RESEARCH INSTITUTE OF
INDUSTRIAL ECONOMICS

The Economic Microgeography of Diversity and Specialization

Martin Andersson, Johan P Larsson and Joakim Wernberg

The Economic Microgeography of Diversity and Specialization

Martin Andersson$^\varnothing$, Johan P Larsson$^\otimes$ and Joakim Wernberg$^\oplus$

ABSTRACT

As cities increasingly become centers of economic growth and innovation, there is a need to understand their inner workings and organization in greater detail. We use ge-coded firm-level panel data at the sub-city level to assess the long-standing question whether agglomeration economies derive from specialization (within-industry) or diversity (between-industry). We show that these two types of externalities co-exist, but differ in their spatial distribution and attenuation within cities. There are robust positive effects of diversity *and* specialization on firms' TFP growth at the local within-city neighborhood level, especially for firms in high-tech and knowledge-intensive activities. While specialization effects are bound to the local sub-city level, we demonstrate a positive effect of overall diversity also at the city-wide level. The results resonate with the idea that urban economies provide a mix of industrial diversity and specialisation. A location in a within-city industry cluster in a diversified, large city appears to let firms enjoy the benefits of local industry-specific externalities, while reaping the general city-wide benefits of a diversified city.

$^\varnothing$Department of Industrial Economics Blekinge Institute of Technology (BTH), Research Institute of Industrial Economics (IFN), Stockholm and the Swedish Entrepreneurshup Forum. E-mail: martin.andersson@bth.se
$^\otimes$Swedish Entreprenurship Fourm, CIRCLE Lund University and Jönköping International Business School (JIBS), Jönköping University; e-mail: johan.p.larsson@ju.se
$^\oplus$CIRCLE, Lund University; e-mail: joakim.wernberg@circle.lu.se

1. INTRODUCTION

As cities across the world increasingly become focal points of both their national economies and the global economy, there is a growing need for both researchers and policymakers to understand their inner workings in greater detail. A majority of the world's population live in urban areas and an estimated 80% of global GDP is produced in cities (McKinsey Global Institute 2011). Large cities are also increasingly hotbeds for knowledge-intensive and technology-driven startups (Guzman and Stern 2016, Duvivier and Polèse 2016), sometimes described as startup ecosystems. While this emphasizes the role cities play in driving growth and innovation, the current understanding of their inner organisation and workings is still limited.

In the urban economics literature, urban growth is attributed to agglomeration economies driven by microfoundations such as sharing of infrastructure, matching in thick labor markets and learning through the spread of ideas and information (Duranton and Puga 2004). A rich body of empirical studies illustrate that city size and density explain differences in productivity, growth and innovation between cities of different sizes. Recent findings in urban scaling analysis similarly indicate that a variety of economic output indicators grows superlinearly with city size, yielding increasing returns to population size (Bettencourt et al 2007). When it comes to the particulars of urban growth and innovation, cities are still in many respects treated as 'club goods' or black boxes (Rosenthal and Strange 2003). Agglomeration economies are, often implicitly, assumed to operate across entire cities. Yet, interactions between individuals or the flow of information and ideas may still be subject to friction even within cities, for example through geographical distance or links in social or organisational networks.

As cities grow in size, they also become more complex in terms of interactions, flows and networks (Batty 2013). This begs the question of how the inner landscape of urban economies, i.e. their economic microgeographies, relate to their aggregate performance. A growing number of studies move beyond the city-wide picture by employing disaggregated data to explore within-city differences in density externalities (Arzaghi and Henderson 2008, van Soest et al. 2006, Andersson et al 2016, Larsson 2014). The combined picture from this line of research is that there seem to be relevant externalities that operate at spatial scales much smaller than entire cities, and that each urban economy exhibits a rich variation in agglomeration economies that may not be fully captured at the city-wide level alone.

We contribute to this line of inquiry by investigating how industry structure relates to agglomeration economies not only at the city-wide level, but also at sub-city (neighborhood) scale. The issue of how local industry structure influences agglomeration economies is a long-standing and central question in several disciplines including urban economics, innovation studies and economic geography. It is commonly framed as a comparison between industry specialisation in localisation economies and

economic diversity in urbanisation economies (Glaeser et al 1992, Henderson et al 1995). Put differently, do positive density externalities arise within or between industries? Most approaches in this subject rely on inter-city comparisons, and the extant literature is surveyed in amongst others Beaudry and Schiffauerova (2009).

In this paper, we shift focus from comparing cities to comparing spatial scales within cities. We employ highly disaggregated geo-coded firm-level data for the cities of Stockholm, Gothenburg and Malmö, and estimate the simultaneous roles of diversity and specialization economies on the sub-city (neighbourhood) level, as well as on the city-wide level. This method allows for the possibility that different types of externalities can co-exist rather than compete within a city, but also that they may operate at different spatial scales. Our results are consistent with the idea that firms, especially in high-tech and knowledge-intensive industries, benefit from economic diversity at both the local and city-wide levels of analysis, combined with highly localised density externalities derived from industrial specialisation at the sub-city level.

First, we present robust empirical evidence of diversity externalities across specifications, both at the neighborhood and city-wide levels of analysis. Neighborhood-level employment outside the firm's own industry has a positive influence on the firm's TFP. We also show that city-wide diversity exerts a positive and economically and statistically significant effect in most specifications. Second, we document a small but significant effect of specialization externalities at the neighborhood level, but not at the city-wide level. In our baseline regression model, own-industry employment at the neighborhood level has a robust positive effect on firms' TFP, but the same variable defined at the city-wide level is virtually unassociated with productivity. The observed specialization externalities are indeed bound to the local sub-city neighborhood.

In addition, we find no influence of neither diversity nor specialization in first-order neighbors to a given neighborhood. We also show that the positive influence of neighborhood-level specialization on firms' TFP is robust across industries. It is significant for firms in both manufacturing and services, as well as for firms in both high-tech and low-tech industries. We further find that the gains from location in a within-city industry cluster in a diversified city is primarily a story that pertains to larger metropolitan regions and cities. Outside Sweden's main metropolitan regions, we find no robust effect of neighborhood-level specialization. In sparser areas, it is instead general employment size that appears to matter for firms' TFP.

Overall, the results are consistent with the idea that there are highly localised density externalities associated to the concentration of industries at the sub-city neighbourhood level. While the empirical strategy does not explicitly inform about underlying mechanisms or microfoundations, the observed

local specialization effects are consistent with the idea of industry-specific knowledge flows that attenuate sharply with distance. Our results suggest that a location in a within-city industry cluster in a diversified, large city may allow firms to enjoy the benefits of local industry-specific externalities, while also reaping the general city-wide benefits that a diversified city provides. The diversity of a productive and innovative urban economy is then neither spatially random nor homogenous. Instead, it appears to be able to house several highly specialised parts that make for a diversified whole. This resonates with Desrochers (2001), who argues that *"… a diversified city made up of many specialized clusters"* should constitute an environment conducive for productivity and innovation, and that "… this is historically what most important cities have been" (ibid, p.388).

The rest of the paper is organized as follows. Section 2 presents the background and conceptual motivation. Section 3 presents data and provide examples of observed within-city industry clusters. The empirical strategy is presented and motivated in Section 4, while Section 5 reports the results from our baseline regression model as well as the robustness checks. Section 6 concludes.

2. BACKGROUND AND CONCEPTUAL MOTIVATION

Agglomeration economies are often invoked as an economic rationale for the existence of cities, since they are believed to increase the productivity of local firms and workers. Ample empirical evidence supports that productivity is higher in cities. For example, estimates of the overall elasticity between employment density and average wages are generally in the order of 2-7 percent (Andersson et al 2014, Ciccone and Hall 1996, Combes et al 2008, Yankow 2006) and local variations in density adds to the production function of firms (Henderson 2003, Tveteras and Battese 2006, Andersson and Lööf 2011).

