FRONT PAGE

Universitetet i Stavanger

UIS BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAM:	THESIS IS WRITTEN IN THE FOLLOWING SPECIALIZATION/SUBJECT:
MSc Business Administration – Business Innovation	Economics of Innovation IS THE ASSIGNMENT CONFIDENTIAL? No (**NB**! Use the red form for confidential theses)

TITLE:

The Effects of Related and Unrelated Variety on Firm Performance in Norway

AUTHOR(S) Yves Matthes		SUPERVISOR: Prof. Ragnar Tveterås
Candidate number: 4028 ………………… …………………	Name: Yves Matthes …………………………………… ……………………………………	

University of Stavanger

Handelshøgskolen

Master Thesis

Yves Matthes

June 2019

The Effects of Related and Unrelated Variety on Firm Performance in Norway

Candidate Number: 4028

Abstract

This Master thesis addresses the question whether the co-location of economic sectors increases firm performances in that region. The empirical study is based on the theoretical concepts of related and unrelated variety. Using a firm level panel data set covering more than 280.000 Norwegian enterprises and a time span of 10 years, I apply random and fixed effects models. The main finding is that related variety has a positive effect on regional firm performance. On the other hand, no evidence for an effect of unrelated variety can be found. Furthermore, different industrial sectors are compared, to show that the degree of variety has a differently strong effect on different industries.

Acknowledgements

I would like to thank my supervisor, Professor Ragnar Tveterås, for his support, encouragement and enthusiasm throughout my writing process. He allowed me to develop the thesis according to my own interests, but provided continuous guidance and suggestions whenever needed. Most importantly, he helped me aquire practical skills to organize and analyse large amounts of data, which will be invaluable for my further professional career.

My special thanks go to Professor Ron Boschma from the University of Utrecht, whose research serves as a theoretical foundation of this thesis. I feel very privileged that he took his time to provide me with valueable and inspiring ideas and suggestions.

Finally, I would like to take this opportunity to express my deep gratitude to my family and friends for the uplifting feedback und conversations, but even more for the emotional support and encouragements during an uneasy time.

Contents

1 Introduction

Does the structure of economic sectors in a region have an effect on its economic growth?

This question, raised by many authors in the fields of economic geography and regional studies, challenges the classical growth model by Solow [1956], which explains growth as a result of increases in input factors and exogeneous technological progress. The so called 'new growth theory' introduces a spatial dimension and states that not only the stock of inputs affects growth, but also its distribution on a regional level. Following Glaeser et al. [1992] the rationale is as follows: Geographical proximity between firms in a region leads to spillover effects, in which knowledge and ideas are transferred to other firms; due to spying, imitation and movement of skilled labour. By combining these new ideas, product and process innovations lead to improvements in firm productivity and growth, respectively[1]. This theoretical framework would then endogenize technological change in the Solow growth model.

However, while the existence of spatial factors is widely recognized and accepted, researchers present ambiguous suggestions whether industrial specialization or diversification are more beneficial. Early studies, such as that of Marshall [1920] propose that a region must specialize into a certain industry to incite knowledge spillovers and growth. This view is opposed by Jacobs [1969]. Her concept, popularized as 'Jacobs externalities', is based on the assumption that economic growth is being facilitated by a sectoral diversity of co-located industries [Glaeser et al., 1992]. Jacobs' thought is taken further by Frenken et al. [2007], who argue that knowledge spillovers are not incited by diversity per se, but rather by a variety amongst firms that already share certain competences. According to the concept

[1]This goes in accordance with Schumpeter's definition of innovation as new re-combinations of already existing resources

of cognitive proximities [Boschma, 2005] the underlying idea is that inter-industry learning only works on basis of a set of common knowledge. Therefore, Frenken et al. [2007] introduce the concept of related and unrelated variety. Related variety is a measure for the degree of diversification amongst firms with common markets, products or technologies. On the other hand, unrelated variety measures the variety amongst firms with a cognitive distance. The development of these measures has incited many further studies that help understanding the impact of related and unrelated variety on economic performance.

This thesis is aimed to contribute to the discussion, by applying the concept of unrelated and related variety empirically. Basis for this is a cross-sectional panel data set covering more than 280.000 Norwegian firms and a time range from 2007 to 2017. In particular, I apply fixed and random effects models to test whether related and unrelated variety have an effect on firms' value-added growth. I hypothesize that related and unrelated variety both have a positive effect on firm performance. Furthermore, the study is extended by applying the model on the manufacturing sector only. Section 2 discusses the theoretical framework. The literature review in Section 3 gives a brief overview over some of the empirical findings so far. The subsequent chapters then describe the procedure and findings of my own study.

2 Theoretical Framework

2.1 Agglomeration Economies

To understand the effects of industrial patters on economic growth, the framework of agglomeration economies seems to be a suitable starting point. Aggomeration economies refer to the notion that firms benefit from being located in spatial proximity to other firms [Frenken et al., 2007]. Scholars have identified various channels through which the location of firms translates to economic growth, such

Source	Localization economies	Urbanization economies
Labor market pooling	Access to specialized labor market	Cost benefits from access to large labor market
Input/output sharing	Access to specialized suppliers	Cost benefits from access to customers
Knowledge spillover	Industry-specific	Between different industries or from scientific environment
	Marshall externalities	*Jacobs' externalities*

Figure 1: Agglomeration Economies, source: De Bok and Van Oort [2011]

as the possibility of knowledge spillovers, the access to existing labour markets or easier transport of goods.

To integrate these channels into a theoretical framework, De Bok and Van Oort [2011] distinguish between localization economies and urbanization economies. For the sake of clarity, I follow the framework presented by De Bok and Van Oort [2011] and visualized in Figure 1. Additionally, I introduce an example from the cities of Stavanger and Sandnes in Rogaland, which are home of the Norwegian oil and gas industry.

Localization economies describe cost savings or productivity gains due to the density of the firm's sector in a region[Van der Panne, 2004]. This goes in line with the Marshall externalities, named after Alfred Marshall, who first introduced it to assess the emergence of industrial clusters in 1890. Localization economies can be divided into three channels:

Firstly, the co-localization of related firms in a region facilitates labour market pooling. Andini et al. [2013] argue in this context that a specialized labour market allows firms and employees to find each other faster. Moreover, the matches are better in terms of skills, knowledge and experience. An extensive empirical survey within the Norwegian petroleum industry by Sasson and Blomgren [2011] emphasizes the need of specialized engineers and economists in the Stavanger region, mainly for the specific fields of operations, drilling and wells. Firms locating within this region have easier access to those specialized professionals, since they are already living in the area.

Secondly, a firm profits from easier transport of goods due to the proximity to specialized industries along their supply chain. This not only minimizes transport costs but also allows close partnerships and knowledge spillovers. In the case of the Norwegian oil and gas sector the co-location of specialized suppliers around Stavanger and in the region of Agder has been actively driven by the government owned petroleum firm Equinor (formerly Statoil), establishing supplier contracts in the fields of mechanical engineering and with former shipyards in the 1980s [Nerheim, 1996]. The reason why these local firms were chosen instead of higher qualified international suppliers might be primarily politically motivated [Sæther et al., 2011]. Either way, they generated a co-located supplier industry that minimizes transport costs[2].

The third aspect of localization economies are knowledge spillovers. Knowledge and ideas can be passed on between companies 'because people talk and gossip, products can be reverse engineered, and employees move between firms'[Glaeser et al., 1992]. It is further argued that the spatial proximity of firms facilitates this spillover effect. The concept of Marshall externalities assumes that these effects occur mostly between firms from the same sector. This corresponds to the concept of cognitive proximities, as coined by Boschma [2005]. He argues that successful knowledge spillovers require a firm to possess technological and market competencies to 'identify, interpret and exploit the new knowledge [Boschma, 2005]. In other

[2]Today, Rogaland is also home of international oil service firms such as Schlumberger (Stavanger), Halliburton (Tananger), BakerHughes (Tananger) and Seawell (Stavanger) [Sasson and Blomgren, 2011]

words, the co-location of related industries facilitates knowledge transfers, which in turn fosters innovation activities and firm growth. For the case of the Norwegian oil and gas industry, Sæther et al. [2011] argue that it is mostly a 'symbiotic relationship' between oil companies and their suppliers, which is responsible for the development of innovations. Therefore, the clustering of various oil firms in the Stavanger region facilities the creation of cooperative networks. Sasson and Blomgren [2011] add to this that also R&D institutions play an important role. The International Research Institute of Stavanger (IRIS) and the Univerisy of Stavanger (UiS) are specializing on petroleum technologies and therefore provide an environment which can facilitate knowledge spillover.

Urbanization economies assume that the sheer size of an agglomeration stimulates firm performance, irrespective of the industry structure. Glaeser et al. [1992] argue that a firm's location within an agglomeration has the advantage that it is close to its customers. This way, the costs of transportation to markets can be reduced. In addition to that, Frenken et al. [2007] recognize that urbanization economies also have an impact on knowledge spillovers by referring to the function of trade associations, universities as well as social, political and cultural organizations. In the case of Stavanger one can mention the existence of a university (UiS), which creates a scientific environment, partly independent from the existing industry structure in the region.

