

RESEARCH

Open Access

Exploring a leading and lagging regions dichotomy: does entrepreneurship and diversity explain it?



Thomas Ejdemo*  and Daniel Örtqvist

* Correspondence: thomas.ejdemo@itu.se
Entrepreneurship and Innovation,
Luleå University of Technology,
SE-971 87 Luleå, Sweden

Abstract

There has been a long debate about the role of industry structure in the literature on why some regions successfully achieve economic growth, while other regions stagnate or decline. This paper provides an empirical analysis in which we, based on a cluster analysis, develop a taxonomy for regional growth. In a second part of the study, we explore how specialization and entrepreneurship are meaningful to discriminate between the different types of regions. Our results suggest that regional entrepreneurship and industry diversity characterized by relatedness are key elements in understanding why some regions are leading while others lag behind. The suggested taxonomy is argued to contribute with a nuanced perspective that can enhance discussions about improvements of regional development policies and to further empirical analysis on the topic.

Keywords: Regional development, Entrepreneurship, Diversity, Taxonomy

Introduction

Already, Adam Smith (1776) made an inquiry into the nature and causes of the wealth of nations. Substantial, conceptual, and empirical research have thereafter focused on the matter of why geographic regions differ in their economic development. Economic historians have documented a strong correlation between regional growth and geographic agglomeration of economic activities (see for instance Hohenberg & Lees, 1985). In this tradition, a vast number of empirical studies have tried to understand the role of industrial concentration and specialization, in terms of the so-called Marshall-Arrow-Romer (MAR) externalities (Marshall, 1890), and economic and social diversity leading to knowledge spillovers between different sectors, in terms of the so-called Jacobs externalities (Jacobs, 1969). Later research has integrated arguments to demonstrate how agglomeration and regional growth reinforce each other (see for instance Feldman & Audretsch, 1999; Acs, 2002), but at the same time concluded that the importance of aforementioned externalities remains unsolved (Glaeser, 2000). Following the theoretical development in the field of studies about regional growth, we learn that economic progress rarely is evenly distributed in space and that regional conditions might differ dramatically (Hirschman, 1958). Still, most attempts to

empirically investigate these questions presume the potential to use regression models and to assume linearity.

In this paper, we first set out to develop a taxonomy for regional growth. Following the EU policy discourse (e.g., European Commission, 2017a) and by applying cluster analysis, we provide an empirical classification of Swedish functional regions in terms of leading, stagnating, and lagging regions. Besides being frequently adopted in policy contexts, the concepts of leading and lagging regions also appear in theoretical economic research (Batabyal & Nijkamp, 2014a, 2014b) as well as in empirical studies (Terluin, 2000; Terluin, 2003). We argue that empirical analysis could gain from using a regional measure that is categorical instead of continuous since we argue that the quantitative scale of regional growth reflects meaningful qualitative differences which have value for theoretical and policy development. We also argue that a categorical analysis of regional growth can bring new detail and understanding which can be useful for further development of linear models.

In a second part of the paper, we provide an empirical analysis that explores how different regional characteristics underpin our proposed taxonomy. We draw on perspectives from economic geography and regional economics, in particular work that is rooted in the endogenous growth theory (Romer, 1986, 1990) which emphasizes knowledge creation and spillovers as a source of growth. This literature suggests that knowledge externalities (e.g., De Groot, Poot, & Smit, 2016; Frenken, Van Oort, & Verburg, 2007; Glaeser, Kallal, Scheinkman, & Shleifer, 1992) and regional capacity for entrepreneurship (e.g., Acs, Braunerhjelm, Audretsch, & Carlsson, 2009; Audretsch & Keilbach, 2004a) may influence the observed regional growth disparities, through harnessing human capital and new ideas. The relevance of a region's industry structure has been examined at length in previous work on the so-called "MAR vs. Jacobs" controversy (see De Groot et al., 2016). The MAR abbreviation refers to theories by Marshall (1890), Arrow (1962), and Romer (1986), which emphasize spatial clustering and within-industry knowledge spillovers as beneficial for innovation and growth, thus favoring regional specialization. In contrast, Jane Jacobs (1969) argued that important knowledge spillovers occur between rather than within industries, and she thus favored a diversity of regional industries instead of specialized clusters. As these theories have conflicting implications for regional innovation and growth, an empirical literature has emerged which considers if diversified or specialized industry structures are more conducive to knowledge spillovers that enhance innovation and growth. An extensive review by De Groot et al. (2016) shows that the available empirical evidence is mixed. Entrepreneurship research has contributed with the insight that the entrepreneur serves as a conduit for the spillover and commercialization of knowledge (e.g., Acs et al., 2009; Acs, Audretsch, & Lehmann, 2013; Braunerhjelm, Acs, Audretsch, & Carlsson, 2010). In this paper, we contribute to the literature on regional knowledge externalities and entrepreneurship using an unconventional empirical approach. We employ multiple discriminant function analysis to predict whether a region is lagging or not, based on regional characteristics such as regional specialization and entrepreneurship. We argue that regional industry structure and regional entrepreneurship together are important determinants for the regional growth groups extracted in the first part of the study.

The paper is structured as follows. The “Study one: A taxonomy for regional growth” section provides the first study where we develop and empirically assess a taxonomy for regional growth. We then proceed to our second study in the “Study two: Linking regional economic specialization and entrepreneurship to the regional growth taxonomy” section where we examine how regional industry structure and regional entrepreneurship together are important determinants for the regional growth clusters developed. We discuss our results and findings in the “Results” section and conclude with implications for research and policy.

Study one: A taxonomy for regional growth

Although interest for regional growth dates back, there has been a substantial development of theories starting in the 1950s. Some theories were conceptualized with the aim to investigate the economic determinants of development and the mechanisms that enable a system to grow and achieve higher rates of output, greater levels of per capita income, lower unemployment rates, and higher levels of wealth. These theories made analytical modeling of the growth path possible, assuming that it is possible to express and model linear relationships of supply and demand conditions with economic outputs. A large number of factors have been revealed to trigger regional growth processes. Capello (2011) mentions for instance: increased demand for locally produced goods; greater local production capacity; a more abundant endowment of local resources and production factors; and a larger amount of savings available for investments in infrastructures and technologies intended to increase the efficiency of production processes.

While this development has been important for furthering our knowledge about what enables a system to grow and achieve higher rates of output, there is still a potential that we overlook the qualitative differences between a region that grows from a region that declines. As a response, we suggest an alternative to linear models of regional growth that could further our understanding of what differentiates between groups of regions and thereby contributes to knowledge and policy design for regional development and growth. We propose and perform a cluster analysis of Swedish functional analysis regions to provide an example of how a typology could be generated.

Research methods and data

Our sample consists of all 60 functional analysis (FA) regions in Sweden. Our main source of data is a database on Swedish regions (rAps-RIS) which is provided by the Swedish Agency for Economic and Regional Growth. FA regions are “subregions” that essentially reflect local labor markets. An important aspect of this regional classification is that FA regions are not administrative units; the classification is instead based on functional aspects and mainly reflects commuting behavior (Tillväxtanalys, 2015). The classification is used in regional analyses by for example the Swedish Agency for Growth Policy Analysis (Tillväxtanalys, 2013) and in the official regional statistics. The current classification was introduced in 2015. It replaced a version from 2005 in accordance with the strategy to revise the FA regions every 10 years to maintain their analytical relevance (Tillväxtanalys, 2015).

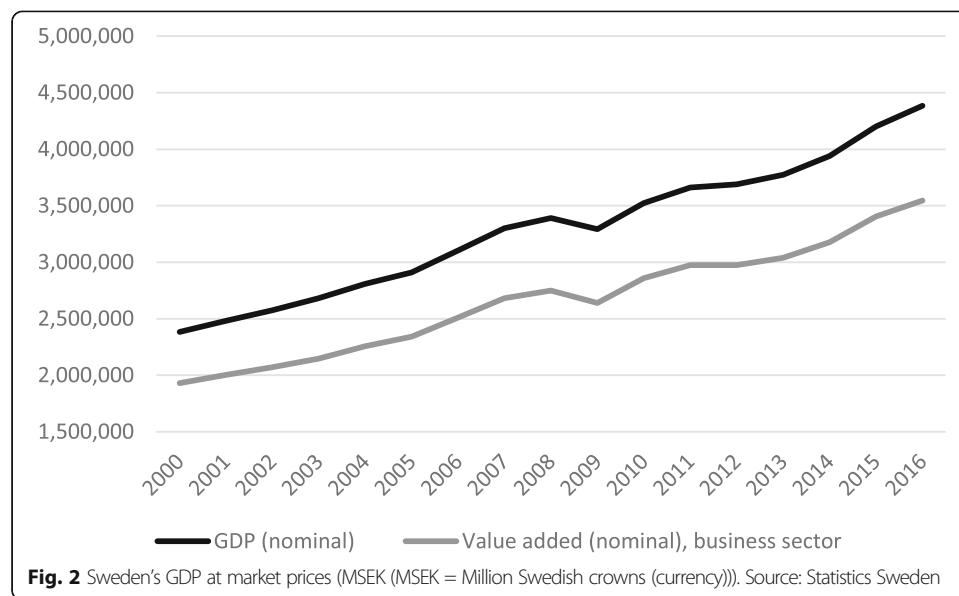


Official statistics on FA regions are thus currently supplied in accordance with the 2015 classification. Figure 1 illustrates a map of the current FA regions.

