Received: 11 September 2017 | Revised: 27 February 2018 | Accepted: 5 March 2018
DOI: 10.1111/pirs.12366

FULL ARTICLE

Location decisions of enterprise R&D investments as a function of related and unrelated regional industry structures: A multilevel study

Jarle Aarstad | **Olav Andreas Kvitastein** | **Stig-Erik Jakobsen**

Western Norway University of Applied Sciences, Bergen, Norway

Correspondence
Jarle Aarstad, Western Norway University of Applied Sciences, PO Box 7030, NO-5020 Bergen, Norway.
Email: jarle.aarstad@hvl.no

Funding information
Research Council of Norway

JEL Classification: D22; O18; O32; R12; R58

Abstract

Despite that previous studies have examined factors that affect location decisions of enterprise R&D investments, they have not investigated if regional industry structures play a role. Responding to this research gap, we analyse data from Norway and find that location in regions with unrelated, diversified, and fragmented industry structures increases both the probability and amount of enterprise R&D investments. Location in regions with related and complementary industry structures, on the contrary, has no effect. We further find that location in populous regions spanning a large geographical area increases the probability of enterprise R&D investments.

KEYWORDS

location decisions, R&D investments, regional industry structures, related and unrelated variety

1 | INTRODUCTION

Over the past decades, academic research has paid increasing attention to factors affecting location decisions of enterprise R&D investments. Studies have, for instance, examined how regional agglomeration (e.g., Belderbos, Roy, Leten, & Thijs, 2014; Siedschlag, Smith, Turcu, & Zhang, 2013; Smith, Broberg, & Overgaard, 2002; Yang & Hayakawa, 2015) and dependence on regional markets (Odagiri & Yasuda, 1996) play a role. Other studies have shown how regional knowledge infrastructure, national talent pool, and political risk explain location decisions of R&D investments (Demirbag & Glaister, 2010), and Alcácer (2006) found that the location of R&D infrastructure is a function of regional concentration of similar activities. Yet, despite that numerous studies have examined location decisions of enterprise R&D investments, previous research has not explicitly examined if regional industry structures play a role.

Responding to this research gap, we study in this paper if location decisions of enterprise R&D investments are a function of related or unrelated regional industry structures. Related regional industry structures are complementary having an overlapping knowledge base, whereas unrelated structures are fragmented with limited potential for complementarity and knowledge overlap (Frenken, Van Oort, & Verburg, 2007). Scholars have argued, in particular within the field of economic geography, that regional industry structures play a crucial role as carriers of innovation performance and productivity in the local economy (e.g., Aarstad, Kvitastein, & Jakobsen, 2016a). It is therefore surprising that location decisions of enterprise R&D investments have not received explicit attention within this broad field of research. We thus argue that our study adds an important contribution to research streams examining location decisions of enterprise R&D investments and characteristics of regional industry structures.

In addition to be of academic and scholarly interest, we furthermore argue that our study has relevance for practitioners and industry stakeholders. Analysing factors that affect location decisions of enterprise R&D investments can increase our knowledge about strategic management behaviour and explain where crucial resources are channelled. R&D investments moreover create employment opportunities for a highly skilled workforce, which can have further local spillover effects. Specifically, research has shown that R&D investments are a strong driver of productivity growth and innovation performance in the regional economy (e.g., Vogel, 2015). Overall, there is a large volume of research emphasising that innovation performance is a function of R&D investments (e.g., Bhattacharya & Bloch, 2004; Britton, 2003; Moreno, Paci, & Usai, 2005; Shefer & Frenkel, 2005; Sun, 2002; Tsai, 2001; Yang & Lin, 2012). Gaining knowledge about how regional industry structures may affect location decisions of enterprise R&D investments is accordingly crucial for both managers, policy-makers, and other stakeholders, we argue.

This is a multilevel study where location decisions of enterprise R&D investments, as a lower-level unit, is a function of regional industry structures, as a higher-level unit. The scholarly literature on multilevel research was pioneered by Lawrence Jr. and Iversen (1979), and the approach has since received much attention among researchers, both theoretically, conceptually and methodologically (e.g., Hundt & Sternberg, 2016; Klein, Danserau, & Hall, 1994; Leslie, 2017; Raudenbush & Bryk, 2002; Rousseau, 1985). While most social science studies normally focus at one particular level of analysis, for example, enterprise level or regional level, respectively, multilevel studies, on the other hand, empower the ability to assess how lower-level units respond, behave or act as a function of the context, that is, characteristics of higher-level units, in which they are embedded. In our econometric modelling, we follow Statistics Norway's division of the country into 89 economic regions, which corresponds to the EU's classification of local administrative units (LAU-1). Other studies follow a similar division of economic regions (e.g., Kowalski & Marcinkowski, 2014; Ogdul, 2010; Ottaviano & Pinelli, 2006). In the multilevel analyses, the industry structure is a constant for all enterprises that operate within a particular economic region, while it varies between them.

Below, we first review the concepts of related and unrelated variety, and next, we elaborate and argue how these two dimensions of regional industry structures may affect location decisions of enterprise R&D investments. In the following sections, we test our theoretical arguments empirically by analysing a large-scale dataset from the 2010 Norwegian Community Innovation Survey (CIS). In the final section, we discuss the empirical findings, address the study's limitations, and suggest avenues for future research.

2 | THEORY

2.1 | Regional industry structures and the concepts of related and unrelated variety – a condensed literature review

Marshall (1890) was a pioneer in the study of regional industry structures as a potential vehicle for local development. He emphasized that regional industry specialization is optimal for value creation because it induces economies of scale and efficient local economic transactions. Jacobs (1969) has later challenged Marshall's view and argued that industry diversity is preferable because it provides variety and diversification enabling knowledge externalities that can result in innovative products and services in the marketplace. In the scholarly literature, both the perspectives of Marshal and Jacobs have received much attention, yet, there is no consensus as to whether regional industry specialization or diversification is preferable for the local economy (see for instance Beaudry & Schiffauerova, 2009; Groot, Poot, & Smit, 2016; Glaeser, Kallal, Scheinkman, & Shleifer, 1992; Porter, 2000).

Frenken et al. (2007) have, in our opinion, played a crucial role in the debate concerning regional industry structures by emphasizing that diversification is not a one-dimensional, but a two-dimensional construct labelled related and unrelated variety or diversity. Following Frenken et al., we have defined related variety as regional industry structures having complementary and overlapping knowledge base, while unrelated variety means fragmented structures with limited potential for complementarity and knowledge overlap.

Aarstad et al. (2016a) found that a regional industry of related variety increased the probability of enterprise product innovation, and other publications report similar findings (Castaldi, Frenken, & Los, 2015; Tavassoli & Carbonara, 2014). A likely explanation of positive innovation effects from related variety is the possibility to recombine resources from different complementary and overlapping perspectives (Burt, 1992, 2004). Boschma, Minondo, and Navarro (2012) and Frenken et al. (2007) moreover showed that related variety increased economic growth in Spanish and Dutch regions, respectively, and in a recent review, Content and Frenken (2016) likewise concluded that related variety has an overall positive performance effect.

