A multilevel approach to geography of innovation

Martin Srholec[*]

Centre for Technology, Innovation and Culture; University of Oslo

Version of October 1st 2007

TIK Working Papers on Innovation Studies
October 2007

Abstract

The aim of this paper is to demonstrate how research on geography of innovation can benefit from multilevel modeling. Using explanatory factors operating at different levels of the analysis, we assess the hypothesis that regional innovation systems influence the firm's likelihood to innovate. We estimate a logit multilevel model of innovation on micro data from the third Community Innovation Survey in the Czech Republic. The results indicate that the quality of the regional innovation system directly determines firm's likelihood to innovate and mediates the effect of some firm-level factors. Also structural problems in the region influence innovation in firms.

Keywords: innovation, multilevel modeling, regional innovation system, Czech Republic.

JEL: O30, R15, D21.

[*] I am grateful to the Czech Statistical Office for providing access to the firm-level data. Special thanks go to Martin Mana and Josef Plandor for supporting access to the data. Financial support from the post-doc project on "Innovation, Fragmentation and Economic Development" supported by the University of Oslo is gratefully acknowledged. All usual caveats apply. E-mail: martin.srholec@tik.uio.no.

1. Introduction

Many kinds of data have hierarchical structure. Researchers in education science, human geography or biology have long recognized this issue. Offspring from the same parents and environment tend to be more alike that those chosen at random from the population. School performance is not only given by the amount of study time of a child, but also by higher-level factors such as characteristics of the class, school or national educational system. Similarly the innovation process in firms is influenced by factors operating at the micro level as well as by characteristics of innovation systems at sectoral (Malerba and Orsenigo, 1995, 1997), regional (Cooke, 1992; Asheim and Isaksen, 1997 and Morgan, 1997) and national levels (Lundvall, 1992; Nelson, 1993 and Edquist, 1997).

An appropriate approach to analyse relations identified at different levels is multilevel modeling (Goldstein, 2003; Hox, 2002 and Luke, 2004). Single-level models assume that observations are independent from each other. If a nested structure of data exits, however, the independence assumption is likely to be violated. By relaxing this assumption, multilevel modeling provides a tool for analysis of units grouped at regional and other levels. A proper recognition of data hierarchies allows us to analyse the extent to which specific differences between regions are accountable for outcomes at the firm level. Unlike any other method, multilevel modeling also enables the researcher to explore mechanics by which these regional factors operate at the micro level and the extent to which these effects differ for different kinds of firms.

It is well understood in the literature on geography of innovation that a vibrant regional innovation system is instrumental for firm's capability to generate new products and processes (for recent overviews see Cooke, 2004; Asheim, Gertler, 2004; Doloreux and Parto, 2005). An important insight from this literature is that spatial concentration of the relevant actors, resources and other environmental factors conducive to learning influences firms' innovative performance as much as their individual characteristics, such as the size, age or ownership of firms. Although a firm embedded in a regional innovation system is the best example of a hierarchical structure, empirical research in this tradition continues to use single-level models that are severely restricted to handle multilevel hypotheses.

The main aim of this paper is to help in filling this gap. We demonstrate how research on geography of innovation can benefit from multilevel modeling. Using explanatory factors operating at different levels of the analysis, we provide a formal assessment of the hypothesis that regional innovation systems influence the firm's likelihood to innovate. Section 2 introduces a basic multilevel model and addresses related conceptual and methodological issues. Application of the model is illustrated on a large sample of micro data from the third Community Innovation Survey (CIS) in the Czech Republic. Section 3 presents the micro dataset and constructs regional variables with the help of factor analysis. Section 4 specifies a bivariate logit multilevel model of innovation and provides results of the estimates. Since to my best knowledge this is the first time a multilevel model is used to study innovation, the last section outlines agenda for future research along these lines.

2. Multilevel modeling

A hierarchy refers to units clustered at different levels (Goldstein, 2003). For example, firms may be the level-1 units nested within a higher-level structure, where these higher levels are sectors, regions or countries. A multilevel model sometimes also called a hierarchical, random coefficient or mixed-effect model is then defined a statistical model that relates a dependent variable to predictor variables at more than one level (Luke, 2004).

Suppose a multilevel model has 2-level structure with firms at level-1 located in regions at level-2. A standard linear 2-level model with one explanatory variable at each level is the following:

(1) Level-1 model:

$$y_{ij} = \beta_{0j} + \beta_{1j}\, x_{ij} + e_{ij}$$

Level-2 model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, w_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, w_j + u_{1j}$$

where y_{ij} is the dependent variable, x_{ij} is the level-1 predictor, w_j is the level-2 predictor, e_{ij}, u_{0j} and u_{1j} are random effects (normally distributed residual terms for each equation), i is the firm ($i = 1\ldots n$) and j is the region ($j = 1\ldots m$).

At the level-1 the equation refers to a firm-level relationship, which is defined separately for each region. If the level-2 equations were not specified, the level-1 relationship might have been estimated as a standard single-level model. A multilevel model emerges, if we let the intercept β_{0j} and slope β_{1j} to become random variables. Since the level-2 effects are identified by the subscript j, we have a system of equations at different levels, where we are allowing each region to have a different average outcome and a different effect of the level-1 predictor on the outcome.

Although a different firm-level model is being estimated for each region, the level-2 equations tell us that the intercept and slope are influenced by the regional effects. The model indicates that γ_{00} is average of the level-1 dependent variable after controlling for the level-2 predictor, γ_{01} is the effect of the level-2 predictor on the level-1 intercept, γ_{10} is average of the level-1 slope after controlling for the level-2 predictor and γ_{11} is the effect of the level-2 predictor on the level-1 slope.