The 2015 revision resulted in a decreased level of geographic disaggregation, from 72 to 60 FA regions in the current version. This may potentially increase intraregional disparities, particularly as some regions comprise several agglomerations that are dispersed over large areas. While acknowledging this, we choose to rely on the official classification as the Swedish Agency for Growth Policy Analysis emphasizes that FA regions are “...intended to facilitate regional analyses,” by providing an “...appropriate geographic classification” based on local labor markets (Tillväxtanalys, 2015).

Time period

We consider the time period of 2010 to 2015. The period is restricted to these years for two reasons. First, our empirical approach requires access to detailed sectoral employment data, which is affected by a break in the time series since



Statistics Sweden introduced a new industry classification (SNI2007) in 2008. Second, since we are interested in studying regional economic growth, we consider the implications of the financial crisis in 2008/2009. Figure 2 illustrates the development of Sweden's GDP, measured at market prices. As shown, there was a clear downturn in overall output in 2008/2009. The Swedish Agency for Growth Policy Analysis (Tillväxtanalys, 2013) finds that the global financial crisis had a notable effect on Swedish regional economies, which is reflected in the fact that only one FA region achieved positive employment growth between 2008 and 2009. The negative impacts were particularly pronounced in regions with large shares of manufacturing industry. However, according to the same source (ibid), there has been a rapid recovery in the majority of Sweden's FA regions, which appears to be reflected in the GDP development displayed in Fig. 2. In their analysis from 2013, the Swedish Agency for Growth Policy Analysis found that every FA region had achieved growth in the rate of employment since 2009.

Against this background, we argue that it is appropriate to restrict the time period under study so that the very turbulent years of the financial crisis (2008 and 2009) are excluded. We thus opt to use 2010 as our base year, and we consider regional growth trajectories until 2015. Admittedly, this leaves us with a short time period which is still characterized by macroeconomic turbulence following the crisis, as well as potentially exacerbated and accelerated rates of structural change (Tillväxtanalys, 2013). In this regard, regional economic growth trajectories immediately following the crisis may reflect resilience.

Clustering procedure

In this paper, one of our objectives is to explore a lagging-leading region dichotomy. The European Commission (2017a) defines two specific types of lagging regions: low-growth regions and low-income regions. Both types are classified based on regional GDP per capita at PPS (purchasing power standards) relative to the EU average. Since

Table 1 Variables included in the cluster analysis, $n = 60$

Variables	Measured as	Mean ^a	Std. dev.	Min	Max
•GRP per capita growth (nominal) 2010-2015	%*	0.115	0.169	-0.411	0.438
•Population growth 2010-2015	%*	0.008	0.033	-0.084	0.081

*Note that percentages are reported in decimal form here, i.e., 1 = 100%

^aMean change across all regions during 2010-2015.

our focus in this paper is on the relationship between Swedish FA-region growth trajectories and regional characteristics such as industry structure, we adopt a growth-based measure¹, which is computed to reflect the relative differences between Swedish regions. Thus, our classification does not indicate whether an FA region would be classified as lagging or not in an EU context. We compute two common indicators of regional performance and use them as clustering variables: (1) the growth of gross region product (GRP, i.e., regional GDP) per capita between 2010 and 2015 and (2) the population growth in percent, again reflecting the change between 2010 and 2015. In a sense, this indicator controls for per capita changes in regional GDP that are due to demographic development rather than change in aggregate economic output. Regional population growth is also intended to capture qualitative differences between regions, such as the different amenities that make certain regions attractive places to live in. Table 1 summarizes the clustering variables.

The clustering procedure is performed in two steps using SPSS. We follow the common advice to use hierarchical and non-hierarchical procedures in tandem (e.g., Hair Jr., Black, Babin, & Anderson, 2014; Ketchen Jr. & Shook, 1996). First, we employ hierarchical clustering with Ward's method to explore an appropriate number of clusters. We use the squared Euclidean distance as a measure of similarity. We then examine a cluster solution using non-hierarchical (k means) clustering and perform one-way ANOVA tests for significant differences between clusters for each variable.

Analysis

We begin by performing a cluster analysis to explore if groups of FA regions can be identified and understood in a meaningful way based on their different growth trajectories between 2010 and 2015. Our results suggest that three groups can be distinguished in the data (agglomeration coefficients are provided in Appendix). We then proceed to non-hierarchical clustering (k means) where we impose a three-cluster solution. A stable solution (i.e., further iterations do not change cluster centroids) is achieved in five iterations. The three clusters consist of $n_1 = 31$, $n_2 = 14$, and $n_3 = 15$ FA regions. Table 2 reports the cluster means and the results from an ANOVA which indicates significant differences of means across the clusters. Post hoc tests (provided in Appendix) show that means are significantly different from each other across all clusters and variables.

Figure 3 illustrates the distribution of cluster means in the solution. GRPGr represents the growth of regional GDP (gross region product) at market prices between 2010 and 2015 and PopGr represents population growth in the same period. As a

¹We note that the literature contains example of alternative measures for classification. Terluin (2000) and Terluin (2003) use a classification of leading and lagging EU regions based on their non-agricultural employment growth.

Table 2 Results of *k*-means cluster analysis

Variable	Mean values by cluster number			<i>F</i>	<i>p</i>
	1	2	3		
GRPGr	0.148	-0.142	0.28	140.155	0.000
PopGr	0.006	-0.020	0.038	17.410	0.000
Cluster sample size	31	14	15		

reference, the growth of Sweden's GDP calculated from the producer side was approximately 19%.

The clustered groups of FA regions can be described as follows:

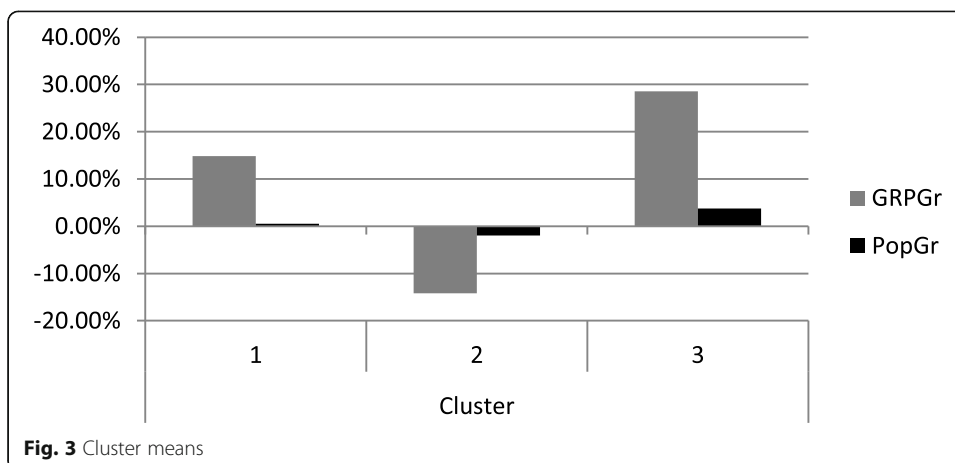
1. A group of slow-growing, *stagnant* FA regions ($n = 31$), where production values as measured by gross region product increased, but at a slower pace than GDP. Their population development has been almost stationary during 2010-2015.
2. A group of declining, or *lagging* FA regions ($n = 14$), whose economies have contracted during the period, in addition to exhibiting a negative population development.
3. A group of fast-growing, *leading* regions ($n = 15$) who are outperforming national GDP-growth. These FA regions have also enjoyed population growth during the period.

Figure 4 illustrates the distribution of mean group values for GRP- and population growth between 2010 and 2015. The bubbles are scaled to reflect the sum of gross region product in 2015 within each group. As the figure shows, the leading regions generated approximately 75% of Sweden's GDP in 2015. The regions classified as stagnant generated about 22%, while the lagging regions generated a mere 3%.

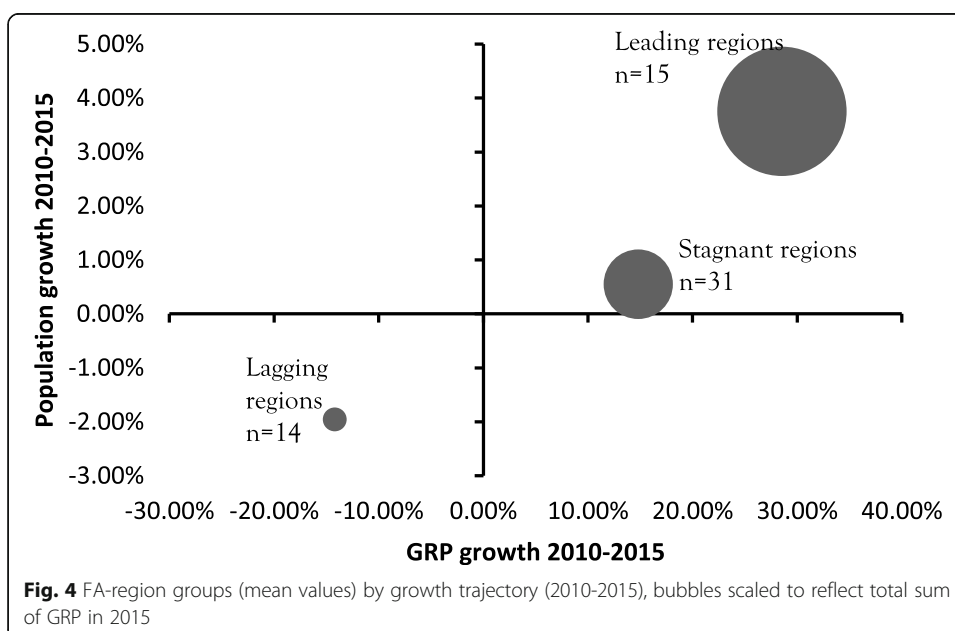
Figure 5 provides a map where the regional growth trajectories of Swedish FA regions between 2010 and 2015 are plotted. We emphasize that the labels of stagnant, lagging, and leading regions should not be interpreted as qualitative assessments. This terminology simply reflects quantitative outcomes in nominal gross region product and population growth during the observed period.

