. Both ω_{jt} and η_{jt} are unobserved, but ω_{jt} is a state variable, and thus affects firm's choices – the investment demand and the decision to exit, while η_{jt} has zero expected mean given current information, and hence does not affect decisions.

Firm's single period profit function is $\pi(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}) - c(i_{jt}, r_{jt})$, where both restricted profit, π(.) and adjustment cost, c(.) depend on r_{jt}, which represents the economic environment that firms face at a particular point in time; r_{jt} captures input prices, demand conditions, industry characteristics. As in Olley and Pakes (1996) all these factors are allowed to change over time; importantly, in our extension, we allow the factors to also vary by narrowly defined spatial units.

The incumbent firm maximizes its expected value of both current and future profits according to:

$$V(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}) = \max \begin{cases} \Phi(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}), \\ \max_{i_{jt} \geq 0} \{\pi(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}) - c(i_{jt}, r_{jt}) + \\ \beta E[V(k_{jt+1}, a_{jt+1}, l_{jt+1}, \omega_{jt+1}, r_{jt+1}) \mid k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}, i_{jt}]\}. \end{cases} \qquad (2)$$

The Bellman equation explicitly considers two firm decisions. First is the exit decision; $\Phi(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt})$ represents the sell-off value of the firm. Second is the investment decision, i_{jt}, which solves the interior maximization problem. Under the assumption that equilibrium exists and that the difference in profits between the firm continuing and exiting is increasing in ω_{jt} we can write the optimal exit decision rule as

$$\mathrm{X}_{jt} = \begin{cases} 1 \quad if \quad \omega_{jt} \geq \overline{\omega}_t(k_{jt}, a_{jt}, l_{jt}) \\ 0 \qquad\qquad otherwise \end{cases} \qquad (3)$$

and the investment demand function as

$$i_{jt} = i_t(k_{jt}, a_{jt}, l_{jt}, \omega_{jt}, r_{jt}). \qquad (4)$$

Olley and Pakes (1996) assume that productivity, ω_{jt} follows an exogenous Markov process and is determined by a family of distributions conditional on the information set at time t-1, J_{jt-1}, which includes past productivity shocks. Given this set of distributions, both the exit and investment decisions will crucially hinge upon the firms' perceptions of the distribution of future market structure conditional on current information (and past productivity). The decisions that the firms take will in turn generate a distribution for the future market structure (Maskin and Tirole, 1988). In our behavioral framework we explicitly introduce spatial information in r_{jt}. Decisions on whether to exit the market and how much to invest will depend explicitly on firm's location.[9]

Furthermore, since we deflate value added with an industry-wide PPI, we do not control for the fact that output and factor prices might be different across firms and/or evolve differently over time. Therefore we drop the assumption of industry homogeneity and incorporate the location information in the investment and survival equilibrium equations. More formally, we explicitly allow that demand conditions, market structures and factor prices affecting firm decisions about investment and exit to differ by narrowly defined spatial units.

Following ideas in Roback (1982) and Ottaviano and Peri (2006) we introduce in r_{jt} information on wages and land rents. This helps disentangling the effects of consumption amenities from productive advantages across locations. Higher wages make workers better off whereas higher rents make them worse off. Thus, greater consumption amenities will make workers willing to accept both lower wages and higher rents. For firms, both higher wages and higher rents mean increased costs. Thus, localized productive advantages will make firms willing to accept higher wages and higher rents. Consequently, both consumption

[9] Note that the investment policy function in Olley and Pakes (1996) is a solution to a complicated dynamic programming problem and depends on all the primitives of the model like demand functions, the specification of sunk costs, form of conduct in the industry and other factors as discussed by Ackerberg et al. (2007). All these factors are allowed here to be different across narrowly defined spatial units and evolve differently over time.

amenities and productive advantages should be associated with higher rents. However, consumption amenities should be associated with lower wages whereas productive advantages should be associated with higher wages. Thus, if big cities are characterized by both better amenities and higher productivity, the net effect on wages will be ambiguous.

3.2 Econometric implementation

Because productivity ω_{jt} is not observed directly in our data estimating Equation (1) is affected by simultaneity and selection biases. Simultaneity means that estimates for more variable inputs such as labour will be upward biased if an OLS estimator is used, assuming a positive correlation with unobservable productivity. Selection (exit) depends on productivity as well as on the capital stock representing fixed cost. Thus, the coefficient on capital is likely to be underestimated by OLS as higher capital stocks induce firms to survive at low productivity levels (Olley and Pakes, 1996). Besides these two biases, a potential problem afflicting productivity measure is associated with the spatial dependency of observations within a geo-space. Spatial dependency leads to the spatial autocorrelation problem in statistics since - like temporal autocorrelation - this violates the standard statistical assumption of independence among observations (Anselin and Kelejian, 1997).

To deal with the biases, we explicitly build the productivity and location relationship into a (structural) model of the unobservable productivity following the Olley and Pakes (1996) approach by inverting the investment demand function to generate a proxy for unobserved productivity.[10] Pakes (1994) discusses conditions under which the investment

[10] The invertability of the investment function requires the presence of only one unobservable which Olley and Pakes (1996) refer to as *scalar unobservable* assumption. This assumption means that there can be no measurement error in the investment function, no unobserved differences in investment prices across firms, and no unobserved separate factors that affect investment but not output. The monotonicity needed in Olley and

demand function (Equation 4) is strictly monotonic in ω_{jt}. Under such conditions investment can be inverted to generate the productivity function:

$$\omega_{jt} = h_t(i_{jt}, r_{jt}, k_{jt}, a_{jt}, l_{jt}) \;. \tag{5}$$

Thus, we specify productivity of a firm, *j*, at a point in time, *t* as a function of the firm's state variables (capital, k_{jt}, labour, l_{jt}, age, a_{jt}), investment, i_{jt}, and the economic environment characteristics that the firm faces at a particular point in time, r_{jt}, and treat the function non-parametrically in our estimation algorithm. Investment demand traces expected productivity and it thus serves as a main control variable.

Given that our strategy is to control for unobservable productivity while estimating production functions, rather than explicitly identifying effects, we use in r_{jt} as proxies of agglomeration effects the land price and annual wage at municipality level.[11] As argued by Voith (1991), Graham (2009), Drennan and Kelly (2011) and others agglomeration effects are capitalised in immobile factor prices, and analysis based on small spatial units increases the probability of homogeneity of rents within each area.[12] In addition, we include time trend and population density at regional level to control for common effects with respect to time periods and COROP regions. Considering the fact that spatial autocorrelation is mostly due to omitted location-specific variables and spatial heterogeneity, by including land prices, wages and a control for density of economic activity at disaggregated spatial units substantially

---

Pakes (1996) does not depend on the degree of competition in the output market; it just requires the marginal product of capital to be increasing in productivity.

[11] In terms of Ackerberg et al. (2007) land prices and wages are additional observed controls of firm investment choices. Alternatively, land prices and wages can be treated as state variables.

[12] While the distinction between urbanization and localization effects is conceptually valid, it can, as theory indicates, be very difficult to identify empirically and in particular for industries that are prominent in dense urban environments. Thus, the problem of identification is potentially most severe for highly urbanized countries such as the Netherlands. However, Graham (2009) shows that estimations for various industries using generic agglomeration variables present evidence of agglomeration economies with no substantial loss in model fit compared to an estimation, where urbanization and localization effects are separated.

enriches the model of the unobservable and enhances our ability to minimize the spatial autocorrelation bias.