An unrelated and fragmented industry structure (Frenken et al., 2007), on the other hand, may preclude the probability of recombining resources from different perspectives, due to limited complementarity and knowledge overlap in the regional economy. Castaldi et al. (2015) nevertheless showed that unrelatedness could induce radical or disruptive innovations, probably because of the possibility to recombine very different ideas from very different perspectives. In a similar vein, Frenken et al. (2007) showed positive effects from unrelated regional industry structures by inducing regional robustness to increases in unemployment. Aarstad et al. (2016a), on the other hand, found that unrelated variety had a negative effect on enterprise productivity, probably because of the preclusion of local competition (cf. Porter, 2000) and the hampering of economies of scale caused by fragmented regional industry structures (cf. Boschma, 2005).

Taken together, regions with related variety can increase innovation performance (Aarstad et al., 2016a; Castaldi et al., 2015; Tavassoli & Carbonara, 2014) and economic growth (Boschma et al., 2012; Content & Frenken, 2016). Regions with unrelated variety are robust to increases in unemployment (Frenken et al., 2007), can be prone to develop disruptive innovations (Castaldi et al., 2015), but have a negative effect on enterprise productivity (Aarstad et al., 2016a).

2.2 | Location decisions of enterprise R&D investments as a function of related and unrelated variety

Despite numerous contributions having examined regional industry structures in general, and recently also related and unrelated structures in particular, they have not studied location decisions of enterprise R&D investments. We, therefore, lack a comprehensive theoretical or empirical framework that can partake to explain the potential effect of related and unrelated variety in this regard. Other studies that, indeed, have examined location decisions of R&D investments (e.g., Alcácer, 2006; Belderbos et al., 2014; Demirbag & Glaister, 2010; Odagiri & Yasuda, 1996; Siedschlag et al., 2013; Smith et al., 2002; Yang & Hayakawa, 2015), are similarly silent about the role played by regional industry structures. In the following, we will nonetheless aim to address a few arguments that may shed some conceptual light on the research question that we address in this paper.

We have referred to studies finding that regional industry structures of related variety increase innovation performance. In other words, such regional industry structures represent a benign context for the development and introduction of new products or services at the marketplace. This may be attractive for enterprises in their pursuits of gaining maximum innovative performance from their R&D investments. We have also referred to Castaldi et al. (2015) showing that unrelatedness induced radical or disruptive innovations, probably because of the possibility to recombine very different ideas from very different perspectives. In other words, we cannot rule out that a regional industry structure of unrelated variety may also appear to be a benign location for enterprise R&D investments. Having said this, we do acknowledge that innovation in regions with unrelated variety may be a function of accidental events recombining knowledge from very different perspectives instead of a systematic approach to developing new products. In other words, finding that unrelatedness can induce radical innovation may not *per se* affect location decisions of enterprise R&D investments. We, therefore, conclude so far that both regional industry structures of related and unrelated variety may affect location decisions enterprise R&D investments, yet we assume that related variety has a stronger effect than unrelated variety.

Issues other than potential innovation performance can further induce location decisions of enterprise R&D investments in regions with related or unrelated industry structures. Regional population size (in number of inhabitants) is associated with both related and unrelated variety (Aarstad et al., 2016a), but populated regions also affect location decisions of enterprise R&D investments, due to proximity to research-based infrastructure, low transportation costs and abundant demand and supply of factor inputs (cf. Belderbos et al., 2014; Krugman, 1991; Siedschlag et al., 2013; Smith et al., 2002; Yang & Hayakawa, 2015). It is therefore not far-fetched to assume that R&D investments in populous regions may spill over and be partly absorbed by either related or unrelated industry structures. Institutional forces can furthermore cement decisions of R&D investments in regions of related or unrelated variety, independent of whether such decisions are genuinely rational or not (please see Scott, 2014 for a review and explanation of institutional carriers of organizational behaviour).

Political forces may also affect the channelling of enterprise R&D investments into certain regional industry structures. In Norway, parts of enterprises' R&D funding are often publicly financed, either through direct grants or as government provided loans, and certain R&D investments are furthermore eligible for tax deductions. Location decisions of enterprise R&D investments may, therefore, deviate when it comes to being distributed into different regional industry structures depending on whether or not they are privately or publicly funded. Aarstad et al. (2016a) found that unrelated regional industry structures hamper enterprise productivity, we have noted, and which may legitimise a political rationale to channel R&D investments into such regions in an aim to offset their inferior performance.

Taken together, in this section we have theorised in what ways regional industry structures of related and unrelated variety may affect location decisions of enterprise R&D investments. Below, we elaborate the empirical context and the methodology that we apply to test the theoretical arguments.

3 | METHODOLOGY

3.1 | Research context and data

To study our research question, we analyse data from the Norwegian part of the 2010 Community Innovation Survey, collected by Statistics Norway in collaboration with Eurostat. The Community Innovation Survey is comprehensive gathering data from enterprises in most industries from all parts of the country. The survey, moreover, includes data from small-sized enterprises, with a minimum of five employees, to large-sized enterprises. The initial part of the 2010 Community Innovation Survey requests enterprises to report data on R&D investments, which partakes to provide valid data on this critical variable. Data from the Community Innovation Survey have moreover been applied in numerous academic studies in many different national contexts (e.g., Dautel & Walther, 2014; Smit, Abreu, & de Groot, 2015; Wixe, 2016).

Participation in the Norwegian part of the 2010 Community Innovation Survey was mandatory for selected enterprises, which minimises non-responding bias in the data (Armstrong & Overton, 1977). The survey includes all enterprises in Norway with at least 50 employees, and a selected sample of smaller enterprises (Wilhelmsen & Foyn, 2012, explain in detail the sampling procedure for the survey). The sample includes about one-third of all enterprises in Norway employing about two-thirds of the work stock in the private sector in the country.

3.2 | Dependent variable and enterprise level control variables

We model the dependent variable as enterprise R&D investments measured in 1,000 NOK (Norwegian Kroner) in 2010. We control for whether an enterprise is multidivisional or not (dummy) and the size of the number of employees. Some industries may invest more in R&D than others, so we control for random effects of enterprises operating (or being nested) in particular industries. In fact, we control for random effects of enterprises operating (or being nested) in particular industries that are located (or nested in) particular regions. Below, we further explain this issue.

3.3 | Independent and control variables at region level

We have noted that region variables are constant for all enterprises located in a particular region and vary between them. Statistics Norway divides the country into 89 economic regions, and we have also noted that the division corresponds to the European Union's grouping of local administrative units at level one (LAU-1). The 2010 Community Innovation Survey identifies the regional location of each enterprise for the study.

Related and unrelated variety of the regional industry structure (at a regional level) are independent variables, and we use Shannon's (1948a, 1948b) entropy measure of surveyed enterprises' nomenclature of economic activities (NACE) codes to model the concepts. In total, we use NACE codes from 6,589 surveyed enterprises located within each economic region, respectively, to model related and unrelated variety (at a regional level).

