

Lehmann, Robert; Kluge, Jan

Conference Paper

Marshall or Jacobs? Answers to an unsuitable question from an interaction model

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Lehmann, Robert; Kluge, Jan (2012) : Marshall or Jacobs? Answers to an unsuitable question from an interaction model, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/120543>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Marshall or Jacobs? Answers to an unsuitable question from an interaction model

February 15, 2012

– Preliminary Version –

Please do not cite

Abstract: This paper investigates whether localization economies as brought forward by Marshall (1890) or urbanization economies as mentioned by Jacobs (1970) are more decisive for regional gross value added per capita. Our novel approach is to explicitly allow for interdependencies between these two theories and to take into account that the initial levels of specialization and diversification might play a role. We therefore deploy a model with interaction terms and find that these two theories are not mutually exclusive in most of our sectors. In addition, the empirical results show that the initial levels of specialization and diversification do matter as well.

Keywords: localization and urbanization economies, interaction models, regional gross valued added

JEL Code: C33, O18, O40, R11

Jan Kluge
ifo Institute
Branch Dresden
Einsteinstr. 3
01069 Dresden, Germany
Phone: +49(0)351/26476-35
kluge@ifo.de

Robert Lehmann*
ifo Institute
Branch Dresden
Einsteinstr. 3
01069 Dresden, Germany
Phone: +49(0)351/26476-21
lehmann@ifo.de

* Corresponding author.

1 Introduction

One of the major issues that have been reconsidered in the field of regional economics within the last few years is the well known feature of knowledge spillovers. Krugman (2011, p. 2) mentions, for example, that "...a focus on advanced economies might suggest that it is time to downplay the emphasis on tangible factors like transportation costs in favour of intangible factors like information spillover,...".

A lot of work has been done in order to understand these kinds of externalities (for recent surveys of the existing theoretical and empirical literature see Rosenthal and Strange (2004) or Beaudry and Schiffauerova (2009)). The literature is all about the question where knowledge spillovers, which are assumed to be the driving force within this group of externalities, occur (within or between industries) and what the consequences are (specialization or diversification of industries). In particular, there are two major but opposed theories that describe in what way these spillovers are responsible for the creation or diffusion of knowledge and hence foster economic growth: localization and urbanization economies.¹

First, Marshall (1890) argues that companies being surrounded by others of the same industry will grow faster, assuming that knowledge will circulate mainly within the same industry. Companies benefit from being located closely to each other because they can gain from, what we call today, localization economies. This theory was picked up again by Arrow (1962) and Romer (1986), which is the reason why Glaeser *et al.* (1992) refer to it as MAR externalities. The empirical literature finds either positive or negative effects of specialization (see e.g. Henderson *et al.* (1995), Combes (2000), de Lucio *et al.* (2002), Dekle (2002), Blien and Suedekum (2005) for different countries and industries).

The opposite opinion was brought forward by Jacobs (1970). She rejects the conception, that knowledge flows do only occur within the same industry. According to her, companies will gain from facing a diverse environment consisting of different kinds of industries because new ideas will not come from within but from outside the respective industry. The mechanisms by which diversity leads to economic growth are usually called urbanization economies. Empirical evidence for this kind of externality can be found in Glaeser *et al.* (1992), Lee *et al.*

¹These expressions were originally introduced by E. M. Hoover (see Hoover (1937) and Hoover (1948)).

(2005), Blien *et al.* (2006), Fuchs (2011) or Illy *et al.* (2011).

All these prominent studies on this issue have two things in common: First, they assume, that localization and urbanization economies work exclusively from each other. Although half of the studies surveyed by Beaudry and Schiffauerova (2009) find effects for both externalities in a specific industry at the same time, none of them deal with interactions between them since their regression coefficients only measure average effects. We believe, however, that both might be attached to each other: Imagine a manufacturer having advantages from being close to other companies from the same industry nearby, which is in line with the idea of localization economies. At the same time, it is possible that the productivity of these manufacturers depend on companies from other sectors (e.g. input suppliers, universities, tax consultants, bank services, employment agencies or even other manufacturers from different branches) as well, which is in line with urbanization economies: Maybe a manufacturing firm profits from the idea flows of universities (e.g. new and more efficient production processes for photovoltaics) and recombine these ideas for their specific business. Another possibility is the application of working processes between different branches: Imagine that a mechanical engineering firm can adapt a process, maybe with some modifications, from a vehicle production firm. Now our contribution adds to the existing literature in the following way: If we believe that localization and urbanization economies are not mutually exclusive and strengthen or weaken each other, we have to account for this by deploying an interaction term in our empirical model which is further described in the next section.

Second, the studies assume that localization and urbanization economies are independent of the initial degrees of specialization and diversification. We believe, however, that the effects depend crucially on the initial levels because knowledge is not least a matter of critical mass. Knowledge spillovers will only get under way if a sufficient number of knowledge carriers has assembled which is why cities with a couple of firms of a certain industry might benefit more from further specialization than a city without any at all. Marshall (1890) and Jacobs (1970) presume that it is always favorable for the regional economic performance to reach a higher degree of specialization and diversification. However, there might be disadvantages from further specialization/diversification as well, especially in cities which are already

specialized/diversified to a very high extent. To tap such (potentially inverse) U-shaped relationships, we include squared terms of our specialization and diversification measures, which is the same as if the variables "interact with themselves".²

We use official German data and find, that the effects of localization and urbanization economies for specific industries are interdependent in several ways and lead to particular thresholds that must be regarded when interpreting the results. MAR and Jacobs externalities depend crucially on the initial levels of specialization and diversification which is why there cannot be simple answers on the question of whether the former or the latter is better for economic development of sectors. Nonetheless, we have strong evidence in favor of MAR in most sectors while Jacobs externalities seem to work only under very special constellations.

The paper is organized as follows. In section two, estimation approach and data set are described. Section three discusses our regression results for different branches of the economy. The last section concludes.

2 Data and Method

2.1 Estimation Approach

We assume the following Cobb-Douglas production function with constant returns to scale,

$$Y_{z,s,t} = A_{z,s,t}(\mathbf{X}) (K_{z,s,t})^\alpha (L_{z,s,t})^{1-\alpha} , \quad (1)$$

$Y_{z,s,t}$ represents the sector s specific real GVA in city z for a given year t . $A_{z,s,t}$ stands for a function of shifting parameters, dependent on a vector of variables \mathbf{X} , which we will discuss later. Furthermore, output is produced with capital ($K_{z,s,t}$) and labor ($L_{z,s,t}$). α and $1 - \alpha$

²There are studies that use squared terms as well but do not interpret them as interaction terms (see Illy *et al.* (2011) for Germany or de Lucio *et al.* (2002) for Spain). These studies only mention the existence of thresholds where the effects are reversed, but do not give a graphical interpretation if such a threshold is meaningful for their data set.

are the corresponding output elasticities. Expressing (1) per employee, we have

$$y_{z,s,t} = A_{z,s,t}(\mathbf{X}) (k_{z,s,t})^\alpha . \quad (2)$$

Now, let us consider the term $A_{z,s,t}(\mathbf{X})$. Since we cannot observe total factor productivity (TFP) directly, we assume that it is a function of different variables,

$$A_{z,s,t}(\mathbf{X}) = Constant (Spec_{z,s,t})^\beta (Div_{z,s,t})^\gamma (Interactions) (Controls) U_{z,s,t} , \quad (3)$$

especially localization ($Spec_{z,s,t}$) and urbanization ($Div_{z,s,t}$) economies (see e.g. Martin *et al.* (2011)). Since we are interested in interaction effects between the two theories, several interaction terms (*Interactions*) are included. To filter the true effect of specialization and diversification, we include control variables (*Controls*). $U_{z,s,t}$ stands for sector and city specific components, which are not directly observable.