Substituting equation (5) into the production function (1) gives us

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + h_t(i_{jt}, r_{jt}, k_{jt}, a_{jt}, l_{jt}) + \eta_{jt}. \tag{6}$$

Equation (6) can be estimated as in Olley and Pakes (1996) with semi-parametric methods that treat the inverse investment function $h_t(.)$ non-parametrically, using either a polynomial or kernel. The non-parametric treatment, however, results in collinearity and requires the constant, k_{jt}, a_{jt}, and l_{jt} terms to be combined into a function $\phi_t(i_{jt}, r_{jt}, k_{jt}, a_{jt}, l_{jt})$ such that equation (6) becomes

$$y_{jt} = \phi_t(i_{jt}, r_{jt}, k_{jt}, a_{jt}, l_{jt}) + \eta_{jt}. \tag{7}$$

The capital, age and labor coefficients are identified in the third stage of the algorithm. First, note that even though we do not identify any input coefficients in the first stage of the algorithm we are able to estimate $\hat{\phi}_t$ for use in the third stage. We express ω_{jt} as $\hat{\omega}_{jt} = \hat{\phi}_{jt} - \beta_0 - \beta_k k_{jt} - \beta_a a_{jt} - \beta_l l_{jt}$. Second, the first stage is not affected by endogenous selection because ϕ_t fully controls for the unobservable; by construction, η_{jt} represents unobservable factors that are not known by the firm before input and exit decisions are made. In contrast, the third stage of the estimation algorithm is affected by endogenous selection. From Equation (3) it is evident that the exit decision in period *t* depends directly on ω_{jt}. Using the assumption that ω_{jt} follows an exogenous first-order Markov process $p(\omega_{jt} \mid \omega_{jt-1})$, we can decompose ω_{jt} into its conditional expectation given the information known by the firm at *t*-1 and a residual:

$$\begin{aligned}\omega_{jt} &= E[\omega_{jt} \mid J_{jt-1}] + \xi_{jt} \\ &= E[\omega_{jt} \mid \omega_{jt-1}] + \xi_{jt} \\ &= g(\omega_{jt-1}) + \xi_{jt}.\end{aligned} \tag{8}$$

By construction ξ_{jt} is uncorrelated with J_{jt-1} and thus with k_{jt}, a_{jt} and l_{jt} which are functions of only the information set at time t-1. Next using Equation (1) and substituting Equation (8) above we can write:

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g(\omega_{jt-1}) + \xi_{jt} + \eta_{jt}\,. \tag{9}$$

To correct for endogenous selection (exit) we take the expectations of Equation (9) conditional on both the information at t-1 and on X_{jt}=1 (i.e., firms survive till the next period). We can write,

$$\begin{aligned}E[y_{jt} \mid J_{jt-1}, \mathrm{X}_{jt} = 1] &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + E[\omega_{jt} \mid J_{jt-1}, \mathrm{X}_{jt} = 1] \\ &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g(\omega_{jt-1}, \overline{\omega}_t(k_{jt}, a_{jt}, l_{jt})).\end{aligned} \tag{10}$$

We do not directly observe $\overline{\omega}_t(k_{jt}, a_{jt}, l_{jt})$ and to control for it we use information on observed exit. This means that the probability of being in the data at period t conditional on information known at t-1 is:

$$\begin{aligned}\Pr(\mathrm{X}_{jt} = 1 \mid J_{jt-1}) &= \Pr(\omega_{jt} \geq \overline{\omega}_t(k_{jt}, a_{jt}, l_{jt}) \mid J_{jt-1}) \\ &= \varphi_t''(\omega_{jt-1}, \overline{\omega}_t(k_{jt}, a_{jt}, l_{jt})) \quad = \varphi_t'(\omega_{jt-1}, k_{jt}, a_{jt}, l_{jt}) \\ &= \varphi_t(i_{jt-1}, r_{jt-1}, k_{jt-1}, a_{jt-1}, l_{jt-1}) \quad = P_{jt}.\end{aligned} \tag{11}$$

Equation (11) represents the second stage of our estimation algorithm and can be estimated non-parametrically using Probit model with a polynomial or kernel as in Olley and Pakes (1996). Note that we extend the state variable set with location information which is an important determinant of firms' decision to exit. Substituting this equation into Equation (10) and using expressions for estimated values, $\hat{\omega}_{jt-1}$ and $\hat{P}_{jt}$ gives us

$$
\begin{aligned}
E[y_{jt} \mid J_{jt-1}, \mathrm{X}_{jt} = 1] &= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g(\omega_{jt-1}, f(\omega_{jt-1}, P_{jt})) \\
&= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g'(\omega_{jt-1}, P_{jt}) \\
&= \beta_0 + \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g'(\hat{\phi}_{jt-1} - \beta_0 - \beta_k k_{jt-1} - \beta_a a_{jt-1} - \beta_l l_{jt-1}, \hat{P}_{jt}),
\end{aligned}
$$

which after removing the expectations operator becomes

$$
y_{jt} = \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g'(\hat{\phi}_{jt-1} - \beta_k k_{jt-1} - \beta_a a_{jt-1} - \beta_l l_{jt-1}, \hat{P}_{jt}) + \xi_{jt} + \eta_{jt}, \qquad (12)
$$

where the two β_0 terms have been encompassed into the non-parametric function g'. Equation (12) represents the third (last) stage of our estimation algorithm and can be estimated with Non-Linear Least Squares (NLLS), approximating g' with a polynomial series.[13][14]

In line with our goal to estimate unbiased measure of productivity, introducing location-specific controls in the model does minimise the deviation from the original Olley-Pakes scalar unobservable assumption, necessary to invert the investment function, and helps the precision of estimates. However, assuming that firms' location choices are completely independent from ω_{jt} is not satisfactory. Ciccone and Hall (1996), Combes et al. (2009) and the related literature demonstrate the link between productivity, density of economics activity and agglomeration. A sorting equilibrium exists such that *a priori* more productive firms value agglomeration highly and locate in highly urbanized regions where their productivity in turn is further enhanced by agglomeration externalities; less productive firms that have high congestion costs locate in less urbanized regions. To capture this relationships we follow

[13] It is also important to note that the identification of the labor coefficient in the last rather than in the first stage of the estimation algorithm requires making assumptions on how current labor responds to the current realizations of ξ_{jt}. One possible treatment is that labor is fixed before the realization of ξ_{jt}, which is the same assumption as for capital. This implies that current labor input is not correlated with current innovation in productivity and β_l can be identified in the third stage. A second, and more realistic, assumption is that current labor input can respond to current innovations in productivity. We still can obtain estimates of β_l using Equation (12) and relying on the fact that lagged values of labor (l_{jt-1}) should be uncorrelated with ξ_{jt} which follows from the information structure of the model.

[14] Woodridge (2009) presents a concise, one-step formulation of the original Olley and Pakes (1996) approach using GMM estimator which is more efficient but less flexible than the standard Olley-Pakes methodology.

Ackerberg et al. (2007) and Rizov and Walsh (2009; 2011) and model productivity as a second-order Markov process, $p(\omega_{jt} \mid \omega_{jt-1}, \omega_{jt-2})$, where firms operate through time forming expectations of future ω_{jt}s on the basis of information from two preceding periods.