Based on this, Jacobs [1969] argues that urbanization economies provide an environment which stimulates productivity gains due to radical innovations. Her concept, known as *Jacobs externalities*, states that innovative behaviour is mainly caused by knowledge spillovers between sectors that are not closely related. Frenken et al. [2007] argue in this context that the co-location of diversified industries helps to copy or re-combine ideas to generate radical product innovations. Note, that this view opposes Marshall externalities, which propose the co-location of firms of the same sector. An example of Jacobs externalities could be observed in the case of the Norwegian oil and gas industry. When Norway started extracting oil and gas from the North Sea in the late 1960s, it didn't possess prior technological expertise in this field. However, the experience from maritime operations in the local shipping and fishing industry helped to create the radical innovations necessary to develop offshore oil fields [Sæther et al., 2011].

Apart from localization and urbanization economies, the existence of a heterogeneous business landscape with a variety of different sectors benefits from the portfolio effect: Sectoral diversification serves as a risk spreading strategy. According to Attaran [1986], regions with a large variety of industrial sectors are better prepared to compensate external demand shocks and to avoid rapid increases in unemployment. The city of Stavanger serves as a good example for this. As mentioned earlier, Stavanger is highly dependent on the petroleum sector. After the oil price drop in 2014, the city suffered from a rapid rise in unemployment. A more diversified composition of industries would have reduced this effect. Pasinetti [2006] further argues that on long term industries might disappear. A well diversified region would then be able to re-route the workers into new industries.

To sum up, economic growth can arise from sectoral co-location of specialized industries (Marshall externalities), sectoral diversification (Jacobs externalities) or the co-location of firms of any sector (urbanization economies). Following the train of thought of Frenken et al. [2007], a clear distinction of these theories is crucial, as they are connected to different types of benefits. Marshall externalities are expected to generate knowledge spillovers within a specialized industry environment, leading to incremental process innovations. On the contrary, Jacobs externalities stimulate either radical product innovations or incremental process innovations depending on how different the firms are in terms of technologies and markets.

2.2 Related and Unrelated Variety

According to the theory of Jacobs externalities, sectoral diversification facilitates knowledge spillovers. However, inter-industry knowledge spillovers are stronger and more likely, when the firms have some similarities in terms of products, technologies and markets to be able to learn from each other [Boschma and Iammarino, 2009]. Therefore, it can be argued that innovations and growth are not stimulated by variety per se, but by variety amongst related firms. To be able to test for this concept empirically, Frenken et al. [2007] propose a method to measure the degree of related and unrelated variety. Related variety shows the degree of diversification amongst firms that share cognitive proximities. Unrelated variety is the degree of variety amongst unrelated firms. The procedure works as follows:

Firstly, Frenken et al. [2007] identify sectors that share some degree of relatedness. For this, they utilize five-digit NACE[3] codes, which describe each firm's sector. Table 1 illustrates the composition of the NACE code of a firm that manufactures hand tools. The first two digits define a more general description of the firms activity. Adding further digits leads to a more detailed description. Now, all firms with the same first two digits can be considered related. In this example, we assume that all firms that manufacture fabricated metal products have a cognitive proximity and are therefore related, regardless of whether they produce hand tools or other types of metal products. On the other hand, firms that do not share the same first two digits are considered unrelated.

NACE-Code	Description
25	Manufacture of fabricated metal products
25.7	Manufacture of cutlery, tools and general hardware
25.73	Manufacture of tools
25.73.1	Manufacture of hand tools

Table 1: Example for the decomposition of NACE-codes

[3]Nomenclature of Economic Activities code. NACE is the industry standard classification system of the European Union

Secondly, the distribution of economic activity amongst all sectors in a region must be measured. This could be based on the number of employees in each sector, but also other economic variables such as value-added or wages could be used. The NACE code now allows to determine the share of total economic activity for each sector.

Thirdly, Frenken et al. [2007] calculate the degree of related and unrelated variety based on the shares of economic activity. To do so, Frenken et al. [2007] use entropies, as proposed by Attaran and Zwick [1987]. This method allows to determine variety measures at any sectoral digit level [Frenken et al., 2007]. A more detailed explanation and the mathematical derivation of the variety measures will be provided in section 6.2.

	Specialized industry structure	Diversified industry structure
Type of externalities (Conceptualized by Marshall 1920, Jacobs 1969)	MAR externalities	Jacobs externalities
Agglomeration externalities (Conceptualized by Hoover, 1937)	Localization externalities ↓	Urbanization externalities ↓
Sources and ways of knowledge transfer	Intra-industry learning ↓	Inter-industry learning ↙ ↘ RV UV ↓ ↓
Possible growth drivers	Incremental innovation	Incremental innovation (RV) / Radical innovation (UV)

Figure 2: Conceptual Framework, source: Kublina [2015]

A low degree of related variety would then indicate a region that is highly specialized. According to Boschma [2005] this would result in a cognitive lock-in which prevents innovation. High related variety, on the other hand, is expected to facilitate economic growth. Applying the example, this means that within the sector of 'Manufacture of fabricated metal products' there is a high variety amongs its sub-sectors. High unrelated variety then indicates that a region is very diverse, which could facilitate structural stability through the portfolio effect or incite radical product innovations. Figure 2 embeds related and unrelated variety into the conceptual framework. In this thesis, it will be investigated whether related and unrelated variety have an effect on firm performance in Norway.

3 Literature Review

Frenken et al. [2007] first introduce the concept of related and unrelated variety. Using data from 40 regions in the Netherlands, they find that improvements of related variety in a region increase employment growth while unrelated variety lowers unemployment growth as a result of the portfolio effect. This paper resulted in many subsequent studies, as discussed in this section. The literature review is divided into two parts: The first part focuses on macro-level studies, where the dependent variables are macroeconomic indicators such as employment or GDP. The second part discusses studies about the effects of variety on the firm level. For each part, a table is provided, which summarizes the main findings of the reviewed studies.

3.1 Macro-Level Studies

Most studies regarding related and unrelated variety are on a country level. The main findings of the discussed papers are summarized in Table 2[4].

Boschma and Iammarino [2009] estimate the effect of related and unrelated variety on regions' value-added growth and labour productivity in Italy. Unlike

[4]a more detailed literature review on macro-level studies is provided by Content and Frenken [2016]

Frenken et al. [2007], they measure variety based on regional trade data. Boschma and Iammarino [2009] argue that in a world of global integration trade profiles can provide a good picture of regional industrial structures and the existing degree of specialization or diversification. They find evidence for a positive impact of related variety on regional value-added growth.

A similar approach supporting the findings of Boschma and Iammarino [2009] is presented by Saviotti and Frenken [2008] who apply the concept of export varieties on a national level. For 156 OECD countries the authors measure related and unrelated export variety based on trade data and estimate its effect on the countries' GDP growth. The long time range in consideration from 1964 to 2003 allows to notice changes in the global economy with increasing exports and growing international interrelations. Their findings suggest that related variety leads to immediate growth, while unrelated variety only has a positive effect on long term (which is in line with the theory of portfolio effects).

Further studies that find a positive effect of related variety on the economy are provided by Boschma et al. [2012] for Spain, Falcıoğlu [2011] for Turkey, Kublina [2015] for West- Germany, as well as Cortinovis and Van Oort [2015] for Europe. However, the effects of unrelated variety seems to be unclear. As can be seen in Table 2, the findings suggest both positive and negative effects, while most estimations are not significant.

Assuming that related and unrelated variety facilitate knowledge spillovers, one can argue that certain industries benefit more than others, depending on the degree of knowledge intensity. The most obvious differentiation would be between service and manufacturing sectors, as done by Mameli et al. [2012] for Italy and Bosma et al. [2011] for the Netherlands. Both studies suggest that in service sectors related variety has a positive impact on employment growth, while no effect can be found in the manufacturing sector. However, Hartog et al. [2012] take a

Paper	Region	Dep Var	RV	UV	Urb.
Frenken et al. [2007]	Netherlands	Employment growth	+	0	-
		Productivity growth	-	0	-
		Unemployment growth	0	-	-
Boschma and Iammarino [2009]	Italy	Employment growth	M	0	0
		VA growth	+	+	0
		Productivity growth	M	0	0
Saviotti and Frenken [2008]	OECD countries	GDP growth	+	-	NA
		Labour Prod.	+	-	NA
Boschma et al. [2012]	Spain	VA growth	+	0	+
Falcıoğlu [2011]	Turkey	Productivity growth	+	NA	0
Quatraro [2010]	Italy	Productivity growth	M	0	M
Kublina [2015]	West-Germany	Employment growth	+	+	0
Cortinovis and Van Oort [2015]	Europe	Empoyment growth	+	0	0
Mameli et al. [2012]	Italy	Empoyment growth	+	+	M
Bosma et al. [2011]	Netherlands	Productivity growth	0	NA	M
Hartog et al. [2012]	Finland	Employment growth	+	0	-
Bishop and Gripaios [2010]	UK	Employment growth	M	M	M

Table 2: Macro level studies. Partly adapted from Content and Frenken [2016]. Extended with urbanization economies. + and - indicate positive and negative effects, 0 no significant results, M mixed results and NA not analysed

closer look into manufacturing industries and find that the effects of varieties depend on the specific sector. Using a panel data set for industries in Finland, they find that in knowledge intense sectors related varieties have a positive effect on regional employment growth. A finer division of industrial sectors is provided by Bishop and Gripaios [2010], who discover for the UK that the effects of varieties are heterogenous across industries. Surprisingly, they find that over all unrelated variety has a stronger effect than related variety, which diverges from most findings in literature.

Almost all studies also test for the effects of population density as an indicator for urbanization economies. While Boschma et al. [2012] find evidence for a positive effect, Hartog et al. [2012] and Frenken et al. [2007] suggest negative effects. However, most studies come to insignificant or mixed results.