While the study of agglomeration economies has a long history, far less is known about their nature and sources. Studies of agglomeration economies often employ highly aggregated data, resulting in the implicit assumption that agglomeration economies are 'club goods' that operate at rather large spatial scales such as metropolitan areas (Rosenthal and Strange 2003). This is also true for the large literature on how local industry structure, and in particular the relative roles of industry diversity and specialization, influence urban growth. In a meta-analysis of 73 research articles, de Groot, Poot and Smit (2016) find that there is significant heterogeneity in results across studies in this vein, even though diversity appears to have a more robust positive influence on urban growth and productivity. The vast majority of studies covered by their meta-analysis follow the tradition of Glaeser et al (1992) and use whole cities, regions or local labor market areas as observational units. Variables that reflect diversity and specialization are then set up to 'compete' in order to determine which one better explains growth or productivity at that spatial scale.

This type of comparison implicitly assumes that both types of externalities operate at a similar spatial scale. It follows that any type of externality that attenuates sharply with distance and is bound to a spatial scale significantly smaller than the unit of analysis is unlikely to be properly identified. For example, consider firms located within a large metropolitan city. They may benefit from city-wide sharing of infrastructure and a wide diversity of business service suppliers throughout the city. Assume that they also benefit from being in the same block or district as other firms from the same industry because it enables locally bound flows of knowledge and information between firms. A city-wide analysis is then more likely to capture the city-wide externalities than the district-specific externalities, especially if only some districts exhibit such specialization externalities.[1] This is partly due to the fact that any aggregate analysis will by definition omit information about the parts being aggregated, in this case within-city location patterns and their characteristics. de Groot et al (2016) indeed conclude from their review of the literature that "conventional lines of inquiry in this literature may now have reached strongly diminishing returns" (ibid, p.776). They call for new analyses that use micro-level data and that pay greater attention to the spatial scope of agglomeration externalities.

A growing number of studies exploit high-resolution geographic data to analyze the distribution and attenuation of agglomeration externalities. These studies show that the the sub-city level is of significant importance. Arzaghi and Henderson (2008) use disaggregated spatial data on advertising agencies on Manhattan in New York and analyze the role of localized within-city spillover effects. They document considerable within-industry local externalities associated with co-location, but these effects dissipate within less than a kilometer. Rosenthal and Strange (2003) study the birth of new establishments with respect to location and initial size in six different industries using data at the ZIP-code level.[2] They find that that new firm entry is positively influenced by employment in the own industry, corresponding to specialization externalities. Also, within the first mile this effect is orders of magnitude larger than when the same variable is measured 2 to 5 miles away. The effects of employment in other industries, capturing diversity or urbanization effects, were less clear-cut. Their conclusion is that "agglomeration should ideally be studied at a much more refined level than has been the norm" (ibid, p. 378). van Soest et al (2006) arrive at a similar conclusion using data on new establishments at the level of within-city ZIP-codes in the Netherlands. Recent analyses also establish that elasticities of wages with respect to employment density varies considerably between different parts of the the city, and that there is a premium associated with denser districts or neighborhoods within cities (Larsson 2014, Andersson, Klaesson and Larsson 2016). Although the literature on within-city attenuation of agglomeration

---

[1] If, theoretically, every block within a city exhibited their own specific specialization externalities, we would in fact observe them as equivalent to diversity externalities at a city-wide level, but as specialization externalities only at the block level.

[2] Zip-code-level data corresponds to a subdivision of metropolitan areas in the U.S.

economies is still limited, the available empirical evidence supports the proposition that there are externalities operating at the sub-city level which play a significant role in agglomeration economies.

Conceptually, highly localized specialization externalities could be assumed to reflect domain-specific knowledge and information externalities. Knowledge and information spillovers are often employed as a theoretical motivation for benefits that arise from close proximity between firms. They are also considered, theoretically and by empirical estimations, to be bound to areas much smaller than the entire city and to depreciate sharply with distance – more so than benefits arising from sharing labor and inputs (Ellison, Glaeser and Kerr 2010, Rosenthal and Strange 2008).

Moreover, recent research suggests that such knowledge spillover phenomena also tend to be domain-specific. The conceptual argument for domain-specificity is that the extent to which knowledge and information spills over between firms become productive depends on the 'economic proximity' between them (cf. Moretti 2004, Boschma 2005, Frenken, van Oort and Verburg 2007). The main idea is that spillovers require some similarity in terms of knowledge bases, competencies or skills, to ensure absorptive capacity (Cohen and Levinthal 1990, Boschma 2005).[3] In other words, just as Williamson (1985, pp. 18-19) claimed that "transaction costs are the economic equivalent of friction in physical systems", proximity between knowledge domains could be thought of as reducing frictions in the transmission of information and ideas. This means that close geographical proximity (for example in terms of intracity co-location) *and* economic proximity (for example in terms of industries) bode for productive information and knowledge flows between firms.

There is growing evidence in favor of this line of argument. For example, Greenstone et al (2010) analyze the extent to which spillover effects from new large plants depend on economic and geographic proximity between new and incumbent plants. They find that incumbents that are geographically and economically linked to the new plant experience larger spillovers relative to incumbents that are geographically close but economically distant.[4] Several studies in the literature on related variety also show that it is not merely industrial diversity but also how related different industries are, i.e. related rather than unrelated variety, that matter for explaining city growth and productivity (Frenken et al 2007, Wixe and Andersson 2016).[5]

---

[3]This is in fact a key conceptual underpinning in the economic geography and regional science literature on *relatedness* and *related variety* (Boschma 2005, Frenken et al 2007, Boschma and Frenken 2011).

[4]To measure economic proximity they use several measures including input and output flows, measures of the degree of sharing of labor pools, and measures of technological linkages

[5]In this paper, we limit economic proximity to comparing the relative roles of within- and between industry employment density in explaining firms' productivity. Although industries constitute a core perspective in the research on specialization and diversity externalities, we recognize that there are arguments in favor of that it is not industries *per se* that matters. Instead, it is the underlying similarities in terms of knowledge bases, labor pools and technologies that are important. As Porter (2003) puts it: "clusters are important because of the externalities that connect the constituent industries, such as common technologies, skills, knowledge and purchased inputs".

Our hypothesis is that externalities pertaining to specialization and diversity co-exist within cities, but that they differ in their attenuation. Localized specialization externalities are likely to operate at the small sub-city scale where firms benefit from close proximity to other firms in the same industry, while diversity externalities are likely to also operate on larger spatial scales, such as entire cities. In this way, localized specialization effects may co-exist with diversity effects. Being located in a specialized district or neighborhood within a large diversified city may allow firms to enjoy the benefits of both local specialization externalities and city-wide diversity externalities. For the purpose of the model and analysis in this paper, we test the existence of density externalities related to diversity and specialisation measured in terms of industry employment, and in particular whether there are differences in how these externalities are expressed at city-wide and sub-city neighbourhood levels respectively.

3. DATA

3.1 A geo-coded firm-level panel dataset

We employ geocoded firm-level panel data for Sweden spanning 14 years (1997-2010). The data are register data (audited by Statistics Sweden) covering the population of single-plant firms in Sweden in both manufacturing and services industries. The data include balance-sheet information such as wages, employment, and book value of capital.

Another data set for the same time period contains geo-coded plant-level data, where each plant is associated with a geocoded cell in a larger country-wide grid of 1×1 km squares. Plants are grouped into squares based on their physical addresses. Henceforth, we refer to these squares as neighborhoods.[6] The area and location of the squares are exogenously determined, meaning that they are unaffected by social or economic conditions of any kind. This alleviates issues of endogeneity with regard to geographic delineation.

By construction of the square grid, each square, n, has eight first-order neighbors. This is illustrated in Figure 1. In the analysis the sub-city scale will be captured by both the level of squares and the level of first-order neighbors. While there are theoretical arguments favoring the sub-city scale as such, it is a prioi unclear what the relvant sub-city scale is. By including neighborhood and first-order neighbors, we test for attenuation of effects at the sub-city scale, i.e. what is the relevant sub-city scale?