The non-parametric function g' in the third stage, Equation (12), has to be modified and becomes $g''(.) = g''(\hat{\phi}_{jt-1} - \beta_k k_{jt-1} - \beta_a a_{jt-1} - \beta_l l_{jt-1}, \hat{\phi}_{jt-2} - \beta_k k_{jt-2} - \beta_a a_{jt-2} - \beta_l l_{jt-2}, \tilde{\hat{P}}_{jt})$. The lagged $\hat{\phi}$ variables are obtained from the first stage estimates at *t*-1 and *t*-2 periods. Because the conditional expectation of ω_{jt} given J_{jt-1} now depends on ω_{jt-1} and ω_{jt-2}, we need to use estimates of $\hat{\phi}$ from two prior periods. Controlling for endogenous selection (exit) has to be modified accordingly as well by using information from two preceding periods; note the change of notation ($\tilde{\hat{P}}_{jt}$ instead of $\hat{P}_{jt}$).

In the empirical analysis that follows we use the production function coefficients $\hat{\beta}_k$ and $\hat{\beta}_l$ consistently estimated from the specification with second-order Markov process and back out unbiased and consistent firm-specific productivity (TFP) measures, calculated as residuals from the production function:

$$q_{jt} = \omega_{jt} + \beta_a a_{jt} + \eta_{jt} = y_{jt} - \hat{\beta}_k k_{jt} - \hat{\beta}_l l_{jt}. \quad (13)$$[15]

4 Data and estimation results

4.1 Data and summary statistics

We apply the estimation algorithm developed in Section 3 to the AMADEUS sample of Dutch firms. AMADEUS of the Bureau van Dijk is a comprehensive, pan-European database

[15] As explained in Ackerberg et al. (2007), including age in the specification helps control for cohort effects on firm productivity which improves precision of coefficient estimates; we do not net out the age contribution from the TFP measure.

containing records for more than 250,000 small, medium and large firms in the Netherlands over the period 1997-2006.[16] For each firm there is detailed information from unconsolidated financial statements as well as about ownership structure, location (post code) and indistrial activity. The coverage of the data compared to the aggregate statistics reported by the CBS is very good as for sales it is 73 per cent and for employment – 70 per cent. The industry sectors are identified on the bases of the current NACE Rev.1 classification at the 2-digit level and cover agriculture, manufacturing, construction, trade and services (codes range from 01 to 74) - 28 industries in total. All nominal monetary variables are converted into real values by deflating them with the appropriate 2-digit NACE industry deflators provided by CBS. We use PPI to deflate sales and cost of materials, and asset price deflators for capital and fixed investment variables.[17]

In this paper, our goal is to estimate unbiased and consistent TFP measures at firm level, within industries, and to document the aggregate productivity gaps between less urbanized, urbanized, and highly urbanized regions. The strategy of our empirical analysis is to run regressions within industries and we apply our estimation algorithm to the 28 largest 2-digit industries, with sufficient number of observations. After lags are applied and

[16] The TOP-1.5-million AMADEUS module contains firms which must satisfy one of the following criteria: i) operating revenue > €1 million; ii) total assets > €2 million; iii) number of employees > 15. There is also a TOP-250,000 module which contains only large firms which must satisfy one of the following criteria: i) operating revenue > €10 million; ii) total assets > €20 million; iii) number of employees > 150. Financial information is reported from unconsolidated firm financial statements. We note, however, that for some firms information is incomplete or only for a single year.

[17] A number of studies (e.g., Katayama et al., 2009; Del Gatto et al., 2008) point that production functions should be a mapping of data on inputs and outputs. Most studies tend to use revenue and expenditure data and apply industry level deflators for output, raw materials and capital to get back the quantity data needed. However, inputs and outputs can be priced differently for different firms within narrowly defined industries. This results in inconsistency discussed by Klette and Griliche (1996) in the case of common scale estimators. To deal with the problem some studies (e.g., Del Gatto et al., 2008) introduce average industry sales as an additional regressor in the production function. We note, however, that introducing detailed location information in the state space will control for persistent pricing gap across firms in their use of inputs and their outputs within each industry. Furthermore, Foster et al. (2008) find that productivity estimates from quantity and deflated revenue data are highly correlated, and that the bias vanishes on average such that estimated average productivity is unaffected when aggregate deflators are used.

observations with missing values deleted, there are 13,897 remaining observations for 4,220 firms. The estimated sample accounts for about 51 per cent of the sales and 47 per cent of the employment in our AMADEUS data. The correlations between the CBS aggregate statistics series and the estimated sample series are as follows: value added (used in the regressions as dependent variable) - 0.91, employment - 0.94 and capital – 0.88.

The descriptive statistics calculated from the estimated sample of firms are reported in Table 2. We compare average firm characteristics across less urbanized, urbanized, and highly urbanized regions. Firms in highly urbanized regions, compared to their counterparts in urbanized and less urbanized regions are larger in terms of value added, employment, and capital, and invest more. These characteristics are in accord with the socio-economic measures, proxying density of economic activity, reported in Table 1. The land prices differ in a similar manner, however, the difference between highly urbanized and urbanized regions is relatively small suggesting that much higher density of population does not correspond to proportionate increase in land price. Interestingly, regional industry concentration characterised by market share of the top four 2-digit industries (C4) as well as by the Herfindal Index does not show substantial differences across categories of urbanization although higher degree of urbanization appears to be associated with higher concentration. There are differences in the composition of the top four industries dominating each urbanization category. In all categories the most dominant are wholesale trade (51) and construction (45). Less urbanized regions are the only part of the country where manufacturing of machinery and equipment (29) has an important presence while food industry (15) is also important in urbanized regions. Sales and maintenance of automobiles and automotive fuel sales (50) are important in urbanized and highly urbanized regions. Only in highly urbanized regions retail sales (52) make part of the C4 industries. Overall, there are differences across urbanization categories regarding industry composition while differences in concentration are rather modest, therefore in the following analyses we report results for the total sample of firms as well as by separate manufacturing and services samples to demonstrate clearly heterogeneity in behaviour.[18]

- Table 2 about here -

4.2 Productivity estimates

Weighted average coefficients (using value added as weight) over the estimated 28 industry production functions, by urbanization category and industry type are reported in Table 3. Coefficient estimates from all 28 industry regressions, number of observations and test statistics are reported in Appendix 1. The reported coefficients demonstrate expected differences across urbanization categories and industry types (manufacturing vs. services) with respect to the shares of labour and capital in output. For the total sample both coefficients, on labour and capital, decline systematically across urbanization categories as the value of labour coefficient is 0.647 for less urbanized regions while it is 0.623 for highly urbanized ones. The pattern of the capital coefficient is similar – 0.178 for less urbanized regions and 0.162 for highly urbanized regions.

- Table 3 about here -

Weighted average productivity measures by urbanization category clearly show that highly urbanized regions are the most productive; for the total sample the average firm TFP in these regions is 4.382, while it is 3.599 and 3.329 - in urbanized and less urbanized regions respectively as all differences are significant at the 1 percent level. Services are more

[18] Recent studies on the effects of specialization and diversity of economic activity suggest that there are likely differential responses by manufacturing firms compared to firms in services for which local interactions are more relevant (Duranton and Overman, 2005; Ellison et al., 2010; Drennan and Kelly, 2011). Our manufacturing sample contains all available 2-digit manufacturing industries in the range 15 – 36 and the services sample – in the range 50 – 74.

productive than manufacturing everywhere and, in terms of location, the patter is the same as for the total sample.