By substituting the equations for β_{0j} and β_{1j} in the level-1 model we arrive to a single equation or "mixed" formulation of the model:

(2) $$y_{ij} = \gamma_{00} + \gamma_{01}\, w_j + \gamma_{10}\, x_{ij} + \gamma_{11}\, w_j\, x_{ij} + (u_{0j} + u_{1j}\, x_{ij} + e_{ij})$$

where the dependent variable becomes the sum of a fixed part and a random part of the model (in the brackets). Since this notation shows the multilevel model in a familiar linear regression format, it draws attention to the main differences from the standard single-level model. These are the higher-level predictor, the cross-level interaction term and the random part. Since more than one residual term is present, the traditional estimation procedures such as ordinary least squares are inapplicable and specialized maximum likelihood procedures must be used to properly estimate these models (Raudenbush, et al. 2004).

It should be noted that equation (1) shows a typical structure of a multilevel model, but there is a variety of specifications that can be estimated depending on the research question. Of course, we can include more than one predictor variable at each level. It is possible to specify models with only the intercept as a function of level-2 predictors or models with random effects only in selected level-2 equations. Also models with more than 2 levels can be formulated.

So why should we use multilevel modeling? A major assumption of single-level models is that the observations (and hence residuals) are independent from each other. If a nested structure of data exits, units belonging to the same group tend to have correlated residuals and the independence assumption is likely to be violated. By relaxing this assumption, multilevel modeling provides statistically more efficient estimates of regression coefficients and estimate correct standard errors, which are more "conservative" than those ignoring the hierarchical nature of data (Goldstein, 2003). Statistically significant relationships that have been established in the literature by using the standard methods may come out not significant in the multilevel analysis. A lot that we have learned empirically about innovation in firms from research on data at the aggregate level might appear different in the multilevel framework.

Economic geographers have been long interested in comparing regions and countries. Aim of such comparisons may be to examine issue that are distinctly macroeconomic, such as institutions and policies, where the aggregate level is appropriate unit of analysis. If the prime interest is how regional differences influence firm performance, however, it might be problematic to project statistical inferences discovered at a higher level to occur at a lower level. Analysis that assumes relationships observed in groups to hold for individuals may suffer from so-called aggregation (or ecological) fallacy (Luke 2004). An example is correlation between consumption patterns and diseases epidemiology, which for some diseases holds at the aggregate level but is not very strong for individuals. Even if the aggregate level is the prime interest, a multilevel approach is useful because it allows us to integrate into the analysis possibly important variables that are measured at the micro level, which may considerably improve predictive power of the model.

Studies that use exclusively micro data to account for the effects of environment on firms often suffer from issues of endogeneity. A good example is the set of variables on obstacles to innovation from CIS. Even though most of these obstacles, such as excessive regulation or lack of customer interest, refer to factors external to the firm, they fail to measure the environmental effects. More innovative firms systematically report more severe obstacles to innovation, because they are more aware of what is hindering innovation than the less innovative firms. An inevitable outcome of a single-level analysis is highly positive correlation between innovativeness and these external obstacles to innovation. Innovation influences firm's perception of the obstacles, not the other way round (Veugelers and Cassiman, 2004). A multilevel model should be used to analyse the role of external factors, where we can safely assume that the arrow of causality goes from actual regional characteristics to the firm's performance.

Apart from the statistical consequences, a proper recognition of data hierarchies allows us to examine new lines of questions. Using the example of firms in regions, the multilevel approach enables the researcher to explore the extent to which specific differences between regions are accountable for outcomes at the firm level. It is also possible to investigate the mechanics by which the regional factors operate at the firm level and the extent to which these effects differ for different kinds of firms. For example, we may analyse whether differences in innovation systems across regions are more important for smaller than larger firms. Such research questions can be straightforwardly examined by multilevel modeling, but can be neither easily nor properly examined by the standard methods.

As already mentioned above, moreover, another important reason for using multilevel modeling to study innovation is theoretical. A central argument of the literature on innovation systems is that firms are embedded locally and therefore the theory implicitly predicts a nested structure of micro data. In other words, the basic assumption of the standard multiple regression models on independent residuals is expected to be violated from the outset. Empirical research in this tradition that uses single-level models to study how characteristics of innovation systems influence innovation in firms suffers from a methodological contradiction. Anytime a researcher aims to test hypotheses that are operating at different levels, a multilevel statistical model is appropriate.

A common approach to avoid having all of the contextual effects pooled into the single error term is to ignore the random variability associated with the higher-level factors and include a fixed effect that corresponds to the hierarchical structure into a single-level regression. For example, dummies for the higher-level units are often used to control for the compositional effects. Although we may detect rough patterns of the structure, it is only a partial solution. Using dummies might be a useful quick-fix solution, if we are interested exclusively in the level-1 relationships, but it is of a little help if the prime interest is in effects of the higher-level factors or cross-level interactions. A dummy is a "catch-all" variable stripped of the context for which we can only speculate what it really represents. And if there are many higher-level units, models with these dummies will have many more parameters, decreasing the degrees of freedom and resulting in reduced parsimony. After all, if the higher-level dummies significantly improve predictive power of the model, which indeed is often the case in the literature; a multilevel model should be given priority.

It should be noted, finally, that not only multilevel modeling relaxes the standard independence assumption on residual terms. Spatial autocorrelation techniques have been developed to produce valid statistical inferences if errors tend to be correlated regionally. Also survey design and analytical tools recognize the need to take into account the hierarchical structure of the population. Although these procedures are deemed to be necessary to obtain efficient estimates, the higher-level effects typically do not merit a serious interest themselves. Only multilevel modeling allows us to look closely at the patterns and consequences of hierarchical structure of the phenomena in question.