3.2 Micro-Level Studies

While the research focus has been on macro-level studies, Aarstad et al. [2016] point out that the firm level has so far been widely ignored. However, the micro-level might reveal some interesting insight, as it allows to capture the direct effects of knowledge spillovers on firms. The rationale is that enterprises located in a region with a high degree of related variety benefit from knowledge spillovers, which in turn leads to innovation activities. Innovations are expected to lead to productivity improvements and firm growth. So, if it can be shown empirically that high related variety leads to improvements in innovation activity or productivity, it would suggest that related variety facilitates knowledge spillovers. The empirical findings are discussed in this section and are summarized in Table 3.

One study provided by Castaldi et al. [2015] investigates this hypothesis based on US patent data. They find that related variety increases firms' innovation activities. Furthermore, they can find that unrelated variety leads to breakthrough innovations, which emphasizes the existence of Jacobs externalities. This is further supported by Tavassoli and Carbonara [2014] who find that variety (related more than unrelated) incites innovation activities in Sweden. Nevertheless, Antonietti and Cainelli [2011] do not agree with these findings. Using survey data for Italian manufacturing firms, they apply a probit model with a firm's decision to innovate as a binary dependent variable. They cannot find evidence for an effect of any type of variety on innovation activity.

Besides innovation, researchers investigate the effect of variety on firm productivity. Wixe [2015] analyses the effects of different externalities using a micro-level panel data set for Sweden, covering almost 40.000 industrial firms. She finds that urbanization economies have a positive effect on labour productivity. Varieties have a positive long term effect, but no significant short term effect. However, she does not differentiate between related and unrelated variety.

Paper	Region	Dependent Variable	RV	UV	Urb.
Castaldi et al. [2015]	US	Patents	+	0	NA
	US	Breakthrough Innovations	0	+	NA
Tavassoli and Carbonara [2014]	Sweden	Patents	+	+	+
Antonietti and Cainelli [2011]	Italy	Choice to innovate	0	0	0
Wixe [2015]	Sweden	Labour Productivity	M*	M*	+
Aarstad et al. [2016]	Norway	Productivity	0	-	+
	Norway	Product Innovation	+	0	-

Table 3: Micro-level studies. + and - indicate positive and negative effects, 0 no significant results, M mixed results and NA not analysed. *Wixe [2015] does not differentiate between RV and UV

The study that comes closest to my own paper is that of Aarstad et al. [2016], who address this issue on a firm-level in Norway. The data used is a survey covering 6595 enterprises. Each firm provides information on whether they had any product innovation in the last year. Aarstad et al. [2016] find that regions with high related variety incite product innovation, while unrelated variety has no significant effect. Regarding productivity, Aarstad et al. [2016] find no significant effect of related variety, but a negative effect of unrelated variety. Furthermore, they argue that population density has a positive effect on productivity, but a negative effect on product innovation. For my own thesis, the contribution of Aarstad et al. [2016] serves as a useful point of reference, because it does not only cover Norway, but also studies the enterprise level.

4 Data

Each private stock-based firm in Norway must report their yearly business data to the Norwegian government. I obtain this data from Proff Forvalt[5], who make these firm-level data available. I extract all firms with the enterprise form Aksjeselskap (AS), referring to a stock-based company with a limited liability for their owners. Since the AS form is by far the most commonly used in Norway, it

[5]https://forvalt.no

can be assumed that this selection reflects the corporate landscape in the country adequately. To prepare the data set for statistical analyses, I carry out several modifications. Firstly, firms lacking crucial information are excluded. This implies enterprises that don't reveal the sector they are operating in or firms that cannot be geographically located. Furthermore, obvious false observations (such as firms with negative wage costs) and the year 2018 (too few observations) are excluded. After these adjustments the data contains a sum of 280.374 firms that either still exist or have existed at some point during the last 10 years[6]. For each firm, the data contains the address of its headquarters, the economic sector they are operating in and the number of employees in 2017. Furthermore, it provides business data from 2007 to 2017, including yearly operating profits and wage costs. After this, the the data is restructured to a panel form containing a total of 1.902.591 observations.

The data allows to determine the location of each firm and to link it to administrative units. Norway is divided into 18 counties (Fylker), corresponding to a NUTS 3 classification. The counties are then further subdivided into a total of 422 municipalities (Kommuner) on a NUTS 5 level. In the past years Norway has undergone administrative reforms. The latest reform on 1. January 2018 has merged the counties Nord-Trondelag and Sor-Trondelag to Trondelag. Furthermore, municipalities in Vestfold and Trondelag have been reorganized. For the sake of uniformity, I adjust the data sets by rearranging former municipalities and counties into their new entities. Moreover, the islands of Svalbard in the Arctic Sea are excluded due to its too little population to provide any significant results. Consequently, each firm is connected to one municipality and one county. Furthermore, I group the municipalities in 90 so called economic regions as proposed by Statistics Norway, which allows me to also generate variables on a NUTS 4 level. In order to determine these regions, Statistics Norway uses commuting tables that

[6] Descriptional statistics of the panel data set are presented in Figure 11 in the Annex

show where employees live and where they work. This can be taken as an indicator for the geographical concentration of labour markets. Additionally, they use sales data and information on the spread of local newspapers, which gives an idea about what region the inhabitants feel connected to. It is important to note that in accordance with NUTS classifications, economic regions are not allowed to overlap county borders.

Each firm's sector is described by the Nomenclature of Economic Activities (NACE) code. NACE is the industry standard classification system of the European Union. This five-digit number groups each firm into an industrial group, based on its products, technologies or markets. The first two digits provide general information on the firms activity. By increasing the number of digits, the described sector becomes more specific. This information allows to determine the composition of economic sectors in different regions. An example for the composition of NACE codes was given in Table 1.

Moreover, I use data provided by Statistics Norway to create variables for regional centrality, population density and education levels. The exact methods are described in Section 6.

5 Hypotheses

The aim of this thesis is to investigate which regional industry structures facilitate firm growth. Thereby, the following four hypotheses will be tested.

Hypothesis 1: The main hypothesis in this thesis is that related variety has a positive effect on firm performance, measured in value-added. This would be in line with both economic theory and the empirical findings in the large body of literature. The rationale is that on a regional level the existence of related industries leads to knowledge spillovers, which result in innovations and thus in enterprises' productivity improvements and growth.

Hypothesis 2: Unrelated variety can affect economic growth in two ways: Firstly, it can serve as a region's risk-spreading strategy to dampen external demand shocks. However, my research design might not be able to capture this effect since it focuses only on a 10-year time range and on firm level data (the portfolio effect can be better observed looking at unemployment rates). Secondly, unrelated variety can spur breakthrough product innovations and thus facilitate firm growth. Therefore, I hypothesize a positive effect of unrelated variety on value-added growth.

Hypothesis 3: Based on the theory of urbanization economies I expect that densely populated regions or regions with a high degree of centrality benefit firm performance.

Hypothesis 4: Following empirical studies such as Hartog et al. [2012] and Bishop and Gripaios [2010] it can be assumed that certain industries benefit more from related variety than others. In particular, I expect that knowledge-intense industries are more in need of knowledge spillovers and thus depend more on changes in related variety. Focussing on the manufacturing sector, it is therefore hypothesized that related variety has a stronger effect on the performance of high-tech than on medium & low-tech firms.

6 Variables

In order to test the hypotheses, suitable regression models must be developed. This section discusses the choice and construction of dependent and independent variables. Descriptive statistics for all variables are summarized in Table 4 and histograms can be found in the Annex.

6.1 Value-Added Growth

According to the theory discussed in Section 2, economic variety in a region results in knowledge spillovers that incite the combination of ideas, leading to innovation. Incremental process innovations and disruptive product innovations are expected to increase firm productivity. However, the lack of historical employment data does not allow to test for labour productivity. Therefore, I choose value-added growth as a measure for firm performance, as has been done before by Boschma et al. [2012] and Boschma and Iammarino [2009]. Value-added (VA) can be defined as the sum of a firms operating profits and wage costs.

$$VA_{i,t} = Operating\, Profits_{i,t} + Wage\, Costs_{i,t} \tag{1}$$

The growth rate in consideration is then a one-year change in percent:

$$\Delta VA_{i,t} = ln\frac{VA_{i,t}}{VA_{i,t-1}} = ln(VA_{i,t}) - ln(VA_{i,t-1}) \tag{2}$$

Histograms for the growth rates of value-added, wage costs and operating results are presented in the Annex.

6.2 Variety

To test the hypotheses regarding the effects of variety in different regions, an appropriate measurement of related and unrelated variety is crucial.

In this context Frenken et al. [2007] introduce entropy measures that have been adapted by fellow researchers. The main advantage over other techniques (like for example the Herfindahl Index) is its decomposable nature. Suppose that each firm is assigned a five-digit NACE code which describes the sector it is operating in. Based on this, the firm can also be linked to a group of similar industries (two-digit level). The entropy between all sectors on a five digit level can now be decomposed to a variety within the sub-group (related variety) and between the sub-groups (unrelated variety). Furthermore, Boschma and Iammarino [2009] argue that these properties imply that variety at different digit levels can be included in a regression analysis without necessarily generating collinearity. In this section I first explain the theoretical concept of entropies and apply it to my case. Afterwards, the measures calculated for Norway will be presented.

The entropy framework has been first introduced by Shannon [1948], who applied a concept of thermodynamics in the field of information technology. Attaran and Zwick [1987] formalize this concept to measure industrial diversity D:

$$D(P_1, P_2, ..., P_n) = -\sum_{i=1}^{n} P_i log_2 P_i \tag{3}$$

where n denotes the number of economic sectors and P_i the proportion of some total quantity. In the context of diversity of sectors this can be the share of employees on total employees in the nation or region[7]. Due to the lack of historical employment numbers, this thesis uses the share of a firm's wage costs on total wages paid in the region. The wage costs then reflect changes in employment as well as workers' skills and changes in the labour market.