Furthermore, not only the mean but the whole distribution of firm TFPs in highly urbanized regions dominates the corresponding distributions of firm TFPs in urbanized and less urbanized regions. Figure 2 illustrates the distributions of firm TFPs across urbanization categories and industry types by the means of kernel density estimates. The Kolmogorov-Smirnov two-sample tests for stochastic dominance are statistically significant at the 1 percent level for most distribution pairs and confirm the fact that firms in highly urbanized regions are the most productive. The test results are repprted in Table 4, Panel A. We also plot the distributions and test the differences for subsamples of manufacturing and service industries confirming the findings from the total sample. It worth noting that productivity levels and differences appear larger for services compared to manufacturing.

- Figure 2 about here -

- Figure 3 about here -

- Table 4 about here -

When we consider the distributions of TFP annual changes presented in Figure 3, however, we see that the largest changes for the total sample as well for the subsamples of manufacturing and services are exhibited by firms in urbanized regions rather than in the highly urbanized ones. The Kolmogorov-Smirnov two-sample tests reported in Table 4, Panel B are again significant at the 1 percent level for the pairs of interest and confirm the fact that productivity of firms in urbanized - rather than in highly urbanized - regions grows the fastest. The finding that highly urbanized regions lag behind urbanized regions in terms of TFP growth we interpret as evidence of congestion due to too a high density and degree of urbanization, i.e., the net agglomeration effect on productivity is non-linear, with an inverted-U shape.

5 Spatial variation in aggregate productivity: Is there a limit to agglomeration?

The evidence and discussion in previous sections suggest that there is a systematic relationship between productivity and the degree of agglomeration as measured by the levels of urbanization and density of economic activity. In this section we analyse differences in regional productivity across urbanization categories and industry types by applying decompositions of productivity levels and changes following Rice et al. (2006) and Oosterhaven and Broersma (2007). Given our analytical strategy to build into the estimated model of (unobservable) productivity all relevant factors affecting it, to demonstrate the link between agglomeration and productivity it is sufficient to use unconditional shift-share type decomposition. Saito and Gopinath (2009) and Combes et al. (2009) identify and compare the importance of agglomeration factors and firm (and industry) selection for regional productivity. Therefore in the decompositions we consider these two factors as sources of the spatial variation in regional productivity (productivity changes). First, differences in individual firm productivities (productivity changes) within each industry, resulting in different average productivities (productivity changes) across industries depend on the strength of various agglomeration economies. Second, differences in the industry composition within each region depend on firm (and indistry) location choices driven by selection.[19]

We calculate aggregate industry productivity, q_u^n by urbanization category, u and industry, n as weighted average of individual firm TFPs (q_{jt}) using firm value added as

[19] The firm (and industry) selection is an outcome from a sorting equilibrium - that is, firms that value agglomeration highly locate in urbanized regions, firms that have high congestion costs are found in less urbanized regions.

weight.[20] The total value added in urbanization category, u is denoted by $S_u = \Sigma_u s_u^n$ and the share of industry, n in the total value added in urbanization category, u is $\lambda_u^n = s_u^n/S_u$. The average productivity of industry, n for the economy as a whole (i.e., aggregating across all urbanization categories, u) is given by $\bar{q}^n = \sum_u s_u^n q_u^n / \sum_u s_u^n$, while $\bar{\lambda}^n = \sum_u s_u^n / \sum_u S_u$ is the share of industry, n in total value added for the whole economy. Aggregate regional productivity, q_u is weighted average of industry productivities in urbanization category, u, using industry value added shares as weights.

Regional productivity (a) may be decomposed as follows:

$$\underset{(a)}{q_u \equiv \sum_n q_u^n \lambda_u^n} \quad = \underset{(b)}{\sum_n q_u^n \bar{\lambda}^n} \quad + \underset{(c)}{\sum_n \bar{q}^n \lambda_u^n} \quad - \underset{(d)}{\sum_n \bar{q}^n \bar{\lambda}^n} \quad + \underset{(e)}{\sum_n (q_u^n - \bar{q}^n)(\lambda_u^n - \bar{\lambda}^n)} . \qquad (14)$$

The first term on the right-hand side of Equation (14) is the average level of productivity in urbanization category, u conditional on industry composition being the same as for the whole economy; we refer to this as *productivity index* (b). The second term is the average level of productivity of urbanization category, u given its industry composition but assuming that the productivity of each industry equals the economy-wide average for that industry. It is referred to as the *industry composition index* (c). Remaining terms (d) and (e) measure the *residual covariance* between industry productivities and industry shares in urbanization category, u. It is important to point out that comparison between productivity and industry composition indexes, while taking into account the residual covariance terms, in Equation (14) can provide useful information about the net impact of aglomeration and selection factors on regional productivity. The decomposition of productivity changes is analogous to the decomposition of productivity levels described above.

[20] Note that industry productivity is determined by individual firm productivities and firm market shares, within the industry, as discussed by Olley and Pakes (1996) and Rizov and Walsh (2009), among others. Thus, there could be two sources of industry productivity – within-firm productivity increases and reallocation of market shares towards more productive firms.

We report decomposition results for the highly urbanized, urbanized and less urbanized regions in the Netherlands, for the total sample and for the separate manufacturing and setvices samples, in Table 5, Panel A. While variation in aggregate productivity by urbanization category reflects differences in both productivity and industry composition, the spatial variation observed in the productivity index derives entirely from spatial variation in firm (industry) productivity and is independent of differences in industry composition. A higher value of the productivity index in a given urbanization category would suggest that industries in this category are more productive. The spatial variation in the industry composition index derives entirely from differences in the industry composition across urbanization categories and is independent of variation in industry productivity. A higher value of the industry composition index in a given category implies that the more productive industries are represented by larger industry shares in that urbanization category. The last covariance term in Equation (14) provides information about the link between industry shares and productivity; a positive sign of the term in a given urbanization category means that the more productive industries are also relatively larger indicating a positive regional specialisation effect.

- Table 5 about here -

The results in Panel A are computed as averages for the 2000-2006 period and confirm that highly urbanized regions, with the highest density of economic activity, have the highest aggregate productivity. The urbanized regions lag behind in aggregate productivity by 19.7 percent, while less urbanized regions are the least productive, with aggregate productivity lower by 25.6 percent compared to the highly urbanized regions. Productivity index and industry composition index also are lower for urbanized and less urbanized regions compared to the highly urbanized ones as the differentials for the productivity index are 10.3 percent and 11.4 percent, while the differentials for the industry composition index are 12.3

percent and 13.7 percent respectively. The magnitudes of the differentials demonstrate that urbanized and less urbanized regions are characterised by monotonically lower productivity and industry composition indexes. The covariance term is positive for all urbanization categories but its magnitude is the largest for the urbanized regions suggesting a substantial unexplained allocation of industry shares to more productive industries or alternatively higher productivity of larger, clustered industries. Comparing the separate manufacturing and services samples we find evidence that service industries benefit more from high levels of urbanization both in terms of productivity and industry composition, as argued in several previous studies. Otherwise, the pattern of decomposition indexes across urbanization categories is similar to the one for the total sample.