Inserting (3) into (2) and taking natural logarithm, brings forward our empirical model

$$\begin{aligned} \ln(y_{z,s,t}) = & c + \alpha \log(k_{z,s,t}) + \beta \log(Spec_{z,s,t}) + \gamma \log(Div_{z,s,t}) + \\ & + interactions + controls + a_{z,s} + v_t + \epsilon_{z,s,t} . \end{aligned} \quad (4)$$

GVA per employee, which is a specific form of labor productivity, in real terms ($y_{z,s,t}$) is described by a constant (c), a measure for localization ($Spec_{z,s,t}$) and urbanization economies ($Div_{z,s,t}$), capital intensity in that sector ($k_{z,s,t}$), specific interaction terms, a set of control variables as well as city and industry specific time invariant fixed effects ($a_{z,s}$), year dummies (v_t) and an error term $\epsilon_{z,s,t}$.

The choice of appropriate indicators for localization and urbanization economies is critical. Beaudry and Schiffauerova (2009) argue, that the results of studies depend crucially on the choice of how to measure specialization and diversification. They recommend using one separate indicator for each of the two hypotheses, since both types can occur at the same time. Using only one indicator will make it impossible to distinguish between the two types of externalities or analyze interactions. Although literature does not seem to have reached a

consensus on the validity of MAR and Jacobs economies yet, there are a few indicators that are widely accepted.

Following the tradition of Glaeser *et al.* (1992), more than 40% of the studies regarded by Beaudry and Schiffauerova (2009) measure specialization using the relative location quotient

$$Spec_{z,s,t} = \frac{labor_{z,s,t}/labor_{z,t}}{labor_{s,t}/labor_t} . \quad (5)$$

This relative location quotient compares the degree of specialization of an industry s in city z to the national degree of specialization of that industry, using employment figures ($labor_{z,s,t}$). It takes values greater than one, if this share is above German average. The most specialized industries in our sample are advanced services in Frankfurt am Main and manufacturing in Wolfsburg (see descriptive statistics in the Appendix).

When searching for evidence of MAR economies, it is important to look at the respective industry and to find a measure for its relative specialization. The concept of Jacobs economies, however, requires to focus on the environment of that industry and to measure the degree of diversification of that environment. Most of the studies covered in Beaudry and Schiffauerova (2009) do that by using a relative version of the Hirschman-Herfindahl-Index (see e.g. Blien *et al.* (2006))

$$Div_{z,s,t} = \frac{1/ \sum_{s'=1, s' \neq s}^S (labor_{z,s',t}/(labor_{z,t} - labor_{z,s,t}))^2}{1/ \sum_{s'=1, s' \neq s}^S (labor_{s',t}/(labor_t - labor_{s,t}))^2} . \quad (6)$$

The index becomes larger the more diversified the environment of industry s is in comparison to the national average. In our sample, the most diversified environments are faced by manufacturers in Herne and Bottrop (see the Tables in the Appendix for descriptive statistics). Since we expect, that the impacts of localization and urbanization economies depend on the actual level of specialization or diversification and that the two theories are linked to each

other, our interaction terms take the following form:

$$\begin{aligned}
\text{Degree of specialization (DOS):} & \quad (\log(\text{Spec}_{z,s,t}))^2 \\
\text{Degree of diversification (DOD):} & \quad (\log(\text{Div}_{z,s,t}))^2 \\
\text{Mutual interaction (MI):} & \quad (\log(\text{Spec}_{z,s,t})) * (\log(\text{Div}_{z,s,t})) .
\end{aligned}$$

With these terms, our interaction model is fully specified (see Brambor *et al.* (2006)),

$$\begin{aligned}
\ln(y_{z,s,t}) = c + \alpha \log(k_{z,s,t}) + \beta \log(\text{Spec}_{z,s,t}) + \gamma \log(\text{Div}_{z,s,t}) + \\
+ \delta \text{DOS} + \zeta \text{DOD} + \eta \text{MI} + \text{controls} + a_{zs} + v_t + \epsilon_{z,s,t} .
\end{aligned} \tag{7}$$

Now, the estimated coefficients of $\text{Spec}_{z,s,t}$ (β) and $\text{Div}_{z,s,t}$ (γ) are no longer interpretable as average effects (see Brambor *et al.* (2006)). This would only be the case if and only if either the conditioning variables or the regression coefficients of the interaction terms take a value of zero (e.g. $\partial y / \partial \log(\text{Spec}) = \beta$ if $\log(\text{Spec}) = 0$ or $\delta = 0$ and $\log(\text{Div}) = 0$ or $\eta = 0$).³ We see that the marginal effect of either localization or urbanization economies on GVA per employee in such interaction models,

$$\frac{\partial y}{\partial \log(\text{Spec})} = \beta + 2\delta \log(\text{Spec}_{z,s,t}) + \eta \log(\text{Div}_{z,s,t}) \tag{8}$$

$$\frac{\partial y}{\partial \log(\text{Div})} = \gamma + 2\zeta \log(\text{HHI}_{z,s,t}) + \eta \log(\text{Spec}_{z,s,t}) , \tag{9}$$

takes other expressions than in a regression analysis without interaction terms,

$$\begin{aligned}
\frac{\partial y}{\partial \log(\text{Spec})} & = \beta \\
\frac{\partial y}{\partial \log(\text{Div})} & = \gamma .
\end{aligned}$$

The former two expressions capture the effects of localization (urbanization) economies conditional on the degree of specialization and diversification. Now it is possible to describe the

³Since we look at logarithmic variables, a value of zero means, that a particular city is as specialized or diversified as the German average.

effects of these externalities for the whole data set and not only for a single point.

We add several control variables: the balance of commuters to measure knowledge movements between a city and the surrounding periphery, the number of firms per capita to capture the effects of different firm sizes.⁴ Since gross value added is produced with capital and labor, capital intensity is a necessary control variable. As described above, year dummies are used, especially to capture business cycle effects or other, more temporal shocks to the economy. Furthermore we include two more dummy variables, one for cities in Eastern Germany⁵ and the other to capture effects of structural breaks in the German Classification of Economic Activities.⁶

The coefficients of (7) are estimated with linear panel model techniques. Because this approach accounts for cross-section and time information, we have to deal with several problems inherent to these two dimensions.

First, to avoid heteroscedasticity, we use robust standard errors estimates. Second, we have to account for multicollinearity. We calculated variance inflation factors and find that multicollinearity is not problematic for our variables of interest, namely the location quotient and the Hirschman-Herfindahl-Index. One can argue, that there has to be a strong problem of multicollinearity in such interaction models. Brambor *et al.* (2006) state that the omission of important variables is much more problematic than multicollinearity, because omitted variable bias causes the coefficients to be wrong. Third, a necessary condition for regression results is stationarity. We run several stationarity tests for panel data and conclude that all series are stationary in levels or trend-stationary. If the latter is the case, we include a time variable to capture an underlying linear trend. A fourth issue is arising from the correlation of the error term $\epsilon_{z,s,t}$ over time. With a test presented by Wooldridge (2002), we are able to detect autocorrelation in the error terms and correct the standard errors to get robust results. Finally, a problem which is often ignored in empirical economics is cross-sectional or

⁴It is not possible for us to observe the firm size directly, e.g. with the number of employees. Therefore, we use firms per capita as a proxy variable.

⁵The economy in Eastern Germany is predominantly described by a business organization, which is divided into small sections. In addition to this, it is difficult to realize advantages from lot sizes. So the productivity in Eastern Germany has to be naturally smaller than in cities in Western Germany.

⁶The German Classification of Economic Activities is the national classification scheme for all branches of the economy in Germany. The aim of this classification is to describe statistical units consistently and make different statistics comparable with each other.