Still, industries presumably capture some dimension of economic proximity. However, combining the analysis of disaggregated spatial data with more detailed data on various dimensions of relatedness between industries provides a promising venue for future research.

[6] Not to be confused with a functional or perceived definition of a neighborhood.
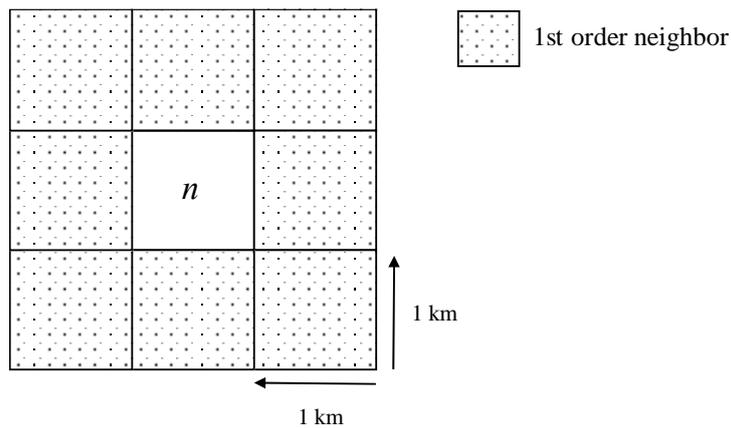
Figure 1. *Squares and neighbors.*

The disaggregated sub-city neighborhood level of analysis is then supplemented by a functional city-region labour market level[7], defined by Statistics Sweden based on commuting patterns. We limit our primary analysis to the labor markets in three main metropolitan regions: Stockholm, Gothenburg and Malmö. In an extension of the anlysis, we include data all labor market regions in the entire country in order to get a tentative indication of how the results vary between large dense city regions and less dense regions.

The plant-level data include a firm-level identifier, assigning each plant to a firm. Since our objective is to analyze the influence of the local environment on firms, we focus our analysis on single-plant firms, i.e. firms for which it is possible to define one unique location as given by the address of their plant. We restrict the analysis to firms with at least one employee in 2-digit NACE industries 15-74, which means that we exclude agriculture, fishing and mining as well as service industries dominated by the public sectors, e.g. schools and healthcare. The resulting dataset comprises 190,114 single-plant firms with a total of 634,114 firm-year observations.

3.2 Industry sub-city clusters

To explore the characteristics and potential of the disaggregated spatial data and the square grid setup, consider the Stockholm region. This region is the largest in Sweden in terms of population. The total area amounts to about 8,000 km^2. There are about 3,500 squares (roughly 44% of the surface area) that belong to Stockholm and that host at least one firm with at least one employee. Furthermore, the variation between squares in terms of standard measures like employment density is sizeable. There are also signficant differences between industries. As an example, Figure 2 illustrates the within-city

---

[7]Local labor market regions constitute a common unit of analysis in studies of spatial externalities (see e.g. Beaudry and Schiffauerova 2009).

employment densities for financial services and construction, respectively, in all neighborhoods with at least 10 employees per km^2. It is evident that although both sectors are centrally clustered, the distrbution and concentration differs notably between them. This in turn implies large variations beyond aggregated size in the characteristics of the local environment for firms located in in different parts of the same city.

Banking (SIC 65)   Construction (45)
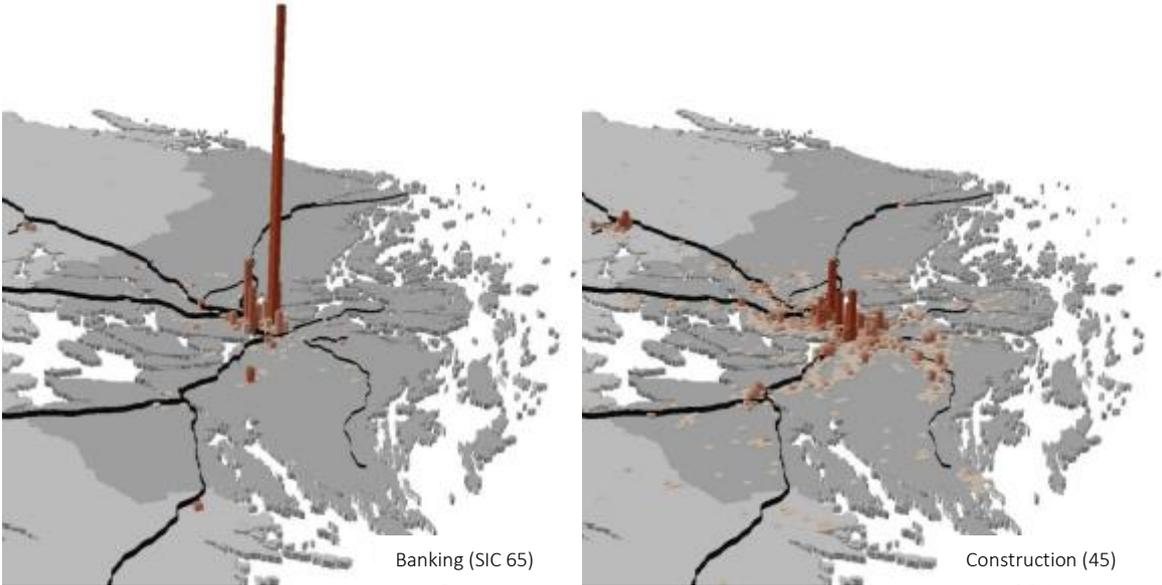
Figure 2. Neighborhood level concentration per industry in the Stockholm labor market region (blue shaded area). The selected industries are banking and financial services (65) and construction (45).

The fact that firms in similar industries cluster together in the city is not surprising. Standard bid-rent models in urban economic theory based on von Thünen (1826) and Alonso (1964) suggest that, to the the extent that they have similar willingness to pay for office space or land, firms in similar industries will have similar bid-rent curves. Therefore, they will locate in similar areas in the city (see e.g. Fujita 1989). However, our focus is not on location patterns per se, but rather on the effects on productivity

We test empirically if firms benefit from proximity to other firms within and outside of its industry, respectively. If cities are fully integrated so that any externalities are evenly spread or exhibit small or no decay inside cities, then within-city location will not affect a firm's producitivty. For instance, in the special case of zero attenuation within the city, sub-city location becomes irrelevant and firms located outside the central business district or outside a within-city industry cluster will enjoy the same producitivty benefits as those located within those areas. If, on the other hand, there are specialization externalities that are unevenly spread and that exhibit sharp attenuation even at small distances within cities, we would expect own-industry clustering at the sub-city neighborhood level to play an important role.

9

4. EMPIRICAL STRATEGY AND VARIABLES

4.1 Empirical model

We estimate a production function where firm-level total factor productivity (TFP) is modeled as a function of diversity and specialization indices at the sub-city neighborhood and city-wide levels. We control for a wide set of pertinent firm attributes, including firm fixed effects. In addition, we undertake a host of robustness tests by examining differences in the relative roles of specialization and diversity at the given spatial scales between firms in different industries as well as for firms in sparser, non-metropolitan areas.[8]

The baseline is a Cobb-Douglas production function:

$$(1) \qquad Y_{its,nR} = A_{its,nR} L_{it}^{\alpha} K_{it}^{\beta}$$

where $Y_{its,nR}$ denotes the value-added of firm i in year t operating in sector s. Subscript n refers to the neighborhood the firm is located in and R the wider city-region. L_{it} is the labor force and K_{it} the capital stock of the same firm i in year t. $A_{its,nR}$ is the total factor productivity (TFP). Dividing both sides by $L_{it}^{\alpha} K_{it}^{\beta}$ and taking logs leave us with a simple expression for TFP:

$$(2) \qquad \ln(A_{its,nR}) = \ln\left(Y_{its,nR} \Big/ L_{it}^{\alpha} K_{it}^{\beta} \right)$$