So, if for example there is only one existing sector that employs workers, then $P_i = 1$ and $D(P_1) = -\sum_{i=1}^{n} log_2(1) = 0$. On the other hand, when all firms have the same share of wages ($P_1 = P_2 = ... = P_n$), the diversity value reaches its maximum at $D = log_2(n)$. Higher values indicate a higher degree of diversity. The maximum values increase with the number of observed sectors. Applying all 5-digit employee shares of one region into Equation 3, the total variety is obtained, which can now be further decomposed. Therefore, each 5-digit industry is grouped

[7] other possible variables are value-added or traded goods (Boschma and Iammarino [2009] for instance measure export diversification for Italy)

into a 2-digit industry.

The wage share W_g of each 2-digit set S_g is then the sum of its 5-digit shares W_i:

$$W_g = \sum_{ieS_g} W_i \tag{4}$$

with $g = 1, ..., G$.

Based on those sets, Attaran and Zwick [1987] demonstrate how to disaggregate the entropy measure into its between-set and within-set. The between-set describes the diversity *between* all 2-digit industries (indicating unrelated variety), while the within-set characterizes the industrial patterns *within* a 2-digit sector (related variety). By applying Equation 3 the entropy within a 2-digit set S_g is obtained:

$$D_{withinS_g} = -\sum_{ieS_g}(W_i/W_g)log_2(W_i/W_g) \tag{5}$$

To get an entropy measure for related variety (RV), the expression has to be weightened by the relative share of each individual set on the total sets W_g/W_s. The resulting RV measure determines the average diversity within the 2-digit level sets:

$$RV = \sum_{g=1}^{G}(W_g/W_s)[-\sum_{g=1}^{G}(W_g/W_s)log_2(W_g/W_s)] \tag{6}$$

Accordingly, the entropy measure for unrelated variety (UV) is given by

$$UV = -\sum_{g=1}^{G}(W_g/W_s)log_2(W_g/W_s) \tag{7}$$

which describes the degree by which the employment is distributed between the different 2-digit sets. Furthermore, the decomposable nature of the entropy measure implies

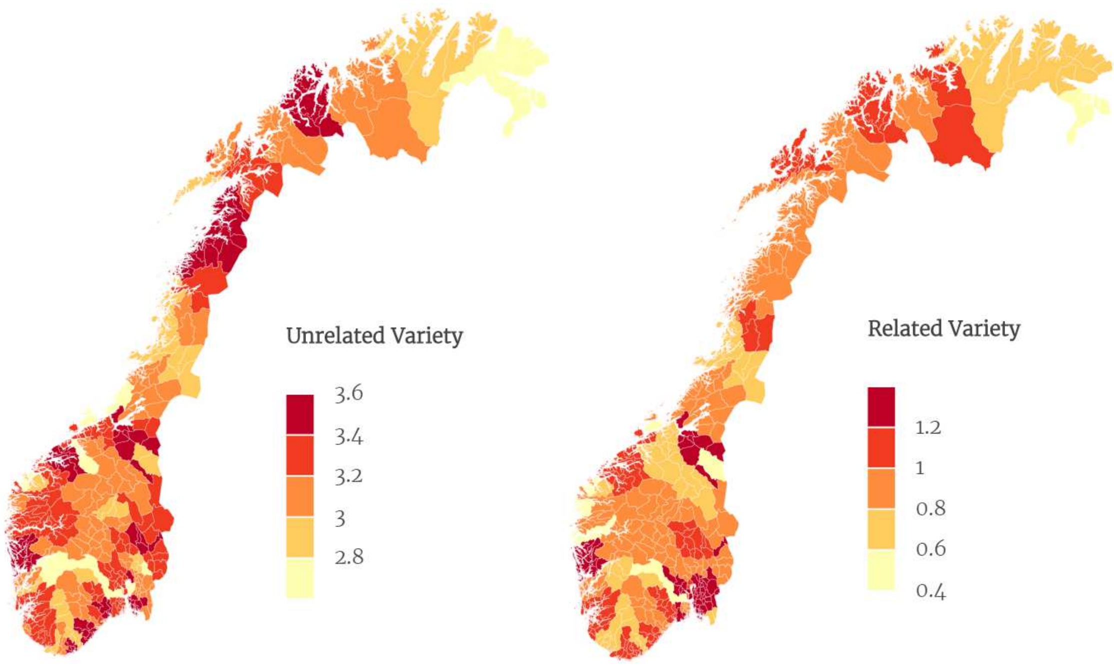
$$RV + UV = TV. \tag{8}$$

To compare varieties in different parts of Norway, it is crucial to divide the country into appropriate regions. Municipalities seem to be too small to identify spillover effects, since firm interactions and labour markets go beyond their borders. On the other hand, the county level is too large, since it covers different urban areas that might not be integrated at all. Therefore, I use the classification of 90 Economic Regions by Statistics Norway, as described in Section 4. This classification has already been used in this context by Aarstad et al. [2016].

The related and unrelated variety measures are computed for each economic region and each year, using Stata 13 software. Figure 14 in the Annex reveals that the measures stay quite stable over time. This is understandable, as industrial patterns don't change much within one decade[8]. Observable changes, such as a variety decrease in Kirkenes (located in Norways North-east corner), can be explained by firm entries or exits.

The maps in Figure 3 illustrate the values of related and unrelated variety in 2017. Note in this context that the variety measures are calculated for economic regions, which go beyond municipality borders. The highest variety measures can be observed around the largest agglomerations, while peripheral regions have rather low values. This observation is further emphasized by Figure 4, which displays the related and unrelated variety for each economic region in 2017. For the sake of clarity only the most striking regions are labelled in the graph. The plot reveals that related and unrelated variety are moderately correlated (the correlation

[8]Significant variety changes over time are captured by Saviotti and Frenken [2008] who cover the years 1964 to 2003 on a national level



(a) Unrelated Variety (b) Related Variety

Figure 3: Variety Measures 2017

coefficient is 0.4). Moreover, it is striking that large cities such as Oslo, Bergen, Trondheim, Fredrikstad, Sarpsborg, Molde, Alesund and Tromso have high levels of related and unrelated variety. This is no surprise, since places with higher population naturally offer a greater variety of firms. However, the interpretation of low unrelated variety is somewhat more ambiguous. On one hand, it can be explained by sparsely populated areas with few companies. This can be observed for the economic region of Kirkenes in the very north of Norway. On the other hand, it can indicate that a region specializes on certain industries. This seems to be the case for Kongsberg, which is home of a mechanical engineering cluster accommodating highly skilled workers in 30 firms that offer a total of 4000 jobs [Asheim et al., 2011]. Furthermore, Odda and Ulsteinvik are industrial regions focusing only on few sectors [Jakobsen et al., 2005]. The highest levels of related variety are found in the economic regions Follo, Lillestrom and Drammen in southeast Norway.

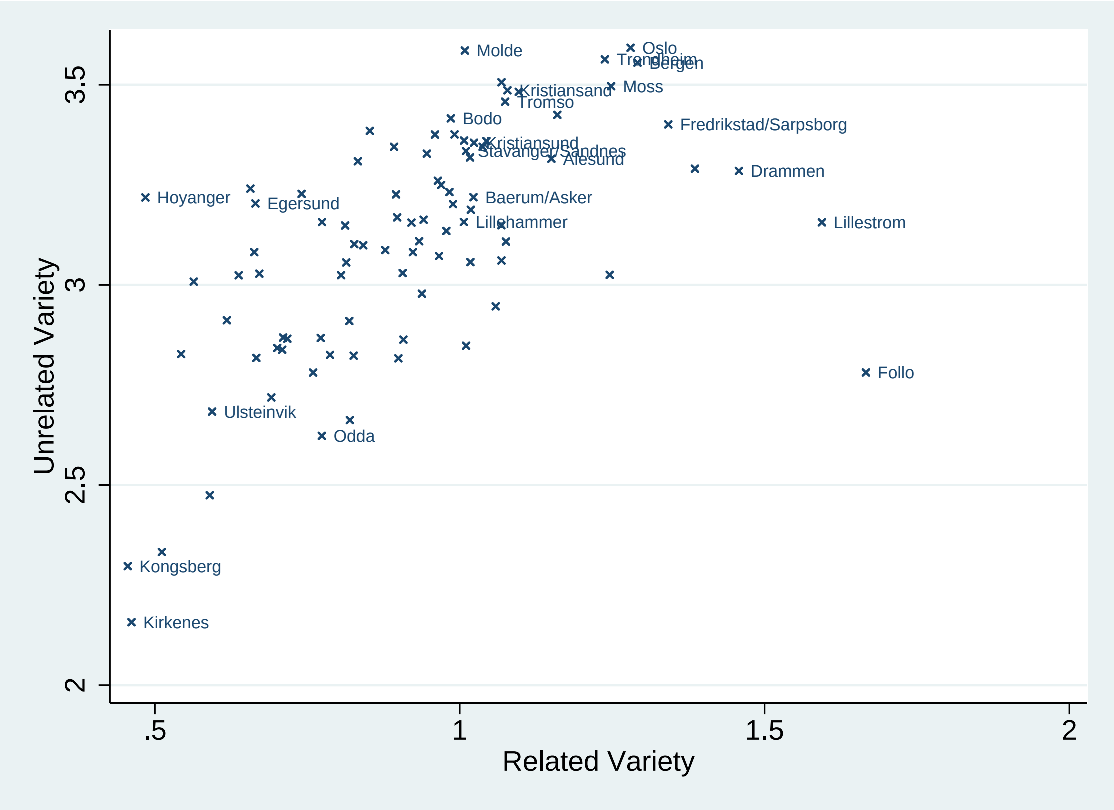


Figure 4: Related and Unrelated Variety in 2017

6.3 Urbanization

The concept of urbanization economies states that firms benefit from the accessibility of institutions [Hartog et al., 2012]. Researchers commonly measure urbanization economies with population density. This seems reasonable, as it can be assumed that more densely populated areas host more institutions and a better infrastructure than other regions. Therefore, I compute the population density for each municipality, using data from Statistics Norway and area data from Kartverket. In line with the existing literature I measure the logarithm of the number of inhabitants for each square kilometre. This procedure reflects the decreasing marginal benefit for each additional inhabitant in a region [Frenken et al., 2005].