"spatial" dependence (see Hoechle (2007)). This means that the error term $\epsilon_{z,s,t}$ is correlated between the cities, which causes the standard errors of the coefficients to be biased. A test, detecting such a problem, was developed by Pesaran (2004). In presence of very general forms of spatial correlation, Hoechle (2007) suggests to use the approach presented by Driscoll and Kraay (1998), which is able to correct such spatial correlations. If there is no evidence for cross-section dependence, Hoechle (2007) prefers using estimation procedures which produce heteroscedasticity- as well as autocorrelation-robust standard errors for the coefficients. To decide whether to use a random-effects or fixed-effects estimator, we apply a heteroskedastic and cluster-robust form of the Hausman-test. This module was developed by Schaffer and Stillman (2010). We perform tests for the overall significance of the year dummies as well as a F-Statistic for model appropriateness. See Table 1 for information about the deployed estimation techniques.

Table 1: Test results and preferred estimation methods

	Manufacturing D	Construction F	Basic services G-I	Advanced services J,K
Heteroscedasticity	yes	yes	yes	yes
Autocorrelation	yes	yes	yes	yes
Cross-Section-Dependence	yes	no	no	yes
Preferred estimation method	Driscoll&Kraay	FE-Cluster	FE-Cluster	Driscoll&Kraay

Source: Author's illustration.

2.2 Data set

We examine our empirical question for the case of Germany. Our data set includes the 70 biggest German cities for the period between 1998 and 2008. Different sources were necessary to construct our data set.

First, to obtain sectoral and regional indicators for specialization ($Spec_{z,s,t}$) and diversification ($Div_{z,s,t}$) we use all full-time employment persons subject to social security ($labor_{z,s,t}$).⁷ These regional figures are provided by the Federal Employment Agency of Germany (2010). Second, the Working Group Regional Accounts VGRdL (2011) provides data on gross value

⁷For example, civil servants and self-employed are not part of these figures.

added (GVA) for all German NUTS-3-districts.⁸ Using this source, it is possible to calculate figures for GVA per employee ($y_{z,s,t}$) as an indicator for productivity. Because of the lack of price indices on the regional level, we deflate the data of GVA in nominal terms with the deflator from the specific German state.⁹ Regional GVA figures are only available for the one-digit level, so we can just distinguish between eight different branches.¹⁰ Nonetheless, we run our regression analysis only for four (manufacturing (D), construction (F), basic services (G–I) and advanced services (J+K)) of these industries.¹¹ We use data on GVA even though most of the studies use employment as dependent variable. This, however, has often been criticized in the literature for several reasons: First, if employment growth has to be a good approximation for economic growth, then labor has to be a homogeneous input factor which is regional mobile within or between countries (see Beaudry and Schiffauerova (2009)). This is, as Almeida (2007) mentions, not the case, because migration costs are not equal across countries and constant over time. Second, because of the substitutive relationship between capital and labor and that technological progress can be a labor-saving process, employment growth is not a good proxy for economic growth (Paci and Usai (2005)). Third, if employment is a reasonable measure for regional economic development, one has to assume that the capital stock remains constant over time. As noted by Dekle (2002), these three assumptions do not seem to be realistic. Few studies made robustness checks for their results using employment growth as well as other indicators. Dekle (2002) as well as Cingano and Schivardi (2004) have found reversed results for localization and urbanization economies using either employment or total factor productivity growth.

Furthermore, our data set contains comprehensive information about establishments. We know the number of companies in a specific region as well as in which branch of the economy this enterprise is operating (Federal Employment Agency of Germany (2011)). With this information, we calculate the number of firms per capita, so it is possible to capture regional

⁸NUTS-3-districts in Germany contain *Landkreise* as well as *kreisfreie Städte*.

⁹This calculation of the specific deflators is consistent with the official statistics.

¹⁰Table 2 in the Appendix shows the acronyms for the different industries.

¹¹We exclude the sectors agriculture, forestry and fishing (A+B); mining and quarrying (C); energy and water supply (E) as well as the public sector and household related services (L–P). It is hard to believe that knowledge spillovers are at work in these sectors. Furthermore, companies in these sectors cannot choose their location as freely as companies in other sectors.

effects from different firm sizes.

Finally, we have information on the capital stock for all German Länder (see Working Group Regional Accounts VGRdL (2010)). Since there are no figures for the capital stock available on the NUTS-3-level, we have to generate viable proxies. Shares of the city's nominal GVA are used to calculate regional capital stocks.

3 Regression results

As already discussed in section 2.1, it is not sufficient to look at the standard estimation output when interaction terms are deployed. We will therefore concentrate on a graphical analysis following Brambor *et al.* (2006). Nonetheless, the numeric results are shown in Table 7 in the Appendix. The figures that we use in order to show the results all have the same form as briefly described below using the example of manufacturing (see Figure 1).

3.1 Manufacturing

The left panels depict specialization ($\log(\text{Spec})$) while the right ones concentrate on diversification ($\log(\text{Div})$). The three panels on each side respectively have the same codomain (from min to max in our data set). What is shown on the y-axis is described in the titles.

Panel A (B) shows the constitutive effect of specialization (diversification). This is the marginal impact of $\log(\text{Spec})$ ($\log(\text{Div})$) on GVA per employee, assuming that whether there is no interaction between the two variables ($\eta = 0$) or the level of $\log(\text{Div})$ ($\log(\text{Spec})$) is zero (see equation 8 (9)).¹² Hence, panel A (B) shows the interaction that we refer to as DOS (DOD), i.e. the respective variable's "interaction with itself".

These panels provide two pieces of information: First, the intercept at $\log(\text{Spec}) = 0$ ($\log(\text{Div}) = 0$) depicts the respective estimation coefficient which we can take from the traditional estimation output in the Appendix (see Table 7). In the case of manufacturing, this is 0.1966 for $\log(\text{Spec})$ and -0.1657 for $\log(\text{Div})$. Both effects are significant at the 5%-

¹²Since our variables are normalized on the German average and logarithmized, values of zero represent the German average. This is a very helpful property. Keep in mind, that the variable that is not shown in a panel is set to the average. Actually, this assumption is necessary due to the fact that 3D-figures showing both variables at the same time are not suitable.