"To examine empirically the effect of related or unrelated variety [at a regional level] is not a trivial matter…," "[yet the] main advantage of the entropy measure… is that entropy can be decomposed at each sectoral digit [or NACE code] level," according to Frenken et al. (2007, p. 689). To model unrelated variety at a regional level, we use NACE codes at level two, which identifies enterprises operating within a particular industry division (e.g., the manufacture of textiles, food products, furniture, and computers, respectively) in a particular region. Level two is a relatively crude measure to distinguish between enterprises that operate in industries that differ substantially from each other. It, therefore, captures regional unrelated variety or unrelated diversity, we argue, and Frenken et al. (2007) have suggested a similar approach. Formally, we define the entropy concept of regional unrelated variety as follows:

$$\text{Regional unrelated variety (UV)} = \sum_{k=1}^{p} s_{kr} \ln\left(\frac{1}{s_{kr}}\right),$$

in which s_{kr} is the share of enterprises in class k (NACE code at level two) in region r. If $s_{kr} = 0$, it means that $\ln(1/s_{kr}) = 0$. The p is the number of observed NACE codes at level two.[1]

---

[1] Frenken et al. use binary or base 2 logarithm ($\log_2$) while we use natural or base e logarithm (ln). The relationship between $\log_2$ and ln is as follows: $\log_2 X = (1/\ln 2)^* \ln X \approx 1.44^* \ln X$, i.e. a constant of about 1.44, which does not alter any statistical conclusion. Frenken et al. moreover use the share of employment within each NACE code, but we argue that such an approach discriminates against the prevalence of small and medium sized enterprises in the regional economy, namely, an enterprise with 1,000 employees will contribute as much to (and deflate) the entropy measure as will 100 enterprises with 10 employees each that are located in the same region. In particular, this is critical in a Norwegian context, due to the relatively strong presence of small and medium sized enterprises in the regional economy. However, we do acknowledge that, anything else being equal, enterprises with many employees should play a larger role than enterprises with few employees when modelling concepts of regional industry structures, and the sampling procedure for the CIS 2010 survey takes account of this issue: all enterprises with 50 or more employees were sampled, we have noted, 43% of enterprises with 20–49 employees were randomly sampled, 25% of enterprises with 10–19 enterprises were randomly sampled, and 19% of enterprises with 5–9 employees were randomly sampled (for details, see Wilhelmsen & Foyn, 2012, p. 37). In other words, larger enterprises have a larger share in the sample than smaller enterprises have. Aarstad et al. (2016a) use the same measures of regional industry structures as is used in this study. Fitjar and Timmermans (2017) have moreover suggested refined measures of regional industry structures and that differ from the Frenken et al. methodology.

To model related variety at a regional level, we first use NACE codes at level five, which identifies the finest grained distinction of industry classes, and carry out similar entropy measures as described above (except that we apply NACE codes at level five instead of at level two). Next, we subtract unrelated variety from the equation (to “take out” this effect). Formally, we define the entropy concept of related variety as follows:

$$\text{Regional related variety (RV)} = \sum_{m=1}^{P} l_{mr} \ln\left(\frac{1}{l_{mr}}\right) - \text{UV},$$

in which l_{mr} is the share of enterprises in class m (NACE code at level five) in region r. If $l_{mr} = 0$, it means that $\ln(1/l_{mr}) = 0$. The P is the number of observed NACE codes at level five. Frenken et al. (2007) follow a similar procedure to model related variety (but please see footnote 1, which informs about how we have modified the measure).

Related and unrelated variety embody the total regional industry variety or diversity (TV) and is, conversely, decomposed into the two components of related variety (RV) and unrelated variety (UV); TV = RV + UV (also see Aarstad, Kvitastein, & Jakobsen, 2016b for further explanations). Thus, low values of related and unrelated variety depict regional industry specialization.

We control for each region's population size in number of inhabitants (multiplied by 1,000) as of the first quarter of 2008 and each region's geographical area in square kilometres. Populated regions, we have argued, can attract enterprise R&D investments because of abundant supply and demand factors, access to research-based infrastructure, and low transportation costs. Geographically large regions decrease, anything else being equal, geographical proximity, which may decrease the overall efficiency of the regional economy (Boschma, 2005). As a consequence, this may induce a disincentive for enterprises to invest in R&D in geographically large regions. To take further account of regional heterogeneity in our data, we also model an interaction term between regional population size and regional geographical size. Some populous regions in central parts of Eastern Norway are relatively small in geographical size, while some populous regions in Western Norway, Mid-Norway, and partly also Southern Norway, are relatively large in geographical size. Populous, but geographically small regions in central parts of Eastern Norway, have a relatively large service sector, while populous but geographically large regions Western Norway, Mid-Norway, and Southern Norway, have a relatively large manufacturing industry sector. We, therefore, argue that modelling an interaction term between regional population size and regional geographical size may be a relevant proxy parameter to take account of possible heterogeneity in our data, and which NACE codes may not take full account of. Concerning regions with relatively low population, some span a relatively large geographical area, in particular in Northern Norway and North-East Norway, while others span a relatively small geographical area. The latter group is mainly located in central parts of Eastern Norway and are in geographical proximity to other more populous regions. This combination of regions with relatively low population, but differing in geographical size, further legitimises the modelling of an interaction term between regional population size and regional geographical area, we argue. Statistics Norway provided the data on regions' population size and geographical area.

3.4 | Econometric modelling

In the econometric modelling, we apply a random-intercept mixed-effects (or hierarchical) model with three nested levels. Level-one units are enterprises, level-two units are industries in which enterprises are nested (NACE codes at level two), and level-three units are regions in which industries are nested. The model is displayed as follows:

$$Y_{ikr} = +\beta_{0kr} + \beta_{0r} + \sum_{h=1}^{s} \beta_h x_{hikr} + R_{ikr}, \tag{1}$$

where Y_{ikr} is the dependent variable – enterprise R&D investments – for (level-one unit) enterprise i, operating or nested in (level-two unit) industry k, and located or nested in (level-three) region r. The β_{0kr} indicates that the intercept varies between industries nested in regions, and the β_{0r} indicates that the intercept varies between regions.

The x_{hikr} are independent or explanatory variables, and the R_{ikr} are enterprise (level-one) residuals (or error terms). Assuming that the β_{0kr} and β_{0r} follow a normal distribution with variance τ_0, we have:

$$\beta_{0r} + \beta_{0kr} = \beta_0 + U_{0kr} + V_{0r} \tag{2}$$

where the $U_{0kr} \sim N(0,\tau_0)$ and $V_{0r} \sim N(0,\tau_0)$. Equation 2 accounts for the enterprise variation within industries nested in regions (U_{0kr}) and enterprise variation within regions (V_{0r}). Substituting Equation 2 into Equation 1, we get:

$$Y_{ikr} = \beta_0 + \sum_{h=1}^{s} \beta_h x_{hikr} + R_{ikr} + U_{0kr} + V_{0r}, \tag{3}$$

where the β_h are fixed effects (or "ordinary") regression coefficients, whereas the R_{ikr}, U_{0kr}, and V_{0r} are random effects (for further readings on mixed-effects or hierarchical multilevel models, see for instance Raudenbush & Bryk, 2002; Snijders, 2011).

Due to many enterprises with zero R&D investments, the dependent variable is left-censored. We, therefore, primarily use multilevel mixed-effects tobit regression, which is now possible to estimate in Stata 15 (StataCorp., 2017).[2] Tobit regression with censored data was pioneered by Tobin (1958) and has received much attention in the econometric literature (e.g., Cameron & Trivedi, 2010; Wooldridge, 2010). In a left-censored multilevel mixed-effects tobit model with three nested levels and zero as the lowest (left-censored) value, the observed dependent variable, Y_{ikr}, pertains to the partly unobserved latent dependent variable, Y^*_{ikr}, as follows:

$$Y_{ikr} = \begin{cases} Y^*_{ikr} & \text{if } Y^*_{ikr} > 0 \\ 0 & \text{if } Y^*_{ikr} \leq 0. \end{cases}$$

We can assume an enterprise has a latent (unobserved) need or demand for R&D investments, labelled as Y^*_{ikr}, but which is only observable for positive values. I.e., an enterprise with zero R&D investments may or may not have a latent (unobserved) need or demand for R&D investments, but without any positive observation whatsoever (cf. a person with a strong need or demand to buy a house, but unable to purchase, is as much having zero purchase as another person with a low or absent need or demand). (Please see for instance Cameron & Trivedi, 2010.)