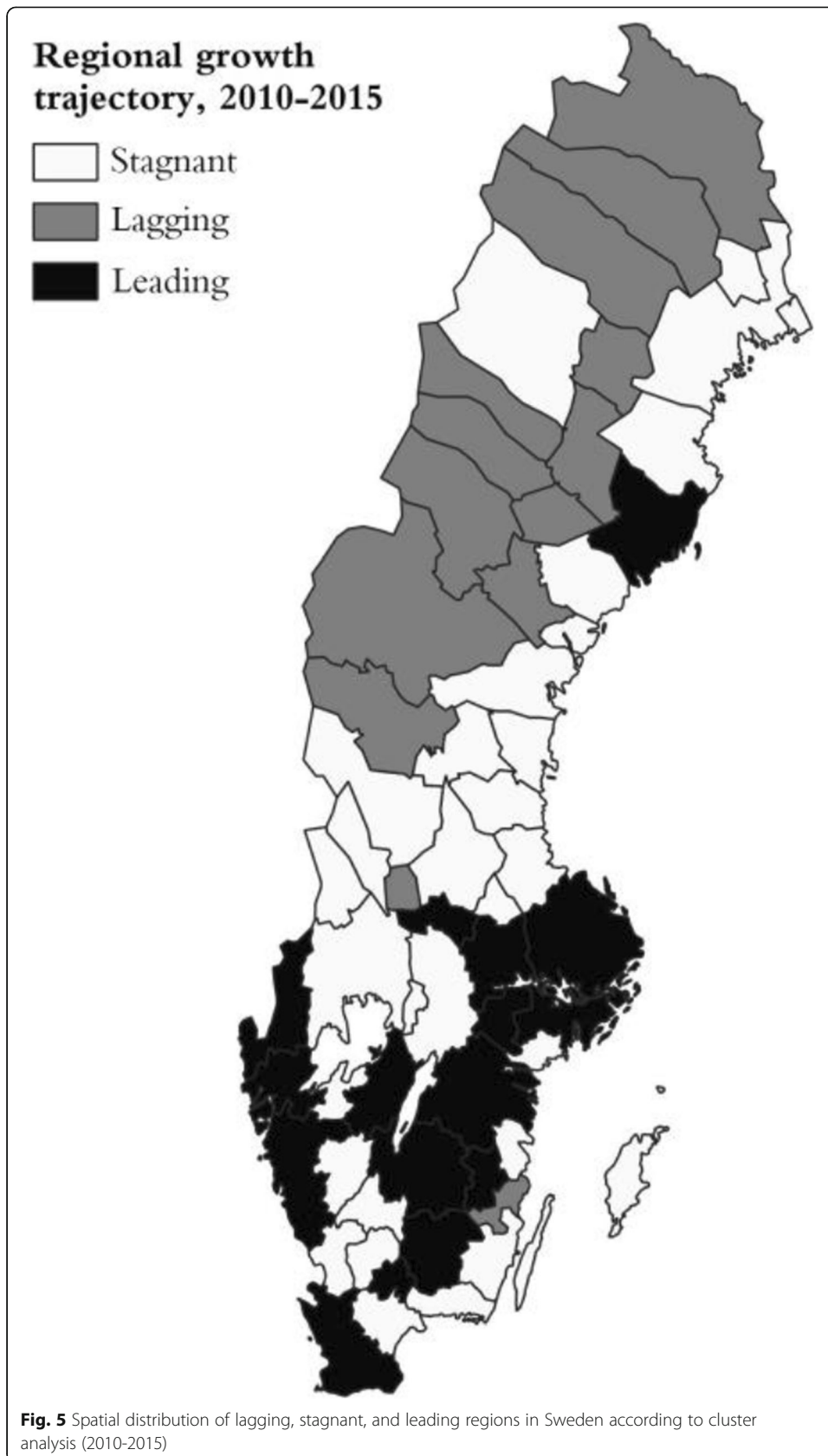
The second stage of the analysis aims to explore why we observe these disparities in regional growth. We can however note generally that the majority of the regions that are classified here as "lagging" are located in what is sometimes referred to as the northern periphery². These regions tend to be small in terms of population size with limited economic diversity, often relatively specialized on primary production such as mining, forestry, agriculture, and energy production. Their gross region product is therefore often sensitive to price fluctuations such as the downturn in metal prices between 2011 and 2016 (see for example commodity price data reported by the IMF, 2020) and their labor markets are limited. According to the results of our cluster analysis, Lycksele in northern Sweden is a typical representative of such a region. In contrast, the "leading" regions include Sweden's three large

²The EU has used this terminology, for example in their "Northern Periphery and Arctic Programme 2014-2020."



metropolitan regions Stockholm, Gothenburg and Malmö, and the regions adjacent to them. These three metropolitan regions are characterized by economic diversity including large knowledge-intensive sectors. About 49 percent of Sweden’s population resided in these metropolitan regions in 2010. We also note that Umeå is the only northern region which is classified as “leading” according to the cluster analysis. Umeå is a comparatively large urbanization with a sizeable knowledge-intensive sector, not least due to its university and hospital which are both large relative to other northern regions in Sweden. Unlike most northern regions, Umeå achieved a population growth of several percent during the observed period 2010-2015. Lastly, the regions classified as “stagnant” are more difficult to characterize in terms of key differences. They tend to be mid-sized and have relatively diversified economies, but are typically not located adjacent to a metropolitan region.





Karlstad is a typical representative of such a region, according to our cluster analysis.

To proceed to the second stage of our analysis, we require that (1) our cluster analysis identifies groups that are significantly different from each other, and that (2) groups (clusters) can be understood in a meaningful way. Since these criteria are fulfilled, we can now examine the differences between groups further to explore the role of various regional characteristics in understanding differences in regional development outcomes. In the second part of the study we examine to what extent regional specialization and entrepreneurship are important to discriminate leading, stagnating, and lagging regions from each other.

Study two: Linking regional economic specialization and entrepreneurship to the regional growth taxonomy

Regional growth theory offers different perspectives on specialization as either advantageous or potentially detrimental to regional growth. Broadly, these theories consider the presence and importance of different types of agglomeration economies (McCann, 2009). In this section, we do not intend to provide a comprehensive review but we briefly discuss some of these perspectives. Our empirical analysis draws mainly on arguments that are emphasized in the literature on endogenous growth determinants—the implications of industry specialization vs. diversity for knowledge externalities, and the role of entrepreneurship for regional growth (Capello, 2009). Accordingly, this section first considers agglomeration economies that are associated with industry composition, followed by a brief discussion of the role of entrepreneurship for regional growth.

Specialization, diversity, and regional growth

The notion that specialization is important for regional growth is not new³. Arguments for such a relationship traces back to Adam Smith (1776) and to David Ricardo (1817) who contributed with the theory of comparative advantage. Krugman (2009) succinctly defines comparative advantage as meaning that “countries trade to take advantage of their differences.” The notion of comparative advantage has important implications for regional analyses. Drawing on comparative advantage, the work of Heckscher (1919) and Ohlin (1924) became the foundation of the Heckscher-Ohlin model (see for ex. Jones, 1956), which predicts that regions will specialize in producing and exporting products that are intensive in their relatively abundant factors of production (e.g., Kim, 1995). The local availability of inputs is however only one of several factors that may influence the location decision of firms. It is not our ambition to provide a comprehensive review of location theory, nor to review the literature on regional growth in this paper. Instead, we draw on for example Frenken et al. (2007) who note that the economics of agglomeration

²The EU has used this terminology, for example in their “Northern Periphery and Arctic Programme 2014–2020.”

³Ideas about the relationship between specialization and economic development can (at least) be traced back as far as Plato’s *Republic*, where he emphasizes that people’s skills are heterogeneous and therefore, division of labor within a state leads to greater efficiency (e.g., Sandelin, Trautwein, & Wundrak, 2001).

posits that economic activity tends to cluster because firms experience benefits from locating near other firms. Frenken et al. (2007) outline four sources of such benefits, or agglomeration economies, which we discuss briefly here.

First, the field of New Economic Geography (NEG), which in particular builds on the work of Krugman (1991), emphasizes increasing returns to scale as a benefit of agglomeration. According to NEG, the presence of large economies of scale, low transport costs, and enough mobile production can lead to a self-sustaining spatial concentration of production (i.e., specialization), and its location can be a result of initial conditions or mere historical “accident” (Krugman, 2009).

The second source of agglomeration economies discussed by Frenken et al. (2007) is localization economies. These reflect Marshallian externalities that emanate from sectoral specialization, which enables labor market pooling, creation of specialized suppliers, and fosters knowledge spillovers⁴. Localization economies are thus available to all same-industry firms in the area. Stigler (1951) emphasized the relationship between size and degree of specialization: the creation of specialized suppliers requires that the industry or market is sufficiently large.

The third type of agglomeration benefit that Frenken et al. (2007) identify consists of external economies that arise from urban size and density. These urbanization economies are available to all local firms. This concept reflects advantages such as a dense presence of diverse economic, social, political, and cultural organizations and actors, typically associated with large urbanizations (cities). Florida, Adler, and Mellander (2017) argue that cities have distinct advantages as they provide “the enabling infrastructure where connections take place, networks are built and innovative combinations are consummated.”

