

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

Marian Rizov<sup>1,2</sup>, Arie Oskam<sup>2</sup> and Paul Walsh<sup>3</sup> \*

<sup>1</sup> Middlesex University Business School

<sup>2</sup> Wageningen University

<sup>3</sup> University College Dublin

### Abstract

We compute aggregate productivity of three categories of regions, classified by level of urbanization 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, inverted U-shape 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

Address:

M. Rizov

Department of Economics and Statistics

Middlesex University Business School

London NW4 4BT

E-mail: [m.rizov@mdx.ac.uk](mailto: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 financial support from the Mansholt Graduate School of Social Sciences is also 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 new entrants 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, characterized by high density of economic activity, 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 well known formalisation of monopolistic competition of 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 may offset the productivity advantages thus acting as centrifugal forces.<sup>1</sup> 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.<sup>2</sup> 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 impact of 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 still exhibits sufficient diversity in the degree of urbanization. Three main categories of regions can be distinguished according to their level of urbanization and population density. Our approach extends the analysis of labour productivity and productivity growth in Dutch regions by Broersma and Oosterhaven (2009) by applying an advanced TFP estimation technique following modelling ideas in Akerberg et al. (2007) and an application by Rizov and Walsh (2009). We explicitly model unobservable productivity using unique land price data at postcode level and incorporate directly effects of this and other location characteristics into the structural estimation algorithm. The computed measure of TFP is more comprehensive than a labour productivity measure. Our results add robust empirical evidence to the small but growing literature on the limits of agglomeration. In line with Broersma and Oosterhaven (2009) results we find a non-linear effect of agglomeration (density of economic activity) on 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 implement 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 growth rates by the means of decompositions. 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 typology based on degree of urbanization 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, on the basis of population density, 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. 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.

A comparative analysis of main characteristics of the three urbanization categories, for the 2002-2003 period, summarised in Table 1 reveals that employment growth in all three categories was positive as the growth rates were the highest in the less urbanized regions. 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 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.<sup>3</sup> Therefore, next, we focus on the relationship between population density reflecting the strength of agglomeration and land prices and find a nonlinear relationship which is depicted graphically 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 agglomeration on labour productivity growth in the Dutch regions. We need to 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.

- 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 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 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 quite 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).<sup>4</sup> 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.<sup>2</sup> Furthermore, the assumptions of the model that productivity levels are region specific, but not specific either to industries or factors, give the 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 estimate aggregate productivity across urbanization categories and then decompose it into productivity and industry composition effects.

### 3 Estimation framework: Location characteristics and firm productivity

To estimate productivity we employ a semi-parametric estimation algorithm in the spirit of Olley and Pakes (1996) following extensions in Akerberg et al. (2007) and application by Rizov and Walsh (2009). As in Olley and Pakes (1996) we specify a log-linear production function,

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

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,  $l_{jt}$ . Investment demand,  $i_{jt}$  determines the capital stock at the beginning of each period. The law of capital accumulation is given by  $k_{j,t+1} = (1 - \delta)k_{jt} + i_{jt}$ , while age evolves as  $a_{j,t+1} = a_{jt} + 1$ . The error structure comprises a stochastic component,  $\eta_{jt}$ , with zero expected mean, and a component that represents unobserved productivity,  $\omega_{jt}$ . Both  $\omega_{jt}$  and  $\eta_{jt}$  are unobserved, but  $\omega_{jt}$  is a state variable, and thus affects firm's choice variables – decision to exit and investment demand, while  $\eta_{jt}$  has zero expected mean given current information, and hence does not affect decisions.

