Papers in Evolutionary Economic Geography

12.05

The impact of related variety on regional employment growth in Finland 1993-2006: high-tech versus medium/low-tech

Matté Hartog, Ron Boschma and Markku Sotarauta

Utrecht University
Urban & Regional research centre Utrecht

The impact of related variety on regional employment growth in Finland 1993-2006: high-tech versus medium/low-tech

Matté Hartog *, Ron Boschma * and Markku Sotarauta **

*Urban and Regional research centre Utrecht, Faculty of Geosciences, Utrecht University, P.O. Box 80115, 3508 TC, Utrecht, The Netherlands

**Research Unit for Urban and Regional Development Studies, University of Tampere, FI-33014 University of Tampere, Finland

Abstract

This paper investigates the impact of related variety on regional employment growth in Finland between 1993 and 2006 by means of a dynamic panel regression model. We find that related variety in general has no impact on growth. Instead, after separating related variety among low-and-medium tech sectors from related variety among high-tech sectors, we find that only the latter affects regional growth. Hence, we find evidence that the effect of related variety on regional employment growth is conditioned by the technological intensity of the local sectors involved.

JEL Codes: D62, O18, R11

1 Introduction

In the context of the current economic crisis, the question of what kind of economic composition in regions is best for regional employment growth is more than ever prominent on the political and scientific agenda. Till recently, the key question was whether regions should be mainly specialized, or whether the economic composition of regions should be mainly diversified. Especially, the importance of regional diversity or Jacobs' externalities has been subject to much empirical work from the 1990s onwards (Glaeser et al., 1992; Van Oort, 2004), with mixed results so far. That is, studies have shown positive, negative or no impact of a diversified industrial mix in regions on their economic growth (see for an overview Beaudry and Schiffauerova, 2009). A possible reason for this is the crude way in which variety is often dealt with in the Glaeser-related literature (Iammarino and McCann, 2006).

In recent years, studies have challenged the view that a variety of sectors in a region as such is sufficient for local firms to learn and innovate from knowledge spillovers (Frenken et al., 2007; Boschma and Iammarino, 2009). Particularly, following Cohen and Levinthal (1990), it has been argued that learning from spillovers is unlikely to take place when there is no cognitive proximity between local firms. Recent literature has proposed that knowledge is more likely to spill over between sectors that are cognitively proximate (Nooteboom, 2000; Morone, 2006; Leahy and Neary, 2007). Frenken et al. (2007) have therefore introduced the notion of related variety, in order to underline that not regional variety *per se* matters for urban and regional growth, but regional variety between sectors that are technologically related to each other. Recent studies in The Netherlands (Frenken et al., 2007), Italy (Boschma and Iammarino, 2009; Quatraro, 2010) and Spain (Boschma et al., 2011) have indeed confirmed that related variety tends to contribute positively to regional employment growth.

This study investigates the impact of related variety on regional growth in Finland between 1993 and 2006. Recent studies have argued that sectoral specificities might matter in this respect. We investigate whether related variety among high-tech sectors has affected regional growth in Finland in the period 1993-2006, during which the Finnish economy changed into a high-tech economy. Some scholars (Heidenreich, 2009; Kirner et al., 2009; Santamaria et al., 2009) have argued that inter-industry knowledge spillovers and product innovations are especially relevant for high-tech sectors. The relationship between related variety and regional employment growth is examined by means of dynamic panel regressions using generalized method of moments (GMM) estimators, which allow us to take into account the possibility of reverse causality between related variety and regional growth over time. This makes the estimated effects dynamic in comparison to existing studies, which have been mainly cross-sectional.

The structure of this study is as follows. Section 2 elaborates on how agglomeration economies are linked to economic growth in regions, particularly related variety. Section 3 contains the empirical framework that describes the evolution of the Finnish economy from 1993 onwards in greater detail, and then elaborates on the data and the methods used. Section 4 presents and discusses the results. A conclusion follows in the final section that also describes the challenges for future research on this topic.

2 Related variety and regional growth

Agglomeration economies refer to external economies of scale that arise from firms being concentrated close to one another in physical space, and from which firms can profit. In particular, agglomerations are an important source of increasing returns to knowledge (Rosenthal and Strange, 2004; Storper and Venables, 2004; Audretsch and Aldridge, 2008). Agglomeration economies are usually linked to three different sources: urbanisation economies, localisation economies and Jacobs' externalities.

The first source of agglomeration economies are urbanisation economies. These relate to external economies from which all co-located firms can benefit regardless of the industry they operate in. A dense environment in terms of population, universities, trade associations, research laboratories and so on, facilitates the creation and absorption of new knowledge, which in turn may lead to innovative performance (Harrison et al, 1996). As Lucas (1993) argues, productivity increases due to urbanization economies also result from increasing returns to scale to firms, for example due to the presence of larger labour markets in agglomerations. There are, however, also urbanisation diseconomies, such as higher factor costs, higher land prices and higher living costs. Furthermore, there may be negative externalities caused by pollution or congestion (Quigley, 1998). Thus, a dense environment provides advantages in terms of knowledge production and productivity increases, but may also be more costly to doing business than a scarcely occupied area.

The second source of agglomeration economies are localisation economies (Glaeser et al., 1992). They differ from urbanisation economies in that they refer to external economies that are available only to firms that operate within the same industry. In addition to labour pooling and the creation of specialized suppliers, MAR externalities arise from knowledge spillovers that occur between firms that are cognitively similar (Henderson, 1995). An often cited example of the effects of these externalities is the uprising of the semiconductor industry in Silicon Valley, which was characterized by a process of self-reinforcing knowledge accumulation due to spatial proximity between specialized suppliers and customers, universities, venture capital firms and so on (Saxenian, 1994).

The third source of agglomeration economies are Jacobs' externalities. Named after the work of Jacobs (1969), these externalities originate from a variety of sectors in a region and are available to all local firms. The basic line of argument is that a regional economy characterized by a varied industrial mix spurs innovation because local firms are able to recombine knowledge stocks from different industries (Van Oort, 2004). Hence, the existence of regional variety itself is regarded as a source of knowledge spillovers. As such, Jacobs' externalities are likely to lead to regional employment growth because the recombination of knowledge from different industries fosters radical innovations that lead to the creation of new markets.

Studies on the effects of Jacobs' externalities on regional growth have produced mixed results so far. Some studies find either positive or negative effects, whereas others find no evidence for the presence of Jacobs' externalities (overviews are given in Beaudry and Schiffauerova, 2009; De Groot et al., 2009). Hence, there is ambiguity as to whether the presence of a diversity of industries is best for regional economic growth. In dealing with this, Frenken et al. (2007) and Boschma and Iammarino (2009) have recently argued that for Jacobs' externalities to occur in a region, the industries in the region have to be cognitively related to some extent. It is argued that learning between local firms is unlikely to take place when there is no cognitive proximity between them

Incorporating the notion of cognitive proximity into Jacobs' externalities, Frenken et al. (2007) make a distinction between related variety and unrelated variety. Related variety is defined as industries that share some complementary capabilities, while unrelated variety refers to sectors that do not. As expected, they find that it is related variety that mainly contributes to regional employment growth, whereas unrelated variety mainly acts as a local stabilizer, dampening regional unemployment growth. The latter result is expected because unrelated variety is unlikely to facilitate effective learning between firms due to the lack of cognitive proximity, and because it protects regions from negative sector-specific demand shocks. Similar findings of the impact of related and unrelated variety on regional growth have been found in the case of Italy (Boschma and Iammarino, 2009) and Spain (Boschma et al., 2011).