We then model the total factor productivity of each firm as a function of eight components (see equation 3). The first two are specialization and diversity at the level of the neighborhood, i.e. the level of 1×1 km within-city squares. We denote these by $N_{ts,n}^{SPEC}$ and $N_{ts,n}^{DIV}$. The next two are equivalent indices, but measured at the level of first-order neighbors (see figure 1). We denote these by $N(1)_{ts,n}^{SPEC}$ and $N(1)_{ts,n}^{DIV}$. We also include specialization and diversity at the level of the wider city-region, $R_{ts,R}^{SPEC}$ and $R_{ts,R}^{DIV}$. This formulation implies that externalities are assumed to shift the production function of individual firms. We also assume that the total facor productivity depends on firm-, sector- and year-specific effects, contained in matrix $\mathbf{X}$. The final component, $\varepsilon_{its,nR}$, captures unobserved stochastic productivity shocks, $\varepsilon_{its,nR} \sim IID(o, \sigma_{\varepsilon}^2)$. The result is an estimation equation with which we can assess the respective influence of diversity and specialization at the two spatial levels:

---

[8]For example, the meta-analysis by de Groot et al (2016) underscore the need to account for such heterogeneity.

$$\ln\left(A_{its,nR}\right) = \phi_1 \ln\left(N_{ts,n}^{SPEC}\right) + \phi_2 \ln\left(N_{ts,n}^{DIV}\right) + \gamma_1 \ln\left(N(1)_{ts,n}^{SPEC}\right) + \gamma_2 \ln\left(N(1)_{ts,n}^{DIV}\right) + ..$$

(3)

$$.. \theta_1 \ln\left(R_{ts,R}^{SPEC}\right) + \theta_2 \ln\left(R_{ts,R}^{DIV}\right) + X'\boldsymbol{\beta} + \varepsilon_{its,nR}$$

To test our hypothesis we compare the estimates of the parameters associated with the variables reflecting neighborhood and city-wide externalities, i.e. $\phi_1, \phi_2, \gamma_1, \gamma_2$ and θ_1, θ_2. For instance, if $\phi_1 > 0$, $\theta_2 > 0$ and other parameters are zero, the results are consistent with specialization effects at the level of own local sub-city neighborhoods, and diversity effects at the level of the wider city. If, on the other hand, $\phi_1 = \phi_2 = \gamma_1 = \gamma_2 = 0$ but $\theta_1 > 0$ and $\theta_2 > 0$, then the neighborhood level matters little for either type of externality. Instead, both specialization and diversity would appear to operate at the level of the wider city. In principle, all estimated parameters may come back positive, indicating that both types of effects operate at all spatial resolutions. In this case, the magnitude of order of the estimated parameters may still be different at the different spatial scales, which inform about the extent to which the strength of effects are related to proximity within cities.

3.2 Variables

Total Factor Productivity (TFP)

Following Martin, Mayer and Mayneris (2011) we estimate TFP by the semi-parametric technique developed by Levinsohn and Petrin (2003), henceforth LP. In a survey of different methodologies to compute firm-level TFP, van Beveren (2012) argues that semi-parametric methods, such as LP, are preferable in general. He also notes that in many cases the TFP estimates derived from different methodologies are highly correlated: "simple correlations between the different TFP measures generally amount to more than 0.80" (ibid, p.102).

In the computation of TFP, LP's method uses intermediate inputs to control for correlations between input levels and unobserved firm-specific productivity.[9] It is designed to deal with simultaneity and selection bias in the estimation of production functions at the firm level. TFP is computed as the residual of a Cobb-Douglas production function, where firms' value-added is the dependent variable.[10]

---

[9] This is different from Olley and Pakes (1996) who use investments as a proxy to control for the part of the error term correlated with inputs

[10] We implement the procedure using the LEVPET command in STATA (Petrin et al 2004), with labor input as the freely variable input and intermediate inputs as proxy variable.

Specialization

To capture externalities associated with specialization, we follow Martin et al. (2011) and compute, for each firm, the number of employees working in the same industry but outside of the firm, for each neighborhood square and for the regional labor market, respectively. Industry is here taken to mean the same 2-digit NACE code, i.e. s refers to a unique 2-digit NACE code.[11] For firm i in sector s, in year t and located in neighborhood n our measure of neighborhood specialization $N_{its,n}^{SPEC}$ is:

$$(4a) \qquad N_{its,n}^{SPEC} = EMP_{ts,n} - L_{it} + 1$$

where $EMP_{ts,n}$ is the total number of employees in sector s in neighborhood n in year t.[12] L denotes labor force in firm i. Note that at the neighborhood level, the uniform 1km^2 square grid implies that a direct measure of employment in sector s is also an exact measure of employment density per km^2 in that sector.

Specialization at the level of first-order neighbors for neighborhood is calculated by summing the number of employees working in the eight first-order neighbors (see figure 1). For firm i in sector s, in year t and located in neighborhood n our measure of first-order neighborhood specialization, $N(1)_{its,n}^{SPEC}$, is:

$$(4b) \qquad N(1)_{its,n}^{SPEC} = EMP_{ts,n(1)} + 1$$

where $EMP_{ts,n(1)} = \sum_{n*=1}^{8} EMP_{ts,n*}$, i.e. the sum of employees in sector s in the eight squares that are first-order neighbors to neighborhood n. These neighbors are denoted by $n*$.

At the city-wide regional level the measure aimed at specialization externalities is defined as the number of employees in the same sector but outside the neighborhood square in which firm i is located as well as outside the first-order neighbors of that square. This provides an on-average separation between localized (sub-city scale) and wide-spread specialization externalities affecting firm i.[13] This translates to the following expression for firm i in sector s and neighborhood n at time t:

---

[11] Each 2-digit NACE industry often comprises a plethora of firms and activities. Firms in a given two-digit industry may employ different business models, focus on different customers or sub-markets and employ workers with different education and experience profiles, and so on. Therefore, this measure of specialization is likely to at least to some extent capture within-industry diversity (cf. Frenken et al. 2007).

[12] In the previous section we stated that our empirical analysis on single-plant firms for which we can identify a unique address in space. While this is true for the sample of firms used in the econometric analysis, the variables reflecting the external environment of firms are based on 'complete' data and thus include all activities. For example, $Emp_{st,n}$ include employment of plants that do belong to firms with many plants in different locations.

[13] The assumption is somewhat crude in that it assumes that all firms located in the middle of their square, making it an on-average separation rather than a strict one.

(4c) $$R_{ts,R}^{SPEC} = Emp_{ts,R} - Emp_{ts,n} - EMP_{ts,n(1)} + 1$$

City-region refers to local labor market regions, typically composed of a main city with surrounding smaller areas' labor markets. We employ the definition of local labor markets defined by Statistics Sweden that distinguishes between 81 regions.[14]

These two measures of specialization externalities both capture the classic idea of localization economies: They are *external to the firm but internal to the industry and the geographic area* and that depend on the *absolute scale of the industry* in the area (Ohlin 1933, Hoover 1937). While many studies use relative measures to assess specialization (see e.g. Beaudry and Schiffauerova 2009), there are in fact several arguments to use more straightforward absolute measures (Kemeny and Storper 2015). Absolute measures come closer to the core theoretical arguments, which emphasize scale. For small spatial units, relative measures can in principle be high even though the underlying scales of the industries in the regions are very limited. Empirically, shares or ratios are also more difficult to interpret since change may be driven by denominator or numerator.[15] Changes over time in absolute measures, as in (4a), (4b) and (4c), are also easier to interpret.

Diversity

Diversity externalities at the neighborhood level are defined as the total employment outside the industry that a firm belongs to. For firm *i* in sector *s*, in year *t* and located in neighborhood *n*, $N_{ts^*,n}^{DIV}$, is given by:

(5a) $$N_{ts,n}^{DIV} = Emp_{t,n} - Emp_{ts,n} + 1$$

where $Emp_{t,n}$ is total employment in neighborhood *n* in year *t*. This is thus a measure of the scope of economic activity outside the own industry, in the neighborhood where firm *i* is located.