Because productivity  $\omega_{jt}$  is not observed directly in our data estimating Equation (1) is affected by simultaneity and selection biases. Simultaneity means that estimates for variable (non-dynamic) 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 type 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. We specify productivity of a firm,  $j$ , at a point in time,  $t$  as a function  $\omega_{jt} = h(i_{jt}, k_{jt}, a_{jt}, l_{jt}, r_t)$  of the firm's 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,  $r_t$ , and treat the function non-parametrically in our estimation algorithm. Olley and Pakes (1996) derive the function for productivity by inverting the investment demand function of the firm which itself is a solution to the firm's maximization problem.<sup>6</sup> The economic environment control vector,  $r_t$ , could capture characteristics of the input markets, characteristics of the output market, or industry characteristics like the current distribution of the states of firms operating in the industry. Note that Olley-Pakes formulation allows all these factors to change over time, although they are assumed constant across firms in a given period.

As in Rizov and Walsh (2009) we extend the Olley-Pakes model of (unobservable) productivity in two ways. First, we extend the information content of the economic environment control vector to vary at narrowly defined firm location and denote this by,  $r_{jt}$ , where a subscript index  $j$  is added. Following ideas in Roback (1982) and Ottaviano and Peri (2006) we combine data on wages and rents. This helps disentangle the consumption amenities from the productive advantages in areas with high density of economic activity. For workers, higher wages make them 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 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. Note, however, that this raises an additional concern if looking for agglomeration effects only

through wages. If big cities are associated with both better amenities and higher productivity, the net effect on wages may be ambiguous.

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 at 4-digit postcode level and annual wage at municipality level - in the function of productivity.<sup>7</sup> As argued by Voith (1991), Graham (2009) and others agglomeration effects are capitalised in immobile factor prices, and analysis based on very small spatial units increases the probability of homogeneity of rents within each area.<sup>8</sup> 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. By including a regional control we also address to some extent the problem of spatial autocorrelation.

In line with our purpose to better estimate productivity, introducing location-specific factor price variation in the state space does minimise the deviation from the original Olley-Pakes scalar unobservable assumption, necessary to invert the investment function, and helps the precision of estimates. Second, we relax the scalar unobservable assumption altogether. We model productivity as an exogenous second-order Markov process,  $p(\omega_{jt} | \omega_{jt-1}, \omega_{jt-2})$ , where firms operate through time forming expectations of future  $\omega_{jt}$ s on the basis of information from two preceding periods.<sup>9</sup> The function of productivity can be written as

$$\omega_{jt} = h(i_{jt}, k_{jt}, a_{jt}, l_{jt}, r_{jt}). \quad (2)$$

Next we briefly summarise our estimation algorithm which is outlined in detail in Rizov and Walsh (2009). Substituting Equation (2) into the production function, Equation (1) and combining the constant,  $k_{jt}$ ,  $a_{jt}$ , and  $l_{jt}$  terms into function  $\phi(i_{jt}, e_{jt}, k_{jt}, a_{jt}, l_{jt}, r_{jt})$  gives

$$y_{jt} = \phi(i_{jt}, e_{jt}, k_{jt}, a_{jt}, l_{jt}, r_{jt}) + \eta_{jt}. \quad (3)$$

Equation (3) is the first step of our estimation algorithm and can be estimated as in Olley and Pakes (1996) with OLS and applying semi-parametric methods that treat the function  $\phi(\cdot)$  non-parametrically, using a polynomial.<sup>10</sup> The capital, age, and labour coefficients are identified in the second step of our estimation algorithm. We substitute into Equation (1) the non-parametric function  $g'(\cdot)$  for the unobservable productivity  $\omega_{jt}$  which gives

$$y_{jt} = \beta_k k_{jt} + \beta_a a_{jt} + \beta_l l_{jt} + g'(\cdot) + \varepsilon_{jt}, \quad (4)$$