Hence, related variety as such seems to matter for growth, but to what extent do sector specificities matter in this respect? Henderson et al. (1995) already indicated that variety in general is more important for young and technologically advanced industries,.Paci and Usai (2000) found that variety in general is more important for high-tech industries in urban regions. As for related variety, the results of the empirical study of Bishop and Gripaios (2010) suggest that the impact of related variety on growth differs for different sectors.

Relatedly, Buerger and Cantner (2011) studied innovativeness in two science-based and two specialized supplier industries and found that for all four industries technological relatedness to other local industries is beneficial. Hence, it may be that the impact of related variety on growth depends on certain specificities of local sectors concerned, but empirical studies that have investigated this issue are yet scarce.

In this paper we explicitly relate one sector specificity, namely the technological intensity of local sectors, to the impact of related variety on regional growth. Scholars (Heidenreich, 2009; Kirner et al., 2009; Santamaria et al., 2009) have argued that inter-industry knowledge spillovers and product innovations are especially relevant for high-tech sectors. We investigate regional growth in Finland between 1993 and 2006, a period during which the economy of Finland changed into a high-tech economy, with an increasing variety within the high-tech sector. Inspired by the approach taken by Frenken et al. (2007), we investigate by means of a dynamic panel regression whether the impact of related variety among high-tech sectors on regional growth in Finland is different from the impact of related variety among low-and-medium-tech sectors.

3 Methodology

3.1 Data

We employ annual data by industry at the regional level in Finland from 1993 to 2006. Regions are defined according to the NUTS-4 classification of the European Union, the borders of which approximate local labour market areas, which are commonly used in studies on local knowledge spillovers. The data have been obtained from Statistics Finland, which is the official statistics authority for the Finnish government. In the data, there have been changes in regional borders and industrial classifications over time, and the way in which those changes have been dealt with in this study is described in Appendix 1. There are 67 different regions in total.

The economy of Finland is very diversified at the regional level in terms of its industrial composition and technological intensity. Finland experienced a huge economic recession in the period 1990-1993, during which real GDP dropped by more than 10% and unemployment rose from about 4% to nearly 20% (Honkapohja and Koskela, 1999; Rouvinen and Ylä-Anttila, 2003). From 1993 onwards, the Finnish economy recovered dramatically: the average annual growth rate in GDP was 4,7% between 1993 and 2000 and the unemployment rate went down from nearly 20% in 1993 to around 9% in 2000. The economic boom was characterized by the upcoming of high-tech industries, especially those indulged in manufacturing electronic products related to telecommunication. Some firms, such as Nokia, played an important role in this respect (Ali-Yrkkö and Hermans, 2004). Whereas Finland had a large trade deficit in high-tech products in the early 1990s, it had a significant surplus in 2000, when exports of electronic equipment and other high-tech products accounted for more than 30% of the country's exports (Blomstrom et al., 2002). Hence, the data cover a time period (1993-2006) that contains an economic boom with a prominent presence of high-tech sectors.

3.2 Variables

3.2.1 Dependent variable

The dependent variable in this study is annual employment growth (EMPGROWTH) at the regional level (NUTS4) in Finland between 1993 and 2006. A limitation of employment growth is that it does not measure industry growth as accurately as growth in productivity, which relates more directly to learning from knowledge spillovers through related variety, but data on output is unfortunately unavailable at this spatial scale in Finland.

3.2.2 Independent variables

To measure the different indicators of variety at the regional level, regional establishment data are used which are classified according to the Finnish Standard Industrial Classification 1995 (SIC). This classification is derived from and corresponds with few exceptions to the European Community NACE Rev. 1. Classification. Establishment data are available for all industries in every region at any digit level of the SIC classification.

Regarding the measurement of variety, we use an entropy measure on the regional establishment data. The advantage of using an entropy measure is that it can be decomposed at every sectoral digit level of the SIC classification. Hence, variety can be measured at several digit levels, and subsequently these different variety measures can enter a regression analysis without necessarily causing multicollinearity.

We first measure variety in general that represents the degree of variety of establishments in a region as a whole. In turn, variety in general is decomposed into unrelated variety (UNRELVAR) and related variety (RELVAR), in a similar vein as in Frenken et al. (2007) and Boschma and Iammarino (2009). Subsequently, the contribution of this study is to further decompose related variety (RELVAR) into high-tech related variety (RELVARHTECH) and low-and-medium-tech related variety (RELVARLMTECH).

First, let p_i be the five-digit SIC share of establishments, then variety in general is measured as the sum of entropy at the five-digit level:

$$V = \sum_{g=1}^{G} P_i \log_2\left(\frac{1}{Pi}\right) \qquad \text{Eq. (1)}$$

This measure thus represents regional variety in general, or Jacobs' externalities not further specified. The higher its value, the more diversified the industrial composition of a region is. To take into account the degree of cognitive proximity between sectors, and hence learning opportunities for industries, this measure is split into an unrelated and related part. First, one can derive the two-digit shares P_g by summing the five-digit shares p_i:

$$P_g = \sum_{i \in S_g} p_i \qquad \text{Eq. (2)}$$

Then, unrelated variety (UNRELVAR) is measured by the entropy at the two-digit level:

$$UV = \sum_{g=1}^{G} P_g \log_2 \left(\frac{1}{P_g} \right) \qquad \text{Eq. (3)}$$

Hence, this variable UNRELVAR measures unrelated variety by means of variety at the two-digit level. We thus assume that sectors that belong to different two-digit classes are unrelated from one another. Hence, the higher the value of this variable, the more variety there is at the two-digit level, and thus the more a region is endowed with very different industries. It is expected that effective knowledge spillovers do not occur when the degree of UNRELVAR is high, because it is unlikely that sectors in different 2-digit classes can effectively learn from each other because they are not cognitively proximate.

We also measure related variety (RELVAR). Following Frenken et al. (2007), this is done by taking the weighted sum of entropy within each two-digit sector:

$$RV = \sum_{g=1}^{G} P_g H_g \qquad \text{Eq. (4)}$$

where

$$H_g = \sum_{i \in S_g} \frac{p_i}{p_g} \log_2 \left(\frac{1}{p_i / p_g} \right) \qquad \text{Eq. (5)}$$

Hence, this variable RELVAR measures the degree of variety *within* every two-digit class in a region, and sums that for all the two-digit classes in that region. We thus assume that sectors that belong to the same two-digit class are related to one another technologically, and hence we assume that they can effectively learn from one another through knowledge spillovers. And, the higher the degree of RELVAR is, the higher the number of technologically related industries in the region, the more innovation opportunities there are.