At the level of first-order neighbors, the equivalent diversity measure is:

(5b) $$N(1)_{ts,n}^{DIV} = Emp_{t,n(1)} - Emp_{ts,n(1)} + 1$$

---

[14]Local labor market regions are delineated based on commuting flows and consists of municipalities that form an integrated labor market.

[15]This is an issue that is particularly relevant in studies employing location quotients (LQ) to measure localization economies. LQs are in essence the ratio of two ratios, making interpretation of changes very difficult.

where $Emp_{t,n(1)}$ is total employment in the eight first-order neighbors surrounding neighborhood n and $Emp_{ts,n(1)}$ is total employment in sector s in the same set of neighbors to neighborhood n.

Diversity externalities at the city-wide level are defined as the total employment outside of firm i's industry in the city-region, excluding those in the firm's own neighborhood and its eight neighbors. Excluding employment in the neighborhood where firm i is located and in its first-order neighbors again allows us to make a distinction between neighborhood and city-wide externalities. This translates to the following expression:

$$(5c) \qquad R_{ts,R}^{DIV} = Emp_{t,R} - Emp_{ts,R} - Emp_{t,n(1)} - Emp_{t,n} + 1$$

This measure thus reflects the extent of economic activities outside of a firm's own industry, but within the same wider region beyond the neighborhood level. The measures in (5a), (5b) and (5c) reflect the idea of urbanization economies, namely scale economies that *are external to the firm but internal to the geographic area*. They are not bound to any specific industry. If diversity effects dominate, then we should observe that it is not necessarily the scale of the industry that matters, but rather the scale of economic activities in general. This is what (5a), (5b) and (5c) are intended to capture by representing the scope of other-industry economic activities.

At the city-region level, we follow Martin et al (2011) and also include an index of the relative distribution of employment across industries.[16] For firm i in sector s, in year t and located in city-region R this measure of industry diversity is given by:

$$(6a) \qquad Div_{ts,R} = {}^{1}\!/\!_{Herf_{ts,R}}$$

where $Herf_{ts,R}$ is defined as:

$$(6b) \qquad Herf_{ts,R} = \sum_{s' \neq s} \left({}^{Emp_{ts',R}}\!\big/\!_{Emp_{t,R} - Emp_{ts,R}} \right)^2$$

This gives the sum of of employment in other industries as shares of the total employment in industries other than s. For each firm, $Div_{ts,R}$ is the inverse of a Herfindahl-index. This is a scale-independent indication of the regional diversity in industries *outside* the 2-digit industry in which it operates. This measure of 'extra-industry diversity' complements the other measures by focusing on distribution of

---

[16]Because many neighborhoods only have a few sectors, we only calculate this index at the level of the wider region.

employment across industries, rather than the total scale of employment outside a firm's main industry. The higher the value of $Div_{ts,R}$, the greater the extra-industry diversity.

Control variables

The matrix $\mathbf{X}$ in (3) contains firm-, industry-, year-, and sector-year effects assumed to influence *A* for each firm. More precisely, we first include a measure of human capital in the firm defined as the share of employees with a university education of at least three years relative to the total number of employees. It is well established in the literature that skilled workers are important for firms' capacity to develop as well as to absorb new technologies, i.e. a capacity crucial for firm productivity (Bartel and Lichtenberg 1987, Cohen and Levinthal 1990). Second, as there are many unobserved factors that could influence firms' productivity, such as managerial practices and routines, we also account for time-invariant unobserved firm heterogeneity through firm-specific fixed effects. Being able to account for unobserved heterogeneity among firms is one of the main virtues of panel data (Hsiao 2002). The fixed effects estimation alleviates the risk that our estimates of the influence of external scale economies at the neighborhood and region level reflect that 'better' firms sort themselves to more central neighborhoods and regions. For instance, recent work on agglomeration economies with heterogeneous firms point to spatial sorting of firms and workers as a both theoretically and empirically important explanation for different levels of productivity across regions (Baldwin and Okubo 2006, Combes et al 2008). We further include year and industry, as well as year-industry effects. Year and industry effects are intended to capture trends over time, and sector-specific characteristics relevant for productivity, respectively. The year-industry effects capture industry-specific shocks over time.

Descriptives

Basic descriptive statistics for the variables included in the empirical analysis are presented in Table 1. These figures are based on 190,105 single-plant firms operating in 2-digit NACE sectors 15-74. The table also reports basic information on how the single-plant firms in the sample are distributed across broad industry categories and size-classes of firms in terms of employees.

About 12 % of the single-plant firms are active in manufacturing sectors, and almost 50 % of the firms are active in sectors that are classified as High-tech manufacturing or knowledge-intensive services.

Table 1. Descriptive statistics, panel data 1997-2010

	Mean	Std. deviation
TFP (log)	5.82	0.77
Fraction of skilled workers	0.26	0.37
Specializaton neighborhood (log)	4.39	2.36
Diversity neighborhood (log)	6.96	1.73
Specialization in the eight 1st-order neighborgs (log)	2.56	2.75
Diversity in the eight 1st-order neighborgs (log)	6.23	1.92
Specialization city-wide (log)	10.05	1.26
Diversity city-wide (log)	12.73	0.66
Extra-industry diversity city-wide (log)	2.73	0.08
Manufacturing	10.8 %	
High-tech or knowledge-based industries	48.9 %	

Note: The table reports descriptive statistics for 190,105 single plant firms in 2-digit NACE sectors 15-74 with at least one employee for years 1997-2010 located in Stockholm, Göteborg or Malmö local labor market region (metropolitan area). The numbers are based on a total of 632,032 firm-year observations. Manufacturing refers to 2-digit NACE sectors 15-36. The rest is services. High-tech or knowledge-based industries refers to sectors classified as High-tech, medium-high-tech manufacturing or knowledge intensive services after Eurostat's *High-tech aggregation* by NACE rev 1.1.

3.3 Estimation

The full empirical model including all variables discussed above is presented in equation (7), where λ_i denotes the time-invariant firm-specific effect and H_{it} is human capital as defined above. We estimate the model with the panel fixed effects (FE) estimator meaning that the parameters are based on a within-transformation (variables are expressed in terms of deviations from their firm-mean values). Parameters are thus identified based on changes in variables over time. Consequently, any permanent heterogeneity at the firm-, industry-, neighborhood- or region-level is wiped out.[17]

---

[17]The parameters for the industry dummies, for example, will be identified based on those firms that change industry over time.

$$(7) \quad \ln A_{its,nR} = \underbrace{\delta + \lambda_i + \delta H_{it}}_{\textit{Firm-level characteristics}} + \underbrace{\phi_1 \ln N_{its,n}^{SPEC} + \phi_1 \ln N_{ts,n}^{DIV}}_{\textit{Neighborhood external scale economies}} + \cdots$$

$$\underbrace{\cdots \gamma_1 \ln N(1)_{ts,n}^{SPEC} + \gamma_2 \ln N(1)_{ts,n}^{DIV}}_{\textit{First-order neighbors external scale economies}} + \underbrace{\theta_1 \ln R_{ts,R}^{SPEC} + \theta_2 \ln R_{t,R}^{DIV} + \theta_3 \ln DIV_{ts,R}}_{\textit{City-wide level external scale economies}} + \cdots$$

$$\cdots + \underbrace{\sum_{t=1}^{T} \gamma_t D_t + \sum_{s=1}^{S} \sigma_s D_s + \sum_{s=1}^{S}\sum_{t=1}^{T} \varphi_{st}(D_s \times D_t)}_{\textit{Industry, year and industry-year effects}} + \varepsilon_{it}$$

Estimation of the model in (7) with the FE estimator implies that the parameters for our five key variables reflecting diversity and specialization at different spatial scales are identified from how their changes over time correspond to changes in firm-level TFP. A positive estimate of ϕ_1, for example, suggests that an expansion of the number of employees in an industry in a given neighborhood is associated with an increase in the productivity of firms belonging to that industry and neighborhood.