4 | RESULTS

4.1 | Log-transforming and descriptive statistics

We log-transform the variables using the natural logarithm (except for the binary variable indicating whether an enterprise is multidivisional or not). For the dependent variable, enterprise R&D investments, we add the constant one before log-transforming it, due to the inclusion of zero values. One region has zero related variety, and for this variable also, we add the constant one before log-transforming it. To model the interaction between regional population size and regional geographical size, we follow Cronbach's (1987) recommendation to mean-centre the log-transformed parameters before multiplying them.

Table 1 reports descriptive statistics and correlations between the variables for the study. Absolute values of skewness and kurtosis do not exceed critical values (since multidivisional enterprise is binary, we do not report skewness and kurtosis for this variable). We observe that some regional level variables, in particular between related variety and regional population size, correlate strongly, and which may cause challenges in the econometric modelling concerning multicollinearity (see for instance O'Brien, 2007). Below, we explain how we deal with this issue.

---

[2] In the Results section, we also report some multilevel mixed-effects logit and linear regressions. Logit regressions distinguish between enterprises that do and do not invest in R&D, while linear regressions measure the amount invested (and therefore only include enterprises with positive R&D investments). In other words, logit models estimate the probability of enterprise R&D investment and linear models estimate the amount invested (among those who actually do invest in R&D).

TABLE 1 Descriptive statistics and correlations

Min.	Max.	Mean	SD	Skew.	Kurt.		1	2	3	4	5	6	7
0	13.8	1.83	3.44	1.45	0.354	Enterprise R&D investments (1)							
0	1	0.171	0.376			Multidivisional enterprise (2)	0.081						
1.61	9.77	3.48	1.24	0.688	0.553	Size/number of employees (3)	0.196	0.538					
8.57	13.2	11.7	1.21	−0.416	−0.783	Reg. pop. size (POP) (4)	0.069	0.096	0.155				
5.44	9.85	7.39	1.05	−0.143	−1.19	Reg. geogr. Size (REG) (5)	−0.053	−0.050	−0.086	−0.368			
−4.27	4.31	−0.469	1.25	−0.164	0.321	POP*REG (6)	0.028	−0.041	−0.008	−0.106	0.228		
0	0.921	0.630	0.245	−0.555	−0.821	Related variety (7)	0.069	0.092	0.145	0.934	−0.474	−0.064	
0.677	1.28	1.15	0.119	−1.39	1.47	Unrelated variety	0.079	0.066	0.121	0.783	−0.020	0.125	0.641

Notes: N = 6,584. All correlation coefficients with absolute values larger than .020 are significant at the 5% level (two-tailed tests of significance). All variables, except for multidivisional enterprise (as a binary variable) are log-transformed using natural logarithm.

4.2 | Multilevel mixed-effects tobit regressions

In Table 2, we present the results of the multilevel mixed-effects tobit regressions with enterprise R&D investments as the dependent variable. Model 1 reports, for illustrative purposes, a multilevel regression excluding independent variables (i.e., $s = 0$, according to Equation 1 and 3). We observe that the parameters for the level-two random effect (enterprises nested in industries) and level-three random effect (industries nested in regions) are considerably higher than their standard errors. In particular, this is the case for the level-two random effect (enterprises nested in industries).[3] The findings are indicative of nested or clustered random effects at the aforementioned levels. The log likelihood ratio (LR) test assesses whether the mixed-effects multilevel tobit model has better model fit than a one-level tobit model excluding the level-two and level-three random effects, and a strongly significant χ^2 shows that this is the case.

Model 2 includes the control variables for this study only. We find that multidivisional enterprises tend to have relatively low R&D investments (yet after controlling for enterprise size in number of employees), while large enterprises regarding the number of employees (unsurprisingly) tend to have relatively high R&D investments.[4] The effects of multidivisional enterprises and enterprise size on R&D investments are consistent with the following models that we report. Location in populous regions has a borderline significant, positive effect on enterprise R&D investments, while the effect of regional geographical size is non-significant. Significant, positive interaction effect between regional population size and regional geographical size implies that location in populous regions spanning a large geographical area has a positive effect enterprise R&D investments. The interaction effect is consistent in the following models where included. The level-three random effect (industries nested in regions) is reduced in model 2, as compared to model 1. It indicates that some of the effects of particular R&D intensive industries located or nested in particular regions are absorbed by the control variables (please also see footnote 3). The level-two random effect (enterprises nested in industries) is relatively unaffected. An LR test reporting a significant χ^2 shows that the mixed-effects multilevel tobit model (still) has better model fit than a one-level tobit model excluding the random effects. We furthermore observe in model 2 that a significant Wald χ^2 is indicative of robust model fit as compared to model 1, which excludes control variables. The Wald χ^2 is furthermore significant and consistent with the following models that we report. The random effects and the LR comparing one-level vs multilevel tobit regressions we observe in model 2 are also consistent with the following models that we report.

Model 3 includes related variety as an independent variable, and the parameter receives negative but non-significant support. The LR test (reported at the bottom of Table 2), comparing model fit between model 3 and model 2, likewise reveals non-significant (n.s.) support. Thus, model 3 is not significantly improved as compared to model 2. We have noted that related variety correlates strongly with regional population size (see Table 1), and which is a possible reason for the increase in the standard error of the regional population size parameter and the non-significant effect that we observe in model 3. Since the parameter of regional population size is non-significant, and possibly causing multicollinearity problems and regional geographical size is also non-significant, we omit them in model 4.[5] The revised model (model 4) shows that the parameter of related variety is positive but non-significant.[6] Model 5

---

[3] In an unreported model that excludes the random effect parameter of enterprises nested in industries, the random effect of enterprises nested in regions is considerably stronger than the random effect of industries nested in regions. This indicates that some of the regional random effect (of industries nested in regions) is absorbed by a prevalence of enterprises operating in particular R&D intensive industries that are located in particular regions.

[4] The substantial interpretation of significant parameters is not straight forward in tobit regression (McDonald & Moffitt, 1980), and we address this issue in subsection 4.3.

[5] We are aware that it may be of concern to omit the regional population size (POP) and the regional geographical size (REG) parameters while retaining the POP × REG interaction in model 4 (and also in later models). However, due to the inclusion of the POP and REG parameters in model 3 and the aforementioned arguments concerning their exclusion in model 4, we find it defendable. The POP × REG interaction is also consistent in all models where included, independent of whether POP and REG are included or not. In addition, we will later present models that exclude the POP × REG interaction altogether.

[6] In unreported models, we have also added an interaction term between regional population size and related variety to the parameters reported in models 3 and 4, but model fit is not improved and the interaction term is non-significant while the interaction term between regional population size and regional geographical size remains significant.