Besides that, Norway's special geographical properties should be taken into account. Compared to other economies Norway has a small population. Most of its citizens reside in the few cities on the south and west shore, so that most parts of the country are rather sparsely populated. On top of that, the presence of geographical features such as mountain ranges, fjords and islands as well as few roads in rural regions emphasize the significance of 'distance'. Regarding the concept of urbanization economies it could be argued that a firm's access to urban institutions and labour markets in Norway is rather determined by commuting distances than population density. Therefore, I also apply an alternative indicator called Centrality Index (CI). This index is introduced by Statistics Norway, which measures the degree of centralization for each municipality based on two factors: The average commuting time for each person in a municipality to its working place and the number of different service functions one can reach by car within 90 minutes. Oslo obtains the index value (1000) and all other municipalities are being assigned a respective value. Since this index has not been used by any researchers in this context yet, it must be seen with caution.

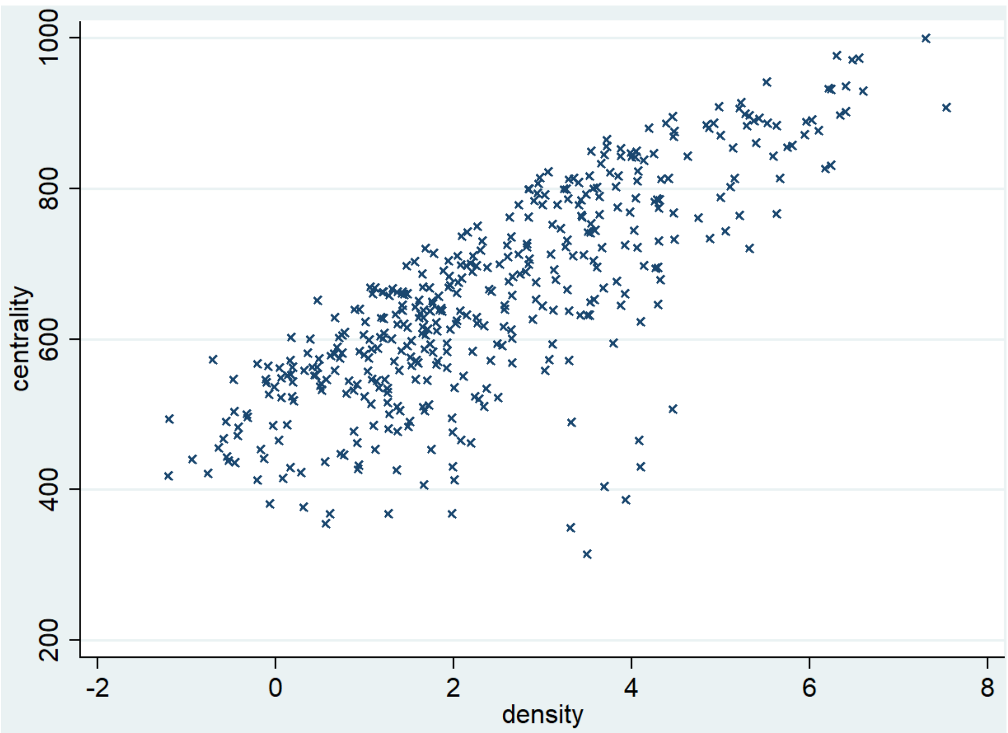
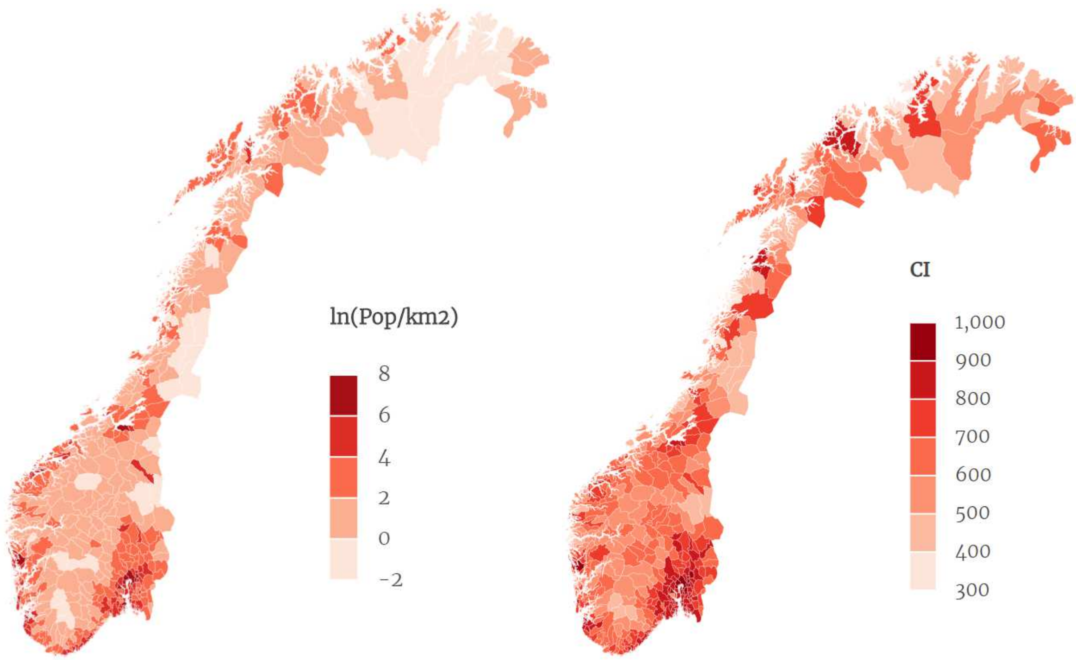


Figure 5: Relationship between centrality and population density on municipality level

Figure 5 shows the population density and CI for each municipality. As could be expected, both indicators are strongly positively correlated (correlation coefficient = 0.909). The outliers in the bottom right of the graph are all remote places. Even with a relatively high population density they lack connection to important institutions and thus lack the benefits of urbanization economies. The centrality measurement would take this more into account than population density. The maps in Figure 6 show the two indicators on a municipal level. As can be expected, the highest values for both measurements are around Oslo in the Southeast and around coastal towns such as Stavanger, Bergen, Trondheim and Tromsø. In the regression models both indicators are estimated separately to avoid multicollinearity issues.



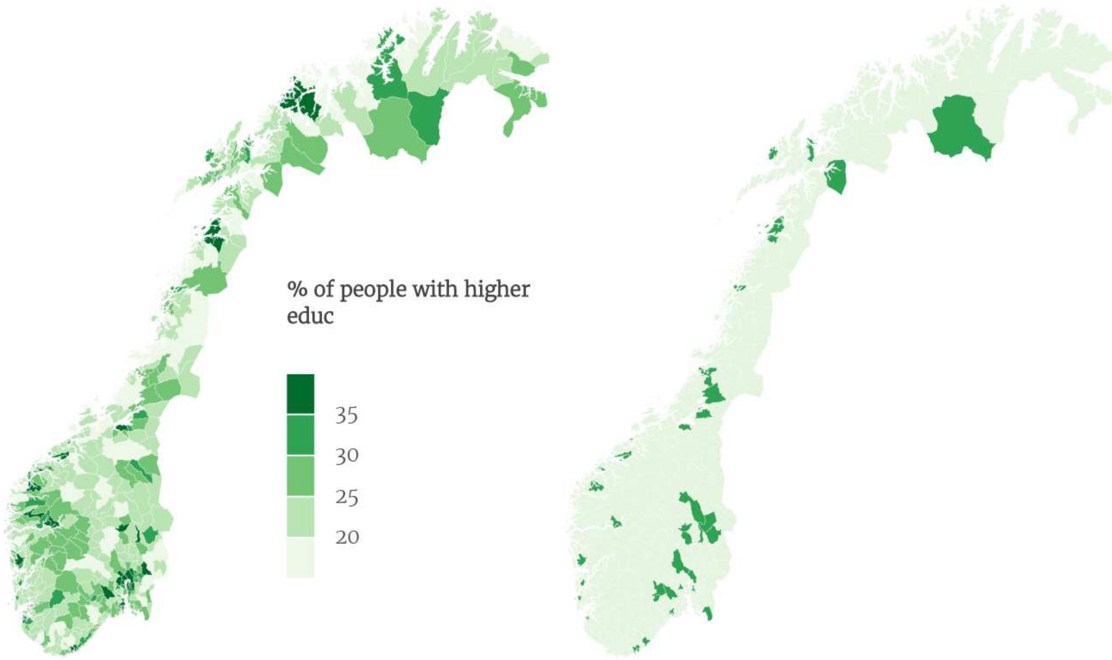
(a) Density

(b) Centrality

Figure 6: Urbanization Economies

6.4 Human Capital

Frenken et al. [2005], Hartog et al. [2012] and Boschma et al. [2012] include human capital in their models for the Netherlands, Finland and Spain to control for the