3. Overview of the data

The analysis is based on micro data from a compulsory survey organized by the Czech Statistical Office, which asked firms about their innovative activities over 1999-2001. It was conducted as a part the third wave of CIS organized by Eurostat and therefore fully harmonized with the methodology of the Oslo Manual (OECD, 1997). A representative sample of 5,829 enterprises with more than 10 employees was surveyed of which about 65% responded. After omitting firms with incomplete records the survey provides a dataset of 3,801 firms in both industry and market services (10-74 codes according to NACE, rev. 1.1).

Our dependent variable is "INNOV", which is a dummy with value 1 if the firm successfully introduced a new product or process. About 37% respondents innovated over the period. Besides evidence on innovation activities, the dataset provides information on size, age, ownership and location. "SIZE" of the firm refers to the number of employees (in logs) at the beginning of the period. "AGE" refers to the number of years since registration of the firm in the business register (also in logs) until the end of the period.[1] About half of the sample consists of small firms with less than 50 employees, while roughly one fifth represents large firms with more than 249 employees. A typical age is 8 years and about 7.5% of the firms were newly registered during the period. "FOREIGN" is a dummy variable with value 1 for firms with more than 50% share of non-residents in equity. About one fourth of the respondents were foreign affiliates, which broadly corresponds to the official statistics of foreign ownership in the Czech economy.

The location of firms is identified by the NUTS4 code.[2] At this level the Czech Republic is divided into 91 units of which 15 are in the capital city of Prague. Since the district borders within the Prague agglomeration are rather artificial and regional statistics are reported only for the whole capital city, we have combined these into a single Prague region. This leads to 77 regions with median population of 109 thousand people and median area of 1,030 km^2. About fourth of the firms is located in Prague. Median number of observations is 32 per region. Regional distribution of the sample is highly representative with regards to concentration of business activity in the Czech Republic.

To account for the higher-level effects, we select the following 11 indicators from the Czech regional statistics: 1) Log of population density per km^2 in 2000; 2) Urbanization defined as the percentage of population living in towns in 2000; 3) University attainment given by the percentage of people with a university degree in population more than 15 years old in 2001; 4) Agglomeration of specialized business services refers to the percentage of business services (70-74 codes according to NACE, rev. 1.1) in employment in 2001; 5) Average monthly wage in 2001; 6) Log of acquired tangible fixed assets per capita (excl. investment in housing construction

[1] Since firms without employees at the beginning of the period and/or firms established in the final year were present in the sample, we had to add one to SIZE and AGE before the log-transformation (in order to prevent natural logarithm of zero).

[2] It should be stressed that the NUTS4 regions embrace historical (and sometimes even distinct cultural) homogeneity dating back at least to the land reform introduced by Maria Theresia in the first half of the eighteenth century. For example, local labour markets (and local job offices) or representation to the lower house of the parliament are primarily based on the NUTS4 basis.

and environment protection) in 2001; 7) Log of intangible acquired intangible fixed assets per capita in 2001; 8) A dummy for presence of a technical university in the region; 9) Long-term unemployment rate identified by job applicants registered more than 12 months at the labour office relative to the labour force in 2002; 10) Number of divorces per 100 marriages in 2001; and 11) Log of SO_2 emissions (REZZO 1) in tonnes per year per km^2 in 2000.[3]

Since many of these indicators are highly correlated to each other, it would be problematic to include all of them in the estimate. Fortunately, there is a well-established method of multivariate analysis, so-called factor analysis, which allows us to condense correlated indicators into a smaller number of latent variables. The idea is that highly correlated indicators are likely to reflect the same underlying dimension and therefore can be combined together without loss of much information. For example, Fagerberg and Srholec (2006) show in a large sample of countries that many indicators of economic, technological and social development can be reduced with the help of factor analysis into a few principal factors jointly explaining almost three fourths of the total variance. Likewise, there might be only a few underlying dimensions in the regional data.

The aim of factor analysis is to identify a limited number of factors that account for most variance in the dataset. In the first step, a set of factors is iteratively generated based on patterns of correlation between the variables. After consulting the so-called eigenvalues, only the most significant factors are retained. In the second step, the solution is rotated to maximize differences between the retained factors in order to improve interpretability of the results. Orthogonal rotations, such as the most widely used varimax normalized rotation, are constrained to produce factor scores that are uncorrelated. More complex and recently developed oblique rotations do not impose this restriction. Since the assumption of orthogonality often leads to biased results, we use the more flexible (and realistic) oblique oblimin rotation (for details see Basilevsky, 1994). It should be noted, however, that the different rotations produced very similar outcomes.

Table 1 gives results of the factor analysis for the dataset of 11 indicators in 77 regions. Only two factors with eigenvalue higher than one were detected and therefore retained for the rotation. About three fourths of the total variance is jointly explained by these two dimensions of the data. Other eigenvalues were marginal with value of 0.45, 0.14 and less and therefore also the scree-test confirmed the solution. Our main interest is in factor scores on these two retained factors, which are linear combinations of the underlying variables with weights generated by the procedure. Interpretation of the factors is given by the so-called factor loadings (pattern matrix) reported in the table.[4]

---

[3] Information from the nearest available year to the initial period of the CIS survey was used. Older data are not available for most of the indicators. Statistics of R&D expenditure and employment is available only at the NUTS3 level in the Czech Republic.

[4] Factor loadings are correlation coefficients between the original variables and the factors. It should be noted, however, that in the case of oblique rotations one should examine the pattern matrix as well as the so-called structure matrix. Coefficients in the pattern matrix represent unique contributions of the variables, while structure matrix (as in orthogonal rotations) contains both the unique and common contributions. For the sake of brevity and space, we report only the pattern matrix, because interpretation of the structure matrix is not different.