To explore further the factors affecting aggregate productivity, by urbanization category and indusrty type, we analyse average annual productivity change over the 2000-2006 period following the decomposition defined in Equation (14) and report results in Table 4, Panel B. The period of analysis is generally characterised by stable economic and trade conditions after the implementation of the single currency, the Euro in the beginning of 1999. We are able to establish the magnitudes of contributions by both industry productivity and industry composition changes to the aggregate productivity of highly urbanized, urbanized and less urbanized regions. The results in Panel B show substantial heterogeneity in productivity growth by urbanization category. Aggregate productivity in urbanized regions increases with the highest annual rate of 1.7 percent followed by the rates in highly urbanized and less urbanized regions - 1.4 and 1.3 percent respectively. This finding is of high importance and demonstrates that the net effect of agglomeration (net of congestion) on regional productivity growth is non-linear and picks up at intermediate level of urbanization – a result similar to findings by Broersma and Oosterhaven (2009) for the case of labor productivity growth in the Netherlands. The pattern is also evident when we compare separate manufacturing and services samples as it is more pronounced in the case of services.

The sources of aggregate productivity growth seem to vary by urbanization category and industry type. For the highly urbanized and less urbanized regions improvements in both productivity and industry composition indexes are evident. For urbanized regions the growth in productivity index is the most important while contributions by the industry composition index are less relevant. The contribution of the industry composition index is the most significant in less urbanized regions suggesting that reallocation of industry shares towards more productive industries is taking place in those regions. There is also evidence of relative improvement in the industry composition in highly urbanized regions over time. The effects appear large in the services sample compared to the manufacturing sample. The negative change in the residual covariance terms, however, provides a relative evidence of deteriorating productivity in important clustered industries in the highly urbanized regions and increasing productivity in small dispersed industries in the less urbanised regions.

6 Conclusion

The focus of the paper is on evaluating the net impact of agglomeration on productivity (TFP) in the Dutch regions classified by level of urbanization. We build a structural model of the unobservable productivity incorporating land price (and wage) information to proxy for the effects of agglomeration and adapt the semi-parametric estimation approach proposed in Olley and Pakes (1996) to estimate the parameters of production functions using firm data, within 2-digit industries, for the period 1997 - 2006. We use information on land prices available at 4-digit postcodes and allow market structure to differ at disaggregate municipality (gemeente) level. We model the unobservable productivity as a second-order Markov process which enhances our ability to obtain unbiased and consistent estimates of the production function parameters and thus, back out unbiased and consistent firm-specific TFP measures.

We aggregate the firm TFPs by urbanization category and industry type and find that aggregate productivity systematically differs across highly urbanized, urbanized and less urbanized regions as the magnitudes of the differentials from the productivity of highly urbanized category are -19.7 percent and -25.6 percent, respectively. Our results confirm findings of previous studies that productivity and agglomeration are positively correlated. Further, analysing productivity changes reveals important differences across urbanization categories. The main finding is that there is a tendency of urbanized regions – exhibiting annual growth rate of 1.7 percent - catching up with highly urbanized regions – with annual growth rate of 1.4 percent - in terms of aggregate productivity over the period of analysis. Similar to Broersma and Oosterhaven (2009) we find evidence that the growth of productivity is the highest in urbanized rather than in highly urbanized regions in the Netherlands pointing to the fact that agglomeration has led to congestion and negatively affected productivity growth at high levels of urbanization and density of economic activity.

We also decompose aggregate productivity into productivity index and industry composition index. The productivity index is the highest in highly urbanized regions suggesting that (firm and industry) productivity is positively influenced by agglomeration. The industry composition index captures the extend to which production in different urbanization categories is allocated to industries that are more or less productive compared to the average for the Dutch economy. Changes in industry composition index are more important in highly urbanized and less urbanized regions where selection forces seem to dominate. Changes in productivity index are larger in urbanized regions indicating stronger agglomeration effects.

References

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes. 2007. Econometric Tools for Analyzing Market Outcomes, in Handbook of Econometrics, Eds. J.J. Heckman, and E.E. Leamer, Vol. 6A, pp. 4171-4276, Elsevier: North Holland.

Adsera, A. 2000. Sectoral spilovers and the price of land: A cost analysis, Regional Science and Urban Economics, 30, 565-585.

Alonso-Villar, O. 2008. A model of economic geography with demand-pull and congestion costs, Papers in Regional Science, 87(2), 261-277.

Anselin, L. and H. Kelejian. 1997. Testing for spatial error autocorrelation in the presence of endogenous regressors, International Regional Science Review, 20, 153-182.

Broersma, L., and J. Oosterhaven. 2009. Regional labor productivity in the Netherlands: Evidence of agglomeration and congestion effects, Journal of Regional Science, 49(3), 483-511.

Carlino, G. and R. Voith. 1992. Accounting for differences in aggregate state productivity, Regional Science and Urban Economics, 22(4), 597-617.

Ciccone, A. 2002. Agglomeration effects in Europe, European Economic Review, 46(2), 213-228.

Ciccone, A. and R. Hall. 1996. Productivity and the density of economic activity, American Economic Review, 86(1), 54-70.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux. 2009. The productivity advantages of large cities: Distinguishing agglomeration from firm selection, Discussion Paper 7191, Centre for Economic Policy Research.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux. 2010. Estimating agglomeration effects with history, geology, and worker fixed-effects, In E. L. Glaeser (ed.), Agglomeration Economics. Chicago, IL: University of Chicago Press.

Davis, D. and D. Weinstein. 2008. A search for multiple equilibria in urban industrial structure, Journal of Regional Science, 48(1), 29-65.

Del Gatto, M., G. Ottaviano, and M. Pagnini. 2008. Openness and trade and industry cost dispersion: Evidence from a pannel of Italian firms, Journal of Regional Science, 48(1), 97-129.

Dixit, A. and J. Stiglitz. 1977. Monopolistic competition and optimum product diversity, American Economic Review, 67(3), 297-308.

Drennan, M. and H. Kelly. 2010. Measuring urban agglomeration economics with office rents, Journal of Economic Geography, 11(3), 481-507.

Duranton, G. and H. Overman. 2005. Testing for localization using micro-geographic data, Review of Economic Studies, 72, 1077-1106.

Ellison, G., E. Glaeser, W. Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration pattersn, American Economic Review, 100, 1195-1213.

Ericson, R. and A. Pakes. 1995. Markov-perfect industry dynamics: A framework for empirical work, Review of Economic Studies, 62, 53–82.

Forslid, R. and G. Ottaviano. 2003. An analytically solvable core-periphery model, Journal of Economic Geography, 3, 229–240

and C. Syverson. 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review, 98(1), 394-425.

Fujita, M. 1988. A monopolistic competition model of spatial agglomeration: Differentiated product approach, Regional Science and Urban Economics, 18(1), 87-124.

Glaeser, E., H. Kallal, J. Scheinkman, and A. Schleifer. 1992. Growth in cities, Journal of Political Economy, 100, 1126-1152.

Graham, D. 2009. Identifying urbanisation and localisation externalities in manufacturing and service industries, Papers in Regional Science, 88(1), 63-85.

Hanson, G. 1996. Localization economies, vertical organization, and trade, American Economic Review, 86(5), 1266-1278.