We further decompose related variety (RELVAR) into high-tech related variety (RELVARHTECH) and low-and-medium-tech related variety (RELVARLMTECH) to assess whether they have a different impact on regional employment growth. We use the SIC 1995 classification which separates low-and-medium-tech sectors from high-tech sectors according to their technological intensity, based on their R&D intensity (R&D expenditures over value added) and their share of tertiary educated persons employed. The latter also accounts for sectors that do not necessarily have a high R&D intensity, i.e. knowledge- and innovation-intensive sectors. This classification is commonly used to separate high-tech from low-and-medium-tech sectors in Finland (e.g. Simonen and McCann, 2008). Following this classification, high-tech related variety (RELVARHTECH) is measured in the same vein as related variety (RELVAR), but is applied only to establishments in high-tech sectors, all of which are listed in Table 1. Low-and-medium-tech related variety (RELVARLMTECH) measures related variety within all of the remaining industries. Because of the decomposable of the entropy measure that is used to measure both types of related varieties, they do not necessarily correlate with each other and hence can enter a regression at the same time. In Appendix 2 we elaborate on this issue in greater detail and also describe how the empirical construction of both types of varieties differs from the traditional distinction between related and unrelated variety as in Frenken et al. (2007).

Table 1. High-technology industries based on SIC classification (1995)

Manufacture of pharmaceuticals, medicinal chemicals and botanical products	244
Manufacture of office machines and computers	30
Manufacture of radio, television, communications equipment and apparatus,	32
Manufacture of aircraft and spacecraft	353
Telecommunications	642
Computers and related activities	72
Research & Development	73
Architectural and engineering activities and related technical consultancy	742

3.2.3 Control variables

We include a number of control variables. First, regional population density (POPDENS) from 1993 to 2006 is used as a proxy for urbanisation economies. This variable represents the amount of economic activity in every region regardless of its industrial composition. Second, to measure the effect of human capital (HUMCAP) in a region, we take the percentage of the total population (1993-2006) with a university bachelor degree or higher. This way of measuring educational attainment is in line with most of the literature on human capital and regional growth. Third, Research & Development (R&D) expenditures (R&DEXP) are measured per capita from 1995 to 2006 (excluding 1996). This indicator plays a central role in endogenous growth models, and is also often used to measure the ability of regions to adapt to innovations produced elsewhere (Crescenzi and Rodriguez-Posé, 2008). These variables are some of the variables that are most often included in growth models, but we lack data on some other variables that are also known to influence growth (e.g. variables reflecting capital-labor ratios or competition). Hence, we are not able to estimate a conventional regional growth model with all of the 'usual suspects' included, but we are able to investigate whether the different variety measures have different regional employment effects. The control variables that we include are log transformed, and time dummies are included in the model as well.

3.3 Model specification

To determine the impact on regional employment growth, we adopt a dynamic panel approach using generalized method of moments (GMM) estimators developed by Arellano and Bond (1991) and Arellano and Bover (1995). The growth equation we wish to estimate has the following form:

$$y_{i,t} = \beta' X_{i,t} + \eta_i + \varepsilon_{i,t} \qquad \text{Eq. (6)}$$

where *y* denotes employment growth, *t* denotes 1-year intervals (from 1993 to 2006), *i* denotes the region, *X* denotes the set of explanatory variables, η denotes an unobserved region-specific effect of time-invariant determinants of growth and ε denotes the error term. The variety regressors may be endogenous because growth may also influence the variety in a region (e.g. growth may take place through a process of diversification into related industries as found in Neffke et al., 2011). Normally, one would deal with this issue by using external instruments that are correlated with $X_{i,t}$ and yet uncorrelated with $y_{i,t}$. Suitable external instruments, however, are unavailable in this case, which is a common problem in studies on regional growth (Henderson, 2003). Therefore, we use internal instruments based on lagged levels and lagged differences of $X_{i,t}$ generated with a GMM procedure.

Holtz-Eakin et al. (1988) and Arellano and Bond (1991) were the first to develop a GMM estimator with internal instruments for dynamic panel models such as Eq. (6). First, they take first differences to eliminate fixed effects:

$$y_{i,t} - y_{i,t-1} = \beta'(X_{i,t} - X_{i,t-1}) + (\varepsilon_{i,t} - \varepsilon_{i,t-1}) \qquad \text{Eq. (7)}$$

Afterwards, they instrument potentially endogenous variables with their own levels, lagged twice or more. The estimator assumes that the error term, ε, is not serially correlated and that the explanatory variables, X, are weakly exogenous (uncorrelated with realizations of the error term in the future).

The estimator above, however, does not allow one to study cross-region differences between growth and the explanatory variables as this relationship is eliminated, which is problematic in the context of this study for two reasons. First, from a conceptual point of view we would be interested in studying this relationship as well. Second, lagged levels are weak instruments for the first-differenced equation, Eq. (7), when the explanatory variables are persistent over time, which is likely the case with the different variety measures (as the sectoral composition of regions changes only slowly over time). This finite-sample bias may produce biased coefficients for first-differenced regression equations (Blundell and Bond 1998).

In dealing with this issue, Arellano and Bover (1995) developed a system-GMM estimator. It combines in a system the regression in levels, Eq. (6), with the regression in differences, Eq. (7), where levels are instrumented on lagged first differences (as above) and first differences are instrumented on lagged levels (assuming that past changes in *y* are uncorrelated with the current errors in levels or differences). Blundell and Bond (1998) show with Monte Carlo simulations that in small samples this estimator yields great improvements over the original Arellano and Bond estimator.

In this study we use the two-step variant of the system-GMM estimator and instrument the variety regressors with their lagged values. The two-step variant is asymptotically more efficient than the one-step variant in estimating coefficients but also tends to be severely downward biased when applied to the original Arellano-Bond-Blundell estimators, which we address by applying the finite-sample correction to the standard errors by Windmeijer (2005). We consider the different variety regressors (VARIETY, RELVAR, UNRELVAR, RELVARHTECH, RELVARLMTECH) as potentially endogenous and therefore instrument them with their lagged values. The other regressors are considered exogenous and hence are not instrumented as there is no direct theoretical concern to do so. Also, instrumenting them as well would overfit the model with instruments (a rule of thumb is not to exceed *N* with the number of instruments – derived from Arellano and Bond, 1998).

The extent to which the system-GMM estimator generates reliable parameters depends on whether the instruments used (in levels and differences) are valid instruments, which we assess as follows. First, we report for every model the results of the Hansen (1982) J test for overidentifying restrictions, which is robust for the two-step variant of the system-GMM estimator. Failure to reject its null hypothesis, that the instruments are exogenous as a group, supports the model. The only risk with this test is that it can be weakened by instrument proliferation (Bowsher, 2002), which we take into account by limiting the number of instruments to *N* as suggested by Roodman (2009a). We also report the results of separate difference-in-Hansen tests that assess the validity of the particular subsets of instruments (i.e. levels and differences, both with and without the other exogenous variables included) and have a similar null hypothesis as the Hansen J test.