This line of reasoning is connected to the fourth source of agglomeration economies outlined by Frenken et al. (2007)—*Jacobs externalities*. This concept draws on the work of Jane Jacobs (1969), who emphasized variety and diversity of industries as conducive to innovation and growth due to knowledge spillovers between firms in different industries. Frenken et al. (2007) argue that knowledge spillovers between industries should facilitate radical innovation and product innovation in particular, as knowledge and technologies are recombined in new ways.

Whether regions benefit more from specialization or diversification remains a contested issue which is summarized in the “MAR vs. Jacobs” controversy described by for example Content and Frenken (2016). The empirical evidence on these theories is mixed (e.g., De Groot et al., 2016). Frenken et al. (2007) made an important contribution to this literature with the notion of related variety, which implies that when a diversified economy is characterized by industries that are related in a technological or market sense, it is especially supportive of innovation and regional economic growth, as it fosters Jacobs-type knowledge externalities. This notion has gained increasing empirical support (e.g., Boschma & Iammarino, 2009; Da Silva, Goncalves, & De Araújo, 2019; Miguelez & Moreno, 2018; Östbring & Lindgren, 2013; Tavassoli & Carbonara, 2014; van Oort, de Geus, & Dogaru, 2015).

⁴Generally, externalities that arise from knowledge spillovers within a specialization are referred to in the literature as Marshall-Arrow-Romer (MAR) externalities (e.g., Glaeser et al., 1992; McCann, 2009).

In summary, this limited account of some basic perspectives found in regional growth theory indicates that the field offers a diverse set of explanations for the sources of advantages which lead some regions to prosper while others lag behind. These advantages may include the presence of comparative advantages, economies of scale, and knowledge spillovers between or within industries, facilitated by urban size and density. In this paper, we are particularly interested in the role of industry composition and hence, our empirical analysis will focus mainly on two types of external economies: localization economies and Jacobs externalities. In that sense, we contribute to the literature on “MAR vs. Jacobs.”

Entrepreneurship and regional growth

Solow (1956) argued that economic growth is determined by the stocks of capital and labor. Later, researchers have emphasized that investments in knowledge is an important source of growth since it also spills over for use of third parties (Romer, 1986). Audretsch and Keilbach (2004a) in turn argued that such spillovers require economic agents which create and make use of diversity of knowledge. As such, entrepreneurship has been deduced as an important mechanism facilitating the spillover of knowledge through creating diversity of knowledge (Audretsch & Keilbach, 2004b).

There are several empirical studies examining the relationship between regional levels of entrepreneurship and economic growth. Holtz-Eakin and Kao (2003) performed an empirical test of variations in the birth rate and the death rate for firms across different states in the USA and they found this measure of entrepreneurship to have a positive impact on productivity growth. Audretsch and Fritsch (2002) found similar results where entrepreneurship had a positive impact on regional growth in Germany. Also, Foelster (2000) has found that regional levels of self-employment influence regional growth in Sweden. Callejon and Segarra (1999) performed a study in Spain where they found that new-firm startup rates and exit rates both contribute positively to the growth of total factor productivity in regions. As such, there is ample support suggesting regional entrepreneurship to be positively related to regional growth.

Methods

Since cluster membership is a categorical variable, we employ multiple discriminant analysis to examine how a set of independent continuous variables on regional characteristics such as entrepreneurship and specialization perform in predicting group membership. We aim to examine if the level of specialization and other characteristics in one period seems to influence future growth performance, and therefore we mainly use independent variables covering the year 2010, while the dependent variable (group membership) is constructed based on growth between 2010 and 2015. This approach differs from the extant literature (e.g., Frenken et al., 2007) which typically employs multiple regression analysis and treats economic outcomes in a direct fashion as continuous measures of dependent variables, for example, growth in regional employment or GDP. Our approach enables us to conceptualize economic outcomes as regional growth trajectories that combine economic and demographic development in categorical region types (i.e., lagging,

leading, etc.) which may provide multidimensional and nuanced perspectives to the literature. We do not claim that this approach is necessarily better than typical multiple regression analyses, but rather it is complementary. The results shed light on the role of entrepreneurship, industry structure, and other aspects of regional economies.

Operationalizing specialization

In a study on regional industrial specialization and risk sharing in the EU15, Basile and Girardi (2010) argue that there exists no optimal measure of specialization. In addition, they find that the conventional indices are strongly correlated. They instead assert that a more critical issue is the choice of variable. In their view, specialization can be observed from the input side through employment or from the output perspective by using measures such as value added or exports. Basile and Girardi (2010) favor employment data, as it is less sensitive to valuation problems compared to production data. A third option is proposed by for example Nakamura (2009), who suggest that the number of firms or plants may provide a different perspective than the more conventional expression in terms of employment.

A common approach is to operationalize regional economic specialization through some relative measure of employment specialization. We follow this conventional approach and operationalize three different conceptualizations of specialization or diversity through the use of employment data. We choose the Herfindahl-Hirschman index (HHI) as a measure of the overall level of regional specialization, i.e., measuring the variety of industries in a region. The HHI is computed as follows. Let S_i denote the number of industries in region i , e_{si} is employment in region i in industry s , e_i is total employment in region i . The HHI for each region is then computed as follow:

$$\text{HHI}_i = \sum_{s=1}^{S_i} \left(\frac{e_{si}}{e_i} \right)^2$$

A value near zero indicates a highly diversified regional economy and a value of 1 means that the region is completely specialized in a single sector. Following Kemeny and Storper (2015), we use disaggregated data and compute regional Herfindahl indices using five-digit SNI⁵ 2007 employment data. These data are fully compatible at the four-digit level with the NACE⁶ Rev.2 classification used by Eurostat. We use the same data to compute an indicator of related variety (RV), following the formula proposed by Frenken et al. (2007). It is computed as follows: employment in a five-digit industry i belongs to an aggregate sector S_g ($g = 1, \dots, G$), which is at the 2-digit level in this study. P_g is the 2-digit sector's share of total employment in the region, and P_i is the 5-digit industry's employment share. We then calculate related variety (RV) as follow:

⁵SNI is an abbreviation for "Standard för svensk näringsgrensindelning," which translates to the Swedish Standard Industrial Classification, used in the official statistics. See SCB (2007).

⁶NACE Rev.2 is revision two of the classification of economic activities used by Eurostat, the statistical office of the EU. NACE is an abbreviation for "Nomenclature générale des Activités économiques dans les Communautés Européennes" (in English: Statistical classification of economic activities in the European Communities). See Eurostat (2008).

$$RV = \sum_{g=1}^G P_g H_g$$

where:

$$H_g = \sum_{i \in S_g} \frac{P_i}{P_g} \log_2 \left(\frac{1}{P_i/P_g} \right)$$

High values of RV relative to other regions suggest that the industry structure is relatively more characterized by industries that are related in a technological sense. We select a five-digit SNI level since high levels of aggregation may lump industries together as comparable specializations, when in reality they are “apples and oranges” (Kemeny & Storper, 2015).

Operationalizing entrepreneurship

There is compelling empirical evidence on the link between entrepreneurship and regional growth and development (Audretsch & Keilbach, 2004a; Fritsch & Wyrwich, 2017). Entrepreneurship is a multidimensional concept that is challenging to measure empirically (Fischer & Nijkamp, 2009). We use the concept of entrepreneurship capital (Audretsch & Keilbach, 2004a), which is measured by the number of startups in a region relative to its population. High values of entrepreneurship capital relative to other regions indicate “a regional milieu of agents that is conducive to the creation of new firms” (ibid). We obtain data on new business formation from a database maintained by the Swedish government agency Growth Analysis (Tillväxtanalys). Fritsch (2015) emphasizes that the impacts of new businesses may take considerable time to manifest, even up to ten years. To account for this, we calculate an average rate of new business formation per 1000 inhabitants in ages 16-64 for the years 2008-2010. This treatment of entrepreneurship capital is similar to the approach used by Audretsch and Keilbach (2004a), who compute their measure based on the number of startups between 1989 and 1992, where 1992 is the base year of their analysis.

Additional independent variables

We include the initial level of economic output in 2010, which we measure as GRP (i.e., regional GDP) per capita. This variable is intended to account for the possibility that regions that were highly productive in 2010 may enjoy competitive advantages, while regions that are lagging behind may struggle to overcome disadvantages that are unobserved in this analysis.

Furthermore, there exists compelling support for the notion that human capital is associated with economic development (e.g., Barro, 2001; Lucas, 1988). We therefore include the share of each FA-region’s population in the ages 20 to 64 years with tertiary (post-secondary) education of at least 3 years (i.e., equivalent of bachelor’s degree), again for the year 2010. The use of educational attainment as a measure of human capital endowment is not without critique (Florida, Mellander, & Stolarick, 2008), yet it remains a common measure of this concept in the empirical literature (e.g., Lee, Florida, & Gates, 2010; Tavassoli & Carbonara, 2014).

There are certainly other regional characteristics that would be relevant to include in the analysis. One example is to include a variable which captures urbanization economies (i.e., region size and/or density). However, our relatively small sample size poses a limitation since we apply discriminant analysis. Hair Jr. et al. (2014) suggest a ratio of 20 observations for each independent variable, with a minimum recommended size of five observations per independent variable. For this reason, we restrict the number of variables in this exploratory analysis.