Henderson, V. 1974. The sizes and types of cities, American Economic Review, 64, 640–656.

Henderson, V. 2003. Marshall's scale economies, Journal of Urban Economics, 53, 1-28.

Hopenhayn, H. 1992. Entry, exit, and firm dynamics in long-run equilibrium, Econometrica, 60, 1127–1150.

Katayama, H., S. Lu, and J. Tybout. 2009. Firm-level productivity studies: Illusions and solutions, International Journal of Industrial Organization, 27, 403-413.

Klette, T. and Z. Griliches. 1996. The inconsistency of common scale estimators when output process are unobserved and endogenous, Journal of Applied Econometrics, 11, 343-361.

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

Leamer, E. 1984. Sources of International Comparative Advantage, MIT Press, Cambridge, MA.

Marshall, A. 1920. Principles of Economics, 8th ed., Macmillan, London.

Maskin, E. and J. Tirole. 1988. A theory of dynamic oligopoly: I&II, Econometrica, 56, 549-600.

Mills, E. 1967. An aggregative model of resource allocation in a metropolitan area, American Economic Review, 57(2), 197-210.

Mitra, A. 1999. Agglomeration economies as manifested in technical efficiency at the firm level, Journal of Urban Economics, 45, 490-500.

Olley, S. and A. Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry, Econometrica, 64(6), 1263-1297.

Oosterhaven, J. and L. Broersma. 2007. Sector structure and cluster economies: A decomposition of regional labour productivity, Regional Studies, 41(5), 639-659.

Ottaviano, G. and G. Peri. 2006. The economic value of cultural diversity: Evidence from US cities, Journal of Economic Geography, 6, 9-44.

Pakes, A. 1994. Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interactions, in: Laffont, J.-J., Sims, C. (Eds), Advances in Econometrics: The Sixth World Congress of the Econometric Society, Vol. II, Chapter 5. Cambridge University Press, New York, pp. 171– 260.

Puga, D. 2010. The magnitude and causes of agglomeration economies, Journal of Regional Science, 50(1), 203-219.

Rappaport, J. 2008. A productivity model of city crowdedness, Journal of Urban Economics, 63, 715-722.

Rice, P., A. Venables, and E. Patacchini. 2006. Spatial determinants of productivity: Analysis for the regions of Great Britain, Regional Science and Urban Economics, 36(6), 727-752.

Rizov, M. and P. Walsh. 2009. Productivity and trade orientation of UK manufacturing, Oxford Bulletin of Economics and Statistics, 71(6), 821-849.

Rizov, M. and P. Walsh. 2011. Is there a rural-urban divide? Location and productivity of UK manufacturing, Regional Studies, 45(5), 641-656.

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

Saito. H. and M. Gopinath. 2009. Plants' self-selection, agglomeration economies and regional productivity in Chile, Journal of Economic Geography, 9(4), 539-558.

Spence, M. 1976. Product selection, fixed costs and monopolistic competition, Review of Economic Studies, 43(2), 217-235.

Terluin, I., L. Slangen, E. van Leeuwen, A. Oskam, and A. Gaaff. 2005. De Plattelandseconomie in Nederland: Een verkenning van definities, indicatoren, instituties en beleid. Report 4.05.04 to LEI, Wageningen University, Wageningen.

Voith, R. 1991. Capitalization of local and regional attributes into wages and rents: Differences across residential, commercial and mixed-use communities, Journal of Regional Science, 3192), 127-145.

Wooldridge, J. 2009. On estimating form-level production functions using proxy variables to control for unobservables, Economics Letters, 104, 112-114.

Table 1: Socio-economic indicators by urbanization category

Indicator	Highly urbanized	Urbanized	Less urbanized
Annual employment growth 1996-2002 (%)	2.4	2.3	2.7
Annual population growth 1996-2002 (%)	0.6	0.6	0.8
Annual economic growth 1996-2002 (%)	2.9	3.0	2.2
Participation rate 2002 (%)	66	66	64
Unemployment rate 2002 (%)	2.4	2.1	2.6
Share of elderly population 2002 (%)	14	13	14
Annual income per capita 2002 (€)	12,000	11,000	10,000
Land price 2003 (€/ha)	221,000	139,000	64,000
Population density 2003 (number/sq. km)	1,580	570	231

Source: Terluin et al. (2005) and CBS.

Table 2: Descriptive statistics of regression variables

Variable	Highly urbanized	Urbanized	Less urbanised
Firm alue added (thousands €)	98,869.2	33,345.3	23,380.6
	(272,741.4)	(73,471.6)	(45,145.4)
Firm total fixed assets (thousands €)	149,226.7	37,822.1	27,072.8
	(359,107.0)	(85,797.9)	(71,405.4)
Firm real investment (thousands €)	41,774.1	8,977.9	5,446.2
	(121,794.7)	(37,686.0)	(14,418.2)
Firm employment (number of full-time-equivalent (FTE) employees)	1,439.9	553.9	520.2
	(4,470.3)	(1,636.1)	(1,358.9)
Firm age (years)	33.0	34.4	37.8
	(39.9)	(32.2)	(35.0)
Regional annual wage (€ per FTE employee)	18,471.6	17,618.4	16.654.1
	(7,614.4)	(8,053.0)	(4,354.9)
Regional land price (€ per hectare)	218,017.0	141,016.8	63,571.3
	(188,687.7)	(147,012.2)	(50,077.0)
Regional population density (number per sq. km)	1,544.6	580.3	232.4
	(493.8)	(204.0)	(37.8)
Market share of top four industries, C4 (%)	56.1	54.7	51.8
Herfindal Index x 100	16.7	15.1	14.0
Number of observations (Total 13897)	3555	7988	2354

Source: AMADEUS, BvD and CBS.
Note: The reported figures are means and standard deviations (in paretheses). Average land price and annual wage are calculated from data for 463 municipalities (gemeenten). Composition of 2-digit NACE C4 industries is as follow: in column (1) 51, 45, 50, 52; in column (2) 51, 45, 50, 15; in column (3) 51, 45, 29, 15. The order of industries is by market share.

Table 3: Production function coefficients and productivity estimates

Coefficient	Highly urbanized	Urbanized	Less urbanized
	Total sample		
Labour	0.623 (0.038)	0.634 (0.038)	0.647 (0.038)
Capital	0.162 (0.018)	0.167 (0.018)	0.178 (0.019)
Age	0.148 (0.054)	0.133 (0.053)	0.114 (0.050)
Adjusted R^2	0.995	0.996	0.995
Number of observations	3555	7988	2354
Regional productivity	4.382 (0.992)	3.599 (0.966)	3.329 (0.979)
	Manufacturing		
Labour	0.662 (0.036)	0.668 (0.040)	0.686 (0.040)
Capital	0.290 (0.024)	0.268 (0.027)	0.254 (0.026)
Age	0.016 (0.035)	0.025 (0.037)	0.018 (0.038)
Number of observations	867	2753	1118
Adjusted R^2	0.996	0.996	0.995
Regional productivity	3.703 (0.981)	3.496 (0.925)	3.275 (0.943)
	Services		
Labour	0.609 (0.038)	0.609 (0.038)	0.606 (0.038)
Capital	0.141 (0.017)	0.141 (0.017)	0.143 (0.017)
Age	0.175 (0.058)	0.177 (0.057)	0.176 (0.057)
Number of observations	2352	3950	868
Adjusted R^2	0.995	0.995	0.995
Regional productivity	4.796 (1.217)	3.717 (1.406)	3.403 (1.213)

Note: The reported coefficients, second-stage R^2, number of observations and regional productivity are weighted averages, using value added as weight, from estimated industry regressions (28 for the total sample, 16 for manufacturing and 10 for services) on firm level data. Standard errors (standard deviations for productivity) are reported in parentheses.