TABLE 2 Multilevel mixed-effects tobit regressions with enterprise R&D investments as the dependent variable

Model →	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Fixed effects										
Constant	−8.89***	−20.4***	−20.9***	−17.1***	−17.9***	−17.5***	−23.2***	−23.4***	−24.0***	−24.5***
	(0.475)	(4.30)	(5.86)	(0.974)	(0.958)	(4.22)	(2.58)	(2.81)	(2.80)	(2.55)
Enterprise variables										
Multidivisional enterprise		−1.05*	−1.05*	−1.05*	−1.09*	−1.03*	−1.05*	−1.04*	−1.09*	−1.08*
		(0.508)	(0.508)	(0.508)	(0.508)	(0.507)	(0.507)	(0.507)	(0.507)	(0.507)
Size/number of employees		2.38***	2.38***	2.38***	2.39***	2.37***	2.37***	2.37***	2.38***	2.38***
		(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)
Region variables										
Regional population size (POP)		0.499†	0.576			−0.620				
		(0.279)	(0.662)			(0.482)				
Regional geographical size (REG)		−0.150	−0.167			−0.530				
		(0.358)	(0.381)			(0.359)				
POP*REG		.641*	.644*	.661**		0.512*	0.656**	0.664**		
		(0.271)	(0.272)	(0.254)		(0.261)	(0.247)	(0.254)		
Related variety			−0.409	2.18	3.15*			−0.220	.717	
			(3.21)	(1.34)	(1.33)			(1.65)	(1.61)	
Unrelated variety						11.1**	6.52**	6.76*	6.70*	7.48***
						(3.99)	(2.24)	(2.84)	(2.83)	(2.22)
Random effects										
Level−one residual (enterprises)	85.1	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2
	(4.05)	(3.57)	(3.57)	(3.57)	(3.58)	(3.58)	(3.57)	(3.57)	(3.57)	(3.58)

(Continues)

TABLE 2 (Continued)

Model →	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Level–two residual (enterprises nested in industries)	44.6	47.0	47.0	46.9	47.1	46.6	46.8	46.8	47.1	47.0
	(4.69)	(4.70)	(4.70)	(4.69)	(4.71)	(4.66)	(4.68)	(4.69)	(4.71)	(4.70)
Level–three residual (industries nested in regions)	3.60	1.05	1.05	1.15	1.42	.640	0.890	0.906	0.932	0.939
	(1.72)	(1.13)	(1.12)	(1.15)	(1.24)	(0.981)	(1.02)	(1.03)	(1.09)	(1.09)
Wald χ^2		280.4***	280.4***	279.5***	273.6***	287.4***	284.6***	284.5***	279.2***	278.9***
Log likelihood	−7653.3	−7505.5	−7505.5	−7505.9	−7509.3	−7501.6	−7503.0	−7503.0	−7506.6	−7506.7
LR test vs. one–level tobit model: χ^2	558.7***	629.1***	629.1***	633.1***	636.2***	618.4***	631.9***	631.2***	630.3***	630.2***
LR test vs. Model 2: χ^2			0.02 (n.s.)			7.76**				

Notes: Two–tailed tests of significance. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Standard error in parentheses. $N = 6584$ (uncensored = 1501, left–censored = 5083). The number of industries nested in regions is 2010. The number of regions is 89. All independent variables, except for multidivisional enterprise (as a binary variable), are log–transformed using natural logarithm.

excludes all region parameters, except for related variety, and we observe that it receives significant positive support. We discuss this finding at the end of the section.

Model 6 includes unrelated variety as an independent variable, and we observe that the parameter has a significant, positive effect on the dependent variable. The LR test (reported at the bottom of Table 2), comparing model fit between model 6 and model 2, likewise reveals significant support. Thus, model 6 is significantly improved as compared to model 2. Including unrelated variety as an independent variable, we moreover observe that regional population size has a negative, albeit non-significant effect, which is the opposite of what we observed in model 2 and 3. Multicollinearity, that is, the high correlation between regional population size and unrelated variety (cf. Table 1), may explain this issue. Following the previous arguing, we omit the parameters of regional population size and regional geographical size in model 7 and observe that unrelated variety still receives significant support (observing that the effect size of unrelated variety is lower in model 7 than in model 6 is possibly due to multicollinearity in model 6). In model 8, we include related variety as an independent variable, and in model 9, we exclude the interaction term between regional population size and regional geographical size. In both models 8 and 9, we observe that the parameter of unrelated variety receives significant support, while related variety is non-significant. In model 10, we exclude all region parameters, except for unrelated variety, and observe that it receives strong and significant support. The effect size of unrelated variety in model 10 does not deviate much from what we observed in models 7–9.

Related variety only receives significant, positive support when included as the single region variable (model 5). In all other models where included, it receives non-significant support (models 3, 4, 8 and 9). We, therefore, conclude that regional industry structures of related variety do not affect location decisions of enterprise R&D investments. Unrelated variety receives significant, positive support in all models where included (models 6–10). We, therefore, conclude that regional industry structures of unrelated variety have a positive effect on location decisions of enterprise R&D investments.

4.3 | Estimation of marginal effects

The substantial interpretation of significant parameters is not straightforward in tobit regression (McDonald & Moffitt, 1980), we have noted (in footnote 4), but estimations of marginal effects in censored tobit regression can provide genuine information about regressors of interest. In Table 3, we replicate for clarity data from model 7 in Table 2 (which includes significant parameters only, and none of which correlate strongly with unrelated variety eliminating a potential problem of multicollinearity). In the two following models, we report average marginal effects for the censored (model 7′) and the truncated outcome (model 7″). Both model 7′ and 7″ are based on regression estimates in model 7 in Table 2 (please see Cameron & Trivedi, 2010, 541–544 for a detailed explanation of censored and truncated marginal effects).

Concerning unrelated variety, model 7′ reports that a 1% change in the parameter changes, on average, the actual enterprise R&D investments with 1.61% for the censored outcome that includes zero values of R&D investments (due to the log-transforming of the variables we can interpret changes in percentage). For the truncated outcome, which omits the zero value of R&D investments, model 7″ reports that a 1% change in unrelated variety changes, on average, the actual enterprise R&D investments with 1.57% (that is, a little lower than for the censored outcome that includes zero values of R&D investments).

While model 7′ and 7″ report average changes in actual enterprise R&D investments as a function of changes in unrelated variety, Figure 1 reports changes in actual enterprise R&D investments for different values for unrelated variety. We observe that for both the censored outcome (the left part of the figure) and the truncated outcome (the right part of the figure) the change in enterprise R&D investments is lowest for the lowest values of unrelated variety and highest for the highest values of unrelated variety. Despite this dissimilarity, all values of unrelated variety have a positive, significant effect on location decisions of enterprise R&D investments, and which is depicted by 95% confidence intervals far above zero in Figure 1.

TABLE 3 Multilevel mixed–effects tobit regressions with enterprise R&D investments as the dependent variable – average marginal effects for the censored (model 7′) and the truncated (model 7′′) outcome (for clarity, model 7 is a replicate of model 7 in Table 2)

Model →	7	7′	7′′
Enterprise variables			
Multidivisional enterprise	−1.05*	−0.256*	−0.250*
	(0.507)	(0.125)	(0.121)
Size/number of employees	2.37***	.581***	.568***
	(0.162)	(0.045)	(0.040)
Region variables			
POP × REG	.656**	.161**	.157**
	(0.247)	(0.060)	(0.059)
Unrelated variety	6.52**	1.60**	1.56**
	(2.24)	(0.571)	(0.546)

Notes: Two-tailed tests of significance. *p < .05, **p < .01, ***p < .001. Standard error in parentheses. N = 6,584 (uncensored = 1,501, left-censored = 5,083). The number of industries within regions is 2010. All independent variables, except for multidivisional enterprise (as a binary variable), are log–transformed using natural logarithm. Model 7 reports regression coefficients (identical with model 7 in Table 2), model 7′ reports average marginal effects for the censored outcome (how the observed variable Y_{ikr}, which includes zero values, changes with respect to the independent variables) and model 7′′ reports average marginal effects for the truncated outcome (how the observed variable Y^{*}_{ikr}, which excludes zero values, changes with respect to the independent variables).