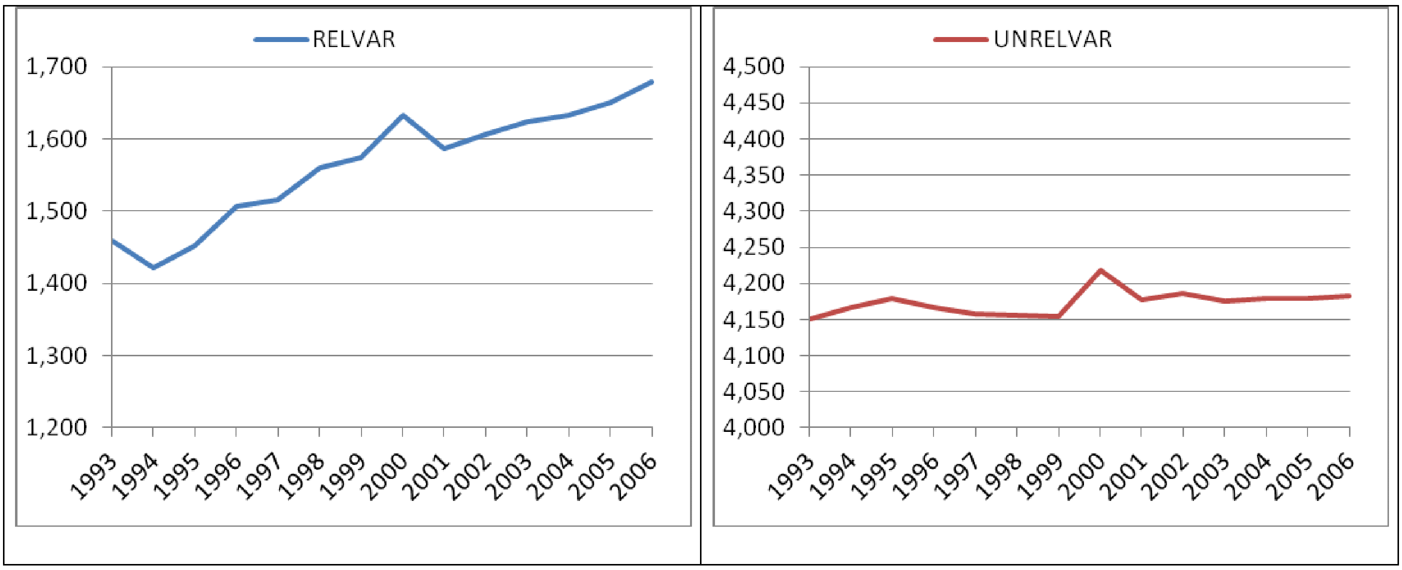
Second, we assess the validity of the instruments by checking for autocorrelation in the error terms. This is done by applying the Arellano-Bond test to the residuals in differences (Arellano and Bond, 1991), which checks whether there is second-order serial correlation in the differenced error term (first-order serial correlation is present by construction because $\varepsilon_{i,t} - \varepsilon_{i,t-1}$ is related to $\varepsilon_{i,t-t} - \varepsilon_{i,t-2}$ because of the shared $\varepsilon_{i,t-t}$ term). If the null hypothesis of no autocorrelation is rejected, it means that the lags of the variety regressors are not exogenous and hence that they are unsuitable for use as instruments.

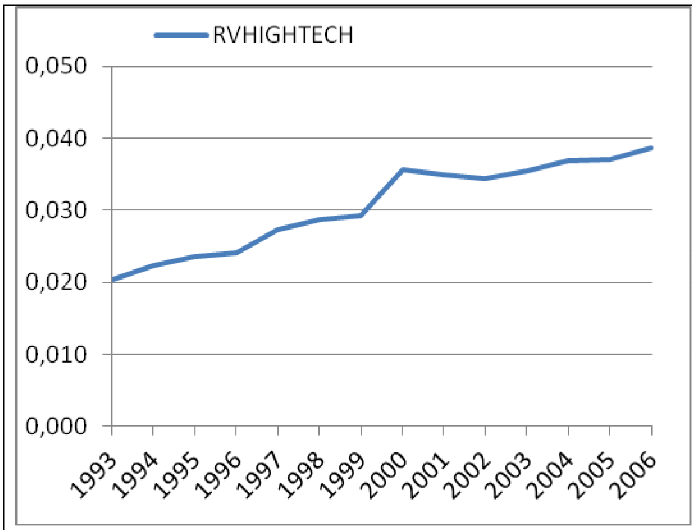
4 Results

As the correlations of some variables are high (see Appendix 2 for the correlations between all independent variables in a cross-section of 1993-2006), we employed a conventional OLS regression on regional employment growth to calculate their variance inflation factor (VIF) score. We find that the different variety measures all score below 5, which suggests that multicollinearity does not substantially bias the results. The dynamic panel framework also renders multicollinearity less of a problem than it would be in a cross-sectional framework.

Figure 1 shows the development of the average related and unrelated variety at the regional level in Finland during the period 1993-2006. A trend is visible of increasing related variety at the regional level in Finland, although slowly evolving, which reminds us that the change of the industrial composition in regions is a slow and gradual process. By contrast, unrelated variety seems to be fairly stable over time. Related variety among high-tech sectors and related variety among low-and-medium-tech sectors both increase over time. Descriptive statistics (mean, standard deviation, minimum, maximum) of these different variety measures, together with descriptive statistics of the other variables, can be found in Appendix 3.

Figure 1: Average related and unrelated variety at regional level in Finland, 1993-2006





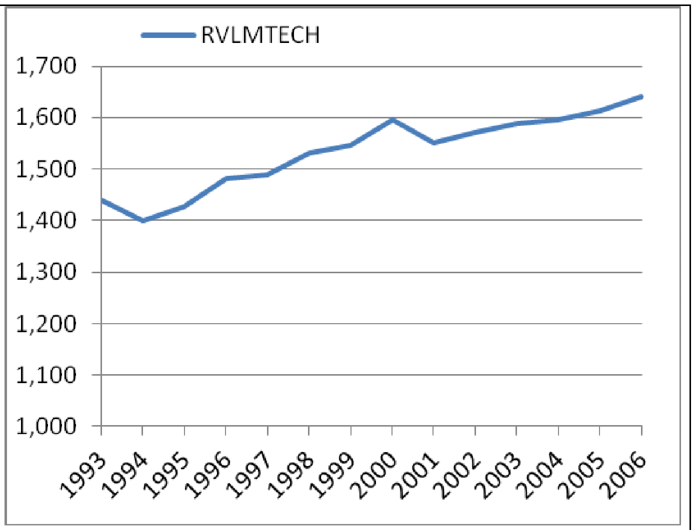


Table 2 shows the results of the system-GMM dynamic panel regression on regional employment growth. Three different models are estimated. Model 1 contains only the control variables. As is often found in the regional growth literature, the amount of human capital is positively related to regional employment growth, whereas population density has a negative impact. No significant effect of R&D expenditures is found.