(a) % of inhabitants with higher education (b) Municipalities with a university

Figure 7: Human Capital

skill level in regional labour markets. In accordance with their approach I compute the percentage of the population with higher education, using data from Statistics Norway. This indicator is only supposed to serve as a rough approximation and is not as elaborated as in the literature. However, the data reveals some useful insight, as can be seen in the map in Figure 7a. Unsurprisingly, municipal education rates are highest in the Oslo region, the large cities on the southwest coast and Tromsø in the north. Most peripheral areas have comparatively low education rates. Nonetheless, some rural municipalities have high education rates, which becomes most apparent in the very north. To find an explanation for this observation, a second map is presented in Figure 7b, highlighting each municipality that hosts at least one university or university college. A visual comparison verifies that the presence of universities accounts for these findings. Including education levels into the regression models controls for the possibility that a firm located in a highly educated municipality can hire skilled labour, which in turn increases the firm's performance. However, based on the map there is a suspicion that education levels are highly correlated with urbanization economies, since skilled workers tend to cluster in the big cities. This will be further investigated in a later section of this thesis.

6.5 Further Control Variables

Norway stretches over a large area, so that it seems to reasonable to split Norway into different geographical areas. I create dummy variables for the north (the counties of Finnmark, Troms and Nordland) and the west (Trondelag, More og Romsdal, Sogn og Fjordane, Hordaland and Rogaland). The southeast serves as the baseline group and is not assigned a dummy to avoid multicollinearity issues.

Furthermore, Norway's economy is strongly dependent on its oil and gas industry. The time range covered in this thesis is characterized by the global oil price drop in 2014, which resulted in a recession from 2014 to 2016. In this time the unemployment rate increased from 3,3% in 2014 to 5,1% in 2016. Furthermore, the GDP per capita decreased drastically (Data from Eurostat). The effects of this economic crisis can also be observed by a decline in total wages paid in the firm level data set. Nonetheless, growth rates in many firms seem to recover quickly. Since the dependent variable in my models is firm growth, it is unclear whether the crisis will be reflected in the regressions. I control for the regression years by applying a dummy for the years 2014 to 2016.

7 Regressions

Descriptive statistics as well as a correlation table are presented in Table 4. Furthermore, histograms of the main variables can be found in the Annex. A closer inspection of the firm-level data reveals that certain modifications must be considered before estimating the models.

Table 4: Descriptive Statistics and Cross-Correlation

Variables	lndva	lunrelvar	lrelvar	density	centrality	north	west	educ	oilshock
Mean	0.074	3.277	1.095	5.029	842.094	0.074	0.349	35.438	0.460
Std. Dev.	0.815	0.2771	0.2561	1.990	129.888	0.262	0.476	10.696	0.498
Min	-11.486	1.991	0.367	-1.207	315	0	0	13.8	0
Max	13.477	3.608	1.79	7.531	1000	1	1	52.5	1
Observations	963040	1607856	1607856	1893985	1893985	1893985	1893985	1893985	1893985
lndva	1.000								
lunrelvar	0.004	1.000							
lrelvar	0.003	0.400	1.000						
density	0.000	0.440	0.530	1.000					
centrality	0.001	0.420	0.612	0.909	1.000				
north	0.005	-0.131	-0.228	-0.387	-0.386	1.000			
west	-0.007	0.001	-0.270	-0.023	-0.244	-0.213	1.000		
educ	0.002	0.423	0.400	0.815	0.815	-0.189	-0.149	1.000	
oilshock	0.003	-0.007	0.029	-0.008	-0.003	0.003	-0.011	-0.006	1.000

lndva is value-added growth

Firstly, it becomes apparent that smaller firms tend to have more erratic growth rates in value-added that larger ones. This is illustrated in Figure 8. Due to this, including all firms leads to insignificant regression results. Therefore, two different methods are applied to fix this issue. The first method would be to simply exclude too small firms. This approach is suggested by Wixe [2015], who excludes all firms with zero employees. By running regressions with different cutting points, I identify 5 employees as the most reasonable threshold, as it leads to the most significant results and excludes the smallest number of observations possible. However, for comparison, the regression results excluding all firms with zero employees is attached in Table 11 in the Annex. The second method is to apply analytic weights, which devaluates smaller firms. The weight is calculated based on the number of employees in 2017. Firms that state to have zero employees get a value of one employee to be not excluded from the analysis. However, this method is only applied in the fixed effects model, as the Stata software does not support it for random effects.

In addition to this, Norways four largest agglomerations Oslo, Bergen, Trondheim and Stavanger/Sandnes are left out. This is due to Norways special geographic properties. The largest urban areas differ strongly from the peripheral ar-

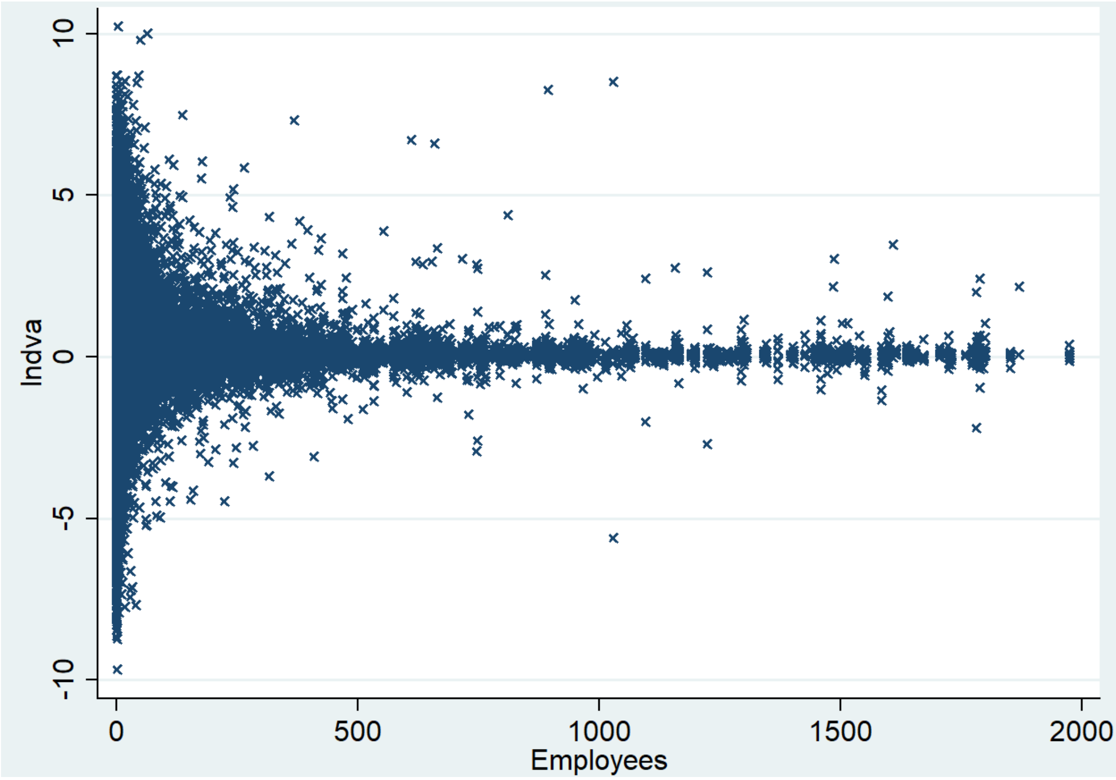


Figure 8: Relationship between a firms value added growth and its number of employees; firms with more than 2000 employees are excluded

eas. Excluding these cities decreases the degree of correlation between independent variables and leads to more significant regression results. Furthermore, a graphical analysis and the Breusch-Pagan test advocate the presence of heteroskedasticity. Therefore, I apply robust standard errors in all regressions.

7.1 All Sectors

Using Stata 13 software, fixed and random effects models are run on the enterprise level. This section will cover all sectors, while the results presented in the next chapter focus on the manufacturing sector. The variables used are described in Table 5.

The estimators of the random effects models are presented in Table 6. Firstly, it can be observed that the firm size measured in value-added (loglva) has a sig-

Table 5: Description of Variables Used in Regression

loglvasq	$ln(ValueAdded^2)$, lagged by one year
loglva	$ln(ValueAdded)$, lagged by one year
lunrelvar	Unrelated Variety, lagged by one year
lrelvar	Related Variety, lagged by one year
density	$ln(population/km^2)$
centrality	Centrality Index
north	Dummy for firms located in the north
west	Dummy for firms located in the west
educ	% of inhabitants with higher education
oilshock	Dummy for the years 2014 to 2016

nificantly negative effect on firm growth. This appears to be reasonable, as larger firms tend to grow slower than smaller firms. This result stays robust through all model specifications.

Related variety has a significantly positive effect on value-added growth in models 1 to 4. This result supports hypothesis 1, after which the co-location of similar industries benefits their performance. This goes in line with a large share of studies as presented in Section 3. Even though the majority of specifications suggest a positive effect of unrelated variety, none of the estimators prove to be significant. Consequently, this finding cannot support hypothesis 2 but is corresponding to many studies that also fail to come to significant results.