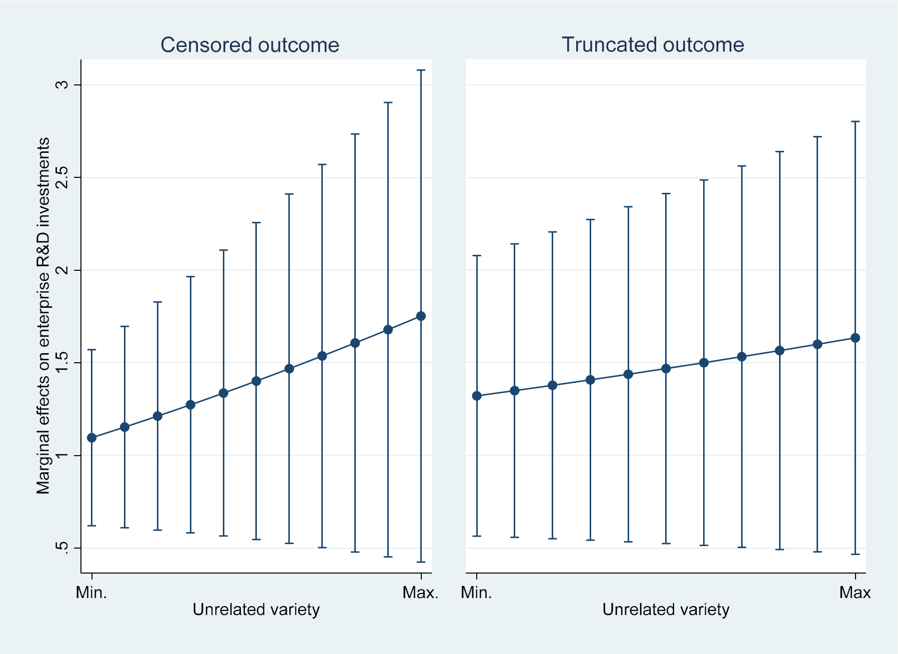

FIGURE 1 Marginal effects on enterprise R&D investments for different values of unrelated variety
Note: 95% confidence interval. Estimates are based on model 7 in Table 2

TABLE 4 Multilevel and mixed–effects logit (models 3a and 6a) and linear (models 3b and 6b) regressions with enterprise R&D investments as the dependent variable

Model →	**3a**	**3b**	**6a**	**6b**
Fixed effects				
Constant	−4.27*** (1.24)	3.08** (1.10)	−3.56*** (0.882)	4.77*** (0.821)
Enterprise variables				
Multidivisional enterprise	−0.220† (0.113)	−0.214* (0.105)	−0.217† (0.112)	−0.208* (0.105)
Size/number of employees	0.485*** (0.037)	0.437*** (0.032)	0.483*** (0.037)	0.435*** (0.032)
Region variables				
Regional population size (POP)	0.111 (0.141)	0.396** (0.127)	−0.142 (0.101)	0.018 (0.095)
Regional geographical size (REG)	−0.035 (0.079)	−0.108 (0.071)	−0.112 (0.074)	−0.133† (0.071)
POP × REG	0.140* (0.057)	−0.017 (0.051)	0.113* (0.055)	−0.040 (0.051)
Related variety	−0.093 (0.683)	−0.983 (0.623)		
Unrelated variety			2.37** (0.852)	1.96* (0.800)
Random effects				
Level–one residual (enterprises)		1.61 (0.071)		1.60 (0.071)
Level–two residual (enterprises nested in industries)	2.10 (0.249)	0.210 (0.059)	2.08 (0.247)	0.209 (0.059)
Level–three residual (industries nested in regions)	0.037 (0.050)	0.063 (0.035)	0.019 (0.043)	0.062 (0.034)
Wald χ^2	226.1***	267.9***	232.8***	272.1***
Log likelihood	−3151.6	−2575.8	−3147.7	−2574.1
LR test vs. one-level model: χ^2	572.5***	48.9***	562.5***	49.4***
N	6584	1501	6584	1501
Number of industries nested in regions	2010	682	2010	682
Number of regions	89	87	89	87

Notes: Two-tailed tests of significance. $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Standard error in parentheses. All independent variables, except for Multidi-visional enterprise (as a binary variable), are log-transformed using natural logarithm.

Returning to Table 3, it moreover informs that multidivisional enterprises invest 25.6% less for the censored outcome (model 7′) and 25.0% less for the truncated outcome (model 7′′) in R&D than single-divisional enterprises do.[7] A likely interpretation is that multidivisional enterprises concentrate R&D investments in one or a few divisions. We moreover observe that a 1% increase in enterprise size in the number of employees increases, on average, enterprise R&D investments with 0.581% for the censored outcome (model 7′) and 0.568% for the truncated outcome (model 7′′). Thus, whereas enterprise size (unsurprisingly) has a genuinely positive effect on R&D investments, we see at the same time that R&D investments per employee decrease as enterprise size in the number of employees increases. A plausible explanation is that large enterprises with many employees may economise on R&D investments or they concentrate R&D investments in one or a few units. Finally, a one-unit increase in the interaction term between regional population size and regional geographical size increases, on average, enterprise R&D investments with 0.161% for the censored outcome (model 7′) and 0.157% for the truncated outcome (model 7′′). The substantial interpretation of the interaction term can be challenging to explain, but it nonetheless shows that location in populous regions spanning a large geographical area increases enterprise R&D investments (please also see footnote 5).

4.4 | Multilevel mixed-effects logit and linear regressions in separate models

In Table 4, we present replications of models 3 and 6 in Table 2, but with the important distinction that model 3a and 6a report multilevel mixed-effects logit regressions and models 3b and 6b report multilevel mixed-effects linear regressions. Logit regressions distinguish between enterprises that do and do not invest in R&D, while linear regressions measure the amount invested (and therefore only include enterprises with positive R&D investments). In other words, the logit models (models 3a and 6a) estimate the probability of enterprise R&D investment and the linear models (models 3b and 6b) estimate the amount invested (for those who do invest in R&D).

We observe in models 3a and 3b that related variety has a non-significant effect on location decisions of enterprise R&D investments, both concerning the probability to invest and amount invested. Models 6a and 6b, on the other hand, report that unrelated variety has a significant positive effect on location decisions of enterprise R&D investments, both concerning the probability to invest and amount invested (p = 0.015 in model 6b). The findings are in line with what we have reported earlier.

Findings concerning the control variables enterprise size in number of employees and whether an enterprise is multidivisional or not are also in line with what we have reported earlier. The control variable regional population size has a positive effect on location decisions of enterprise R&D investment concerning the amount invested (significant effect in model 3b), but the effect is mediated by unrelated variety (non-significant effect in model 6b). Significant positive interaction effects between regional population size and regional geographical size, reported in models 3a and 6a, are in line with what we have reported earlier and implies that location in populous regions spanning a large geographical area increases the probability of enterprise R&D investments. However, the non-significant effect, reported in models 3b and 6b, implies that location in populous regions spanning a large geographical area does not increase the very amount of enterprise R&D investments (among those who *do* invest in R&D). We finally observe in model 6b that regional geographical size has a borderline significant negative effect on location decisions of enterprise R&D investments.