where  $g'(\cdot) = g'(\hat{\phi}_{jt-1} - \beta_k k_{jt-1} - \beta_a a_{jt-1} - \beta_l l_{jt-1}, \hat{\phi}_{jt-2} - \beta_k k_{jt-2} - \beta_a a_{jt-2} - \beta_l l_{jt-2}, \hat{P}_{jt})$  encompasses the constant,  $\beta_0$  and  $\varepsilon_{jt}$  is a composite error term containing  $\eta_{jt}$ . The lagged  $\hat{\phi}$  variables in  $g'(\cdot)$  are obtained from the first step estimates at  $t-2$  and  $t-1$  periods. Because the conditional expectation of  $\omega_{jt}$ , given information in  $t-2$  and  $t-1$  periods, depends on  $\omega_{jt-2}$  and  $\omega_{jt-1}$ , we need to use estimates of  $\hat{\phi}$  from two prior periods.  $\hat{P}_{jt}$  is propensity score which controls for the impact of selection on the expectation of  $\omega_{jt}$ , i.e., firms with lower survival probabilities which do survive to time,  $t$  likely have higher  $\omega_{jt}$ s than those with higher survival probabilities. We estimate  $\hat{P}_{jt}$  non-parametrically using Probit model with a polynomial approximation. Note that we extend the state variable set with land price information which is an important determinants of firm (entry and) exit decision. Equation (4) is estimated with non-linear least squares (NLLS) estimator, approximating  $g'(\cdot)$  with a polynomial.<sup>11</sup>

Having estimated unbiased and consistent production function coefficients we are then able to back out a unbiased and consistent measure (residual) of total factor productivity (TFP) as  $TFP_{jt} = y_{jt} - \hat{\beta}_k k_{jt} - \hat{\beta}_l l_{jt}$ .<sup>12</sup> In the model of unobservable productivity we have explicitly incorporated spatial and time dependencies by merging spatial interactions with disaggregated modeling of productivity at firm level. In terms of verifying whether variations in location make firms more productive, we have controlled in our model of productivity for market-structure specific shocks (such as demand conditions, factor markets, exit barriers) that are different across locations (municipalities). We note that these factors remain constant across firms in the same location within a given industry and a time period.

#### 4 Data and estimation results

We apply the methodology of estimating production functions developed in Section 3 to the AMADEUS sample of Dutch firms. AMADEUS of the Bureau van Dijk is a comprehensive, pan-European database containing information organised in two modules according to firm size. We use the TOP-1.5-million Module which contains more than 250,000 small, medium and large firms for the Netherlands over the period 1997-2006.<sup>13</sup> For each firm there is detailed information on unconsolidated financial statements, ownership structure, location (by post code) and activity description. 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.<sup>14</sup>

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. The estimated sample accounts for about 51 per cent of the sales and 47 per cent of the employment in our data. After lags are applied and observations with missing values deleted, there are 13,897 remaining observations for 4,220 firms. 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 and employment - 0.94.

The descriptive statistics calculated from the estimated AMADEUS 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 do not correspond to proportionate increase in land price. Interestingly, industry concentration characterised by market share of the top four 2-digit industries (C4) does not show substantial differences across categories of urbanization although higher degree of urbanization seems to be associated with slightly 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.

- Table 2 about here -

Summary of the aggregated coefficients, over the estimated 28 industry production functions, by urbanization category is reported in Table 3. Coefficient estimates from all 28 industry regressions, number of observations and test statistics are available from the authors. The aggregated coefficients on labour, capital and age reported in Table 3 are weighted averages using value added as weight. They confirm the differences across urbanization categories with respect to the shares of labour and capital in output. 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 -

Aggregate productivity measures by urbanization category clearly show that highly urbanized regions are the most productive; the TFP of firms in these regions is 4.382, while it is 3.509 and 3.452 - for firms in urbanized and less urbanised regions, respectively. 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 the three urbanization categories 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 all distributions and confirm the fact that firms in highly urbanized regions are the most productive. When we consider distributions of TFP annual growth rates in Figure 3, however, we see that the highest growth rate is exhibited by firms in urbanized regions rather than in the highly urbanized ones. The Kolmogorov-Smirnov two-sample tests are again significant at the 1 percent level for all distributions and confirm the fact that productivity of firms in urbanized regions grows the fastest. Our tests also show that growth rates in highly urbanised regions are the second highest followed by the growth rates in less urbanized regions. The finding that highly urbanized regions lag behind urbanised regions in terms of TFP growth we interpret as evidence of congestion due to too a high density and degree of agglomeration.