Table 1: Results of the factor analysis (factor loadings after rotation)

	Factor 1 Regional innovation system (RIS)	Factor 2 Structural problems (STR)
Population density	0.673	0.415
Urbanization	0.425	0.560
University attainment	0.927	-0.203
Business services	0.866	0.119
Tangible investment	0.630	-0.151
Intangible investment	0.720	0.022
Technical university	0.678	0.091
Average wage	0.869	0.004
Long-term unemployment	-0.281	0.804
Divorces per marriages	0.060	0.789
SO_2 emissions	0.181	0.746
Eigenvalue after rotation	4.876	2.948

Note: Principal factors method; oblique oblimin rotation.

After the rotation, only some of the variables tend to have high loadings on each factor. The first factor loads highly on agglomeration of university educated people, supply of specialized business services (incl. R&D, engineering and consultancy services), investment in intangible assets and the presence of a technical university in the region. It also comes out highly correlated to investment in technology embodied in new tangible assets and not surprisingly to a high wage level suggesting non-price competitive edge of the region, at least as far as the national context in concerned. Since these variables indicate agglomeration of inputs that are deemed critical for innovation, we shall use this factor as a proxy for quality of the regional innovation system “RIS” in the following.

The second factor loads highly on indicators of structural problems in the region reflected in long-term unemployment, volatile social relationships represented by the propensity to divorce and adverse health effects due to high pollution. All of these broader socio-economic problems are no doubt closely intertwined with each other and concentrate in the old industrial regions in the north-western and north-eastern part of the country. It is interesting to note that these are the regions with high concentration of the mature “chimney” industries, such as mining and heavy industry, which used to be the backbone of the Czechoslovak industry, but underwent slow and often painful restructuring after the collapse of central planning. For lack of a better label, we shall use this factor as a measure of structural problems “STR” in the region.

Both of the factor scores are at least modestly correlated to population density and the rate of urbanization. As already hinted above, an important spatial feature of the Czech economy is that both the most developed regional innovation systems as well as the most daunting structural problems concentrate in various urban/industrial areas rather than in the rural/agricultural regions. Another principal factor that would differentiate between urban and rural regions was simply not born out from the data and therefore differences along these lines are not propagated in the analysis.

4. Econometric estimates

Until now we have assumed that the dependent variable is continuously distributed. Since INNOV is binary, we need to specify a non-linear multilevel model. For this purpose, we distinguish between the sampling model (3.1), a link function (3.2) and a structural part of the multilevel model (3.3). The standard linear specification assumes normal sampling model and therefore no link function to transform the level-1 predicted values is necessary. However, a binary specification requires binomial sampling model (the Bernoulli distribution) and a logit transformation. Only the level-1 model differs from the linear case and can be written down as follows:

(3.1) $E\ (y_{ij} = 1 \mid \beta_j) = \varphi_{ij}$

(3.2) $\text{Log}\ [\varphi_{ij} / (1 - \varphi_{ij})] = \eta_{ij}$

(3.3) $\eta_{ij} = \beta_{0j} + \beta_{1j}\ x_{1ij} + \beta_{2j}\ x_{2ij} + \ldots + \beta_{nj}\ x_{nij}$

where η_{ij} is the log of the odds of success (to introduce an innovation, for instance). Although φ_{ij} is constrained to be in the interval (0,1), the logit transformation allows η_{ij} to take any value and therefore can be substituted to the structural model. Note that the predicted log-odds can be converted to an odds by $\exp(\eta_{ij})$ and to the predicted probability φ_{ij} by $\exp\{\eta_{ij}\}/(1+\exp\{\eta_{ij}\})$. Furthermore, there is not a separate term for the level-1 error because for a binary outcome the variance is completely determined by the population mean (for details see Luke 2004, pg. 55).

The aim of the analysis is to explain firm's likelihood to innovate by factors operating at the firm and regional levels. INNOV is the dependent variable, SIZE (in logs), AGE (in logs) and the dummy for foreign ownership FOREIGN are the level-1 predictors and the factor scores on RIS and STR are the level-2 predictors. We build the bivariate logit multilevel model of innovation from bottom up. First, we consider only the level-1 predictors and let the level-2 effects to be random variables. Second, we examine the so-called "intercept-as-outcome" model, which includes the level-2 predictors only for the intercept. And finally, we estimate the full "slopes-as-outcomes" model, which relates the level-2 predictors to both the intercept and slopes.[5]

To improve interpretability of the results, we centre the level-1 predictors SIZE and AGE on zero by deducting mean before the estimation. Also note that factor analysis produces standardized scores, so that the level-2 predictors RIS and STR have mean of zero and standard deviation equal to one. Since zero of the FOREIGN variable implies a domestic firm, all of the predictors have meaningful zero-points, which greatly simplify interpretation of the estimated intercept, as shall be seen below.

[5] A specialized statistical software Hierarchical Linear and Non-linear Modeling (HLM) version 6.04 was used to estimate the equations. See Raudenbush, et al. (2004) for details on the estimation procedure.