Model 2 includes related variety (RELVAR) and unrelated variety (UNRELVAR). Both of them are instrumented with their lagged values. The model passes all the diagnostics tests for the validity of the instruments as none of the Hansen tests and Arellano Bond test are significant in Table 2, which means that the lagged values of related variety and unrelated variety are suitable instruments and that the model is not misspecified. We find that related variety has no significant impact on regional growth. This is contrary to previous studies, but we have to remind that our model cannot replicate other studies due to missing control variables.

In Model 3 related variety is decomposed into high-tech related variety (RELVARHTECH) and low-and-medium-tech related variety (RELVARLMTECH). Both of them, together with unrelated variety (UNRELVAR), are instrumented with their lagged values. The model is not miss-specified as all the Hansen tests and the Arellano bond test are insignificant, which implies that the instruments used are valid instruments. We find that only related variety among high-tech sectors has a positive and significant impact on regional employment growth, whereas related variety among low-and-medium-tech sectors has a negative but insignificant impact. Although the impact of RELVARHTECH is significant at only 10%, its coefficient differs substantially from the coefficient of RELVAR in Model 2. Also, the coefficients of the other variables have hardly changed between Model 2 and Model 3, which implies that it is unlikely that the positive and significant coefficient of related variety among high-tech sectors is a result of interdependencies with the other variables, but instead is the result of separating it from low-and-medium-tech variety (RELVARLMTECH). This may explain why related variety as such has no impact on regional growth: after decomposing it into low-and-medium-tech related variety and high-tech related variety, it turns out that only the latter impacts positively on regional employment growth in Finland between 1993 and 2006.

Table 2: System-GMM dynamic panel regression on annual employment growth in Finnish regions, 1993-2006			
Variables	**Model 1**	**Model 2**	**Model 3**
POPDENS (log)	-0,016 (0,007)**	-0,015 (0,007)**	-0,020 (0,008)**
R&DPERCAP (log)	-0,001 (0,008)	-,001 (0,008)	-0,011 (0,010)
HUMCAP (log)	0,032 (0,011)***	0,030 (0,014)**	0,026 (0,015)*
UNRELVAR		-0,021 (0,021)	0,012 (0,029)
RELVAR		0,008 (0,011)	
RELVARHTECH			0,276 (0,162)*
RELVARLMTECH			-0,019 (0,252)
Time dummies included?	Yes	Yes	Yes
Constant	0,142***	0,206 (0,087)**	0,061 (0,119)
Hansen J-test (p value)		0,700	0,434
Arellano-Bond second order serial correlation test (p value)		0,249	0,240
Difference-in-Hansen test that GMM differenced instruments are exogenous (p value)		0,719	0,695
Difference-in-Hansen tests that system GMM instruments are exogenous (p value)		0,552	0,192
*Significant at 90%; **Significant at 95%; *** Significant at 99%; Standard errors in parentheses; Estimates of time dummies not reported to conserve space; Estimations done with xtabond2 package by Roodman (2009b); The maximum number of lags is restricted to 2 to keep the # of instruments below *N*; Data points with studentized residuals higher than 2,5 are identified as outliers and have been excluded (6 observations out of 881)			

A possible explanation may be the different innovation approaches in high-tech and low-and-medium tech sectors. As high-tech sectors rely heavily on knowledge-related inputs and operate mainly at the technological frontier of their respective markets (Hirsch-Kreinsen et al., 2005; Santamaria et. al., 2009), their competitiveness depends mainly on their ability to produce radical innovations that lead to new products. Recent empirical research by Heidenreich (2009), using European Community Innovation Survey data and EU regional data, finds that the focus on product innovations is a key aspect that differs high-tech sectors from low-and-medium-tech sectors as the latter are instead more focused on process innovations. Santamaria et al. (2009) and Haukness and Knell (2009) also show that really new knowledge and technologies for new products are produced mainly in high-tech sectors. This may explain our finding that only related variety among high-tech sectors in a region enhances regional employment growth. As related variety in high-tech sectors facilitates learning through knowledge spillovers, it may enhance the product innovation capacities of local-high tech sectors, with new products and markets as a result, and therefore more regional employment growth.

The focus of low-and-medium-tech sectors on process innovations, instead, may explain our finding that related variety among low-and-medium-tech sectors has no significant impact on regional employment growth. According to Pavitt (1984), low-and-medium-tech sectors are mostly ‘supplier dominated sectors’ that are heavily dependent on purchased embodied technologies and products. Hence, innovation in low-and-medium-tech sectors is aimed at minimising costs through the improvement of production process technologies (see Kirner et al. 2009 for recent empirical evidence for this). In turn, process innovation concerns productivity improvements which often reduce the amount of labour necessary to produce a single unit of output (Edquist et al., 2001). As low-and-medium-tech related variety may facilitate learning with more process innovations as a result, it will have two opposite effects on employment growth. On the one hand, it may increase the competitiveness of local firms with an increase in the demand of labor as a result, on the other hand it may reduce the amount of labour needed due to labour productivity improvements. The balance between the two depends on various factors, such as demand elasticity (e.g. Combes, 2004), but this all makes it less likely that a positive effect of low-and-medium-tech related variety will be found.

5 Conclusion

The aim of this study is to investigate the impact of related variety on regional employment growth in Finland between 1993 and 2006. Using a dynamic panel framework, we find that related variety in general does not impact on regional growth. Instead, we find that only related variety among high-tech sectors has a positive impact on regional growth. Hence, the technological intensity of local sectors involved matters with respect to the impact of related variety on regional employment growth. We proposed that the different employment effects of related variety may be due to differences in innovation approaches of high-tech sectors and low-and-medium-tech sectors, but we have not investigated this issue in this paper. Future research should shed more light on this matter.

There are a number of other issues that stem from this study that could be addressed in future research as well. The first issue is that Finland might be more of a unique case because its economy experienced a particularly rapid transformation towards high-tech sectors during the time period covered. During this period, particular firms, such as Nokia, have played an important role. Therefore, it may be that the findings of this study may differ from (regional) economies that did not experience such a transformation. Future studies on this topic in other countries could shed light on this issue.

The second issue relates to the measurement of technological relatedness between industries, which has been based on the Standard Industrial Classification (SIC). While this is to some extent defendable since this classification is based primarily on product relatedness, it does not necessarily reflect technological relatedness. More advanced relatedness measures have been developed recently, such as the skill-relatedness measure by Neffke and Henning (2012). They measure relatedness by the degree of labor flows between different industries with the idea that excessive labor flows between certain industries imply that those industries require similar skills and hence are related. Such a measure captures relatedness more directly and is therefore preferable, if such data are available of course, which was not the case for Finland.