Table 4: Kolmogorov-Smirnov tests of distance between TFP distributions

Distribution samples	Total sample	Manufacturing	Services
Panel A: TFP level distributions			
HU-U	+0.116 (0.001)	+0.140 (0.001)	+0.155 (0.001)
HU-LU	+0.217 (0.001)	+0.147 (0.003)	+0.169 (0.001)
U-LU	+0.116 (0.001)	+0.047 (0.386)	+0.070 (0.102)
Panel B: TFP change distributions			
HU-U	-0.077 (0.002)	-0.095 (0.079)	-0.105 (0.001)
HU-LU	+0.065 (0.072)	-0.072 (0.341)	+0.059 (0.108)
U-LU	+0.107 (0.001)	+0.102 (0.029)	+0.090 (0.077)

Note: The two-sample Kolmogorov-Smirnov test checks whether the two data samples come from the same distribution. The distance (D_{KS}) and signifficance (p-value) in pareteses are reported for each pair of TFP distributions by urbanisation category samples. Notation of urbanization categories: HU – highly urbanized; U – urbanisez; LU – less urbanised.

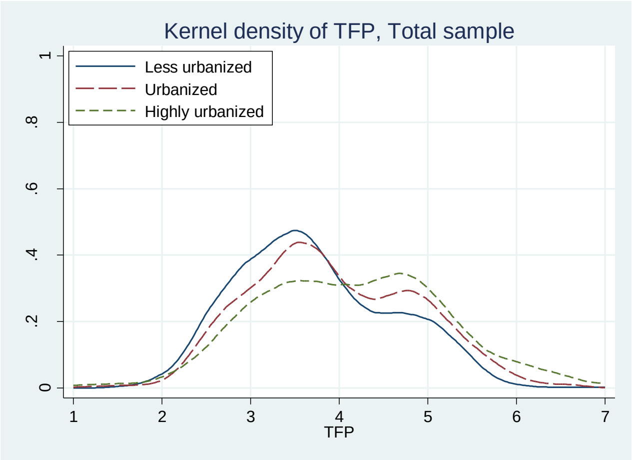
Table 5: Aggregate regional productivity decompositions

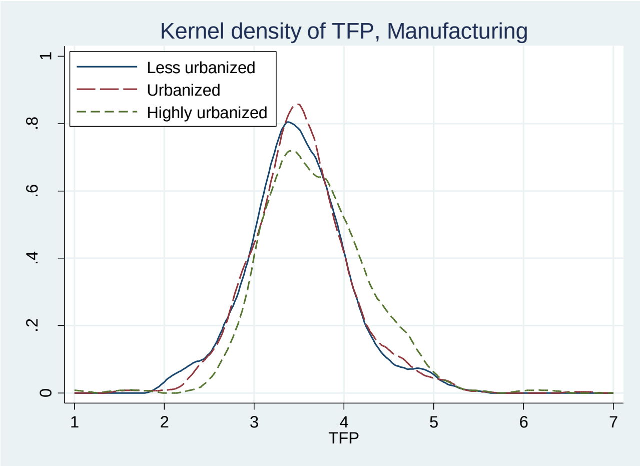
	Regional productivity	Productivity index	Composition index	Residual covariance	
	(a)	(b)	(c)	(d)	(e)
Total sample					
Panel A: Average levels, 2000-2006					
Highly urbanized	1.072	1.019	1.038	1.000	0.014
Urbanized	0.875	0.916	0.915	1.000	0.044
Less urbanized	0.816	0.905	0.901	1.000	0.010
Panel B: Annual changes, 2000-2006					
Highly urbanized	0.014	0.020	0.022	0.014	-0.015
Urbanized	0.017	0.027	0.007	0.014	-0.004
Less urbanized	0.013	0.016	0.039	0.014	-0.028
Manufacturing					
Panel A: Average levels, 2000-2006					
Highly urbanized	1.028	0.993	1.020	1.000	0.016
Urbanized	0.971	0.993	0.977	1.000	0.001
Less urbanized	0.906	0.931	0.941	1.000	0.034
Panel B: Annual changes, 2000-2006					
Highly urbanized	0.011	0.012	0.013	-0.017	-0.030
Urbanized	0.015	0.010	0.001	-0.017	-0.012
Less urbanized	0.013	0.012	0.019	-0.017	-0.035
Services					
Panel A: Average levels, 2000-2006					
Highly urbanized	1.082	1.047	1.032	1.000	0.004
Urbanized	0.835	0.864	0.838	1.000	0.032
Less urbanized	0.767	0.829	0.908	1.000	0.030
Panel B: Annual changes, 2000-2006					
Highly urbanized	0.015	0.025	0.028	0.029	-0.008
Urbanized	0.018	0.037	0.010	0.029	0.000
Less urbanized	0.014	0.018	0.051	0.029	-0.025

Note: For definitions of decomposition components refer to equation (5) in the text. Values reported in Panel A for each sample are normalised by the respective covariance term $\sum_n \bar{q}^n \bar{\lambda}^n$ from Equation (5). Component (d) has a negative sign in the decompositions.