Summary statistics of independent variables

Table 3 lists our independent variables. Our data contained extreme outliers for GRP per capita ($n = 3$ regions dominated by capital-intensive extractive industries). Discriminant analysis is sensitive to outliers Hair Jr. et al., 2014, and for this reason, we eliminate the extreme outliers. After this, several variables still display non-normal distributions, which we address by transforming GRP10 and HHI10 to their natural logarithms.

Discriminant function

We estimate the following discriminant function:

$$Z_{jr} = W_0 + W_1 \ln GRP_r + W_2 Humcap_r + W_3 Entcap_r + W_4 \ln HHI_r + W_5 RV_r$$

Where Z is the discriminant score of function j for region r , W denotes the discriminant weights for each independent variable, $\ln GRP$ is the natural logarithm of GRP per capita in 2010, $Humcap$ is the share of population age 20-64 years with at least 3 years tertiary education in 2010, $Entcap$ is entrepreneurship capital (average for 2008-2010), $\ln HHI$ is the natural logarithm of the Herfindahl-Hirschman index of overall specialization (diversity), RV is the index of industry relatedness, and W_0 is the constant.

We use the simultaneous estimation procedure, which means that the discriminant function is estimated based on the entire set of independent variables in a confirmatory approach to assess the overall performance. Technically, discriminant analysis estimates several functions if the number of groups is more than two, as in our case. The number of estimated functions is equal to the number of groups, less one. Each function will then represent a different dimension of discrimination (Hair Jr. et al., 2014). We employ the leave-one-out classification, which reduces optimistic bias due to small sample sizes (Cox & Wang, 2014).

Table 3 Independent variables

Variable	Definition	N	Mean	Std. dev.	Min	Max
lnGRP	Log of GRP per capita in 2010 ('000 SEK)	57	5.713	0.173	5.29	6.17
Humcap	Human capital, measured as the share of population in ages 20-64 years with > 3 years tertiary education in 2010 (%*)	60	0.149	0.042	0.089	0.281
Entcap	Entrepreneurship capital, measured as new business formation per 1000 inhabitants in ages 16-64 (annual average 2008-2010)	60	8.606	1.570	5.81	13.88
lnHHI	Overall specialization/diversity, measured by the Herfindahl-Hirschman index (HHI) of overall employment concentration for 2010 (log)	60	-3.948	0.400	-4.67	-2.91
Relvar	Related variety: weighted sum of 5-digit entropy within 2-digit sectors (calculated with employment data for 2010)	60	1.777	0.444	0.687	2.491

*Note that percentages are reported in decimal form here, i.e., 1 = 100%

Table 4 Comparison of group means

Variables	Stagnant regions		Lagging regions		Leading regions		ANOVA	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	F	P
lnGRP	5.674	0.145	5.824	0.225	5.715	0.157	3.316	0.044*
Humcap	0.144	0.031	0.121	0.029	0.182	0.057	8.744	0.001*
Entcap	8.300	1.208	8.877	1.994	8.968	1.816	1.180	0.315
lnHHI	-4.023	0.309	-3.552	0.309	-4.185	0.390	12.275	0.000*
Relvar	1.847	0.352	1.285	0.352	2.092	0.329	17.800	0.000*

*Statistically significant difference between group means ($p < 0.05$)

Results

The cluster analysis identified three groups of FA regions which are significantly different in terms of economic and demographic growth between 2010 and 2015. This section reports the results of a multiple discriminant analysis where we examine the role of overall economic specialization (and its flip-side, diversity), industry relatedness, entrepreneurship capital, human capital, and past regional economic performance, in predicting these outcomes.

Table 4 reports descriptive statistics of the group means for each variable, and the results of an ANOVA which tests for equality of group means. We find that there are statistically significant differences between group means for all independent variables except entrepreneurship capital.

The key assumptions of discriminant analysis are multivariate normality of independent variables and equal covariance matrices across groups (Hair Jr. et al., 2014). We addressed departures from normality with appropriate transformations described in the “Study two: Linking regional economic specialization and entrepreneurship to the regional growth taxonomy” section. Equality of covariance matrices is assessed with the Box’s M test. Tabachnick and Fidell (2013) suggest using $p < 0.001$ as a criterion. In our case, the test yields a p value of 0.004, which means that it passes. Furthermore, the pooled within-groups matrices (Table 5) show relatively low correlations, except for our relatedness variable (Relvar) which correlates strongly with lnHHI, which is the diversity indicator. This suggests that multicollinearity may be problematic.

Table 6 reports results for the two discriminant functions. The overall significance of each function is evaluated with Wilks’ lambda, which is significant for functions 1 through 2 ($p = 0.000$) and significant at the 10% level for function 2 ($p = 0.095$). The squared canonical correlations show that function 1 explains 51.8% of variance in group membership and function 2 explains 14.1% of variance.

We follow the approach recommended by Hair et al. (ibid) and focus on interpreting the discriminant loadings, which measure the correlation between each independent variable and

Table 5 Intercorrelation among independent variables (pooled within-groups matrices)

	<i>lnGRP</i>	<i>Humcap</i>	<i>Entcap</i>	<i>lnHHI</i>	<i>Relvar</i>
lnGRP	1.000	0.457	0.268	-0.374	0.366
Humcap	0.457	1.000	0.203	-0.587	0.669
Entcap	0.268	0.203	1.000	-0.202	0.138
lnHHI	-0.374	-0.587	-0.202	1.000	-0.928
Relvar	0.366	0.669	0.138	-0.928	1.000

Table 6 Standardized discriminant loadings and discriminant function statistics

Variable	Discriminant loadings		Discriminant function statistics	
	Function		Function	
	1	2	1	2
lnHHI	-0.649*	-0.044		
lnGRP	-0.278	0.489*		
Humcap	0.466	0.739*		
Entcap	-0.035	0.508*		
Relvar	0.778*	0.190		
Wilks lambda			0.4013 ($p = 0.000$)	0.859 ($p = 0.095$)
Canonical correlation (sq.)			0.518	0.141

*Largest absolute correlation between each variable and any discriminant function

the discriminant function. According to the recommended approach, we interpret loadings of 0.40 or greater as substantive contributions to the respective function.

We find that the majority of our independent variables load stronger than 0.40 on function 1, with the most important being industry relatedness (*Relvar*) in combination with overall industry diversity (*lnHHI*). In function 2, human capital (*Humcap*), entrepreneurship capital (*Entcap*), and productivity (*lnGRP*) display substantive loadings. The only notable cross loading is observed for human capital endowment.

Previous literature on discriminant analysis suggests that a label should be assigned to each function to enhance its interpretation (e.g., Perreault Jr., Behrman, & Armstrong, 1979). We posit that function 1 represents a structural dimension. A region with high levels of industry relatedness and overall diversity will score high on this function. We argue that function 2 represents an entrepreneurial dimension. Regions with high levels of entrepreneurship capital, productivity, and human capital will score high on this function. This is consistent with Audretsch and Keilbach (2004a) who found that increases in entrepreneurship capital and knowledge capital were positively associated with increases in regional labor productivity in Germany. The loadings of

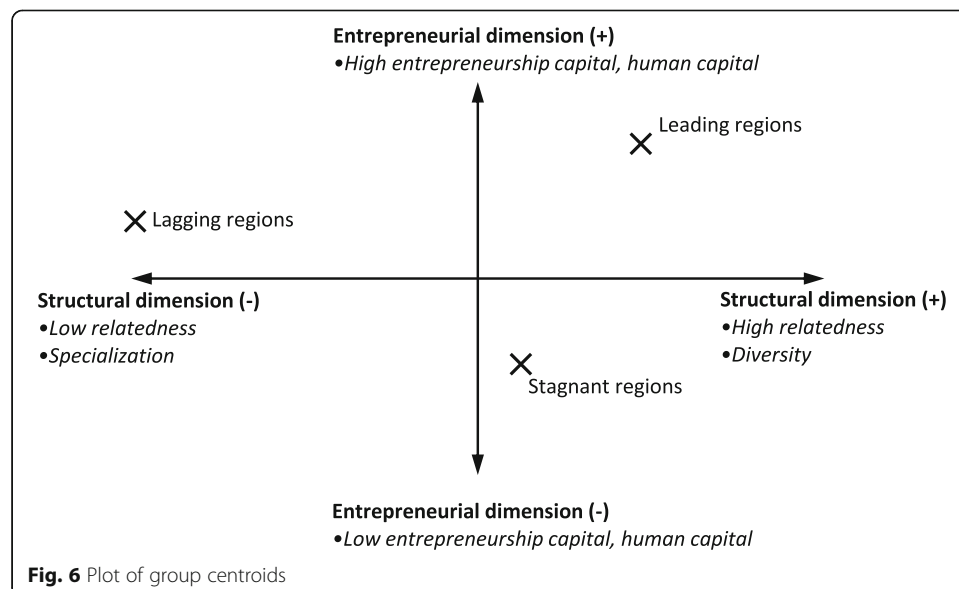


Table 7 Classification results (percent)

Original	Predicted group membership		
	Stagnant regions	Lagging regions	Leading regions
Stagnant regions	90.3%	6.5%	3.2%
Lagging regions	27.3%	63.6 %	9.1%
Leading regions	60.0%	6.7%	33.3%
Cross-validated*			
Stagnant regions	83.9%	9.7%	6.5%
Lagging regions	27.3%	63.6%	9.1%
Leading regions	60.0%	6.7%	33.3%

*Leave-one-out classification

Humcap suggests that human capital is an important aspect in both the structural and entrepreneurial dimensions.

Figure 6 visualizes our results in a plot of group mean scores (centroids) on function 1 (structural dimension) on the X-axis and function 2 (entrepreneurial dimension) on the Y-axis.