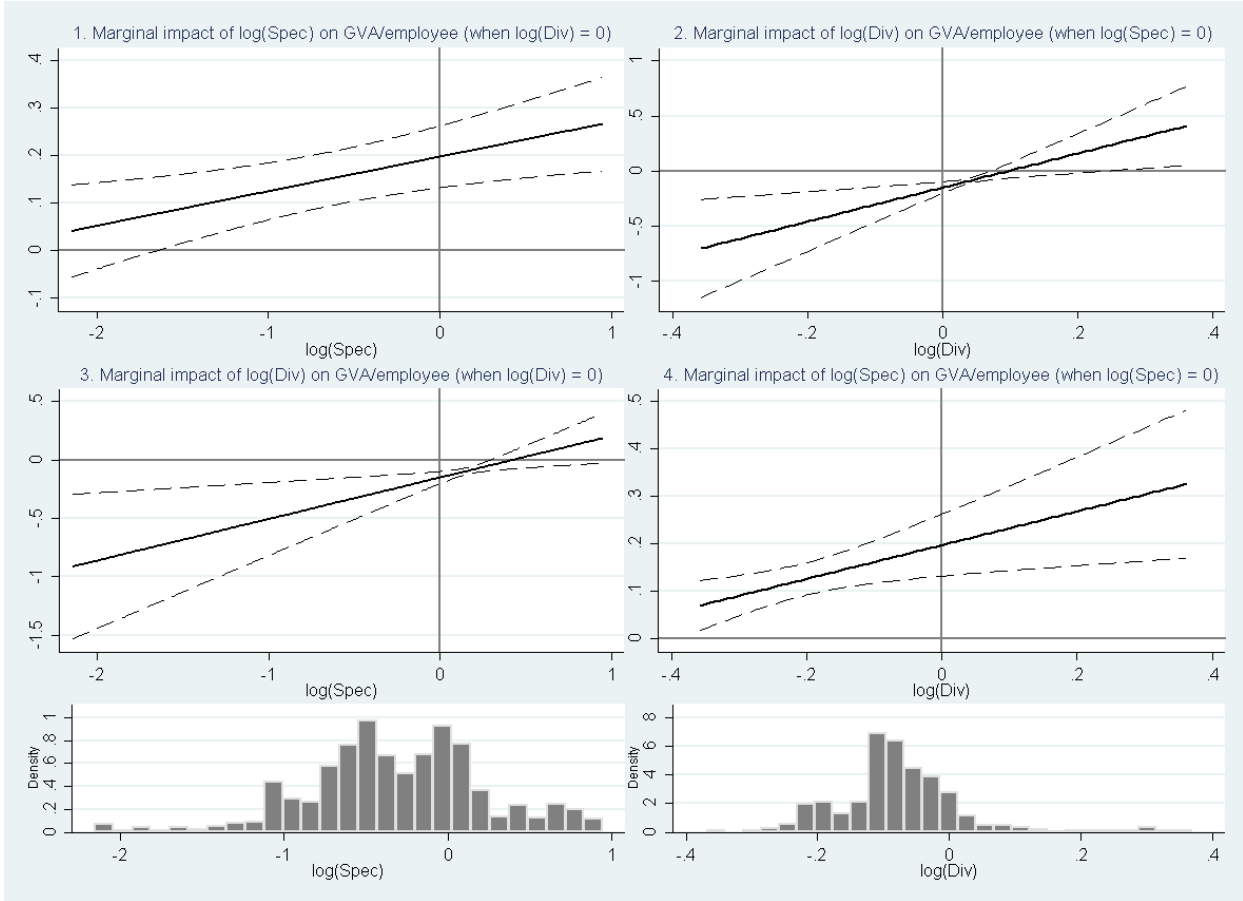


Figure 1: Manufacturing

level (scattered lines represent the 95% confidence interval) which is why we can conclude that there are MAR economies but no (positive) Jacobs economies in this sector. This, however, does only hold for cities in which $\log(\text{Spec})$ and $\log(\text{Div})$ in the sector of manufacturing are indeed zero, i.e. represent the German average. The effects might be completely different in cities that depart from the average. For those cities we need a second piece of information which is not contained in a normal regression output: the position of these curves over the whole range of observations. If no squared terms were deployed, they would have a slope of zero. Hence, it would be assumed that the effect which is shown by the coefficient holds on average for all cities in the data set. However, the slopes are not necessarily zero if we deploy squared terms which is why we can see different effects for different levels of $\log(\text{Spec})$ or $\log(\text{Div})$. Table 7 shows the coefficients for the squared terms which represent the respective slopes in panels A and B. Since they do not tell anything about the actual position of the curve nor of the corresponding confidence bands, we will interpret panels A and B again: In

the case of manufacturing, there is generally a strong positive influence of specialization on per capita GVA over almost the whole range of $\log(\text{Spec})$. However, the impact shrinks the smaller the level of specialization is and it becomes even insignificant for very low levels (see panel A). The effect of diversification is even more negative for lower levels of diversification while it can be even slightly positive in cities in which the sector of manufacturing faces very diversified environments (in this case $\log(\text{Div}) > 0.25$, see panel B).

For a better understanding of these findings, it is helpful to take a look at the corresponding histograms. They show the empirical distribution of our data set with respect to $\log(\text{Spec})$ and $\log(\text{Div})$. In the case of manufacturing, we see that there are few cities which do indeed provide positive effects of diversification. The only exceptions that achieve levels greater than 0.25 are two of the former mining centers in the Ruhr area, Herne and Bottrop.

So far we might conclude that there are MAR economies in most cities while the impact of diversification on per capita GVA is predominantly negative which rejects the theory of Jacobs. This is what we learned from panels A and B.

We now want to add the information about what happens to the curve in panel A (B) if $\log(\text{Div})$ ($\log(\text{Spec})$) departs from zero. For this purpose, we show panel D (C). It shows the same derivation as above. The only difference is, that we now assume that the respective other variable varies and $\log(\text{Spec})$ ($\log(\text{Div})$) is set to zero. Now, bringing panel A (B) into contact with panel D (C) will uncover the whole picture: The slope of curve D (C) makes clear in which direction curve A (B) will shift if the measure for $\log(\text{Div})$ ($\log(\text{Spec})$) varies. This is the mutual interaction between the two variables that we refer to as MI. The intercepts at the zero points in both panels are identical since both variables are set to zero. In the case of manufacturing, curve D has a positive slope, e.g. if $\log(\text{Div})$ departs from zero to the right, curve A will shift upwards and vice versa. The right histogram shows, however, that most of our observations are below zero which means that panel A might be a bit too optimistic. Turning to panel C which has a positive slope as well, we can see that curve B will shift upwards for cities in which manufacturing is specialized to a higher extent than $\log(\text{Spec}) = 0$. Therefore, we have to rethink our conclusion regarding Herne and Bottrop: These cities are characterized by high levels of $\log(\text{Div})$ which is why they might have posi-

tive impacts of diversification. But at the same time they must provide a level of $\log(\text{Spec})$ which is sufficiently high, e.g. at least zero. However, we do not have any observations in our data set that fulfill both requirements. Herne and Bottrop do only achieve below average specialization levels in manufacturing.

While there is strong evidence for MAR economies, Jacobs' idea of urbanization economies must be rejected in manufacturing although they might occur under special constellations that we do not have in our data set.

Taking these findings seriously, it seems to be a good idea for all cities to further specialize in manufacturing while keeping an eye on insuring a satisfyingly diversified environment. A high degree of specialization does promote economic development and reduces the potentially negative effects from lower levels of diversification at the same time. For cities, however, that are already specialized in manufacturing to a very high degree (like e.g. Ingolstadt and Wolfsburg) it seems desirable to enhance the diversification level of the remaining sectors.

3.2 Construction

The effects of specialization on per capita GVA, depicted by panel A in Figure 2, are ambiguous in the sector of construction. The impact is positive for high levels of specialization but negative for lower ones under the assumption of an average degree of diversification ($\log(\text{Div}) = 0$). As we can see, however, from panel D and the corresponding histogram, most cities provide less diversified environments for their construction sectors which supports the positive effects of specialization. Therefore the curve in panel A will shift upwards for most of our observations and the threshold at which the effect gets positive is achieved further left. This is considerable evidence in favor of MAR economies.

The impact of diversification on per capita GVA is again very small. Construction sectors with an average diversified environment are slightly harmed by this and should rather face a less diversified one (see panel B). This is true for most of our observations (see right histogram). Panel C shows that the level of specialization can partly worsen this effect but this is only relevant for a small share of the cities in our data set.