Figure 1: Non-linear net agglomeration effect

Figure 2: TFP level distributions





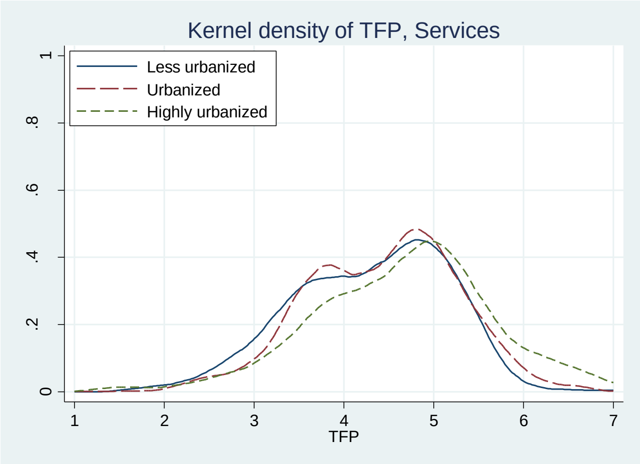
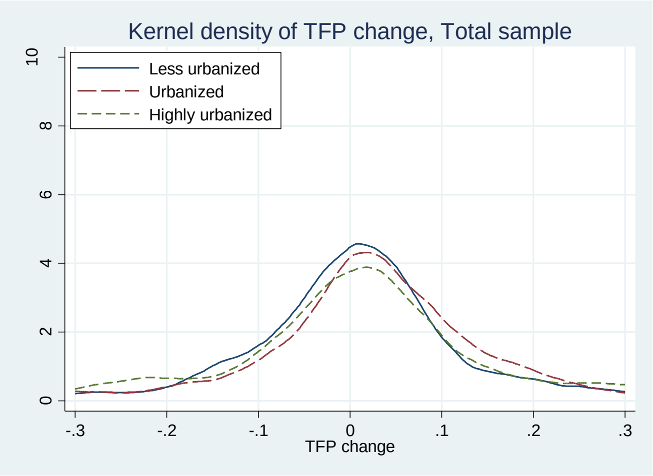
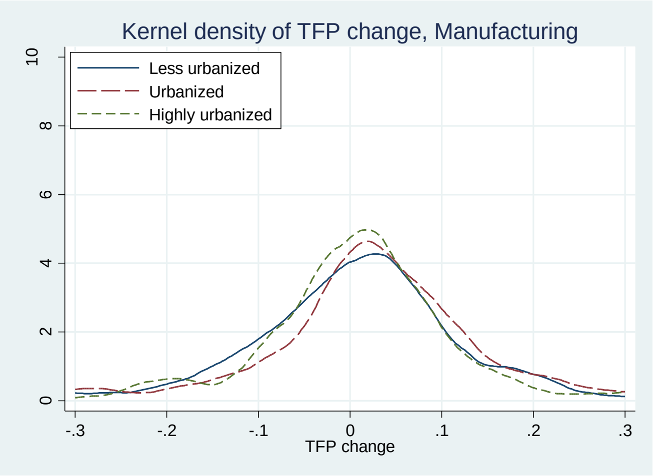
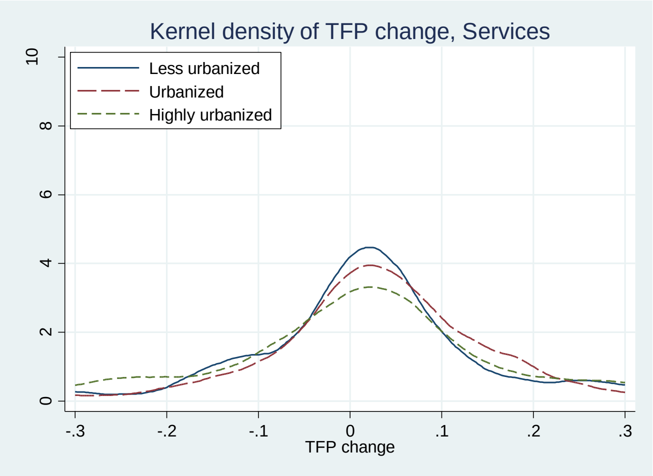


Figure 3: TFP change distributions







APPENDIX 1: Production function coefficient estimates within 2-digit NACE industries

NACE	Parameters		NACE	Parameters		NACE	Parameters		NACE	Parameters	
(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
1	b_l	0.50	15	b_l	0.66	17	b_l	0.63	20	b_l	0.77
	s.e.	0.03		s.e.	0.05		s.e.	0.04		s.e.	0.06
	b_k	0.25		b_k	0.26		b_k	0.24		b_k	0.22
	s.e.	0.02		s.e.	0.03		s.e.	0.03		s.e.	0.02
	b_a	0.01		b_a	-0.05		b_a	0.06		b_a	0.26
	s.e.	0.03		s.e.	0.04		s.e.	0.02		s.e.	0.05
	R^2	1.00		R^2	1.00		R^2	1.00		R^2	1.00
	No	204		No	817		No	145		No	136
21	b_l	0.47	22	b_l	0.60	24	b_l	0.60	25	b_l	0.59
	s.e.	0.06		s.e.	0.02		s.e.	0.02		s.e.	0.09
	b_k	0.34		b_k	0.35		b_k	0.38		b_k	0.25
	s.e.	0.03		s.e.	0.02		s.e.	0.02		s.e.	0.05
	b_a	-0.12		b_a	0.06		b_a	-0.11		b_a	-0.21
	s.e.	0.03		s.e.	0.03		s.e.	0.03		s.e.	0.14
	R^2	0.84		R^2	1.00		R^2	1.00		R^2	1.00
	No	198		No	294		No	404		No	272
26	b_l	0.68	27	b_l	0.59	28	b_l	0.57	29	b_l	0.87
	s.e.	0.04		s.e.	0.06		s.e.	0.03		s.e.	0.03
	b_k	0.21		b_k	0.20		b_k	0.32		b_k	0.15
	s.e.	0.04		s.e.	0.03		s.e.	0.02		s.e.	0.02
	b_a	-0.10		b_a	0.03		b_a	0.04		b_a	0.06
	s.e.	0.04		s.e.	0.05		s.e.	0.02		s.e.	0.02
	R^2	1.00		R^2	1.00		R^2	1.00		R^2	1.00
	No	220		No	188		No	564		No	771
31	b_l	0.62	33	b_l	0.60	34	b_l	0.62	35	b_l	0.87
	s.e.	0.04		s.e.	0.02		s.e.	0.02		s.e.	0.04
	b_k	0.34		b_k	0.33		b_k	0.30		b_k	0.11
	s.e.	0.03		s.e.	0.01		s.e.	0.01		s.e.	0.02
	b_a	-0.09		b_a	0.16		b_a	0.05		b_a	-0.02
	s.e.	0.04		s.e.	0.02		s.e.	0.02		s.e.	0.05
	R^2	1.00		R^2	1.00		R^2	1.00		R^2	0.97
	No	104		No	153		No	141		No	152
36	b_l	0.46	45	b_l	0.86	50	b_l	0.70	51	b_l	0.61
	s.e.	0.02		s.e.	0.03		s.e.	0.04		s.e.	0.04
	b_k	0.44		b_k	0.21		b_k	0.17		b_k	0.13
	s.e.	0.02		s.e.	0.01		s.e.	0.02		s.e.	0.02
	b_a	-0.01		b_a	0.04		b_a	0.04		b_a	0.19
	s.e.	0.02		s.e.	0.04		s.e.	0.04		s.e.	0.06
	R^2	1.00		R^2	1.00		R^2	1.00		R^2	1.00
	No	179		No	1785		No	953		No	3852

APPENDIX 1: Continued

(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
52	b_l	0.64	55	b_l	0.63	65	b_l	0.48	67	b_l	0.58
	s.e.	0.03		s.e.	0.02		s.e.	0.03		s.e.	0.16
	b_k	0.23		b_k	0.15		b_k	0.40		b_k	0.41
	s.e.	0.02		s.e.	0.02		s.e.	0.02		s.e.	0.05
	b_a	0.05		b_a	-0.00		b_a	-0.01		b_a	0.18
	s.e.	0.03		s.e.	0.02		s.e.	0.03		s.e.	0.28
	R^2	1.00		R^2	1.00		R^2	0.99		R^2	0.99
	No	763		No	174		No	254		No	61
70	b_l	0.46	71	b_l	0.42	72	b_l	0.74	74	b_l	0.64
	s.e.	0.02		s.e.	0.03		s.e.	0.04		s.e.	0.04
	b_k	0.42		b_k	0.44		b_k	0.30		b_k	0.23
	s.e.	0.01		s.e.	0.01		s.e.	0.02		s.e.	0.02
	b_a	0.24		b_a	-0.09		b_a	0.05		b_a	-0.08
	s.e.	0.04		s.e.	0.03		s.e.	0.06		s.e.	0.05
	R^2	1.00		R^2	1.00		R^2	1.00		R^2	0.98
	No	164		No	114		No	266		No	569

Note: Reported R^2 statistics and number of observations (No) are from the last step of the estimation algorithm.