- Figure 2 about here -

- Figure 3 about here -

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

The discussion in previous sections and information reported in Tables 1 to 3 and Figures 1 to 3 suggest that there is a systematic relationship between productivity and the degree of agglomeration as measured by the level of urbanization and density of economic activity. In this section we analyse differences in aggregate productivity across urbanization categories by applying a decomposition of productivity levels and growth rates 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. Spatial variation in aggregate productivity (productivity growth rates) derives from two main sources – differences in the individual firm productivities (growth rates) within each industry, resulting in different average productivities (growth rates) across industries, and differences in the industry composition within each urbanization category.

We calculate the weighted average of individual firm productivities (TFPs),  $q_r^k$ , using firm value added as weight, by urbanization category,  $r$  and industry,  $k$ .<sup>15</sup> The total value added in urbanization category,  $r$  is denoted by  $S_r = \sum_k s_r^k$  and the share of industry,  $k$  in the total value added in category,  $r$  is  $\lambda_r^k = s_r^k / S_r$ . The average productivity of industry,  $k$  for the economy as a whole (i.e., aggregating across all categories,  $r$ ) is given by

$\bar{q}^k = \sum_r s_r^k q_r^k / \sum_r s_r^k$ , while  $\bar{\lambda}^k = \sum_r s_r^k / \sum_r S_r$  is the share of industry,  $k$  in total value added for the economy as a whole. Aggregate productivity,  $q_r$  is weighted average of industry productivities in urbanization category,  $r$ , using industry value added as weight. The aggregate productivity (a) may be decomposed as follows:

$$q_r \equiv \sum_k q_r^k \lambda_r^k = \sum_k q_r^k \bar{\lambda}^k + \sum_k \bar{q}^k \lambda_r^k - \sum_k \bar{q}^k \bar{\lambda}^k + \sum_k (q_r^k - \bar{q}^k)(\lambda_r^k - \bar{\lambda}^k). \quad (5)$$

(a)                      (b)                      (c)                      (d)                      (e)

The first term on the right-hand side of Equation (5) is the average level of productivity in urbanization category,  $r$  conditional on industry composition being the same as for the economy as a whole; we refer to this as *productivity index* (b). The second term is the average level of productivity of urbanization category,  $r$  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,  $r$ . It is important to point out that comparison between productivity and industry composition indexes, while taking into account the residual covariance terms, in Equation (5) can provide useful information about the determinants of aggregate productivity in various urbanization categories. The decomposition of productivity growth rates is analogous to the decomposition of productivity levels described above.

We compute the productivity index and the industry composition index as specified above for the highly urbanized, urbanized and less urbanized regions in the Netherlands and report the results by urbanization category, in Table 4, Panel A. Note that values reported are normalised by the term  $\sum_k \bar{q}^k \bar{\lambda}^k$  from Equation (5). 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 industry (firm) 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 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 category. The last covariance term in Equation (5) 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.

- Table 4 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 21.7 percent, while less urbanized regions are the least productive, with aggregate productivity lower by 23.2 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 regions as the differentials for the productivity index are 10.3 percent and 10.4 percent, while the differentials for the industry composition index are

14.4 percent and 13.8 percent respectively. The magnitudes of the differentials suggest that urbanized and less urbanized regions are characterised by similarly low productivity index but the industry composition index for less urbanized regions is higher than the one for the urbanized regions. The covariance term is positive for all urbanization categories but its magnitude is the largest for the urbanized regions suggesting a substantial unexplained reallocation of industry shares towards more productive industries or alternatively increases in productivity of larger industries. From policy view point, efforts to increase firm and industry productivity, through technological innovation and competition, rather than modify industry composition might be more fruitful in less urbanized regions given the larger scope for improvement in the productivity index compared to the industry composition index.<sup>16</sup>

To explore further the factors affecting aggregate productivity, by urbanization category, we analyse productivity growth rates over the 2000-2006 period by the means of the decomposition indexes defined in Equation (5) 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 in productivity 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 1.3 percent followed by the rates in highly urbanized and less urbanized regions - 1.1 and 1.0 percent respectively. This finding is significant and demonstrates that the congestion forces dominate positive agglomeration forces in the highly urbanised regions. Importantly, our result is similar to findings of Broersma and Oosterhaven (2009). This is also an evidence of urbanized regions catching up with highly urbanized regions in terms of aggregate productivity over the period of analysis.