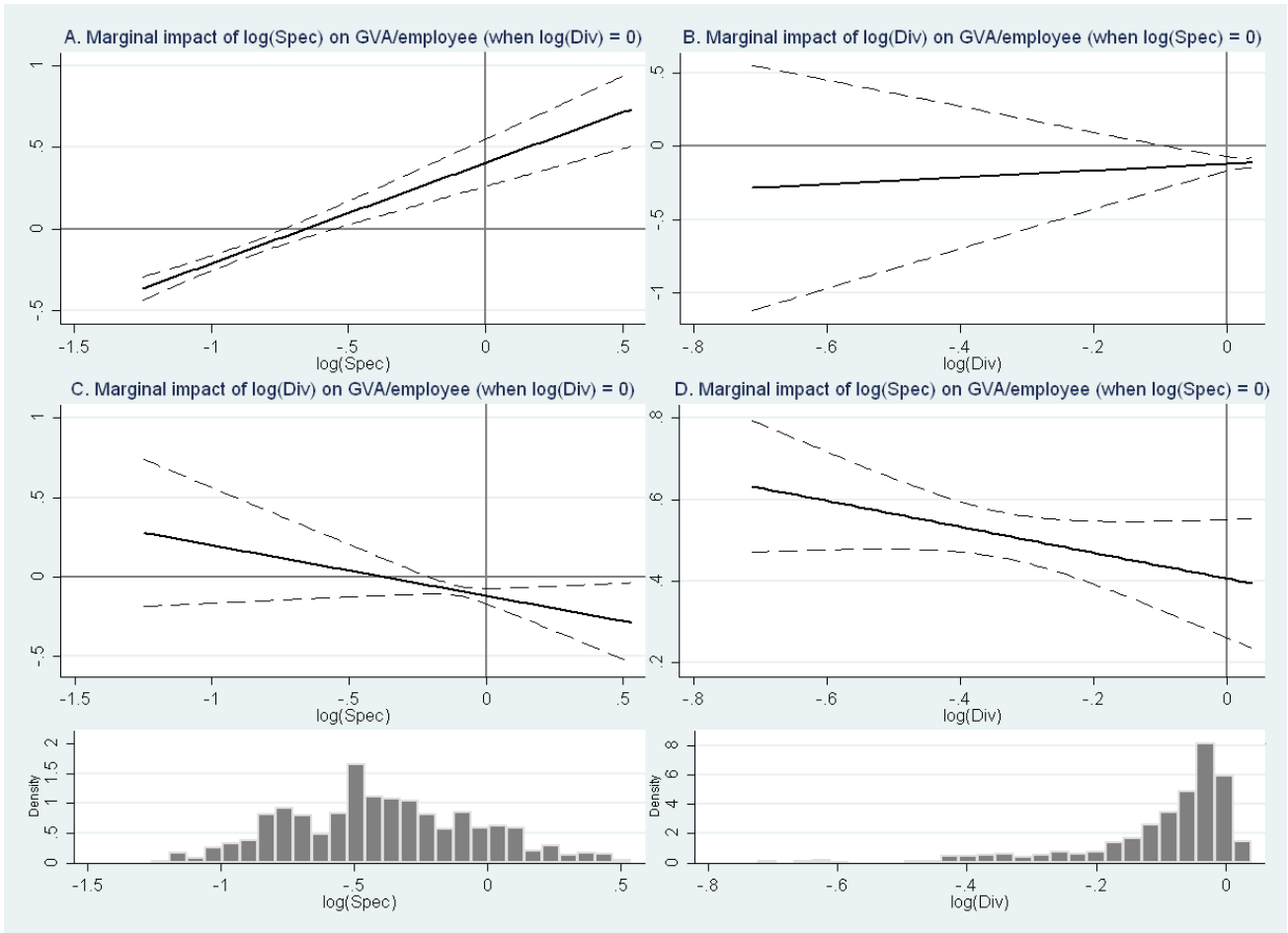


Figure 2: Construction

Since diversification causes (if at all) negative impacts on per capita GVA, we must reject the theory of Jacobs (1970) for the construction sector. However, we have considerable evidence for MAR economies.

These results suggest that it does not matter how diversified the surroundings of construction sectors are since every kind of economic activity has a particular demand for construction. Actually, it is even better if the number of different kinds of construction demands is not too big since these manifold requirements would harm the specialization benefits of construction firms. Increasing the share of construction companies is the only way how to generate advantages for this sector from knowledge spillovers.

3.3 Basic Services

Basic services contain trade and repairation of vehicles, hotel and restaurant industry as well as traffic and communication. As panel A in Figure 3 indicates, there is a slightly positive impact of specialization on per capita GVA in this sector, at least for approximately one half of our data set. Companies do benefit especially in cities in which basic services account for below average shares. This share should not reach the German average and should not be too small since the effect vanishes at some threshold. This result is not influenced by the level of diversification (see panel D).

Although there is a (inverse) U-shaped relationship between the degree of specialization and per capita GVA in the sector of basic services, our evidence is sufficient as to talk about localization economies, at least in some places.

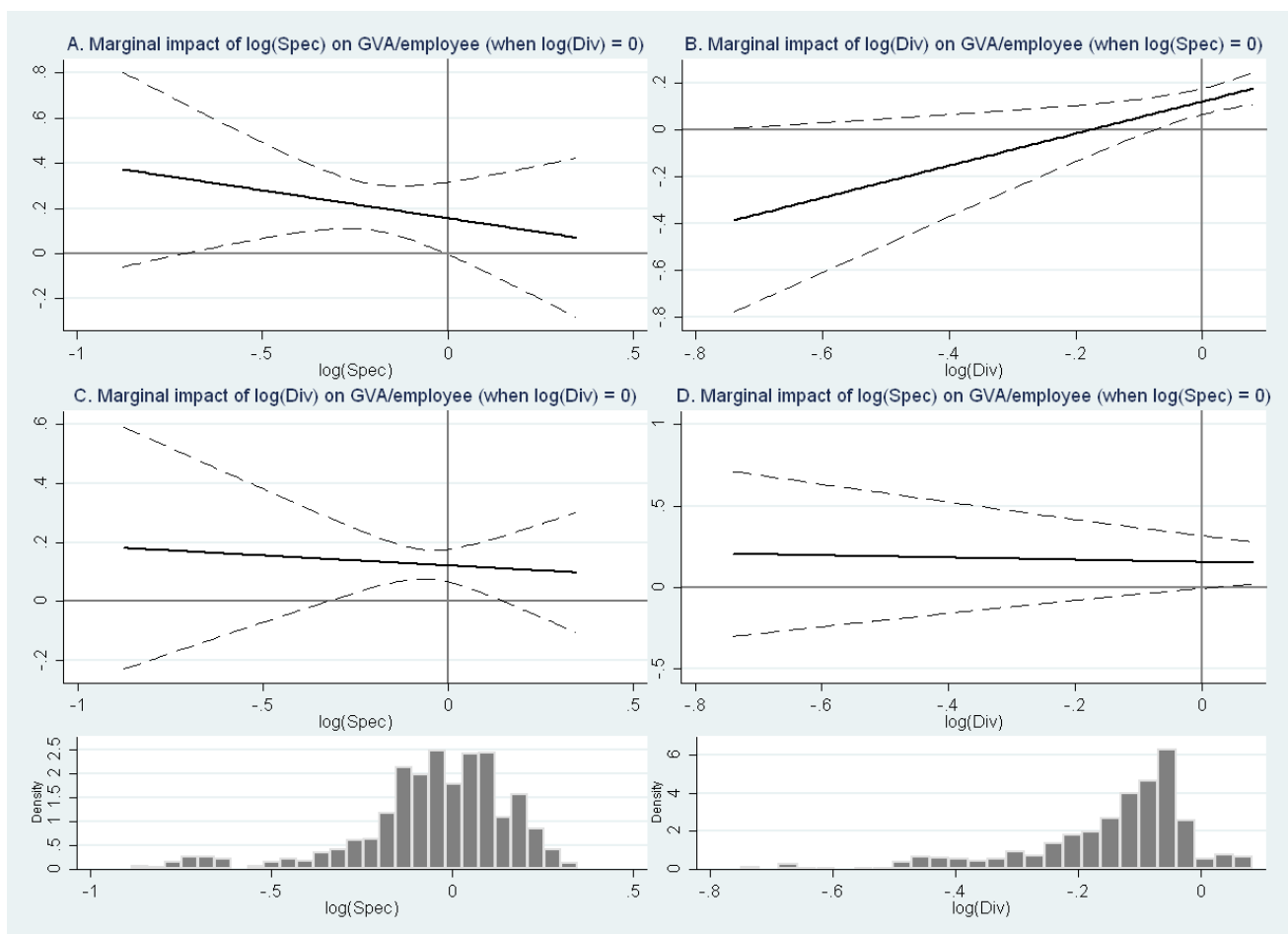


Figure 3: Basic Services

Diversification does seem to have an impact as well. Panel B indicates, that it is slightly positive in cities in which basic services face highly diversified environments. This impact is slightly increased (decreased) in cities in which this sector accounts for a below (above) average share (see panel C). Taking the left histogram into account, the curve in panel B is not far from representing the median observation, i.e. it will shift upwards and downwards for one half of the data set respectively. The cities in which basic services companies do benefit from diversification are again former mining cities like Oberhausen, Mühlheim an der Ruhr and Herne.