Equation (4) specifies the basic model. Although there are no level-2 predictors, we allow the level-1 intercept and slopes to vary across regions by including the random effects:

(4) Level-1 model:

$$E\,(\mathrm{INNOV}_{ij} = 1 \mid \beta_j) = \varphi_{ij}$$

$$\mathrm{Log}\,[\varphi_{ij} / (1 - \varphi_{ij})] = \beta_{0j} + \beta_{1j}\,\mathrm{SIZE}_{ij} + \beta_{2j}\,\mathrm{AGE}_{ij} + \beta_{3j}\,\mathrm{FOREIGN}_{ij}$$

Level-2 model:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + u_{2j}$$
$$\beta_{3j} = \gamma_{30} + u_{3j}$$

where there are four fixed effects (γ_{00} ... γ_{30}) and four random effects (u_{0j} ... u_{3j}) of which γ_{00} is the estimated mean of the log-odds of firms to innovate across regions, u_{0j} tells us that the regions vary around that mean, γ_{10} , γ_{20} and γ_{30} are the estimated slopes across regions and u_{1j}, u_{2j} and u_{3j} indicate that these slopes vary not only as a function of the three level-1 predictors SIZE, AGE and FOREIGN, but also as a function of a unique regional effect (the level-2 residuals are assumed to be sampled from a bivariate normal distribution; with expected zero mean and variance = σ^2_u).

Table 2 provides results of the basic model in the first column. As noted above, we can transform the estimated coefficients from the log-odds back into the expected probability to innovate by using the inverse of the logit link function. For the intercept the conditional log-odds of firms to innovate is estimated at -0.777. Since the intercept is the expected log-odds to innovate for a firm with values of zero on the predictors, the estimated intercept refers to an average age and size domestic-owned firm located in an "average" region (with a random effect of zero). From this follows that for the typical domestic-owned firm the predicted probability to innovate is $\exp\{-0.777\}/(1+\exp\{-0.777\})$, which equals to 31.5%.

SIZE, AGE and FOREIGN are highly significant level-1 predictors. Besides advantages from scale economies of various kinds, size is important to control for due to definition of the dependent variable. Since INNOV is a dummy for introducing at least one innovation, larger firms are by principle more likely to report a positive answer because they comprise more activities or even multiple product lines and plants under a single roof. Age is important because on one hand old firms tend to have more accumulated knowledge and other resources to capitalize on, but on the other hand firms that have been just established are more likely to report "new to the firm" product or process. As the negative coefficient of AGE suggests, the latter effect prevails in the data. Since both of these variables are in logarithms to account for the likely non-linearity involved, the quasi-elasticity of innovation in respect of them is given by $\varphi_{ij}(1 - \varphi_{ij})\beta$. At the estimated sample mean innovation rate of 33.2%, all else equal the probability of a firm to innovate is therefore predicted to increase by one percentage point upon 9.9% increase in SIZE and 16.2% decrease in AGE.

Table 2: Econometric results

	Equation (4)	Equation (5)	Equation (6)
Fixed Effects:			
For Intercept$_{ij}$ (β_{0j})			
Intercept$_{ij}$ (γ_{00})	-0.777 (0.048)***	-0.784 (0.047)***	-0.796 (0.048)***
RIS$_j$ (γ_{01})	..	0.060 (0.012)***	0.075 (0.018)***
STR$_j$ (γ_{02})	..	-0.123 (0.038)***	-0.078 (0.043)*
For SIZE$_{ij}$ slope (β_{1j})			
SIZE$_{ij}$ (γ_{10})	0.455 (0.031)***	0.457 (0.032)***	0.497 (0.030)***
RIS$_j$ (γ_{11})	..	..	-0.054 (0.014)***
STR$_j$ (γ_{12})	..	..	-0.059 (0.020)***
For AGE$_{ij}$ slope (β_{2j})			
AGE$_{ij}$ (γ_{20})	-0.279 (0.068)***	-0.283 (0.074)***	-0.340 (0.100)***
RIS$_j$ (γ_{21})	..	..	0.007 (0.029)
STR$_j$ (γ_{22})	..	..	0.011 (0.060)
For FOREIGN$_{ij}$ slope (β_{3j})			
FOREIGN$_{ij}$ (γ_{30})	0.331 (0.082)***	0.339 (0.077)***	0.330 (0.114)***
RIS$_j$ (γ_{31})	..	..	-0.023 (0.029)
STR$_j$ (γ_{32})	..	..	-0.093 (0.088)
Random effects:			
Intercept$_{ij}$ (u_{0j})	0.026 (86.60)	0.003 (73.28)	0.003 (71.83)
SIZE$_{ij}$ slope (u_{1j})	0.019 (102.07)**	0.019 (102.58)**	0.005 (79.09)
AGE$_{ij}$ slope (u_{2j})	0.004 (51.31)	0.002 (50.53)	0.014 (51.21)
FOREIGN$_{ij}$ slope (u_{3j})	0.032 (85.67)	0.012 (84.62)	0.006 (83.45)
Index of dispersion	0.997	1.006	1.019
Level-1 observations	3,801	3,801	3,801
Level-2 groups	77	77	77

Note: Non-linear unit-specific model with the logit link function; full maximum likelihood (PQL) estimate; coefficients and robust standard errors in brackets reported for the fixed effects; variance components and Chi-square in brackets reported for the random effects; *, **, *** denote significance at the 10, 5 and 1 percent levels.

Using a single-level probit model on the same dataset, Srholec (2005) showed that foreign affiliates tend to engage less in internal R&D compared to domestic-owned firms and Knell and Srholec (2004) found out that foreign affiliates are more likely to cooperate on innovation with external partners abroad, but less likely to venture into innovation cooperation with partners in the Czech economy. A rather bleak picture of the impact of foreign ownership on the local innovation system came from these results. Our model suggests, in contrast, that all else equal foreign affiliates are by 7.3% more likely to innovate compared to their domestic-owned counterparts (0.332*0.668*0.331). Since foreign affiliates can capitalize on knowledge accumulated by their parents abroad, it is not surprising that they are more likely to introduce new products or processes, at least as far as the regional random effects are properly accounted for in a multilevel framework. For the very same reason, this result does not necessarily contradict the previous findings on their lower internal R&D intensity and embeddedness in the Czech innovation system.