The sources of aggregate productivity growth vary by urbanization category. 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 insignificant. 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 negative growth in the residual

covariance terms, however, supports the view that some reallocation of industry shares may lead to deteriorating industry composition which is possibly due to expansion of less productive industries or deterioration in productivity of important industries, especially, in both less urbanized and highly urbanized regions.

## **6 Conclusion**

The focus of the paper is on evaluating the impact of agglomeration on productivity in the Dutch regions classified by level of urbanization. We build a structural model of the unobservable productivity incorporating land prices as a 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 at firm level, within 2-digit industries in the Netherlands, for the period 2000 - 2006. We use information on land prices available for 4-digit post codes and allow market structure to differ at very disaggregate municipality (gemeente) level. We model the unobservable productivity as an exogenous 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 TFP measures at firm level.

We aggregate the firm TFPs by urbanization category and find that aggregate productivity systematically differs across highly urbanized, urbanized and less urbanized regions as the magnitudes of the differentials are 21.7 percent and 23.2 percent, respectively. Our results confirm findings of a number of studies that productivity and agglomeration are positively correlated. Further, aggregating productivity growth rates reveals important differences across urbanization categories. The main finding is that there is a tendency of urbanized regions – exhibiting annual growth rate of 1.3 percent - catching up with highly urbanized regions – with annual growth rate of 1.1 percent - in terms of aggregate productivity over the period of analysis. This is an evidence of negative congestion dominating positive agglomeration effects.

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 strongly influenced by agglomeration. The industry composition index captures the extent 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. Most importantly, we find evidence that productivity growth index is

the highest in urbanized rather than in highly urbanized regions pointing to the fact that agglomeration has led to congestion and negatively affected productivity growth at high levels of density of economic activity. Improvements in industry composition index are more important in highly urbanized and less urbanized regions. Thus, in less urbanized regions, in the light of our decomposition results, efforts to increase firm and industry productivity, through technological innovation and within-industry competition, rather than relying on induced changes in industry composition might be more fruitful, given the larger scope for improvement in the productivity index compared to the industry composition index. In highly urbanized regions preventing further agglomeration is likely to lead to sustainable productivity growth.



## Notes

<sup>1</sup> 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.

<sup>2</sup> 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.

<sup>3</sup> The data source for land prices was the Cadastral Land Sales Database that contains information on land transactions, transaction prices, and the location of each parcel sold in the Netherlands. From the Cadastral Land Sales Database we obtained the transaction prices per hectare in 2003.

<sup>4</sup> 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).

<sup>5</sup> 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).

<sup>6</sup> 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 production. However, the monotonicity needed in Olley and Pakes (1996) does not depend on the degree of competition in the output market; it just needs the marginal product of capital to be increasing in productivity.

<sup>7</sup> In terms of Akerberg et al (2007) land price and wage can be seen as additional observed controls of firm investment choices. Alternatively, land price and wage can be treated as state variables.

<sup>8</sup> While the distinction between urbanisation and localisation 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 urbanised 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 urbanisation and localisation effects are separated.

<sup>9</sup> The fixed effect estimator can be seen as a special case of the Markov process  $p(\cdot)$  where productivity,  $\omega_{jt}$  is set to  $\omega_j$  and does not change over time.

<sup>10</sup> Olley and Pakes (1996) show that kernel and polynomial approximations of the unobservable produce very similar results. In our estimations everywhere we use a computationally easier 4<sup>th</sup>-order polynomial. Also note that even though the first step in our algorithm does not directly identify any of the parameters of the production function, it generates estimates of  $\phi(\cdot)$ ,  $\hat{\phi}_{jt}$ , needed in the second stage.