---

[7] It may be counterintuitive to find that multidivisional enterprises have relatively low R&D investments, but when omitting enterprise size as independent variable in an unreported model, we find that multidivisional enterprises have a positive significant effect on enterprise R&D investments. In other words, multidivisional enterprises tend have larger R&D investments than single-divisional enterprises, but it is not single-divisional vs. multidivisional enterprise form that affects R&D investment, but rather enterprise size in number of employees.

5 | CONCLUSION

5.1 | Discussion of the empirical results

This paper aimed to study if regional industry structures affect location decisions of enterprise R&D investments. Analysing data from Norway, we found that location in regions with unrelated, diversified, and fragmented industry structures increases both the probability and amount of enterprise R&D investments, while the location in regions with related and complementary structures has no effect.

An explanation of the positive findings from regions with unrelated structures is that, because of the possibility to recombine very different ideas from very different perspectives, enterprises anticipate R&D investments to induce radical innovations there (cf. Castaldi et al., 2015). However, we have speculated earlier, that due to the randomness and the rareness of successful innovations in such regional contexts, we do not strongly believe that this in itself will influence location decisions of enterprise R&D investments. The strong positive association between regional population size and an unrelated regional industry structures, which Aarstad et al. (2016a) have described, can be another explanation of unrelated structures' positive effect on location decisions of enterprise R&D investments; R&D investments in populous regions may diffuse or partly spill over into unrelated structures. Institutional factors can, also, affect location decisions of enterprise R&D investments in regions with unrelated industry structures, independent of whether they are rational or irrational (cf. Scott, 2014). A final explanation of the reported findings can be that political factors affect the funnelling of enterprise R&D investments into particular regional industry structures in Norway, as funding often comes from public programmes. Political factors may, for instance, affect the distribution of R&D investments into regions with unrelated industry structures in an attempt to offset or compensate for unfavourable factor conditions concerning enterprise development (cf. Aarstad et al., 2016a who found that unrelated regional industry structures were detrimental to enterprise productivity).

Some of the explanations of unrelated variety's positive effect on location decisions of enterprise R&D investments can also be attributed to related and complementary regional industry structures, yet our analyses showed that this was not the case. We have no clear understanding of why we did not find any effects, but a plausible explanation is that institutional and political factors may play a stronger role in regions with unrelated industry structures than in regions with related structures.

Our empirical findings also showed that location in populous regions spanning a large geographical area increases the probability of enterprise R&D investments. One way to interpret the finding is that enterprise location in unpopulous regions spanning a large geographical area, mainly in Northern Norway and North-East Norway, have a low probability of R&D investments. A likely reason is a combination of limited research-based infrastructure and demand and supply of factor inputs in unpopulous regions along with relatively large geographical distance to other regions (because of regional size). These factors in combination appear to discourage the probability of enterprise R&D investments.

5.2 | Limitations and future research

A limitation of the study is its cross-sectional research design. Wixe (2015, p. 8), nonetheless asserts that “regional characteristics have a tendency to change slowly over time,” which induces that regional industry structures have a stronger effect on enterprise level features than the other way around. Future studies should nevertheless use a longitudinal research approach or use instrumental variables to increase internal validity. Another limitation is that we do not control for public R&D spending at a regional level, which may affect location decisions of enterprise R&D investments. The reason why we do not control for public R&D spending is that the data, unfortunately, does not exist at the regional level (LAU-1) we analyse. We believe, however, that regional population size, regional geographical area, and the interaction between the terms, may be useful, albeit not perfect, proxy variables for public R&D investments. For example, large and R&D intensive universities and colleges tend to locate in populous regions (please see our argument for the inclusion of the interaction term at the end of subsection 3.3 and the discussion at the end of subsection 5.1). A final limitation is that the measure of related variety that we have applied in the study may underestimate the genuine industry structure in some regions. Studying economic regions in Norway, Fitjar and Timmermans (2017, p. 516) assert that using NACE codes “tends to underestimate the level of relatedness in many of Norway's most technologically sophisticated manufacturing regions,” and we, therefore, call for future studies using alternative measures of the concept. Yet, having said this, using the same measure of related variety as applied in this study, Aarstad et al. (2016a) found that the concept had a consistent effect on enterprise product innovation, which indicates strong criterion validity (cf. Cronbach & Meehl, 1955).

We have discussed various reasons why unrelated regional industry structures affect location decisions of enterprise R&D investments (subsection 5.1), and we encourage future research to gain further insight into which of these, or other reasons, that may be genuine explanations. Future research should also aim to investigate why related variety does not seem to have a genuine effect. We have suggested a few reasons, but future research should aim to gain further knowledge about this issue. We finally encourage future research to investigate other regional characteristics than studied here to gain even better knowledge about regional factors affecting location decisions of enterprise R&D investments.

REFERENCES

Aarstad, J., Kvitastein, O. A., & Jakobsen, S.-E. (2016a). Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. *Research Policy*, *45*, 844–856. https://doi.org//10.1016/j.respol.2016.01.013

Aarstad, J., Kvitastein, O. A., & Jakobsen, S.-E. (2016b). Related and unrelated variety in a tourism context. *Annals of Tourism Research*, *57*, 254–256. https://doi.org//10.1016/j.annals.2015.12.002

Alcácer, J. (2006). Location Choices across the Value Chain: How Activity and Capability Influence Collocation. *Management Science*, *52*, 1457–1471.

Armstrong, J. S., & Overton, T. S. (1977). Estimating Nonresponse Bias in Mail Surveys. *Journal of Marketing Research*, *14*, 396–402. https://doi.org//10.2307/3150783

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*, 318–337. https://doi.org//10.1016/j.respol.2008.11.010

Belderbos, R., Roy, V. V., Leten, B., & Thijs, B. (2014). Academic Research Strengths and Multinational Firms' Foreign R&D Location Decisions: Evidence from R&D Investments in European Regions. *Environment and Planning A*, *46*, 920–942. https://doi.org/10.1068/a45536

Bhattacharya, M., & Bloch, H. (2004). Determinants of Innovation. *Small Business Economics*, *22*(2), 155–162. https://doi.org//10.1023/B:SBEJ.0000014453.94445.de

Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, *39*(1), 61–74. https://doi.org//10.1080/0034340052000320887

Boschma, R., Minondo, A., & Navarro, M. (2012). Related variety and regional growth in Spain. *Papers in Regional Science*, *91*, 241–256. https://doi.org//10.1111/j.1435-5957.2011.00387.x

Britton, J. N. H. (2003). Network Structure of an Industrial Cluster: Electronics in Toronto. *Environment and Planning A*, *35*, 983–1006. https://doi.org//10.1068/a35290

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*, 349–399.

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using Stata* (Revised ed.). College Station, Texas: Stata Press.