The third issue relates to the region as the level of analysis. A drawback of focusing on regional growth is that it remains unknown where exactly growth takes place in the region. In the case of related variety, we assume knowledge spillovers to take place between technologically related industries but we do not observe any flows. These are shortcomings that plague almost all agglomeration economies studies, including this one. Future research could address these issues by focusing instead on the effects on growth rates of industries in regions (e.g. Bishop and Gripaios, 2010), and by measuring the effects of actual inter-sectoral knowledge flows at the regional level (as proxied by e.g. labour flows). Also, if one would replace the region by the firm as the unit of analysis, one could also avoid making stylized distinctions between high-tech and low-and-medium-tech at the sectoral level. Any type of sectoral analysis suffers from the fact that there is some 'noise' involved in the sense that all sectors contain, to some degree, different firms, in this case firms that could be classified as either low-and-medium-tech or high-tech (Kirner et al., 2009). Narrower sector definitions may reduce but do not solve this problem – instead, the best solution to this problem would be to draw upon firm data directly, which would enable one to estimate technological intensity more accurately.

When it comes to policy making, the outcomes of this study highlight two important issues. First, when developing regional growth strategies, it is of crucial importance for policy makers to take into account technological relatedness between local firms. Knowledge spillovers between local firms and hence learning only takes place when local firms are to some extent cognitively related. This means that stimulating variety as such, without taking into account relatedness between local firms is unlikely to increase the innovative performance of local firms. Second, policy makers have to consider what kind of regional growth they are aiming for. This is a particularly relevant question for the Finnish innovation and regional development policies that seem to rather be moving towards more focused policies instead of stimulation of cross-sectoral innovation (Edquist et al., 2009). At all events, if policy makers wish to boost regional employment growth, the findings of this study highlight that it is beneficial to stimulate related variety among high-tech firms and to make connections between high-tech industries that are technologically related.

Acknowledgements

Part of the data has been kindly provided by Janne Huovari from the Pellervo Economic Research Institute in Helsinki.

Appendix 1: Regional border changes and Standard Industrial Classification (SIC) changes between 1993 and 2006

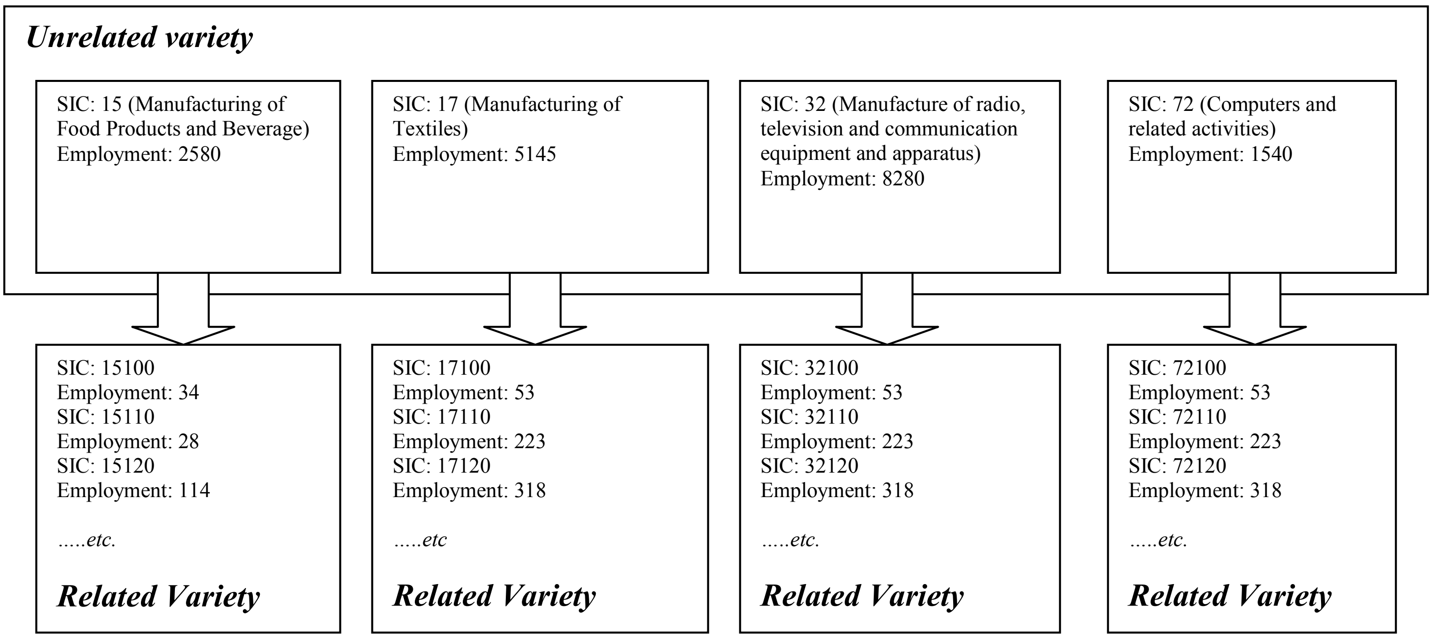
There have been changes in regional borders and the Standard Industrial Classification (SIC) between 1993 and 2006. Changes in regional borders concern the dissolution of some NUTS 4-areas. Those regions have been excluded from the analysis. A comparison of the data before and after the dissolution of these regions shows that regional changes in the data have been very minor.

Regarding changes in the SIC classification, every single 5-digit change between 1993 and 2006 has been checked for and taken into account to make the data comparable over time. The changes concern exclusively the creation of new 5-digit sectors over time (23201 23209 27350 29400 33100 40121 40122 40131 40139 40140 40200 40310 40320 51360 51610 51620 51630 51641 51642 51643 51651 51652 51659 51660 51701 51709 52469 52474 52619 60240 63300 64203 65121 65129 72200 72400 74300 74409 74831 74832 74833 74839 74841 74842 74843 74849 85110 85129 85149 85317 85322 85329 93010 97000 99999). All the data have been recoded so that all the annual employment data from 1993 to 2006 follows the 1993 SIC classification.

Appendix 2: Measuring high-tech related variety, low-and-medium-tech related variety and unrelated variety.

Frenken et al. (2007) make a distinction between related and unrelated variety at the regional level using regional employment data. Their approach is depicted in Figure 2, using an example of a hypothetical region with employment in 4 sectors present. They use sectors that are classified according to the Standard Industrial Classification (SIC codes). Using entropy measures, *unrelated variety* is measured as the entropy at the 2-digit level, and *related variety* is measured as the weighted sum of entropy at the 5-digit level within every 2-digit sector. Hence, variety across 2-digit sectors represents *unrelated variety* and variety within 2 digit sectors represents *related variety*.