<sup>11</sup> Woodridge (2009) presents a concise, one-step formulation of the original Olley and Pakes (1996) approach using GMM estimator which is more efficient than the standard Olley-Pakes methodology.

<sup>12</sup> Estimating the age coefficient is only used to separate out cohort from selection effects in determining the impact of firm age on productivity and therefore we do not net out the contribution of age from TFP.

<sup>13</sup> The TOP-1.5-million 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.

<sup>14</sup> A number studies (e.g., Katayama et al., 2003; Del Gatto et al., 2008) point that production functions should be a mapping of data on inputs and outputs. However, 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.

<sup>15</sup> 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.

<sup>16</sup> The literature on international (and regional) specialization predicts that general technology (Ricardian) and factor supply (Heckscher-Ohlin) differences jointly determine comparative advantage and thus, specialization, measured as industry composition. Recent papers, starting with Harrigan (1997), show that the estimated impact of non-neutral technology differences is large and in accord with the theory, suggesting that Ricardian effects are an important source of comparative advantage, and thus, a determinant of industry composition.

## References

- Akerberg, 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 spillovers 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 panel 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.

- 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.
- Harrigan, J. 1997. Technology, factor supplies, and international specialization: Estimating the neoclassical model, *American Economic Review*, 87(4), 475-494.
- 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.
- Katayama, H., S. Lu, and J. Tybout 2003. Why plant-level productivity studies are often misleading, and an alternative approach to interference, NBER WP 9617.
- 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.
- Mills, E. 1967. An aggregative model of resource allocation in a metropolitan area, *American Economic Review*, 57(2), 197-210.
- 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.
- 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.
- 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*, 31(2), 127-145.
- Wooldridge, J. 2009. On estimating firm-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
Employment growth 1996-2002, % pa	2.4	2.3	2.7
Population growth 1996-2002, % pa	0.6	0.6	0.8
Economic growth 1996-2002, % pa	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
Income per capita 2002, €	12,000	11,000	10,000
Land price 2003, €/ha *	220,684	139,330	63,876
Population density 2003, number/sq. km	1577	570	231

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

Note: \* Average land price and annual wage are calculated from data for 463 municipalities (gemeenten).

**Table 2: Descriptive Statistics of Firm Specific Variables by Urbanization Category, 2000-2006**

Variable	Highly Urbanized, Mean (S.D.) (1)	Urbanized, Mean (S.D.) (2)	Less Urbanised, Mean (S.D.) (3)
Firm alue added (thousands €)	98,869.2 (272,741.4)	33,345.3 (73,471.6)	23,380.6 (45,145.4)
Firm total fixed assets (thousands €)	149,226.7 (359,107.0)	37,822.1 (85,797.9)	27,072.8 (71,405.4)
Firm real investment (thousands €)	41,774.1 (121,794.7)	8,977.9 (37,686.0)	5,446.2 (14,418.2)
Firm employment (number of full-time-equivalent (FTE) employees)	1,439.9 (4,470.3)	553.9 (1,636.1)	520.2 (1,358.9)
Firm age (years)	33.0 (39.9)	34.4 (32.2)	37.8 (35.0)
Regional wage per FTE employee (€ per year)	18,471.6 (7,614.4)	17,618.4 (8,053.0)	16,654.1 (4,354.9)
Regional land price per hectare (€)	218,017.0 (188,687.7)	141,016.8 (147,012.2)	63,571.3 (50,077.0)
Population density (number per sq. km)	1,544.6 (493.8)	580.3 (204.0)	232.4 (37.8)
Market share of top four industries, C4 (%)	56.1	54.7	51.8
Number of observations (Total 13897)	3555	7988	2354

Source: AMADEUS, BvD

Note: 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 Aggregated by Urbanization Category, 2000-2006**

Coefficient	Highly urbanized	Urbanized	Less urbanized
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 Rsq	0.995	0.996	0.995
Aggregate productivity	4.382 (0.992)	3.509 (0.966)	3.452 (0.979)

Note: The reported coefficients and aggregate productivity are weighted averages, using value added as weight, from 28 industry regressions on firm level data. Standard errors (standard deviations for productivity) are reported in parentheses.