Evaluated by the group centroids, leading regions seem to have advantages in terms of entrepreneurship, human capital, industry relatedness, and diversification. Stockholm is a prominent example of such a region, as it had the highest observed entrepreneurship capital among all regions, combined with the second-highest human capital endowment (only bested by Umeå on this measure) and a diversified industry structure characterized by relatedness. Stagnant regions mainly seem to lack entrepreneurship capital and human capital, compared to the leading regions. Examples of such regions include Karskoga, Värnamo, and Oskarshamn, to name a few. They are among the several stagnant regions that have below-average entrepreneurship capital as well as human capital and these aspects seem to be a meaningful difference when comparing them to leading regions. The results for lagging regions show that they are clearly the most specialized, and interestingly, they seem to be more entrepreneurial than the stagnant regions. In fact, 7 out of 14 regions classified as “lagging” had above national average entrepreneurship capital during the observed period. Examples of regions classified as “lagging” with above-average entrepreneurship capital include Åsele, Jokkmokk, and Storuman. A common characteristic is that these regions are located in the northern periphery and are relatively specialized on primary production (i.e., agriculture, raw materials, energy supply). This may suggest that new businesses created during 2008 to 2010 in the declining regions have not exhibited the same growth performance as new businesses in leading regions, possibly due to regional specialization in contracting sectors. This growth trajectory is in line with the concept of path dependence, which recognizes that some regions become locked in to development paths that eventually lose dynamism (Martin & Sunley, 2006). In addition, highly specialized regions become vulnerable to fluctuations in business cycles. This may explain why some regions in our sample have experienced economic decline during the period we examine.

Finally, the classification results are provided in Table 7. Overall, 70.2% of the original group cases were correctly classified. The cross-validated results show that the hit ratio falls to 66.7%. Hair Jr. et al. (2014) suggest comparing the hit ratio with both the maximum and proportional ratio of cases that could be correctly classified by chance,

Table 8 Polar extremes approach: classification results

Original	Predicted group membership	
	Lagging regions	Leading regions
Lagging regions	90.9%	9.1%
Leading regions	6.7%	93.3%
Cross-validated*		
Lagging regions	90.9%	9.1%
Leading regions	20.0%	80.0%

*Leave-one-out classification

multiplied by a threshold value of 25 percent (i.e., the classification should be 25 percent better than pure chance). In our case⁷, the maximum chance criterion is 68% and the proportional chance criterion is 50.3%. Thus, the overall hit rate passes both criteria, while the cross-validated results pass according to the proportional chance criterion.

Our results suggest that the discriminant function performs well in predicting stagnant regions based on their overall level of diversity, relatedness, and additional regional characteristics. Lack of entrepreneurship and human capital seems to be important dimensions in understanding the growth trajectories of these regions. The discriminant function performs reasonably well in predicting lagging regions, but poor in predicting leading regions. Many leading regions were incorrectly classified as stagnant, which suggests that there is some form of overlap between them that is not captured in this analysis. This may be due to the lack of a variable that controls for urbanization economies, and (or) lack of a more elaborate conceptualization of localization economies which distinguishes between different types of specialization.

A “polar extremes” approach

We are particularly interested in examining differences between leading and lagging regions. Therefore, we employ a “polar extremes” approach as suggested by Hair Jr. et al. (2014) and only compare these two groups of regions. As we are now dealing with two groups, only one discriminant function is estimated. The results are evaluated using the same criteria as above. The function is statistically significant ($p = 0.000$) and according to the discriminant loadings, relatedness and diversity in particular and to a lesser extent human capital are important dimensions in discriminating between the lagging and leading regions. Entrepreneurship capital does not seem to be a defining difference between these regions. Group centroids show that lagging regions score low on the function (-1.766), which suggests that they are relatively specialized, with low levels of relatedness and human capital. Although our analysis does not control for urbanization economies, the characteristics of lagging regions are consistent with smaller regions in terms of population size. Leading regions score high (1.295) which implies the opposite—they enjoy high levels of relatedness, diversity, and human capital. Table 8 reports the classification results. Clearly, the discriminant function is able to draw lagging and leading regions apart based

⁷The *maximum chance* criterion is the hit ratio given by assigning all observations to the group with the highest probability. In our sample ($n = 57$, with $n = 31$ or 54.4% stagnant regions), this becomes: $54.4\% \times 1.25 = 68\%$ The *proportional chance* criterion is given by squaring the proportions of each group, adding them, and multiplying by 1.25, which in our case = 50.3%

on their levels of diversity, relatedness, human capital, and other regional characteristics. The overall hit ratio is 92.3%.

Discussion

The present paper set out to develop and empirically test a taxonomy for regional growth and to explore how specialization and entrepreneurship are meaningful in understanding why some regions are leading while others lag behind. Our aim in doing this is to contribute to the still ongoing debate concerning the merits of specialization and entrepreneurship in regional development. Our empirical analysis draws mainly on theories on agglomeration economies and on previous empirical work that emphasizes the creation and spillover of knowledge as a source of regional growth.

Overall, our study suggests that empirical and theoretical contradictions from studies published in the field might be due to a lack of a more fine-grained conceptual development. We offer a categorization of regions into three different types depending on their growth trajectories. The majority of the Swedish functional analysis regions we study appear to be stagnant, and two roughly equal-size groups are polar opposites. These comprise leading regions, who enjoy demographic growth and whose economies have outpaced GDP-development, and lagging regions, where population levels are decreasing and regional economies appear to be contracting.

Our analysis suggests that the overall diversity of industries in combination with high levels of technological relatedness and regional entrepreneurship capital have important roles in explaining different regional development outcomes in Sweden during the period 2010 to 2015. Leading regions tend to be associated with a diversified industry structure characterized by industry relatedness and with a regional milieu that fosters entrepreneurship and benefits from human capital endowment. This is in line with the view of, e.g., Jacobs (1969) and Florida et al. (2017) who emphasize that diversity, which is typically associated with larger agglomerations, generates important knowledge spillovers and other benefits. These results are also in line with the growing empirical evidence on the positive relationship between industry relatedness and economic development (e.g., Content & Frenken, 2016). According to our analysis, the most prominent examples of such regions in Sweden are the three metropolitan regions: Stockholm, Gothenburg, and Malmö. Other regions that are classified as “leading” that seem to benefit from economic diversity, high degree of relatedness and entrepreneurship include Västlandet, Västerås, Linköping, and Växjö. We note that Umeå stands out as the only northern periphery region to be classified as “leading” in our analysis due to its growth in both economic and demographic terms. There may of course be several reasons for such a development. A standout characteristic of Umeå is that it had the highest human capital endowment of all Swedish functional analysis regions in 2010, which is likely due to its large university, hospital, and other knowledge-intensive economic activities. In contrast to leading regions, the lagging regions in our taxonomy tend to have more specialized industry structures. They are on average characterized as relatively entrepreneurial, although they often have comparatively low levels of human capital. An important reason for their decline during the period we observe (2010 to 2015) appears to be their degree of specialization. We do not find room to examine this explicitly in this article, but our data clearly point to specializations on the primary sector, meaning activities such as agriculture, extraction of raw materials, and energy production. Examples of such regions include Åsele, Jokkmokk, and Storuman. It seems that entrepreneurial efforts in these regions have not been successful enough to substantially

impact their growth trajectories and regional specializations on contracting sectors may be an underlying reason which we only address qualitatively in this study. Furthermore, the lagging regions tend to be small and located in the northern periphery-area, which is generally associated with low economic diversity and limited labor markets.

Lastly, stagnant regions mainly seem to lack entrepreneurship and their economies have lower levels of diversity and relatedness, compared to regions that are growing rapidly. Proximity to metropolitan regions also seem to matter, as typical stagnant regions such as Karlstad, Karlskoga, and Sundsvall are not located adjacent to metropolitan regions. In contrast, leading regions are to some extent clustered around metropolitan areas.

Overall, our study elaborates upon conceptual classifications of regions into three regional categories of stagnant, lagging, and leading regions. We operationalize the suggested regional categories so that the roles of diversification, relatedness, and entrepreneurship on a regional level can be examined in a meaningful way. Further, our study elaborates on the role of the aforementioned concepts on regional development outcomes. The results of our study demonstrate that high levels of economic diversity in combination with industry relatedness and entrepreneurship are key elements in understanding why some regions prosper. These results emphasize the importance of knowledge spillovers between industries, as well as the entrepreneurial capacity to transform knowledge spillovers into actual economic growth.

A limitation of our empirical analysis is that we do not explicitly examine urbanization economies, nor do we consider a more disaggregated conceptualization of localization economies (industry-level specialization) other than in qualitative terms. This is mainly due to the small sample size, which limits the number of variables we can use. Future studies that use similar approaches could overcome this limitation by using data that is more disaggregated at the geographical level to increase sample size. In addition, recent research shows that a more disaggregated geographical unit of analysis may reveal the coexistence of different externalities. Andersson, Larsson, and Wernberg (2019) use geo-coded firm-level data and find that firms located in within-city industry clusters enjoy enhanced productivity, while overall density effects operate at the city level. We recognize that such effects may become obscured when the spatial unit of analysis is not as fine grained.

For future studies, we recommend further elaborations of regional classifications as a complement to the current trend focusing mainly on regressions. We suggest altering classifications in more categories than the one used here to elaborate on nuances and differences between for instance different lagging regions. We also suggest examining the relationship between more regional outcomes than the ones presented here to better understand the implications of the different regional classifications. We believe this approach can be of importance for furthering our understanding of the dynamics of different categories of lagging and leading regions.