Population density has a significantly positive effect on a firm's performance. This supports the hypothesis of the existence of urbanization economies, where firms benefit from being located in an agglomeration, regardless of the industry structure. This is further supported by the significantly positive effect for centrality (model 5), which suggests that a firm's proximity to administrative centres is beneficial for its economic performance. However, including centrality instead of population density leads to insignificant variety estimators. A possible explanation may be centrality's high correlation with the variables of related variety

Table 6: Firm Level Random Effects on Value-Added Growth

	(1) lndva	(2) lndva	(3) lndva	(4) lndva	(5) lndva
loglvasq	0.0456***	0.0455***	0.0455***	0.0454***	0.0455***
	(0.000735)	(0.000736)	(0.000736)	(0.000739)	(0.000738)
loglva	−0.937***	−0.936***	−0.936***	−0.935***	−0.936***
	(0.0107)	(0.0107)	(0.0107)	(0.0108)	(0.0108)
lunrelvar	0.00926	0.00732	0.00116	−0.000129	0.00127
	(0.0109)	(0.0109)	(0.0110)	(0.0110)	(0.0110)
lrelvar	0.0585***	0.0358**	0.0370**	0.0318**	0.0133
	(0.0108)	(0.0120)	(0.0126)	(0.0128)	(0.0130)
density		0.00867***	0.0142***	0.0167***	
		(0.00211)	(0.00223)	(0.00281)	
north			0.0810***	0.0833***	0.106***
			(0.0100)	(0.0102)	(0.0110)
west			0.00311	0.00133	0.0258**
			(0.00816)	(0.00822)	(0.00882)
educ				−0.000670	−0.00178**
				(0.000544)	(0.000573)
oilshock				0.00948***	0.00965***
				(0.00191)	(0.00191)
centrality					0.000388***
					(0.0000469)
Constant	4.218***	4.212***	4.198***	4.212***	4.013***
	(0.0503)	(0.0504)	(0.0509)	(0.0525)	(0.0561)
Observations	325139	325139	325139	325139	325139
R^2 within	0.435	0.435	0.435	0.436	0.436
R^2 between	0.0579	0.0578	0.0581	0.0579	0.0580
R^2 overall	0.0904	0.0904	0.0906	0.0905	0.0906

Robust standard errors in parentheses. Excluding firms with 5 or less employees
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

and education[9]. Excluding education or related variety from the model improves the estimations slightly, but does still not lead to significant estimators for related variety (not presented in table). In this context it must be noted that the use of value-added growth as dependent variable might conceal a more ambiguous relationship on a firm level: Aarstad et al. [2016] find for Norway that population density in a region can affect a firm's performance twofold. On one hand, they detect evidence for a positive effect on firm productivity. On the other hand, they discover a negative effect on innovation activity. They explain this result by pointing out the existence of 'dynamic and innovative sectors' in areas of low population density[10]. Nonetheless, the estimations show evidence for the existence of urbanization economies as stated in hypothesis 3.

[9] The analysis of the Variance Inflation Factor (VIF) suggests multicollinearity issues for the centrality variable

[10] In particular, Aarstad et al. [2016] refer to maritime industries, sea farming, mechanical industries as well as oil and gas industries

The introduction of further control variables in the model increases the goodness of fit only slightly. Being located in the north of Norway has a significantly positive effect on firm performance. This result appears to be puzzling, as the peripheral north is mostly rural and not known for industrial activity. One possible explanation is that firms in the north are in average smaller than in the rest of Norway. This can be shown by comparing the mean value-added in the north with the southwest. The negative estimators for loglva suggest that smaller firms grow faster that larger ones, which might explain the positive estimator for the north. The dummy for western Norway only reveals a significant estimator in model 5. It is also positive, but much smaller than the estimator for the north. Education as an indicator for human capital is highly correlated with density. This suggests that highly skilled workers tend to cluster in cities. Therefore, including education and density in one model does not lead to significant estimators for education. However, model 5 reveals that education is significantly negative when tested alongside centrality, suggesting that the presence of highly educated workers has a negative effect on the firm performance in that region. Boschma et al. [2012] come to the same result without further elaborating on that. The dummy 'oilshock' indicating the years from 2014 to 2016 suggest that the crisis years are accompanied with positive growth rates of value-added. This finding seems to be counterintuitive. However, macroeconomic data reveals that even though Norway suffered from the effects of the oil price drop, the GDP growth rates recovered quickly, which might explain this result.

As a next step, fixed effects models with value-added growth as dependent variable are tested. Table 7 presents the estimation results for different model specifications. Model 1 is an unweighted model containing all firms, while Model 2 applies analytic weights as described above. Model 3 excludes all firms with 5 or less employees, as in the random effects model. Model 4 then combines the two

Table 7: Firm Level Fixed Effect on Value-Added Growth

	(1) unweighted	(2) analytic weights	(3) 5+ unweighted	(4) 5+ weighted
loglvasq	0.0376***	0.0356***	0.0414***	0.0342***
	(0.000712)	(0.00341)	(0.000892)	(0.00466)
loglva	−1.190***	−1.126***	−1.256***	−1.095***
	(0.00907)	(0.0638)	(0.0126)	(0.0913)
lunrelvar	0.0152	0.0800	0.0938**	0.0891
	(0.0232)	(0.0730)	(0.0287)	(0.0795)
lrelvar	0.315***	0.632***	0.768***	0.666***
	(0.0293)	(0.0787)	(0.0334)	(0.0891)
Constant	6.121***	6.619***	6.036***	6.546***
	(0.0679)	(0.433)	(0.0890)	(0.572)
Observations	651222	555112	325139	229029
R^2 within	0.427	0.367	0.461	0.353
R^2 between	0.0596	0.0670	0.0476	0.222
R^2 overall	0.0826	0.0821	0.0720	0.148

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

approaches. The fixed effects model does not allow to include variables that remain equal throughout the entire time range in consideration. For this reason, the variables density, centrality, north, west and education are excluded. Furthermore, the control variable 'oilshock' is disregarded as it decreases the significance of the results. The estimators remain similar across the models, which underlines the robustness of the findings. In accordance with the random effects models, firm size has a significantly negative effect on value-added growth. The main finding here is that related variety has a significantly positive effect on firm growth. This corresponds to the outcome of the random effect model. Positive estimators can also be found for unrelated variety. However, the estimator is only significant when firms with 5 or less employees are excluded and no weighting is applied (Model 3).

At this point the use of a Hausman test would be appropriate to determine whether the random effects or fixed effects model suits better. However, this test cannot be applied in this particular case, as the given models violate the asymptotic assumptions of the Hausman test. Nevertheless, the fixed and random effects models come to mostly homogeneous results, which further emphasizes the robustness of the findings.

7.2 Manufacturing

My analysis so far has covered all economic sectors jointly. Yet, there are reasons to assume that variety effects are stronger in certain sectors than in others. According to the study of Hartog et al. [2012], technologically advanced industries benefit more from knowledge spillovers. The effects of related variety can therefore be expected to be stronger in high-tech industries. Accordingly, two questions will be investigated:

(1): Do the findings for all sectors hold when examining manufacturing industries separately?

(2): How does the impact of related variety vary between knowledge-intense and non-knowledge intense manufacturing sectors?

Starting off, all manufacturing firms must be identified. Moreover, they must be divided into high-tech and low & medium tech firms. This is done on basis of two-digit NACE codes and the classifications of Eurostat[11]. Eurostat divides the manufacturing industries into four groups based on the degree of knowledge intensity. For the sake of simplicity, I differentiate between only two groups, as presented in Table 8.

This information is used to compute new measures for varieties in manufacturing industries. The method is equivalent to that described in Section 6.2. The annual measures for related and unrelated variety are individually computed for manufacturing, high-tech and low & medium tech industries. The varieties for all manufacturing sectors are illustrated in Figure 9. As in the previous maps, it becomes apparent that city areas tend to have higher variety measures. However, the map also reveals that related variety within the manufacturing sector is also high in some rural areas (mostly in northern Norway). This might be an indication