So far we have focused only on the fixed effects. Estimates of the level-2 random effects are reported in units of the so-called variance components (square of standard deviation) in the lower part of the table.[6] As outlined in the model, the error term is split into four components. Unexplained variability of firm's innovativeness across regions that is the random effect for the intercept as well as regional variability of the effect of FOREIGN are relatively high, although not significant at conventional levels. A low residual and therefore a strong central tendency across regions in the effect of AGE is not surprising, because a majority of firms in the sample was established or newly registered as joint-stock and limited companies during the early years of transition, which was obviously a national rather than a regional effect. It also suggests that regional differences in the birth of new firms are not important for innovation. Statistically significant and a notable proportion of the variance across regions is accounted for SIZE, which indicates that geography matters a lot for the effect of scale economies.

Since non-zero variance components indicate un-modeled variability, we shall include the level-2 predictors. Equation (5) specifies the intercept-as-outcome model, which incorporates the regional variables for RIS and STR into the model as predictors of the level-1 intercept:

(5) Level-1 model:

$$E\,(INNOV_{ij} = 1 \mid \beta_j) = \varphi_{ij}$$

$$\text{Log}\,[\varphi_{ij} / (1 - \varphi_{ij})] = \beta_{0j} + \beta_{1j}\,SIZE_{ij} + \beta_{2j}\,AGE_{ij} + \beta_{3j}\,FOREIGN_{ij}$$

Level-2 model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\,RIS_j + \gamma_{02}\,STR_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

where we assume that the level-2 predictors explain the different average propensity to innovate across regions, but let the level-1 effects of SIZE, AGE and FOREIGN remain "unconditional" at the regional level.

The hypothesis is that firms located in regions with more developed innovation systems are more likely to innovate because they benefit from all sorts of geographically bounded external economies and agglomeration effects related to localized generation and diffusion of knowledge. On the other hand, structural problems in the region, such as long-term unemployment and concentration of declining industries, should have adverse effects on the frequency of innovation.

A look at the second column in Table 2 reveals that these predictions are firmly supported by the results. Both of the regional variables have the expected sign and are

[6] Since the HLM (version 6.04) package assumes that the variances may not be normally distributed, a chi-square test of the residuals is performed (Raudenbush, et al. 2004).

highly significant explanatory factors of differences in the propensity to innovate across regions; even after controlling for the firm-level factors. A typical domestic-owned firm (all of the level-1 effects of zero) located in the best region, which is the Prague agglomeration with the best RIS and modest STR score, is predicted to have 37.3% probability to innovate, whereas the same firm located in the Karviná region with the worst combination of conditions has only 24.4% probability to innovate. Since the regional variables are in the same units of standard deviation (standardized factor scores), we can directly compare magnitude of their coefficients which suggests that STR has even larger effect than RIS on the firm's propensity to innovate (although the former regional effect is not robust to inclusion of the interaction terms as shall be seen below).

After the level-2 predictors have been included in the model, the random effect for the intercept has decreased substantially, which confirms that a bulk of the unexplained variance across regions has been accounted by the RIS and STR variables. To further investigate their explanatory power, we allow the regional variables to influence also slopes of the level-1 predictors. Equation (6) specifies the full model:

(6) Level-1 model:

$$E\,(INNOV_{ij} = 1 \mid \beta_j) = \varphi_{ij}$$

$$\text{Log}\,[\varphi_{ij} / (1 - \varphi_{ij})] = \beta_{0j} + \beta_{1j}\, SIZE_{ij} + \beta_{2j}\, AGE_{ij} + \beta_{3j}\, FOREIGN_{ij}$$

Level-2 model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\, RIS_j + \gamma_{02}\, STR_j + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}\, RIS_j + \gamma_{12}\, STR_j + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}\, RIS_j + \gamma_{22}\, STR_j + u_{2j}$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31}\, RIS_j + \gamma_{32}\, STR_j + u_{3j}$$

where the new fixed parameters ($\gamma_{11} \dots \gamma_{31}$ and $\gamma_{12} \dots \gamma_{32}$) indicate cross-level interactions between the regional and firm-level predictors. In other words, the full "slopes-as-outcomes" model examines not only whether the regional variables influence the likelihood of firm's to innovate, but also whether these regional factors affect the level-1 relationships.

Innovation may be less elastic to firm's size in regions with stronger innovation systems. Also the environment may influence differently new compared to established firms. For example, internal R&D department is a costly venture and may be out of reach for many small or new firms. More opportunities to tap into external sources of knowledge, such as proximity to a technical university, help to at least partly overcome the "disadvantages of smallness and newness" (Almeida, et al. 2003 and Rodriguez-Pose, Refolo, 2003). It is also reasonable to hypothesize that some regional factors may be more relevant for foreign affiliates than domestic-owned firms or vice-a-versa, whereas other factors may affect all firms regardless of ownership.

A look at results of the full model in the last column in Table 2 shows that a significant cross-level interaction between the regional factors and the size of firms has been detected. As expected the interaction term between RIS and SIZE has a negative sign, which indicates that it is easier to innovate for small firms located in a well-developed innovation system. A small firm with vibrant opportunities to benefit from localized learning is predicted to be more likely to innovate compared to a firm of the same size located in an underdeveloped regional innovation system. All else equal to average, a firm with 10 employees is estimated to be more than two times more likely to innovate in Prague where is the best RIS than in the Jeseník region with the worst RIS score.