Conclusions

The main results of this study suggest that regional entrepreneurship and industry diversity characterized by technological relatedness are key elements in understanding why some regions are leading while others lag behind. Despite its limitations, we argue that our study contributes with relevant results and policy implications. This study

proposes a multi-dimensional method for classifying and analyzing regional growth trajectories. This approach could be adapted and used to examine how policy initiatives in European regions contribute to shifting regional growth trajectories. The proposed taxonomy of leading, stagnating, and lagging regions holds value for policy discussions as it provides a nuanced taxonomy for examining and communicating why regions differ in growth trajectories over time.

The study also holds important implications for the contemporary EU-cohesion policy which encourages regions to set priorities and to concentrate resources in selected fields that show potential, given their existing capabilities (Camagni & Capello, 2010; Foray & Goneaga, 2013; Mancha-Navarro & Garrido-Yserte, 2008). The European Commission (2017b) advises that such priorities could include “...domains, areas and economic activities where regions or countries have a competitive advantage or have the potential to generate knowledge-driven growth...” This policy of *Smart Specialization* aims at stimulating regional innovation and growth through the enhancement of regional capabilities in a few market niches, in order to develop competitive advantages (Foray, 2014). Our results add to the growing body of empirical studies that support the promotion of related variety as an appropriate smart specialization strategy (S3). Lastly, our results emphasize the importance of the entrepreneurial dimension in order to distinguish between regions that prosper and those that lag behind. This reinforces the notion that entrepreneurs serve as important conduits for knowledge spillovers (Acs et al., 2009). Particularly in lagging regions, S3-policies should be accompanied by efforts to promote a regional milieu that is conducive to entrepreneurship.

Appendix

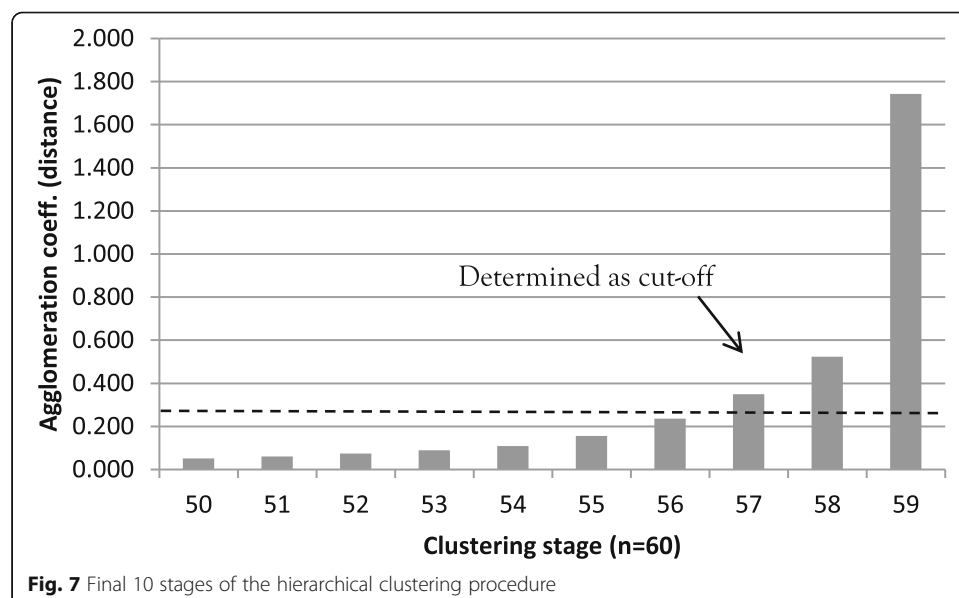


Table 9 Post hoc tests of cluster means from k-means clustering, multiple comparisons

Dependent variable		(I) Cluster number of case	(J) Cluster number of case	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
							Lower bound	Upper bound
GRPGr	Tukey HSD	1	2	0.290204*	0.022719	0.000	0.23553	0.34488
			3	-0.137388*	0.022191	0.000	-0.19079	-0.08399
		2	1	-0.290204*	0.022719	0.000	-0.34488	-0.23553
			3	-0.427593*	0.026219	0.000	-0.49069	-0.36450
		3	1	0.137388*	0.022191	0.000	0.08399	0.19079
			2	0.427593*	0.026219	0.000	0.36450	0.49069
	LSD	1	2	0.290204*	0.022719	0.000	0.24471	0.33570
			3	-0.137388*	0.022191	0.000	-0.18183	-0.09295
		2	1	-0.290204*	0.022719	0.000	-0.33570	-0.24471
			3	-0.427593*	0.026219	0.000	-0.48010	-0.37509
		3	1	0.137388*	0.022191	0.000	0.09295	0.18183
			2	0.427593*	0.026219	0.000	0.37509	0.48010
PopGr	Tukey HSD	1	2	0.025038*	0.008438	0.012	0.00473	0.04534
			3	-0.032060*	0.008242	0.001	-0.05189	-0.01223
		2	1	-0.025038*	0.008438	0.012	-0.04534	-0.00473
			3	-0.057098*	0.009738	0.000	-0.08053	-0.03366
		3	1	0.032060*	0.008242	0.001	0.01223	0.05189
			2	0.057098*	0.009738	0.000	0.03366	0.08053
	LSD	1	2	0.025038*	0.008438	0.004	0.00814	0.04193
			3	-0.032060*	0.008242	0.000	-0.04856	-0.01556
		2	1	-0.025038*	0.008438	0.004	-0.04193	-0.00814
			3	-0.057098*	0.009738	0.000	-0.07660	-0.03760
		3	1	0.032060*	0.008242	0.000	0.01556	0.04856
			2	0.057098*	0.009738	0.000	0.03760	0.07660

*The mean difference is significant at the 0.05 level

Table 10 Classification (*k*-means clustering) of Swedish functional analysis regions according to change in nominal GRP and population between 2010 and 2015

Stagnant regions				Lagging regions		Leading regions	
No.	Name	No.	Name	No.	Name	No.	Name
2	Kristianstad-Hässleholm	33	Avesta-Hedemora	10	Oskarshamn	1	Malmö-Lund
3	Karlskrona	34	Falun-Borlänge	31	Vansbro	4	Älmhult-Osby
5	Ljungby	35	Mora	40	Härjedalen	8	Växjö
6	Halmstad	36	Gävle	41	Östersund	12	Vimmerby
7	Värnamo	37	Bollnäs-Ovanåker	45	Sollefteå	13	Jönköping
9	Kalmar	38	Hudiksvall	46	Strömsund	15	Göteborg
11	Västervik	39	Ljusdal	47	Åsele	16	Trollhättan-Vänersborg
14	Borås	42	Sundsvall	49	Lycksele	18	Skövde-Skara
17	Lidköping-Götene	43	Kramfors	50	Vilhelmina	19	Linköping-Norrköping
20	Gotland	44	Örnsköldsvik	51	Storuman	22	Eskilstuna
21	Nyköping-Oxelösund	52	Skellefteå	53	Arvidsjaur	23	Stockholm
25	Örebro	54	Arjeplog	58	Jokkmokk	24	Västerås
26	Karlskoga	55	Luleå	59	Gällivare	28	Västlandet
27	Karlstad	56	Haparanda	60	Kiruna	32	Ludvika
29	Torsby	57	Överkalix			48	Umeå
30	Malung-Sälén						

Abbreviations

FA-regions: Functional analysis regions; PPS: Purchasing power standards; GRP: Gross region product; NEG: New economic geography; HHI: Herfindahl-Hirschman index; S3: Smart specialization strategy

Acknowledgements

Not applicable (at this point).

Authors' contributions

Both authors have contributed in equal proportions to the work involved in writing the paper. Consequently, both authors have approved the submitted version and agree to be accountable for the work and to ensure that questions related to the accuracy and integrity of any part of the work are appropriately investigated, resolved, and the resolution documented in the literature.

Funding

This research received no specific grant from any funding agency. Open Access funding provided by Lulea University of Technology.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 24 February 2020 Accepted: 19 January 2021

Published online: 11 February 2021

References

- Acs, Z. J. (2002). *Innovation and the growth of cities*. Cheltenham: Edward Elgar.
- Acs, Z. J., Audretsch, D. B., & Lehmann, E. E. (2013). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 41, 757–774.
- Acs, Z. J., Braunerhjelm, P., Audretsch, D. B., & Carlsson, B. (2009). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32, 15–30.
- Andersson, M., Larsson, J. P., & Wernberg, J. (2019). The economic microgeography of diversity and specialization externalities – firm-level evidence from Swedish cities. *Research Policy*, 48(6), 1385–1398.
- Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3), 155–173.
- Audretsch, D. B., & Fritsch, M. (2002). Growth regimes over time and space. *Regional Studies*, 36(2), 113–124.
- Audretsch, D. B., & Keilbach, M. (2004a). Does entrepreneurship capital matter? *Entrepreneurship Theory and Practice*, 28, 419–429.