In this sector, both types of externalities seem to be at work. Actually, the sector of basic services is the only one in our investigation that provides considerable Jacobs economies.

The implication for this sector would be to keep the degree of specialization below a certain threshold (which is incidentally given by the German average) but not too small. At the same time, there should be a highly diversified environment generating a wide range of demand for companies providing basic services.

3.4 Advanced Services

The sector of advanced services contains financial and real estate industry. As panel A in Figure 4 indicates, the constitutive effect of specialization in this sector is not clear. It has a positive impact on per capita GVA at high levels of specialization. Incidentally, cities with above average degrees of specialization in this sector benefit while those with lower degrees do not. Since the majority of our observations is characterized by higher degrees of specialization in advanced services since headquarters of banks and insurance companies are usually located in cities (see left histogram), we might conclude that localization economies do play a role in this sector. The effect is even more positive for cities in which the sector of advanced services faces a below average diversified environment which holds for almost all of our observations (see panel D and right histogram). Taking this into account shows, that the curve in panel A actually lies higher than at an average degree of diversification ($\log(\text{Div}) = 0$). These MAR economies promote the emergence of large service centers like Frankfurt am Main or Düsseldorf.

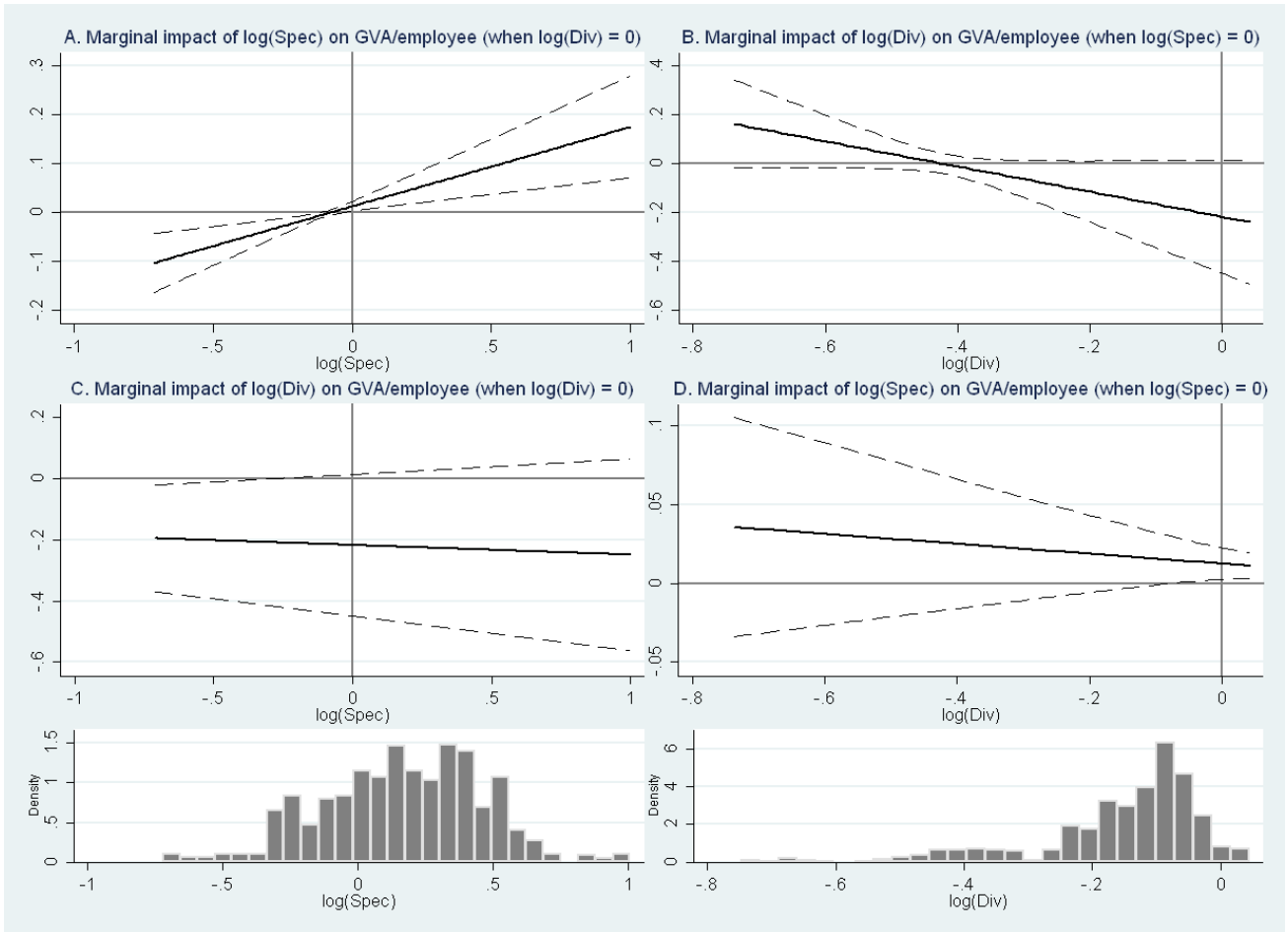


Figure 4: Advanced Services

Turning to the constitutive impact of diversification on per capita GVA (see panel B in Figure 4), this is insignificant for all of our observations at first. It is slightly reduced by higher levels of specialization (see panel C in Figure 4) and might then be negative in high diversified cities.

The small but negative impact of diversification on per capita GVA indicates that Jacobs' theory must be rejected (i.e. if at all, there are negative urbanization economies) but once more we do find strong evidence for MAR economies in the advanced service sector for the majority of our data set.

These results suggest that it would be desirable for those cities with high shares in this sector to focus on increasing this share without paying regard to further diversification. Those cities, however, that are less specialized in advanced services should either try to overleap the threshold (which is incidentally given by the German average degree of specialization) or focus on promoting other sectors.

3.5 Endogeneity

Our regression approach is able to control for several problems (heteroscedasticity, autocorrelation etc.). Endogeneity is a much more striking problem in the field of econometrics and harder to deal with. We concern this problem in the following subsection.

The causes for endogeneity are manifold. Especially mutual causality or omitted variables induce the problem of endogeneity. In our situation, mutual causality could be the cause. This means: Is specialization or diversification the cause for growth or does a city become more specialized/diversified if the specific sector expands? To solve the problem of endogeneity, we apply a two-stage least squares approach (2SLS). For this purpose, we have to find a viable instrument: Because it is impossible that today's growth effects the previous specialization or diversification pattern, we use time-lagged variables as instruments. A valid instrumental variable always requires two characteristics: relevance and exogeneity. The first requirement is comparatively easy to obtain, because there has to be a sufficiently high correlation between the excluded instruments and the endogeneous regressors. To check the relevance of our instruments as a whole and thereby overcoming the "weak instrument" problem (see Baum *et al.* (2003)), we apply a statistical test and present the Cragg-Donald Wald F statistic (CD) from the first stage of our regression. Therefore, we use the critical values obtained from the Stock-Yogo weak ID test to decide if the instruments are weak. Up to a lag of two years, high correlations for our specialization and diversity variable are observable. This is the reason, why we use two lags as instruments. For the second requirement, exogeneity, we present the p-value for the Hansen J statistic (H_p) to check the overall validity of our instruments.