At the first glance, it might seem surprising; however, that also the STR score comes out with a significantly negative interaction term with size. Innovation in large firms appears to be more severely hindered by the structural problems in the region. After the initial rush for reforms, the substantial social problems in these regions have often slowed down privatization and downsizing of the remaining large firms. Smaller firms that emerged from the restructuring became more likely to innovate, while the large firms that have not yet finished or not even seriously started restructuring came out less innovative at the end of the nineties. Since concentration of large firms in the "old" industries is part of the trouble in these regions, it is actually reassuring to find that these variables are intertwined in the same direction. Also small firms might be in a better position to capitalize on the appetite for change in the society that these problems open. Structural changes driven by innovation in small firms may actually hallmark the way forward in these ailing regions.

Any other statistically significant cross-level interactions have not been detected. Although foreign ownership directly matters for the firm's propensity to innovate, there is no evidence that foreign affiliates benefit more (or less) from location in a strong regional innovation system and neither are they differently affected by the structural problems compared to the domestic-owned firms. Similarly the effect of age on the propensity of firms to innovate does not seem to be influenced by the regional characteristics, which confirms the previous conclusion on little differences along these lines across regions in the Czech context.

A comparison with the previous results reveals that the effect of STR on the intercept became much smaller and less significant after the cross-level interaction terms have been included, which suggests that this effect is more relevant for firms with certain characteristics, rather than for the "average" firm. Also the effects of RIS and STR on the intercept came out with very similar magnitude in the full model, although the former remains highly significant, which provides further support to the role of regional innovation systems. Not much has changed in the effects of the level-1 predictors.

A look at the random part of the model confirms that the cross-level effects significantly improved the explanatory power of the model. The variance components for SIZE and FOREIGN decreased substantially to values very close to zero. Since the cross-level interaction terms explained most of the un-modeled variability in the effect of SIZE, the corresponding random effect is no longer statistically significant. The variance component for AGE has actually increased, but it remains highly insignificant. Overall, the very low and statistically insignificant variance components suggest that we might not much improve explanatory power of the model by adding more predictors.

Another diagnostic measure of multilevel models that has not been discussed yet is the so-called index of dispersion. Although logit multilevel models do not have a separate term for the level-1 error, we can calculate a level-1 error variance scaling factor that measures the extent to which the observed errors follow the theoretical binomial error distribution (Luke 2004, pg. 57). Index of dispersion equal to 1 indicates perfect fit between the observed errors and the theoretical assumptions. A significant over- or under-dispersion indicates model misspecification, the presence of outliers or the exclusion of an important level in the model. Less than 5% dispersion is usually seen as satisfactory. The index of dispersion is very close to unity, which confirms that the estimates do not suffer from a major problem.[7]

5. Conclusions

Many researchers have evaluated the link between quality of regional innovation systems and firm's innovativeness, though to the best of my knowledge, none has directly confirmed the relationship on the basis of a formal multilevel analysis so far, as we do in this paper. Using micro data in a multilevel framework, we found that firm's characteristics are important for innovation, but show that geography matters a lot too. Size, age and ownership of firms influence their odds to innovate, so as do benefits from location in a well-functioning regional innovation system. Also the effect of firm's size on innovation is intertwined with the regional factors.

An important implication of the paper is that analyses of innovation should properly account for the hierarchical nature of the micro data. As expected we found strong evidence that the observations are not independent from each other at the regional level. Because so much on innovation is multilevel, we should use analytic methods that are also multilevel. If we do not do it, we may keep missing important part of the picture. Of course, firms can be grouped not only spatially by regions or countries, but also by sectors. Structure of a multilevel model may be more complicated if we wish to include more levels of the hierarchy. Also 3-level models with firms in regions within countries or so-called cross-classified models with firms in sectors and simultaneously in regions (or countries) are feasible.

As more micro data become available for research on innovation, there arises a controversy about the appropriate "unit of analysis" for testing various kinds of hypotheses in the literature. As already noted, the typical approach is to ignore hierarchical data clustering by using only the micro data (and control for sectoral and spatial dummies) or either to conduct the analysis with data only at a higher hierarchical level, such as cross-country regressions. Although there are many relevant hypotheses that are within any of the levels of analysis, there is a host of issues that require a look at relations between the various levels (Feldman, 2000; Boggs and Rantisi, 2003; Overman, 2004 and Lagendijk 2006). Arguably the "unit of analysis" issue might be elegantly resolved, at least in empirical research, by explicit multilevel modeling that would use micro data to study the interaction between firms and their surroundings, such as sectoral, regional and national innovation systems.

It should be mentioned, however, that a major reason why multilevel modeling has not been widely applied in the empirical research on innovation so far is high demand

[7] It should be also mentioned that we have probed more deeply into robustness of the results with regard to a possible effect of outliers and the regional classification. For example, we have estimated the model without firms located in Prague and/or other city-districts (Brno, Ostrava, Plzen) and experimented with regional classification into only 30 regions (and therefore with higher number of level-1 units per region). The presented results are robust to these changes in the sample specification.

on scope and quality of data. To properly estimate a multilevel model, we need harmonized micro data for at least thirty higher level units. Each unit at the higher levels also should have a reasonable number of observations to allow for making meaningful inferences. Ideally we should have at least thirty observations within each higher-level unit, which is especially difficult to achieve in the more complicated models.[8] A lack of such composition of data prevented us from estimating the cross-classified model with both regions and sectors at the higher level in this paper. Furthermore, models with more than two levels of the analysis are naturally more demanding on computational power, which even the state-of-the-art personal computers may not be able to provide yet.

Multilevel models can be used to identify higher-level units with extreme relations between the micro and macro characteristics. For instance, we may identify a region where the relation between foreign ownership of firms interacts quite atypically with attributes of the local innovation system, even when adjusted for other relevant factors in the estimate. Such an "unexplained" relationship can be then selected for a more detailed scrutiny in a qualitative case study research that would dive into issues beyond reach of the econometric estimate. And in turn these findings can influence the future multilevel modeling, so forging the much needed synergy between quantitative and qualitative research on geography of innovation that often remains unexploited due to the lack of the former in the literature (Lorenzen, 2005 and Malmberg and Maskell, 2002).