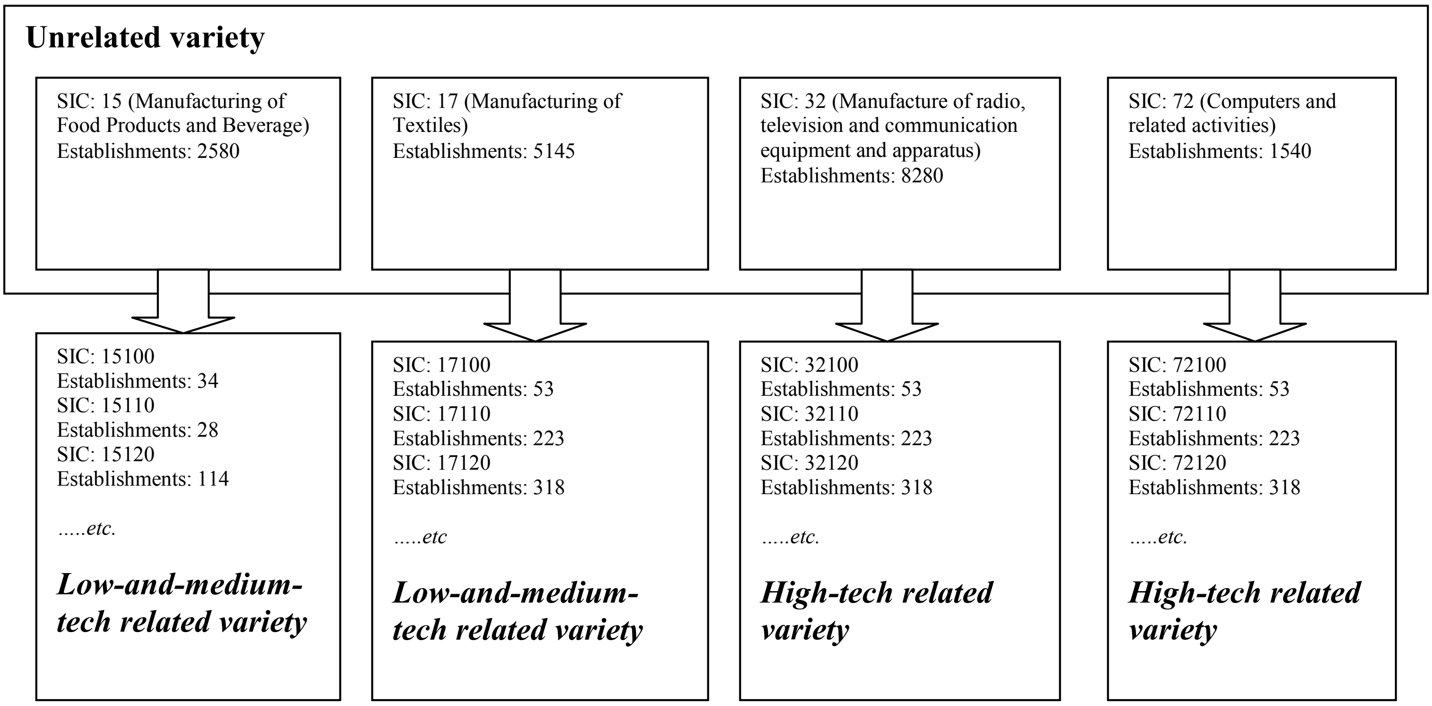
Figure 2: Related variety and unrelated variety as measured by Frenken et al. (2007)



In this paper we further decompose related variety into *high-tech related* variety and *low-and-medium tech related variety*. This approach is depicted in Figure 3. In doing so, we draw upon the Finnish Standard Industrial Classification 1995 (SIC 1995) which is derived from and corresponds with few exceptions to the European Community NACE Rev. 1. classification. This classification separates high-tech sectors from low-and-medium tech sectors on the basis of their R&D intensity (R&D expenditures over value added) and share of tertiary educated persons employed, the latter which also accounts for knowledge- and innovation-intensive sectors (which do not necessarily have a high R&D intensity). Consequently, we measure *high-tech related variety* as the weighted sum of five-digit entropy within all the 2-digit and 3-digit sectors that are classified as high-tech, which are the following: 244, 30, 32, 352, 642, 72, 73 (see Table 1 in the paper). *Low-and-medium tech related variety* is measured as the weighted sum of entropy within all the remaining sectors (15, 17, and so on). Hence, the sum of *low-and-medium tech related variety* and *high-tech related variety* equals related variety (weighted sum of entropy within *all* 2-digit sectors) as depicted in Figure 3.

The entropy measure is decomposable within and across different levels of the SIC-classification, which implies that the different variety measures do not necessarily correlate with each other. Five-digit entropy (which we use to measure *variety* in general) is equal to the sum of the weighted sum of five-digit entropy within each of the 2- and 3-digit high-tech sectors (*high-tech related* variety) and the weighted sum of five-digit entropy within the remaining sectors (*low-and-medium-tech related* variety) and the two-digit entropy (*unrelated variety)*. Hence, it would be possible for a region to have no unrelated variety, no low-and-medium-tech related variety, and much high-tech related variety. An equal amount of unrelated variety, low-and-medium-tech related variety and high-tech related variety would also be possible. Therefore, because the different variety measures do not necessarily correlate with each other, they can enter a regression simultaneously (a mathematical elaboration on the entropy measure and its decomposable nature can be found in, for example, Theil 1972).

Figure 3: High-tech related variety, low-and-medium tech related variety, and unrelated variety as measured in this paper



Appendix 3: Descriptive statistics and correlation matrix independent variables (pooled, cross-section 1993-2006)

Descriptive statistics

Variables	*Mean*	*SD*	*Min*	*Max*
VARIETY	5.7369	0.7388	3.424.053	7.5833
RELVAR	1.5640	0.4606	0.3336	2.9786
UNRELVAR	4.1730	0.3473	2.935	4.7732
RELVARHTECH	0.0306	0.0421	0.000	0.2329
RELVERLMTECH	1.5334	0.4276	.3336	2.8142
POPDENS	42.4381	117.2207	0.6855	918.0660
R&DPERCAP	334.9098	481.6451	4.0880	3523.1140
HUMCAP	0.0569	0.275946	0,01565	0.1847
POPDENS, R&DPERCAP and HUMCAP enter the regression analysis log transformed				

Correlation matrix

Variables	RELVAR	UNREL VAR	RELVAR HTECH	RELVAR LMTECH	POPDENS (log)	R&Dexp (log)	HUMCAP (log)
RELVAR	1						
UNRELVAR	0.6564	1					
RELVARHTECH	0.8074	0.4834	1				
RELVERLMTECH	0.9981	0.6600	0.7693	1			
POPDENS (log)	0.5336	0.4494	0.6643	0.5081	1		
R&DPERCAP (log)	0.4948	0.3406	0.6974	0.4626	0.5949	1	
HUMCAP (log)	0.6565	0.4005	0.7334	0.6338	0.6463	0.7075	1

References

Ali-Yrrkö J, Hermans R (2004) Nokia: A Giant in the Finnish innovation system. In: Schienstock C (ed) Embracing the Knowledge Economy. The Dynamic Transformation of the Finnish Innovation System. Edward Elgar, Cheltenham

Arellano M, Bond S (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies 58: 277-297

Arellano M, Bond S (1998) Dynamic panel data estimation using DPD98 for Gauss: a guide for users. Available at ftp://ftp.cemfi.es/pdf/papers/ma/dpd98.pdf.

Arellano M, Bover, O (1995) Another look at the instrumental-variable estimation of error-components model. Journal of Econometrics 68 (1): 29-52.