Castaldi, C., Frenken, K., & Los, B. (2015). Related variety, unrelated variety and technological breakthroughs: An analysis of US State-Level Patenting. *Regional Studies*, *49*, 767–781. https://doi.org//10.1080/00343404.2014.940305

Content, J., & Frenken, K. (2016). Related variety and economic development: a literature review. *European Planning Studies*, *24*, 2097–2112. https://doi.org//10.1080/09654313.2016.1246517

Cronbach, L. J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, *102*, 414–417.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity for psychological tests. *Psychological Bulletin*, *52*, 281–302.

Dautel, V., & Walther, O. (2014). The geography of innovation in a small metropolitan region: An intra-regional approach in Luxembourg. *Papers in Regional Science*, *93*, 703–725. https://doi.org//10.1111/pirs.12019

Demirbag, M., & Glaister, K. W. (2010). Factors Determining Offshore Location Choice for R&D Projects: A Comparative Study of Developed and Emerging Regions. *Journal of Management Studies*, *47*(8), 1534–1560. https://doi.org//10.1111/j.1467-6486.2010.00948.x

Fitjar, R. D., & Timmermans, B. (2017). Regional skill relatedness: towards a new measure of regional related diversification. *European Planning Studies*, *25*, 516–538. https://doi.org//0.1080/09654313.2016.1244515

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, *41*, 685–697. https://doi.org//10.1080/00343400601120296

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

de Groot, H. L. F., Poot, J., & Smit, M. J. (2016). Which agglomeration externalities matter most and why? *Journal of Economic Surveys*, *30*(4), 756–782. https://doi.org//10.1111/joes.12112

Hundt, C., & Sternberg, R. (2016). Explaining new firm creation in Europe from a spatial and time perspective: A multilevel analysis based upon data of individuals, regions and countries. *Papers in Regional Science*, *95*, 223–308. https://doi.org//10.1111/pirs.12133

Jacobs, J. (1969). *The economy of cities*. New York: Vintage.

Klein, K. J., Danserau, F., & Hall, R. (1994). Level issues in theory development, data collection, and analysis. *Academy of Management Review*, *19*, 195–229.

Kowalski, A. M., & Marcinkowski, A. (2014). Clusters versus Cluster Initiatives, with Focus on the ICT Sector in Poland. *European Planning Studies*, *22*, 20–45. https://doi.org//10.1080/09654313.2012.731040

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99*, 483–499. https://doi.org/10.2307/2937739

Lawrence, J. B. Jr., & Iversen, G. R. (1979). *Contextual analysis: Concepts and statistical techniques*. Belmont, CA: Wadsworth Publishing Company.

Leslie, L. M. (2017). A status-based multilevel model of ethnic diversity and work unit performance. *Journal of Management*, *43*, 426–454. https://doi.org//10.1177/0149206314535436

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

McDonald, J. F., & Moffitt, R. A. (1980). The uses of tobit analysis. *The Review of Economics and Statistics*, *62*, 318–321. https://doi.org//10.2307/1924766

Moreno, R., Paci, R., & Usai, S. (2005). Spatial spillovers and innovation activity in European regions. *Environment and Planning A*, *37*, 1793–1812. https://doi.org//10.1068/a37341

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673–690.

Odagiri, H., & Yasuda, H. (1996). The determinants of overseas R&D by Japanese firms: an empirical study at the industry and company levels. *Research Policy*, *25*, 1059–1079. https://doi.org/10.1016/S0048-7333(96)00892-X

Ogdul, H. G. (2010). Urban and rural definitions in regional context: A case study on Turkey. *European Planning Studies*, *18*, 1519–1541. https://doi.org//10.1080/09654313.2010.492589

Ottaviano, G. I. P., & Pinelli, D. (2006). Market potential and productivity: Evidence from Finnish regions. *Regional Science and Urban Economics*, *36*, 636–657. https://doi.org//10.1016/j.regsciurbeco.2006.06.005

Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly*, *14*, 15–34.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rousseau, D. (1985). Issues of levels in organizational research: Multilevel and cross-level perspectives. In L. L. Cummings, & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 7) (pp. 1–37). Greenwich, CT: JAI Press.

Scott, W. R. (2014). *Institutions and Organizations: Ideas, Interests, and Identities* (4th ed.). Thousand Oaks, CA: Sage.

Shannon, C. E. (1948a). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423.

Shannon, C. E. (1948b). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 623–656.

Shefer, D., & Frenkel, A. (2005). R&D, firm size and innovation: an empirical analysis. *Technovation*, *25*, 25–32. https://doi.org//10.1016/S0166-4972(03)00152-4

Siedschlag, I., Smith, D., Turcu, C., & Zhang, X. (2013). What determines the location choice of R&D activities by multinational firms? *Research Policy*, *42*, 1420–1430. https://doi.org/10.1016/j.respol.2013.06.003

Smit, M. J., Abreu, M. A., & de Groot, H. L. F. (2015). Micro-evidence on the determinants of innovation in the Netherlands: The relative importance of absorptive capacity and agglomeration externalities. *Papers in Regional Science*, *94*, 249–272. https://doi.org//10.1111/pirs.12068

Smith, V., Broberg, A. L., & Overgaard, J. (2002). Does location matter for firms' R&D behaviour? Empirical evidence for Danish firms. *Regional Studies*, *36*, 825–832. https://doi.org//10.1080/0034340022000012270

Snijders, T. A. B. (2011). Multilevel analysis. In M. Lovric (Ed.), *International encyclopedia of statistical science* (pp. 879–882). Berlin: Springer.

StataCorp (2017). *Version 15 (Vol. StataCorp LP)*. College Station, TX: StataCorp LP.

Sun, Y. F. (2002). Sources of innovation in China's manufacturing sector: imported or developed in-house? *Environment and Planning A*, *34*, 1059–1072. https://doi.org//10.1068/a34107

Tavassoli, S., & Carbonara, N. (2014). The role of knowledge variety and intensity for regional innovation. *Small Business Economics*, *43*, 493–509. https://doi.org//10.1007/s11187-014-9547-7

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, *26*, 24–36. https://doi.org//10.2307/1907382

Tsai, W. P. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal*, *44*, 996–1004. https://doi.org//10.2307/3069443

Vogel, J. (2015). The two faces of R&D and human capital: Evidence from Western European regions. *Papers in Regional Science*, *94*, 525–551. https://doi.org//10.1111/pirs.12084

Wilhelmsen, L., & Foyn, F. (2012). *Innovasjon i norsk næringsliv 2008–2010 (Vol. 46)*. Oslo: Statistics Norway.

Wixe, S. (2015). The impact of spatial externalities: Skills, education and plant productivity. *Regional Studies*, *49*, 2053–2069. https://doi.org//10.1080/00343404.2014.891729

Wixe, S. (2016). Neighbourhood related diversity, human capital and firm innovation. *Papers in Regional Science*. *In press*. https://doi.org//10.1111/pirs.12255

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.

Yang, C.-H., & Hayakawa, K. (2015). Localization and overseas R&D activity: The case of Taiwanese multinational enterprises in China. *R&D Management*, *45*, 181–195. https://doi.org//10.1111/radm.12059

Yang, C. H., & Lin, H. L. (2012). Openness, absorptive capacity, and regional innovation in China. *Environment and Planning A*, *44*(2), 333–355. https://doi.org//10.1068/a44182

How to cite this article: Aarstad J, Kvitastein OA, Jakobsen S-E. Location decisions of enterprise R&D investments as a function of related and unrelated regional industry structures: A multilevel study. *Pap Reg Sci*. 2018;1–19. https://doi.org/10.1111/pirs.12366