- Audretsch, D. B., & Keilbach, M. (2004b). Entrepreneurship and regional growth: An evolutionary interpretation. *Journal of Evolutionary Economics*, 14, 605–616.
- Barro, R. J. (2001). Human capital and growth. *The American Economic Review*, 91(2), 12–17. Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association (May 2001).
- Basile, R., & Girardi, A. (2010). Specialization and risk sharing in European regions. *Journal of Economic Geography*, 10, 645–659.
- Batabyal, A. A., & Nijkamp, P. (2014a). Technology, learning, and long run economic growth in leading and lagging regions. *Economic and Political Weekly*, 49, 92–96.
- Batabyal, A. A., & Nijkamp, P. (2014b). Some properties of the technology gap between leading and lagging regions. *Theoretical Economics Letters*, 4, 1–6.
- Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, 85(3), 289–311.
- Braunerhjelm, P., Acs, Z. J., Audretsch, D. B., & Carlsson, B. (2010). The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics*, 34, 105–125.
- Callejon, M., & Segarra, A. (1999). Business dynamics and efficiency in industries and regions: The case of Spain. *Small Business Economics*, 13, 253–271.
- Camagni, R., & Capello, R. (2010). Macroeconomic and territorial policies for regional competitiveness: An EU perspective. *Regional Science Policy and Practice*, 2, 1–19.
- Capello, R., & Nijkamp, P. (2009). Introduction: regional growth and development theories in the twenty-first century – recent theoretical advances and future challenges. In: Capello, R., & Nijkamp, P. (Ed.) *Handbook of regional growth and development theories*. Cheltenham: Edward Elgar.
- Capello, R. (2011). Location, regional growth and local development theories. *AESTIMUM*, 58, 1–25.
- Content, J., & Frenken, K. (2016). Related variety and economic development: A literature review. *European Planning Studies*, 24(12), 2097–2112.
- Cox, R. A. K., & Wang, G. W. Y. (2014). Predicting the US bank failure: A discriminant analysis. *Economic Analysis and Policy*, 44, 202–211.
- Da Silva, G. D., Goncalves, E., & De Araújo, I. F. (2019). Related variety and employment growth: A spatial dynamic model for Brazilian microregions. *Regional Science Policy & Practice*, 2019, 1–19.
- De Groot, H. L. F., Poot, J., & Smit, M. J. (2016). Which agglomeration externalities matter most and why? *Journal of Economic Surveys*, 30(4), 756–782.
- European Commission. (2017a). Competitiveness in low-income and low-growth regions. The lagging regions report. Internet: https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/lagging_regions%20report_en.pdf. (Accessed 30 May 2019)
- European Commission. (2017b). What is smart specialisation? Internet: <http://s3platform.jrc.ec.europa.eu/what-is-smart-specialisation>. (Accessed 10 May 2018)
- Eurostat. (2008). NACE Rev.2. Statistical classification of economic activities in the European Community. Eurostat Methodologies and Working papers. Eurostat, European Commission.
- Feldman, M., & Audretsch, D. (1999). Innovation in cities: science-based diversity, specialization and localized competition. *European Economic Review*, 43, 409–429.
- Fischer, M. M., & Nijkamp, P. (2009). Entrepreneurship and regional development. In R. Capello, & P. Nijkamp (Eds.), *Handbook of Regional Growth and Development Theories*. Cheltenham: Edward Elgar.
- Florida, R., Adler, P., & Mellander, C. (2017). The city as innovation machine. *Regional Studies*, 51(1), 86–96.
- Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development—human capital, the creative class and tolerance. *Journal of Economic Geography*, 8, 615–649.
- Foelster, S. (2000). Do entrepreneurs create jobs? *Small Business Economics*, 14, 137–148.
- Foray, D. (2014). From smart specialisation to smart specialisation policy. *European Journal of Innovation Management*, 17(4), 492–507.
- Foray, D., & Goneaga, X. (2013). *The goals of smart specialisation*. S3 Policy Brief Series No. 01/2013, European Commission, Joint Research Centre, Institute for Prospective Technological Studies. Luxembourg: Publications Office of the European Union.
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5), 685–697.
- Fritsch, M. (2015). Methods of analyzing the relationship between new business formation and regional development. In C. Karlsson, M. Andersson, & T. Norman (Eds.), *Handbook of Research Methods and Applications in Economic Geography*. Cheltenham: Edward Elgar.
- Fritsch, M., & Wyrwhich, M. (2017). The effect of entrepreneurship on economic development – An empirical analysis using regional entrepreneurship culture. *Journal of Economic Geography*, 7, 157–189.
- Glaeser, E. L. (2000). The new economics of urban and regional growth. In G. L. Clark, M. P. Feldman, & M. S. Gertler (Eds.), *The Oxford Handbook of Economic Geography*, (pp. 83–98). Oxford: Oxford 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.
- Hair Jr., J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis*. Pearson New International Edition. Essex: Pearson Education.
- Heckscher, E. F. (1919). Utrikeshandels verkan på inkomstfördelningen. Några teoretiska grundlinjer. *Ekonomisk Tidskrift*, 21(2), 1–32.
- Hirschman, A. O. (1958). *The strategy of development*. New Haven: Yale University Press.
- Hohenberg, P., & Lees, L. H. (1985). *The making of urban Europe (1000-1950)*. Cambridge (Mass.): Harvard University Press.
- Holtz-Eakin, D., & Kao, C. (2003). *Entrepreneurship and economic growth: The proof is in the productivity*. Syracuse University Center for Policy Research, Working Paper No. 50. Syracuse University. Syracuse, New York.
- IMF. (2020). *Primary commodity prices: Indices of primary commodity prices, 2009-2019*. International Monetary Fund. Internet: <https://www.imf.org/en/Research/commodity-prices> (Accessed 13 Nov 2020).
- Jacobs, J. (1969). *The economy of cities*. New York: Vintage Books.
- Jones, R. W. (1956). Factor proportions and the Heckscher-Ohlin theorem. *The Review of Economic Studies*, 24(1), 1–10.
- Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, 49(6), 1003–1018.

- Ketchen Jr., D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. *Strategic Management Journal*, 17(6), 441–458.
- Kim, S. (1995). Expansion of markets and the geographic distribution of economic activities: The trends in U.S. regional manufacturing structure, 1860–1987. *The Quarterly Journal of Economics*, 110(4), 881–908.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99, 483–499.
- Krugman, P. (2009). The increasing returns revolution in trade and geography. *The American Economic Review*, 99(3), 561–571.
- Lee, S. Y., Florida, R., & Gates, G. (2010). Innovation, human capital, and creativity. *International Review of Public Administration*, 14(3), 13–24.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22, 3–42.
- Mancha-Navarro, T., & Garrido-Yserte, R. (2008). Regional policy in the European Union: The cohesion-competitiveness dilemma. *Regional Science Policy & Practice*, 1, 47–66.
- Marshall, A. (1890). *Principles of economics*. London: Macmillan.
- Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, 6, 395–437.
- McCann, P., & van Oort, F. (2009). Theories of agglomeration and regional economic growth: A historical review. In: Capello, R., & Nijkamp, P. (Ed.) *Handbook of regional growth and development theories*. Cheltenham: Edward Elgar.
- Migueluez, E., & Moreno, R. (2018). Relatedness, external linkages and regional innovation in Europe. *Regional studies*, 52(5), 688–701.
- Nakamura, R., & Morrison Paul, C. J. (2009). Measuring agglomeration. In: Capello, R., & Nijkamp, P. (Ed.) *Handbook of Regional Growth and Development Theories*. Cheltenham: Edward Elgar.
- Ohlin, B. (1924). *Handelns teori*. Doctoral dissertation, Stockholms Högskola. Stockholm: Centraltryckeriet.
- Östbring, L., & Lindgren, U. (2013). Labour mobility and plant performance: on the (dis)similarity between labour- and capital-intensive sectors for knowledge diffusion and productivity. *Geografiska Annaler: Series B, Human Geography*, 95(4), 287–305.
- Perreault Jr., W. D., Behrman, D. N., & Armstrong, G. M. (1979). Alternative approaches for interpretation of multiple discriminant analysis in marketing research. *Journal of Business Research*, 7, 151–173.
- Ricardo, D. (1817). *On the principles of political economy, and taxation*. London: J Murray.
- Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5), 1002–1037.
- Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), S71–S102.
- Sandelin, B., Trautwein, H. M., & Wundrak, R. (2001). *Det ekonomiska tänkandets historia*, (3rd ed.,). Stockholm: SNS Förlag.
- SCB. (2007). *SNI 2007. Swedish standard industrial classification 2007. MIS 2007:2*. Örebro: Statistics Sweden.
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*, (2007th ed.,). Hampshire: Harriman House Ltd.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1), 65–84.
- Stigler, G. J. (1951). The division of labor is limited by the extent of the market. *Journal of Political Economy*, 59(3), 185–193.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*, (6th ed.,). NJ: Pearson.
- Tavassoli, S., & Carbonara, N. (2014). The role of knowledge variety and intensity for regional innovation. *Small Business Economics*, 43, 493–509.
- Terluin, I. J., & Post, J. H. (2000). Employment dynamics in leading and lagging rural regions. In: Terluin, I. J., & Post, J. H. (Eds.) *Employment dynamics in rural Europe*. CAB International: Wallingford.
- Terluin, I. J. (2003). Differences in economic development in rural regions of advanced countries: An overview and critical analysis of theories. *Journal of Rural Studies*, 19, 327–344.
- Tillväxtpolitiska utvärderingar och analyser. (2013). *Regional tillväxt 2013 – en rapport om regional tillväxt och utveckling*. Östersund: Myndigheten för tillväxtpolitiska utvärderingar och analyser.
- Tillväxtpolitiska utvärderingar och analyser. (2015). *Funktionella analysregioner – revidering 2015. PM 2015:22*. Östersund: Myndigheten för tillväxtpolitiska utvärderingar och analyser.
- van Oort, F., de Geus, S., & Dogaru, T. (2015). Related variety and regional economic growth in a cross-section of European urban regions. *European Planning Studies*, 23(6), 1110–1127.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)