Audretsch DB, Aldridge TT (2008) The knowledge Spillover Theory of Entrepreneurship and Spatial Clusters. In: Karlsson C (ed) Handbook of Research on Cluster Research on Cluster Theory. Edward Elgar, Northampton MA

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

Bishop P, Gripaios P (2010) Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain. Regional Studies 44 (4): 443-454

Blomstrom M, Ari K, Fredrik S (2002) Growth and innovation policies for a knowledge economy: Experiences from Finland, Sweden, and Singapore. Stockholm: European Institute of Japanese Studies, Working paper no. 156

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 24 (1): 115-143

Blundell R, Bond S (2000) GMM Estimation with persistent panel data: an application to production functions. Econometric Reviews 19 (3): 321-340

Bond SR (2002) Dynamic panel data models: a guide to micro data methods and practice. Portugese Economic Journal 1 (2): 141-162

Boschma RA, Iammarino S (2009) Related variety, trade linkages and regional growth. Economic Geography 85 (3): 289-311

Boschma, R, Minondo, A, Navarro, M (2011) Related variety and regional growth in Spain. Papers in Regional Science, forthcoming, doi: 10.1111/j.1435-5957.2011.00387.x

Bowsher CG (2002) On testing overidentifying restrictions in dynamic panel data models. Economics Letters 77: 211-220.

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87: 115-143.

Buerger M, Cantner U (2011) The regional dimension of sectoral innovativeness: An empirical investigation of two specialized suppliers and two science based industries. Papers in Regional Science 90 2: 373-394.

Cohen WM, Levinthal DA (1990) Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35: 128-152

Crescenzi R, Rodriguez-Pose A (2008) Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. Regional Studies 42 (1): 51-67

De Groot HLF, Poot J, Smit MJ (2009) Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In: Capello R, Nijkamp P (eds) Handbook of Regional Growth and Development Theories. Edward Elgar, Northampton MA

Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies: In: Henderson V, Thisse JF (eds) Handbook of Regional and Urban Economics Vol 4. North-Holland, New York

Edquist CL, Hommen L, McKelvey M (2001) Innovation and Employment, Process versus Product Innovation. Elgar, Cheltenham

Edquist C, Luukkonen T, Sotarauta M (2009) Broad-Based Innovation Policy. In: Oy T (ed) Evaluation of the Finnish National Innovation System - Full report. Helsinki, Helsinki University Print

Frenken K, Van Oort FG, Verburg T (2007) Related Variety, Unrelated Variety and Regional Economic Growth. Regional Studies 41 (5): 685-697

Glaeser EL, Kallal H, Scheinkman J, Shleifer A (1992) Growth in cities. Journal of Political Economy 100: 1126-1152

Hansen, L. (1982), Large sample properties of generalized method of moments estimators. Econometrica 50 (3): 1029-1054.

Harrison B, Kelley MR, Gant J (1996) Innovative firm behavior and local milieu: exploring the intersection of agglomeration, firm effects, and technological change. Economic Geography 72: 233-258

Hauknes J, Knell M (2009) Embodied knowledge and sectoral linkages: An input-output approach to the interaction of high- and low-tech industries. Research Policy 38 (3): 459-569

Heidenreich M (2009) Innovation patterns and location of European low-and-medium-technology industries. Research Policy 38 (3): 483-494

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

Hirsch-Kreinsen H, Jacobson D, Laestadius S, Smith K (2005) Low and Medium Technology Industries in the Knowledge Economy: The Analytical Issues. In: Hirsch-Kreinsen H, Jacobson D, Laesta-dius S (eds) Low-tech Innovation in the Knowledge Economy. Peter Lang, Frankfurt

Holtz-Eakin D, Newey W, Rosen HS (1988) Estimating vector autoregressions with panel data. Econometrica 56: 1371-1395

Honkapohja S, Koskela E (1999) The economic crisis of the 1990 in Finland. Economic Policy 14: 399–436

Iammarino S, McCann P (2006) The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. Research Policy 35 (7): 1018-1036

Jacobs J (1969) The Economy of Cities. Vintage, New York

Kirner E, Kinkel S, Jaeger A (2009) Innovation paths and the innovation performance of low-technology firms - An empirical analysis of German industry. Research Policy 38 (3): 447-458

Leahy D, Neary JP (2007) Absorptive capacity, R&D spillovers, and public policy. International Journal of Industrial Organization 25 (5): 1089-1108

Lucas RE (1993) Making a miracle. Econometrica 61: 251-272

Morone P (2006) The Two Faces of Knowledge Diffusion: The Chilean Case. Journal of International Development 18 (1): 29-50

Neffke F, Henning SM (2012) Skill-relatedness and firm diversification. Strategic Management Journal, forthcoming

Neffke F, Henning M, Boschma R (2011) How Do Regions Diversify over Time? Industry Relatedness the Development of New Growth Paths in Regions. Economic Geography 87 3: 237-265.

Nooteboom B (2000) Learning and innovation in organizations and economies. Oxford, Oxford University Press

Paci R, Usai, S (2000) Technological Enclaves and Industrial Districts: An Analysis of the Regional Distribution of Innovative Activity in Europe. Regional Studies 34 (2), pp. 97-114

Pavitt K (1984) Sectoral patterns of technical change: towards a theory and a taxonomy. Research Policy 13: 343-373

Quatraro, F. (2010) Knowledge coherence, variety and economic growth. Manufacturing evidence from Italian regions. Research Policy, doi: 10.1016/j.respol.2010.09.005

Quigley JM (1998) Urban Diversity and Economic Growth. Journal of Economics Perspectives 12 (2): 127-138

Roodman D (2009a) A Note on the Theme of Too Many Instruments. Oxford Bulletin of Economics and Statistics 71 (1): 135-158.

Roodman D (2009b) How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal 9 1: 86-136.

Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson V, Thisse JF (eds) Handbook of Regional and Urban Economics Vol 4. North-Holland, New York

Rouvinen P, Ylä-Anttila P (2003) Case-study: little Finland’s transformation to a wireless giant. In: Dutta S, Lanvin B, Paua F (eds) The Global Information Technology Report 2003–2004. Oxford University Press, New York

Santamaria L, Nieto MJ, Barge-Gil A (2009) Beyond formal R&D: Taking advantage of other sources of innovation in low- and medium-technology industries. Research Policy 38 (3): 507-517

Saxenian A (1994) Regional Advantage. Harvard University Press, Cambridge MA

Simonen J, McCann P (2008) Firm innovation: The influence of R&D cooperation and the geography of human capital inputs. Journal of Urban Economics 64 (1): 146-154

Storper M, Venables AJ (2004) Buzz: face-to-face contact and the urban economy. Journal of Economic Geography 4 (4): 351-370

Theil, H (1972) Economics and Information Theory. North Holland, Amsterdam

Van Oort, FG (2004) Urban Growth and Innovation. Spatially Bounded Externalities in the Netherlands. Ashgate, Aldershort

Windmeijer F (2005) A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 126: 25-51.