Table 8 shows results of the instrumental variable (IV) approach. Applying a 2SLS approach is not free of difficulties: First, IV estimates with interaction models demand to proceed the two stages separately. Second, since we use different instruments for the two relevant endogeneous regressors ($\log(Spec)$ and $\log(Div)$), it is necessary to perform two IV estimates. Therefore, we show two columns in Table 8 for each branch of the economy. Only with slight differences, our main results stay the same. Almost all of our qualitative conclusions hold when applying an IV approach.

The sign for the location quotient in advanced services is reversed (see Table 7 and Table 8 in the Appendix). In the OLS regression, specialization has on average a positive impact on GVA per employee. Now the average effect becomes negative. But this is less of a problem: As we have seen in Figure 4, our specialization measure has a positive slope as well as a positive sign on average and lies only slightly above the zero line. Now, the slope coefficient (DOS) is still positive but smaller in magnitude, which makes the curve smoother. This is the reason, why the average coefficient for the location quotient becomes negative. In fact, the threshold where the curve of the marginal effects of specialization is reached, shifts a little bit to the right on the axis. Furthermore, we find reversed results for the interaction term between specialization and diversification in basic services. But this is not critical as well, because the coefficient is really small in magnitude and does not change any conclusion drawn from Figure 3.

Test statistics provided by Cragg and Donald show in all regressions that no weak instrument problem is present. From the p-Values for the Hansen J statistic we can see, that in all specifications the null hypothesis of exogeneity cannot be rejected. This emphasizes the validity of our instruments.

4 Conclusion

We have performed a panel regression analysis in order to find evidence for MAR and Jacobs economies in a sample of 70 German cities. We find a strong positive influence of specialization on per capita GVA and therewith evidence for MAR in most of the sectors. Diversification does influence per capita GVA to a smaller extent which is why we must reject the existence of Jacobs economies in the majority of our sample. Only the sector of basic services shows plausible positive effects of diversification in the sense of Jacobs (1970) while other industries provide negative effects of diversification.

So far we do not provide any further insights into the topic since the multitude of existing studies constructs a wide range of different results that our paper might easily fit into. The main finding of our paper, however, is the fact that localization and urbanization economies depend crucially on the initially achieved levels of specialization and diversification. If an

author chooses not to deploy interaction terms, his or her results do only represent the effects for the very special constellation in which the respective other variable takes a value of zero. It is comprehensive that these zero points vary according to the choice of industrial and regional disaggregation. This is why the investigations, that have been carried out yet, might be all right even though they present very different results. Each of these studies might represent one single point in the curves that we presented in the last section. Following the idea of Brambor *et al.* (2006), we would like to encourage to replicate existing studies using interaction terms in order to find out if the published results do actually take place or if they occur far outside of the respective data sets.

Since we find that localization and (in parts) urbanization economies exist, these findings might be helpful for regional policy makers trying to promote economic growth. However, this does not tell anything about the sustainability of certain economic structures. These questions are subjects of further research.

Acknowledgements: We are grateful to Georg Hirte, Marcel Thum and participants of the Brown Bag Seminar at the Dresden University of Technology, Faculty of Business and Economics.

References

- ALMEIDA, R. (2007). Local Economic Structure and Growth. *Spatial Economic Analysis*, **2** (1), 65–90.
- ARROW, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, **29** (3), 155–173.
- BAUM, C. F., SCHAFFER, M. E. and STILLMAN, S. (2003). Instrumental variables and GMM: Estimation and testing. *The Stata Journal*, **3** (1), 1–31.
- BEAUDRY, C. and SCHIFFAUEROVA, A. (2009). Who’s right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, **38** (2), 318–337.

- BLIEN, U. and SUEDEKUM, J. (2005). Local Economic Structure and Industry Development in Germany, 1993-2001. *Economics Bulletin*, **15** (17), 1–8.
- , — and WOLF, K. (2006). Local employment growth in West Germany: A dynamic panel approach. *Labour Economics*, **13** (4), 445–458.
- BRAMBOR, T., CLARK, W. and GOLDER, M. (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, **14** (1), 63–82.
- CINGANO, F. and SCHIVARDI, F. (2004). Identifying the sources of local productivity growth. *Journal of the European Economic Association*, **2** (4), 720–742.
- COMBES, P. (2000). Economic Structure and Local Growth: France, 1984-1993. *Journal of Urban Economics*, **47** (3), 329–355.
- DE LUCIO, J. J., HERCE, J. A. and GOICOLEA, A. (2002). The effects of externalities on productivity growth in Spanish industry. *Regional Science and Urban Economics*, **32** (2), 241–258.
- DEKLE, R. (2002). Industrial concentration and regional growth: evidence from the prefectures. *Review of Economics and Statistics*, **84** (2), 310–315.
- DRISCOLL, J. C. and KRAAY, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics*, **80** (4), 549–560.
- FEDERAL EMPLOYMENT AGENCY OF GERMANY (2010). *Statistics of the Federal Employment Agency – Employed persons subject to social security according to selected branches of the economy, selected regions (place of work)*. Federal Employment Agency Germany, Data upon Request, Nuremberg 2010.
- FEDERAL EMPLOYMENT AGENCY OF GERMANY (2011). *Statistics of the Federal Employment Agency – Establishments according to selected branches of the economy, selected regions (place of work)*. Federal Employment Agency Germany, Data upon Request, Nuremberg 2011.

- FEDERAL STATISTICAL OFFICE OF GERMANY (2003). *German Classification of Economic Activities with explanations, Edition 2003*. Federal Statistical Office of Germany, Wiesbaden 2003.
- FUCHS, M. (2011). The determinants of local employment dynamics in Western Germany. *Empirical Economics*, **40** (1), 177–203.
- GLAESER, E., KALLAL, H., SCHEINKMAN, J. and SHLEIFER, A. (1992). Growth in Cities. *Journal of Political Economy*, **100** (6), 1126–1152.
- HENDERSON, V., KUNCORO, A. and TURNER, M. (1995). Industrial Development in Cities. *Journal of Political Economy*, **103** (5), 1067–1090.
- HOECHLE, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *The Stata Journal*, **7** (3), 281–312.
- HOOVER, E. (1937). *Location Theory and the Shoe and Leather Industries*. Harvard University Press, Cambridge (Mass.).
- (1948). *The Location of Economic Activity*. McGraw-Hill, New York.
- ILLY, A., SCHWARTZ, M., HORNYCH, C. and ROSENFELD, M. (2011). Local economic structure and sectoral employment growth in German cities. *Journal of Economic and Social Geography*, **102** (5), 582–593.
- JACOBS, J. (1970). *The Economy of Cities*. Vintage, New York.
- KRUGMAN, P. R. (2011). The New Economic Geography, Now Middle-aged. *Regional Studies*, **45** (1), 1–7.
- LEE, B., SOSIN, K. and HONG, S. (2005). Sectoral Manufacturing Productivity Growth in Korean Regions. *Urban Studies*, **42** (7), 1201–1219.
- MARSHALL, A. (1890). *Principles of Economics*. MacMillan, London.
- MARTIN, P., MAYER, T. and MAYNERIS, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, **69** (2), 182–195.