This empirical design identifies the parameters of the variables in our model based on short-run year to year variations. Long-run effects may also contribute to externality effects, but these are excluded in this analysis for methodological reasons. There is a trade-off between capturing more long-run (level) effects and accounting for issues of for instance spatial sorting and endogeneity over a longer time span. For example, cross-sectional analyses may capture long-run effects of a variable in the sense that it captures effects accumulated over all years preceding the year of observation (Martin et al. 2011), but it would require the use of instruments and an extensive set of controls to fully deal with sorting and endogeneity. Conversely, such a setup would provide little information about short-run effects and variations. In this analysis, we have chosen to prioritise short-run variations primarily because we study highly heterogeneous firms and location characteristics, but also because it adds to our understanding of more immediate effects of location on firm performance. We also lack appropritate instruments and controls.

5. RESULTS

5.1 Baseline model

Table 2 reports the estimated parameters from (7), using the full dataset of metropolitan single-plant firms 1997-2010 operating in 2-digit NACE industries 15-74. The baseline model is the full model with all variables. The table also presents four other model specifications where we test the robustness of the baseline results to changes in the spatial levels that we choose to represent in the model.

Starting with the baseline results, we see that, at the neighborhood level, specialization as well as diversity are positive and significant at the one percent level. The estimated parameters imply that a one standard-deviation increase (Table 1) in own-industry employment at the neighborhood level is associated with TFP increases of 1.3 percent. A similar change in employment outside a firm's own industry is associated with an increase in TFP of approximately 1.2 percent. We thus find evidence consistent with both specialization and diversity externalities operating simultaneously at the sub-city neighbourhood level inside cities. We find no influence of neither diversity nor specialization in first-order neighbors. At the city-wide level the estimated effect of own-industry employment as well as employment outside a firm's own industry is not statistically significant. However, diversity measured as more evenly spread employment across industries in the city as a whole appears to be conducive for firms' TFP. A one standard deviation increase in extra-industry diversity is on average associated with TFP growth of 2 percent.[18]

Combined, the results suggest that firms can benefit from being located in a more specialised neighborhood within a diversified city. Put differently, the diversity of an urban economy may contain specialised geographical subunits that benfit firms located within such sub-industry clusters. In addition, the results strengthen the assumption that there are externalities that are captured at the neighbourhood level but not at the aggregate city-wide level. The fact that first-order neighbor varibles are insignificant suggests that the relevant sub-city is indeed rather local. Although our estimates are uninformative about the underlying mechanisms, they are consistent with, for instance, localized knowledge and information flows occurring among firms in similar *and* different industries, to the extent that they are closely co-located. In summary, our baseline model supports a positive effect on region-level diversity, but also neighborhood level effects of both specialization and diversity. Also, human capital has the expected

---

[18]It should be noted that since we thus far focus the analysis on short-term variation in the three metropolitan areas in Sweden, and since our data contain few firms that move between these areas, the within-variation in the region-level variables is somewhat limited. As we will see below (Table 4), the results are different when we include observations from the entire country in the analysis.

sign and is statistically and economically significant. Increasing fractions of skilled workers in firms are associated with rising levels of TFP, all else equal.

Table 2. Estimated influence of neighborhood and region diversity and specialization on productivity of single-plant firms in metropolitan areas 1997-2010, panel FE estimates. Dependent variable: Levinsohn and Petrin (2003) TFP.

	Baseline	I	II	III	IV
Firm-level characteristics					
Human capital (fraction of skilled workers)	0.0548***	0.0551***	0.0551***	0.0548***	0.0551***
	(0.00624)	(0.00625)	(0.00624)	(0.00624)	(0.00624)
Neighborhood					
Specialization, emp. within 2-digit industry (log)	0.00524***	0.00536***		0.00822***	
	(0.00112)	(0.00112)		(0.000902)	
Diversity, emp. outside 2-digit industry (log)	0.00635***	0.00646***			
	(0.00156)	(0.00155)			
1ˢᵗ order neighbors					
Specialization, emp within 2-digit industry (log)	-0.000256	-0.000239	-0.000220	-6.79e-05	0.000174
	(0.000308)	(0.000307)	(0.000308)	(0.000288)	(0.000287)
Diversity, emp outside 2-digit industry (log)	0.000733	0.000693	0.00220***		
	(0.000633)	(0.000631)	(0.000611)		
City-wide					
Specialization, emp. within 2-digit industry (log)	-0.00195		-0.00362	-0.000520	3.11e-06
	(0.00655)		(0.00653)	(0.00574)	(0.00574)
Diversity, emp. outside 2-digit industry (log)	0.0119		0.0173		
	(0.0131)		(0.0131)		
Extra-industry diversity (log)	0.249***		0.247***	0.274***	0.279***
	(0.0499)		(0.0499)	(0.0448)	(0.0447)
Firm fixed effects?	Yes	Yes	Yes	Yes	Yes
Year effects?	Yes	Yes	Yes	Yes	Yes
Industry effects?	Yes	Yes	Yes	Yes	Yes
Year-industry effects?	Yes	Yes	Yes	Yes	Yes
No. firm-year obs.	632,032	632,032	632,032	632,032	632,032
No. firms	190,105	190,105	190,105	190,105	190,105
R^2	0.05	0.05	0.05	0.05	0.05

Note: The table reports the estimated influence of neighborhood- and city-level diversity and specialization on TFP among single-plant firms in three Metropolitan areas in Sweden (Stockholm, Göteborg and Malmö) operating in 2-digit NACE industries. Standard errors are presented within brackets. Variables are defined in section 3.2. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

To test whether the baseline results hold up to different specifications, model I excludes the city-wide level whereas model II excludes the neighborhood level. This allow us to test whether results are sensitive to which spatial scales we choose to be represented in the model. Model I shows that results regarding the local neighborhood level do not change if we exclude the city-wide level. Model II shows that specialization at the first-order neighbords as well as at the city-level is insignificant even if we exclude the the local neighborhood level. This suggest that specialization externalities are indeed local and operate at a fine sub-city scale. Methodologically, this implies that studies at aggregate spatial levels, such as entire cities, are likely to not be able to identify specialization effects that operate at the local sub-city level. Model II also shows that the diversity effect at the first-order neighbor level is significant once we exclude the neighborhood level. However, in this case we see that the magnitude of order of the estimated elastictity is much lower than the diversity effect at the neighborhood level in the baseline estimation. This suggests that diversity effects at the sub-city scale dissipate rather quickly with distance (cf. Andersson, Klaesson and Larsson 2016).

In models III and IV we delete the diversity measures in terms of employment outside the industry in which a firm is operating. We then ask if the localized nature of specialization effects hold up. The results in the models show that they do. Even if we only consider the employment scale in the same industry, specialization, we see that we only find a statistically significant effect at the neighborhood level. Even if we exlude the neighborhood level (model IV), it is still the case that specialization is insignificant at the level of first-order neighbors and at the city-wide level. We infer that the main results in the baseline model are robust to alternative specifications and do not depend on the spatial scales represented in the model.

We now proceed to analyze the robustness of the baseline model across industries and whether results depend including the entire country in the regressions, as distinct from the metropolitan areas only.

5.2 Heterogeneity across industries and types of regions

Previous research indicates that there are often significant differences in relative importance of specialization and diversity across industries as well as between cities and regions of different densities and stages of development (e.g. Marrocu, Paci and Usai 2013). This is also one of the conclusions from de Groot's et al (2016) meta-analysis. To further probe the baseline results in Table 2, we therefore perform separate estimations for firms in different industries and for firms located in high- and low-density city regions, respectively.