At last but not least, policy makers should understand and utilize the multilevel perspective if they are to be successful at promoting innovativeness of firms. It comes out from the analysis that smaller firms benefit most from location in a vibrant innovation system. A reasonable strategy to catalyze innovation particularly in small and medium-size firms therefore seems to be to improve the regional innovation system. Although by design of survey the CIS data do not cover micro firms with less than 10 employees, there are good reasons to believe that the regional innovation system is even more important for innovative entrepreneurship. No doubt we need to improve our understanding of the interdependence among different levels of analysis to design more complex and comprehensive innovation policies.

[8] Boschma and Weterings (2005, 576-577) considered using a multilevel model to analyse innovative productivity of software firms in the Netherlands, but concluded that the differences between regions were not statistically significant enough to justify moving beyond a single-level model. Arguably they might have been prevented from using multilevel analysis by a very low number of observations per region (about 169 firms divided into 40 regions) in their dataset.

References

Almeida, P., Dokko, G., Rosenkopf, L. (2003), Startup size and the mechanisms of external learning: increasing opportunity and decreasing ability?, *Research Policy*, 32: 301–315.

Asheim, B.T., Isaksen, A. (1997), Location, agglomeration and innovation: Towards regional innovation systems in Norway?, *European Planning Studies*, 5: 299-330.

Asheim, B.T., Gertler, M. (2004), Regional systems of innovation, in: Fagerberg, J., Mowery, D.C., Nelson, R. (Eds), *The Oxford Handbook of Innovation*, Oxford: Oxford University Press, 291-317.

Basilevsky, A. (1994) *Statistical Factor Analysis and Related Methods: Theory and Applications*. London: John Wiley & Sons Inc.

Boschma, R.A., Weterings, B.R. (2005), The effect of regional differences on the performance of software firms in the Netherlands, *Journal of Economic Geography*, 5: 567-588.

Boggs, J.S., Rantisi, N.M. (2003), The "relational" turn in economic geography. *Journal of Economic Geography*, 3: 109-116.

Cooke, P. (1992), Regional innovation systems: Competitive regulation in the New Europe, *Geoforum*, 23: 365-382.

Cook, P. (2004) Evolution of regional innovation systems: emergence, theory, challenge for action, in: Cooke, P., Heidenreich, M., Braczyk, H-J (Eds), *Regional Innovation Systems*, London: Routledge, 1–18.

Doloreux, D., Parto, S. (2005) Regional innovation systems: Current discourse and unresolved issues, *Technology in Society*, 27: 133-153.

Edquist, C. (1997) *Systems of Innovation: Technologies, Institutions and Organizations*. London: Pinter.

Fagerberg, J., Srholec, M. (2006) Why some countries develop (while other stay poor): The role of "capabilities" in development. Proceedings from the 4th Globelics Conference on "Innovation Systems for Competitiveness and Shared Prosperity in Developing Countries", Trivandrum, India.

Feldman, M.P. (2000) Location and innovation: the new economic geography of innovation, spillovers and agglomeration, in: Clark, G.L., Feldman, M.P., Gertler, M.S. (Eds), *The Oxford Handbook of Economic Geography*, Oxford: Oxford University Press, 373-394.

Goldstein, H. (2003) *Multilevel Statistical Models*. London: Arnold.

Hox, J. (2002) *Multilevel Analysis*. Mahwah (New Jersey): Lawrence Erlbaum Associates, Inc.

Knell, M., Srholec, M. (2004) Innovation Cooperation and Foreign Ownership in the Czech Republic. Proceedings from the 30th European International Business Academy (EIBA) Annual Conference, Ljubljana, Slovenia.

Lagendijk, A. (2006) Learning form Conceptual Flow in Regional Studies: Framing Present Debates, Unbracketing Past Debates, *Regional Studies*, 40: 385-399.

Lorenzen, M. (2005), Knowledge and Geography, *Industry and Innovation*, 12: 399-407.

Luke, A. (2004) *Multilevel Modeling*. London: Sage Publications.

Lundvall, B.Å. (1992) *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter Publishers.

Malerba, F., Orsenigo, L. (1995), Schumpeterian patterns of innovation, *Cambridge Journal of Economics*, 19: 47–65.

Malerba, F., Orsenigo, L. (1997) Technological regimes and sectoral patterns of innovative activities, *Industrial and Corporate Change*, 6: 83–117.

Malmberg, A., Maskell, P. (2002), The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering, *Environment and Planning A*, 34: 429-449.

Morgan, K. (1997), The learning region: institutions, innovation and regional renewal, *Regional Studies*, 31: 491–504.

Nelson, R., (1993) *National Innovation Systems: A Comparative Analysis*. New York: Oxford University Press.

OECD (1997) Oslo Manual. OECD: Paris.

Overman, H.G. (2004), Can we learn anything from economic geography proper?, *Journal of Economic Geography*, 4: 501-516.

Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R., du Toit, M. (2004) *HLM 6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood: Scientific Software International.

Rodriguez-Pose, A., Refolo, M. (2003), The link between local production systems and public and university research in Italy, *Environment and Planning A*, 35: 1477-1492.

Srholec, M. (2005) Innovation Strategies of Multinationals: Firm-level evidence from foreign affiliates in the Czech Republic. Proceedings from the 31th European International Business Academy (EIBA) Annual Conference, Oslo, Norway.

Veugelers R., Cassiman, B. (2004), Foreign subsidiaries as a channel of international technology diffusion: Some direct firm level evidence from Belgium, *European Economic Review,* 48: 455-476.