- PACI, R. and USAI, S. (2005). *Agglomeration economies and growth in Italian local labour systems. 1991 - 2001*. ERSA conference papers.
- PESARAN, M. H. (2004). *General Diagnostic Tests for Cross Section Dependence in Panels*. CESifo Working Paper Series No. 1229.
- ROMER, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy*, **94** (5), 1002–1037.
- ROSENTHAL, S. S. and STRANGE, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In J. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics: Cities and Geography*, Elsevier, pp. 2119–2171.
- SCHAFFER, M. E. and STILLMAN, S. (2010). *xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtiivreg, xtiivreg2, xthtaylor*. <http://ideas.repec.org/c/boc/bocode/s456779.html>.
- STOCK, J. H. and YOGO, M. (2002). *Testing for Weak Instruments in Linear IV Regression*. NBER Technical Working Paper 284.
- WOOLDRIDGE, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge.
- WORKING GROUP REGIONAL ACCOUNTS VGRDL (2010). *Stock of fixed assets in Germany by Bundesland and East-West-Regions 1991 to 2008*. Working Group Regional Accounts VGRdL, Series 1, Regional results Volume 4, Date of calculation: August 2010, Stuttgart 2011.
- WORKING GROUP REGIONAL ACCOUNTS VGRDL (2011). *Gross domestic product, gross value added in Germany on NUTS-3 Level 1992, 1994 to 2009*. Working Group Regional Accounts VGRdL, Series 2, Regional results Volume 1, Date of calculation: August 2010, Stuttgart 2011.

5 Appendix

Table 2: German Classification of Economic Activities

Acronym	Description
A,B	Agriculture, Forestry and Fishing
C	Mining and quarrying
D	Manufacturing
E	Electricity, gas and water supply
F	Construction
G-I	Trade, Hotels and Restaurants, Transport
J,K	Financing, Renting and Business Activities
L-P	Public and Private Services

Source: Federal Statistical Office of Germany (2003).

Table 3: Descriptive Statistics – Manufacturing

Variable	Mean	s.d.	Minimum	Maximum
Spec	0.846	0.469	0.116 Potsdam (2008)	2.588 Wolfsburg (2006)
Div	0.931	0.096	0.691 Bonn (1998)	1.447 Herne (2006)
Firms p.c.	2.273	0.830	0.986 Herne (2006)	6.654 Solingen (1998)
Capital intensity	140762	39451	62567 Gera (2005)	354149 Dresden (2004)
Commuters	38019	48387	-9556 Oberhausen (2008)	260188 Frankfurt a.M. (2001)

Source: Author's calculation.

Table 4: Descriptive Statistics – Construction

Variable	Mean	s.d.	Minimum	Maximum
Spec	0.719	0.262	0.284 Wolfsburg (2008)	1.717 Herne (2005)
Div	0.912	0.100	0.486 Wolfsburg (1998)	1.041 Munich (1998)
Firms p.c.	2.140	0.495	1.307 Erlangen (2008)	4.513 Leipzig (1999)
Capital intensity	38058	10007	10071 Gera (1999)	96223 Herne (1999)
Commuters	38019	48387	-9556 Oberhausen (2008)	260188 Frankfurt a.M. (2001)

Source: Author's calculation.

Table 5: Descriptive Statistics – Basic Services

Variable	Mean	s.d.	Minimum	Maximum
Spec	0.971	0.184	0.413 Wolfsburg (1998)	1.414 Bremerhaven (1998)
Div	0.872	0.107	0.472 Wolfsburg (1998)	1.088 Herne (1999)
Firms p.c.	8.723	1.564	5.080 Herne (2007)	13.129 Düsseldorf (1999)
Capital intensity	99118	29363	56418 Chemnitz (1998)	258807 Fürth (2008)
Commuters	38019	48387	-9556 Oberhausen (2008)	260188 Frankfurt a.M. (2001)

Source: Author's calculation.

Table 6: Descriptive Statistics – Advanced Services

Variable	Mean	s.d.	Minimum	Maximum
Spec	1.231	0.359	0.486 Bottrop (2000)	2.739 Frankfurt a.M. (1998)
Div	0.864	0.103	0.474 Wolfsburg (1999)	1.047 Herne (1998)
Firms p.c.	6.015	1.920	1.855 Salzgitter (1998)	11.493 Munich (2002)
Capital intensity	98266	18832	59088 Leipzig (1998)	158813 Wolfsburg (1999)
Commuters	38019	48387	-9556 Oberhausen (2008)	260188 Frankfurt a.M. (2001)

Source: Author's calculation.

Table 7: OLS-Regression Results

	Manufacturing (D)	Construction (F)	Basic Services (G-I)	Advanced Services (J,K)
	Driscoll&Kraay	FE-Cluster	FE-Cluster	Driscoll&Kraay
<i>Variables:</i>				
log(Spec)	0.1976*** (0.0332)	0.4063*** (0.0676)	0.1558 (0.0946)	0.0124** (0.0051)
log(Div)	-0.1495*** (0.0267)	-0.1196 (0.1576)	0.1212* (0.0971)	-0.2176* (0.1181)
<i>Interactions:</i>				
DOS	0.0365*** (0.0131)	0.3087*** (0.0623)	-0.1235 (0.1799)	0.0812 (0.0244)
DOD	0.7728*** (0.2846)	0.1163 (0.3828)	0.3424 (0.2403)	-0.2566* (0.1392)
MI	0.3557*** (0.1370)	-0.3177 (0.3673)	-0.0678 (0.3198)	-0.0315 (0.0434)
<i>Controls:</i>				
Commuters	-1.24e-06*** (2.77e-07)	1.80e-06 (1.27e-06)	-1.34e-07 (1.01e-06)	-9.71e-07* (7.68e-07)
log(Firms p.c.)	-0.0799*** (0.0112)	0.3732*** (0.0808)	0.0756 (0.1216)	0.1842*** (0.0151)
log(Capital intensity)	1.0172*** (0.0096)	0.8073*** (0.0328)	0.4884*** (0.0865)	1.0387*** (0.0098)
R ²	0.9492	0.8078	0.7574	0.8818
Obs.	770	770	770	770

Source: Author's calculation, Standard errors in parentheses.

***, **, * indicate statistical significance at the 1%, 5%, 10% level.

Table 8: IV-Regression (2SLS) Results

	Manufacturing (D)	Construction (F)	Basic Services (G-I)	Advanced Services (J,K)
<i>Variables:</i>				
log(Spec)	(+) ^{***}	(+)	(+)	(-) ^{**}
log(Div)	(-) ^{***}	(-)	(+)	(-) ^{**}
<i>Interactions:</i>				
DOS	(+)	(+)	(-)	(+) ^{***}
DOD	(+) ^{***}	(+)	(+)	(-)
MI	(+)	(-) ^{**}	(+)	(-) ^{***}
<i>Controls:</i>				
Commuters	(-) ^{***}	(+)	(+)	(-) ^{***}
log(Firms p.c.)	(-)	(+) ^{***}	(-)	(+) ^{***}
log(Capital intensity)	(+) ^{***}	(+) ^{***}	(+) ^{***}	(+) ^{***}
CD	500.363 ^{***}	174.720 ^{***}	169.333 ^{***}	205.575 ^{***}
Hp	0.5044	0.0001±	0.9194	0.4596
R ²	0.9522	0.8708	0.7430	0.8421
Obs.	630	630	630	630
Instrumented	<i>log(Spec)</i>	<i>log(Div)</i>	<i>log(Spec)</i>	<i>log(Div)</i>
Instruments	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>
	<i>log(Spec)</i>	<i>log(Div)</i>	<i>log(Spec)</i>	<i>log(Div)</i>
	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>	<i>t-1 to t-2</i>

Source: Author's calculation, Standard errors in parentheses.

Note: An overview of the critical values for the CD test statistic is found in Stock and Yogo (2002).

± In this case, the Hansen test rejects the exogeneity of our two instruments. Testing only the second lag with an exogeneity test of the Hausman type indicates exogeneity of this instrument (p-value: 0.4457).

Results with reversed signs in comparison to the OLS estimates are printed in boldface.

***, **, * indicate statistical significance at the 1%, 5%, 10% level or 10%, 15%, 20% of maximal IV relative bias (size) for the CD statistic.