Industry

Conceptually, it is often argued that diversity is more relevant for high-tech and knowledge-intensive industries that may draw on cross-fertilizations between industries to boost their innovation activities and ideas (Duranton and Puga 2001, Feldman and Audretsch 1999). Many studies using aggregated data also find that (region-level) diversity is more important in knowledge-intensive industries, whereas specialization appear to matter more in low-tech manufacturing industries (de Groot et al 2016, Beaudry and Schiffauerova 2009). Theses analyses generally use data at a higher level of aggregation than that employed in our analysis, and analyses such as Arzaghi and Henderson (2008) indicate that neighborhood-level specialization certainly can be beneficial in knowledge-intensive industries. To test whether we find such patterns in our empirical context with Swedish firm-level metropolitan data at the sub-city and city-wide scale, we run separate regressions for manufacturing and services firms, as well as distinguish between high-tech knowledge-intensive industries and low-tech industries with low knowledge-intensity. We report results from these estimations in Table 3.

The results suggest rather small differences between high-tech and low-tech industries regarding the relative importance of diversity and specialization at the neighborhood level. Looking at the effect of a one standard deviation increasin in own-industry neighborhood-level employment, however, local specialization effects appear to be more important for services and in high-tech and knowledge-intensive industries. For high-tech and knowledge-intensive industries, as well as for services in general, city-wide diversity in terms of distribution of employment across industries is positive and significant. It is primarily for these industries that it is not only the sub-city scale that matters, but also the diversity of the city at large.

In all industries, employment outside, as well as within, a firm's own industry is positive and significant at the neighborhood level. At the level of first-order neighbors and the wider city-region the estimated influence of employment in the same industry as well as outside the industry a firm operates in is statistically significant. This confirms the baseline in Table 2.

Table 3. Estimated influence of neighborhood and region diversity and specialization on productivity of single-plant firms 1997-2010, panel FE estimates. Dependent variable: Levinsohn and Petrin (2003) TFP.

	High-tech and knowledge intensive	Low-tech and low knowledge intensity	Manufacturing	Services
Firm-level characteristics				
Human capital (fraction of skilled workers)	0.0900***	0.00212	0.0412**	0.0556***
	(0.00849)	(0.00808)	(0.0197)	(0.00662)
Neighborhood				
Specialization, emp. within 2-digit industry (log)	0.00527***	0.00277**	0.00728***	0.00484***
	(0.00166)	(0.00111)	(0.00212)	(0.00133)
Diversity, emp. outside 2-digit industry (log)	0.00855***	0.00651***	0.0116***	0.00602***
	(0.00223)	(0.00168)	(0.00377)	(0.00176)
1st order neighbors				
Specialization, emp within 2-digit industry (log)	-0.000644	0.000412	0.000139	-0.000249
	(0.000427)	(0.000383)	(0.00116)	(0.000324)
Diversity, emp outside 2-digit industry (log)	0.000112	0.00108*	0.000500	0.000771
	(0.000910)	(0.000646)	(0.00150)	(0.000694)
City-wide				
Specialization, emp. within 2-digit industry (log)	-0.00138	0.00158	-0.0176	0.00336
	(0.00505)	(0.00422)	(0.0111)	(0.00872)
Diversity, emp. outside 2-digit industry (log)	0.0202	0.00465	0.0460	0.00733
	(0.0133)	(0.00900)	(0.0316)	(0.0153)
Extra-industry diversity (log)	0.243***	-0.00856	0.155	0.258***
	(0.0583)	(0.0341)	(0.113)	(0.0557)
Firm fixed effects?	Yes	Yes	Yes	Yes
Year effects?	Yes	Yes	Yes	Yes
Industry effects?	Yes	Yes	Yes	Yes
Year-industry effects?	Yes	Yes	Yes	Yes
No. firm-year obs.	322,45	419,368	73,831	558,201
No. firms	93,882	121,756	17,453	174,207
R^2	0.03	0.07	0.06	0.04

Note: The table reports the estimated influence of neighborhood- and city-level diversity and specialization on TFP among single-plant firms in three Metropolitan areas in Sweden (Stockholm, Göteborg and Malmö) by industry group. Standard errors are presented within brackets. Variables are defined in section 3.2. *** p<0.01, ** p<0.05, * p<0.1

High- and low-density regions

Previous literature has documented differences between high- and low-density regions. de Groot et al (2016) conclude that studies of less densely populated as well as less developed cities/regions typically find that specialization is more important than elsewhere. For example, Marrocu et al (2013) compare the role of diversity and specialization externalities between regions in old and new European countries.

22

They find that diversity matters more in dense areas in old Europe, especially in knowledge intensive industries, whereas specialization is more important for less advanced industries in regions in the new Europe. As an extension of our analysis, we employ the the neighbourhood-to-city level of analysis to compare regions with respect to density. The results are presented in Table 4.

Table 4. Estimated influence of neighborhood and region diversity and specialization on productivity of single-plant firms 1997-2010, panel FE estimates. Dependent variable: Levinsohn and Petrin (2003) TFP.

	All of Sweden	Outside metropolitan areas
Firm-level characteristics		
Human capital (fraction of skilled workers)	0.0533***	0.0494***
	(0.00452)	(0.00660)
Neighborhood		
Specialization, emp. within 2-digit industry (log)	0.00365***	0.000994
	(0.000692)	(0.000857)
Diversity, emp. outside 2-digit industry (log)	0.00571***	0.00304**
	(0.00104)	(0.00145)
1st order neighbors		
Specialization, emp within 2-digit industry (log)	-4.57e-05	0.000124
	(0.000228)	(0.000360)
Diversity, emp outside 2-digit industry (log)	0.000776*	0.000823*
	(0.000400)	(0.000498)
City-wide		
Specialization, emp. within 2-digit industry (log)	-0.000135	-0.00572**
	(0.00224)	(0.00238)
Diversity, emp. outside 2-digit industry (log)	0.0128***	0.0212***
	(0.00391)	(0.00687)
Extra-industry diversity (log)	0.0613***	0.00123
	(0.0165)	(0.0189)
Firm fixed effects?	Yes	Yes
Year effects?	Yes	Yes
Industry effects?	Yes	Yes
Year-industry effects?	Yes	Yes
No. firm-year obs.	1,185,556	553,524
No. firms	335,152	146,679
R^2	0.05	0.07

Note: The table reports the estimated influence of neighborhood- and city-level diversity and specialization on TFP among single-plant firms in Sweden as a whole as well as outside Sweden's three metropolitan areas, i.e. outside Stockholm, Göteborg and Malmö local labor market region. Standard errors are presented within brackets. Variables are defined in section 3.2. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

We divide Sweden's city regions in two groups. The first set of regressions is run for all regions. In the second set we analyze only the non-metropolitan areas, i.e. excluding the Stockholm (incl. Uppsala), Göteborg and Malmö/Lund labor market regions. In this second operation we hence exclude the majority of Sweden's highly educated workforce, R&D facilities as well as its high-tech manufacturing and services industries.

From Table 4 it it may be deduced that the effect of neighborhood-level specialization and diversity becomes smaller outside of the metropolitan areas. In fact, when we analyze non-metro areas the coefficient for specialization comes out close to zero, and the effect of diversity is decimated as well. The city-region wide effects of diversity in terms of employment outside the same industry, however, increase as we exclude the metropolitan areas: when we expand the data to deal with larger areas, including sparser locations, the estimated effect of region diversity is positive. The estimated magnitude of the effect of a specialized region, on the other hand, is close to zero for the country at large, and negative in the non-metropolitan areas.

What explains that neighborhood-level variables appear to weigh heavier on productivity in metropolitan areas, while region-level variables appear more important in rural areas? One possible interpretation is that the number of potential day-to-day interactions between firms increases with density, and conversely that they are fewer and to some degree spread across a larger geographical area in sparser regions. The negative parameter for specialisation at the rural regional level could perhaps be interepreted as an indication of sensitivity to industrial specialisation. That is, sparse regions with specialised industry structure might be less able to counter external shocks and changes in the economy, including falling demand. The weaker outcomes on neighbourhood level could be taken as indication that the observed neighborhood level effects are indeed externalities derived in some fashion from overall dense regions, i.e. density externalities. Put differently, it matters whether a industry-specialised neighbourhood is located in a dense city or a sparse region.