**Table 4: Aggregate Productivity Decompositions by Urbanization Category, 2000-2006**

	(a)	(b)	(c)	(d)	(e)
Panel A: Average levels, 2000-2006					
Highly urbanized	1.092	1.019	1.059	1.000	0.014
Urbanized	0.875	0.916	0.915	1.000	0.044
Less urbanized	0.860	0.915	0.921	1.000	0.024
Panel B: Annual changes, 2000-2006					
Highly urbanized	0.011	0.015	0.022	0.013	-0.013
Urbanized	0.013	0.024	0.004	0.013	-0.002
Less urbanized	0.010	0.014	0.037	0.013	-0.028

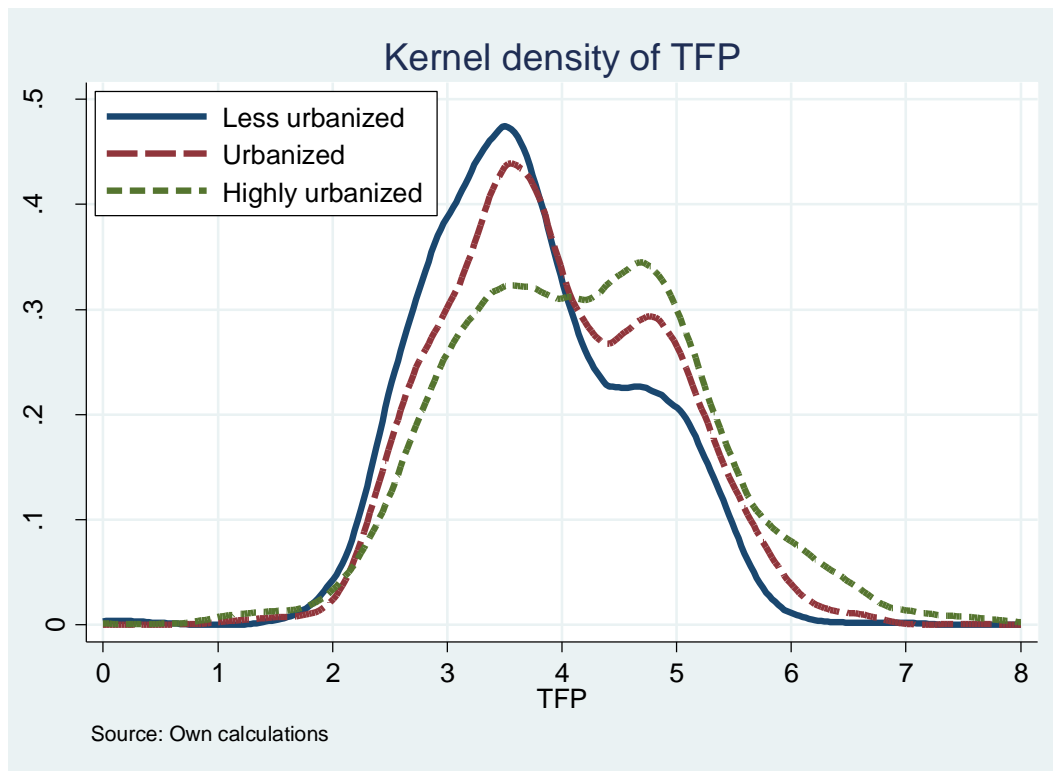
Note: For definitions of decomposition components refer to equation (5) in the text.

Component (d) has a negative sign in the decomposition.

**Figure 1: Non-linear Agglomeration Effects, 2003**



**Figure 2: TFP Distributions by Urbanization Category, 2000-2006**



**Figure 3: TFP Growth Distributions by Urbanization Category, 2000-2006**