6. CONCLUSION

The primary conclusion from the analyses in this paper is that diversity and specialization externalities are not mutually exclusive. Instead, they co-exist within cities but differ in their spatial distribution. Our analyses show that there are significant within-industry specialisation externalities operating predominantly at the neighborhood level within metropolitan cities. We also find positive between-industry diversity externalities operating at both the local neighborhood level and the city-wide level. These results resonate with the idea that firms can benefit from both localised specialisation and overall diversity. Converseley, the diversity of a successful urban economy may not be randomly distributed, but contain highly specialised neighbourhoods. There is a need to "open-up" cities and study their inner organisation and workings. Our results concerning the role of intracity microgeographies in terms of diversity and specialization suggest that the city is not the atom of urban growth analysis, but rather a molecule.

Our results may in addition provide a partial explanation to the recent trend of young high-tech and knowledge-intensive firms to locate in clusters in central districts and neighborhoods in large cities (see e.g. Foord 2013, Polèse 2014, Duvivier and Polèse 2016).[19] For example, in a recent analysis of the state of high-quality start-ups in the US, Guzman and Stern (2016) show that, in recent times, the central cities of Cambridge and areas of Boston, rather than Route 128, have emerged as hotspots for entrepreneurship in the Boston region. They show a similar trend in California, where start-ups appear to increasingly favor San Fransico over Silicon Valley. The reasons for and outcomes of this trend is still not well understood, but the findings presented in this paper suggest that one possible explanation is that a location in a cluster within larger cities allow start-up firms to enjoy both localized specialization externalities as well as city-wide diversity externalities.

Finally, we draw three conclusions relevant to urban policy-making. First, our results indicate that a multitude of specialized business districts seem to be preferable to a city-wide or regional specialization strategy. This does not, however, imply that every firm has "its place" and that there is a predetermined order to the industrial microgeography of a city. On the contrary, a top-down policy approach would appear to be at a disadvantage compared to a bottom-up approach because of the high level of heterogeneity between firms and local charcteristics. Second, conceptually, one could think of a city as a network of neighborhood-level networks (cf. Batty 2013). How each of these parts of the city benefit from localization economies and how they interact with each other and benefit from urbanization economies determines the collective outcome that is the measured urban growth of the city. Accordingly, more fine-tuned intraurban policy measures, taking bottom-up mechanics into account, may prove to be preferrable to top-down city-wide policy initiatives. Finally, knowledge-intensive service firms appear to draw on their surroundings and the associated density externalities to a larger degree than other firms. As this type of firm appears to be among those most affected by urban policy and planning decisions, including such firms in policy process may both inform city leaders and improve urban economies.

REFERENCES

Alonso, W. (1960) A theory of the urban land market. *Papers in Regional Science, 6*(1), 149-157.

Andersson, M., & Lööf, H. (2011). Agglomeration and productivity: evidence from firm-level data. *The Annals of Regional Science*, *46*(3), 601-620

---

[19]An often cited recent example is that firms like Twitter, Uber and Airbnb have chosen to locate in downtown San Francisco rather than the Silicon Valley. The popular press has also coined terms like *Silicon Alley* in Manhattan and *Silicon Roundabout* in London to denote the growing tech-industry in the inner parts of these cities (Duvivier and Polèse 2016).

Andersson, M., J. Klaesson & J. P. Larsson (2014) The Sources of the Urban Wage Premium by Worker Skills - Spatial Sorting or Agglomeration Economies? *Papers in Regional Science, 93*(4), 727-747.

Andersson, M., Klaesson, J., & Larsson, J. P. (2016). How local are spatial density externalities? neighbourhood effects in agglomeration economies. *Regional Studies*, *50*(6), 1082-1095.

Arzaghi, M., & Henderson, J. V. (2008). Networking off madison avenue. *The Review of Economic Studies*, *75*(4), 1011-1038.

Baldwin, R. E., & Okubo, T. (2006). Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography*, *6*(3), 323-346.

Bartel, A. P., & Lichtenberg, F. R. (1987). The Comparative Advantage of Educated Workers in Implementing New Technology. *The Review of Economics and Statistics*, *69*(1), 1-11

Batty, M. (2013). *The New Science of Cities*. MIT Press.

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

Bettencourt, L.M.A, Lobo, J & Strumsky, D (2007). Invention in the City: increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, 36, 107-120.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional studies*, *39*(1), 61-74.

Boschma, R., & Frenken, K. (2011). Technological relatedness, related variety and economic geography. *Handbook of Regional Innovation and Growth*, 187.

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

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128-152.

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

Desrochers, P. (2001). Local diversity, human creativity, and technological innovation. *Growth and Change*, *32*(3), 369-394.

Duranton, G., & Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 1454-1477.

Duvivier, C & Polèse, M (2016). The great urban techno-shift: are central neighborhoods the next Silicon Valleys? Evidence from three canadian metropolitan areas, Inédit working paper 2016-01

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

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review*, *43*(2), 409-429.

Foord, J. (2013). The new boomtown? Creative city to Tech City in east London. *Cities 33*, 51- 60.

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

Fujita, M. (1989). *Urban economic theory: land use and city size*. Cambridge university press.

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

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, *118*(3), 536-598.

Groot, H. L., Poot, J., & Smit, M. J. (2016). Which Agglomeration Externalities Matter Most and Why?. *Journal of Economic Surveys, 30*(4), 756-782.

Guzman, J., & Stern, S. (2016). *The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 15 US States, 1988-2014* (No. w22095). National Bureau of Economic Research.

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

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial Development in Cities. *The Journal of Political Economy*, *103*(5), 1067-1090

Hoover, E. M. (1937). *Location Theory and the Shoe Leather Industries* (Vol. 55). Harvard University Press.

Hsiao, C. (2002). *Analysis of Panel Data*. Cambridge university press.

Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development?. *Regional Studies*, *49*(6), 1003-1018.

Larsson, J. P. (2014). The neighborhood or the region? Reassessing the density–wage relationship using geocoded data. *The Annals of Regional Science*, *52*(2), 367-384.

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

Marrocu, E., Paci, R., & Usai, S. (2013). Productivity growth in the old and new Europe: the role of agglomeration externalities. *Journal of Regional Science*, *53*(3), 418-442.

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, *69*(2), 182-195.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *The American Economic Review*, *94*(3), 656-690.

Ohlin, B. (1933). International and interregional trade. *Harvard Economic Studies, Cambridge, MA*.

Olley, G. S., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica: Journal of the Econometric Society*, 1263-1297.

Petrin, A., Poi, B. P., & Levinsohn, J. (2004). Production function estimation in Stata using inputs to control for unobservables. *Stata Journal*, *4*, 113-123.

Polèse, M. (2014). Why some downtowns are back: Lessons from Urban Resurgence. *City Journal 24*. 84-91.

Porter, M.E. (2003). The economic performance of regions. *Regional Studies*, *37*, 549-578..

Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, *85*(2), 377-393.

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, *64*(2), 373-389.

Tveteras, R., & Battese, G. E. (2006). Agglomeration externalities, productivity, and technical inefficiency. *Journal of Regional Science*, *46*(4), 605-625.

Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of Economic Surveys*, *26*(1), 98-128.

van Oort, F. (2015). Unity in variety? agglomeration economics beyond the specialization–diversity controversy. in Karlsson, C., Andersson, M and Norman, T (eds). *Handbook of Research Methods and Applications in Economic Geography*, 259.

Van Soest, D. P., Gerking, S., & Van Oort, F. G. (2006). Spatial impacts of agglomeration externalities. *Journal of Regional Science*, *46*(5), 881-899.

von Thünen, J. H. (1826). *The Isolated State*. Hamburg: Perthes.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. Simon and Schuster.

Wixe, S., & Andersson, M. (2016). Which types of relatedness matter in regional growth? Industry, occupation and education. *Regional Studies*, forthcoming

Yankow, J. J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics*, *60*(2), 139-161.