[11] https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf

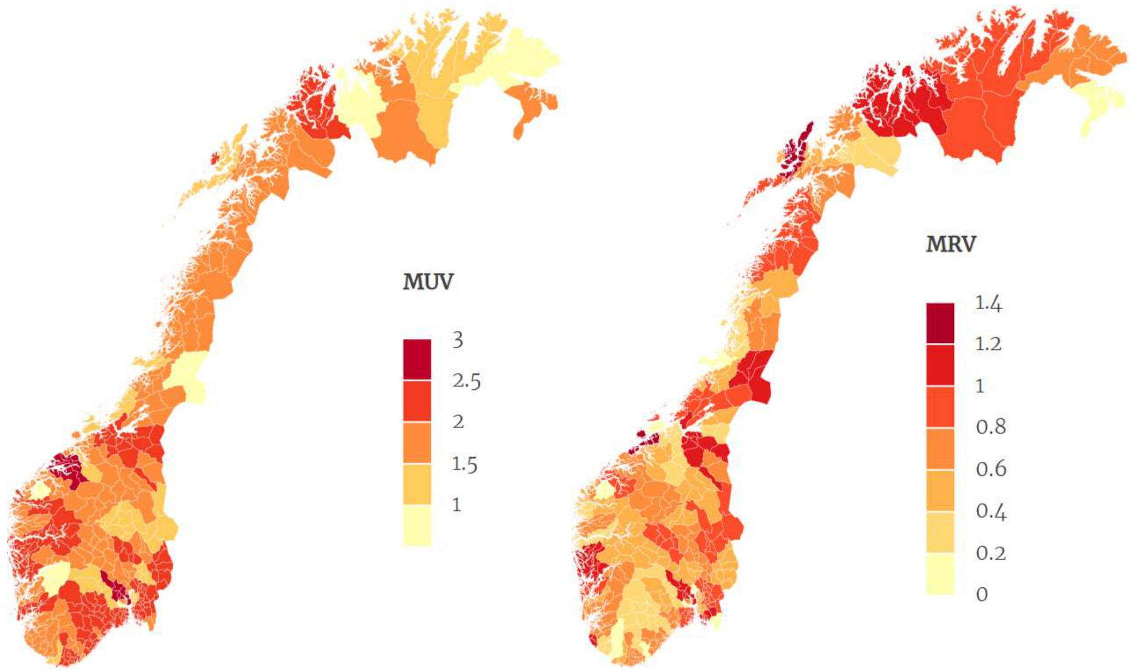
NACE	Description	High-Tech	Low & Medium Tech
10	Manufacture of food products		✓
11	Manufacture of beverages		✓
12	Manufacture of tobacco products		✓
13	Manufacture of textiles		✓
14	Manufacture of wearing apparel		✓
15	Manufacture of leather and related products		✓
16	Manufacture of wood and of products of wood and cork		✓
17	Manufacture of paper and paper products		✓
18	Printing and reproduction of recorded media		✓
19	Manufacture of coke and refined petroleum products		✓
20	Manufacture of chemicals and chemical products	✓	
21	Manufacture of basic pharmaceutical products	✓	
22	Manufacture of rubber and plastic products		
23	Manufacture of other non-metallic mineral products		✓
24	Manufacture of basic metals		✓
25	Manufacture of fabricated metal products		✓
26	Manufacture of computer, electronic and optical products	✓	
27	Manufacture of electrical equipment	✓	
28	Manufacture of machinery and equipment	✓	
29	Manufacture of motor vehicles, trailers and semi-trailers	✓	
30	Manufacture of other transport equipment	✓	
31	Manufacture of furniture		✓
32	Other manufacturing		✓
33	Repair and installation of machinery and equipment		✓

Table 8: Manufacturing sectors based on two-digit NACE codes, grouping according to Eurostat

for manufacturing clusters in peripheral regions.

Figure 10 presents the variety measures for high-tech and low & medium tech sectors. It is striking that the varieties in high-tech sectors are comparably low. This is due to the low number of high-tech sectors. Some economic regions have no firm in the high-tech industry and therefore have a variety measure of zero, as can also be seen in the histograms in the Annex. The maps for low & medium tech varieties are similar to those for all manufacturing firms, as they are firms in consideration overlap.

Having computed the variety measures, regression models can now be run. However, the use of different variety variables for different industry groups would limit the possibility to compare the estimators obtained from the regressions. For this reason, all variety variables are standardized by transforming each observation to z-scores ($z = \frac{X-\mu}{\sigma}$). The underlying rationale is to scale the variables to the same size. Higher estimators then suggest a stronger impact of a change in that particular variety measure. The variables used are listed in Table 9. Descriptive statistics



(a) Unrelated Variety (b) Related Variety

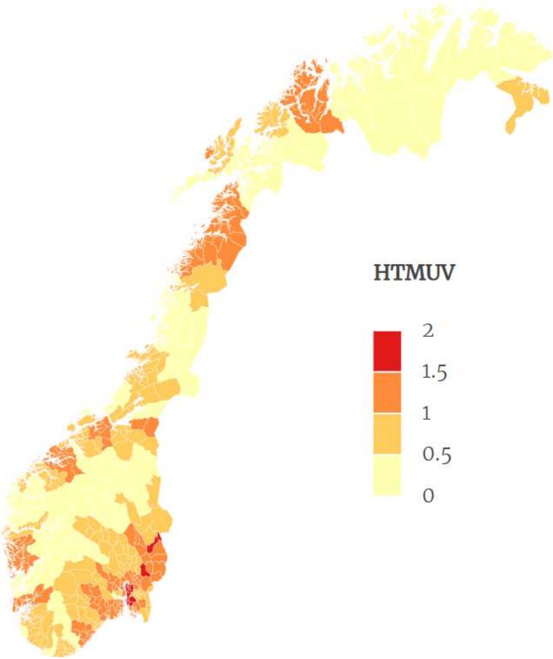
Figure 9: Manufacturing Variety 2017

can be found in Table 15 in the Annex.

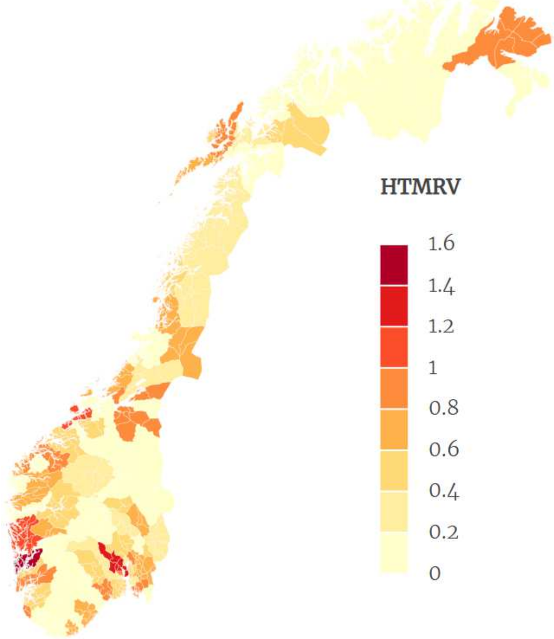
Table 9: Description of Variables Used in the Regression

loglvasq	$ln(ValueAdded^2)$, lagged by one year
loglva	$ln(ValueAdded)$, lagged by one year
stdlunrelvar	Unrelated Variety, lagged by one year (Standardized)
stdlrelvar	Related Variety, lagged by one year (Standardized)
stdlmuv	Manufacturing Unrelated Variety, lagged by one year (Standardized)
stdlmrv	Manufacturing Related Variety, lagged by one year (Standardized)
stdlhtmuv	High-Tech Unrelated Variety, lagged by one year (Standardized)
stdlhtmrv	High-Tech Related Variety, lagged by one year (Standardized)
stdllmtmuv	Low & Medium Tech Unrelated Variety, lagged by one year (Standardized)
stdllmtmrv	Low & Medium Tech Related Variety, lagged by one year (Standardized)

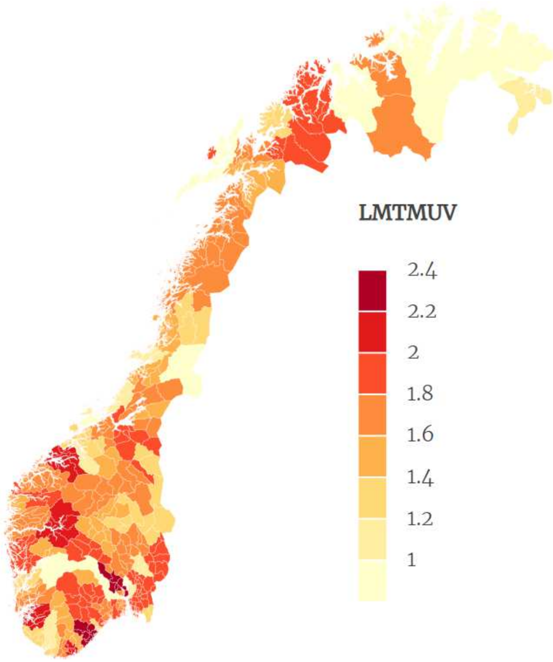
The results for random effects models are not significant and thus cannot support my hypothesis. However, the fixed effects models presented in Table 10 reveal some insight. Model 1 includes all firms and reflects the results from the previous section (the only difference is that the estimators change as results of the scaling). Model 2 includes all manufacturing firms and the respective variety measures.



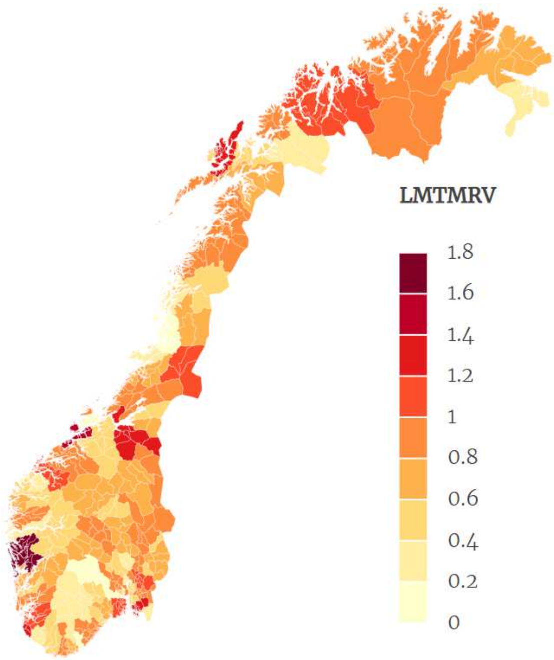
(a) High Tech Unrelated Variety



(b) High Tech Related Variety



(c) Low & Medium Tech Unrelated Variety



(d) Low & Medium Tech Related Variety

Figure 10: High-Tech and Low & Medium Tech Variety 2017

The findings underline that the significantly positive effect of related varieties on firm performance can also be found in the manufacturing sector. Nevertheless, the effect is not as strong as for all sectors. This can possibly be explained by stronger relationships in other industries, such as the service sector. This would be in accordance with the findings of Bosma et al. [2011], who only find positive effects of related variety among service industries. At this point a more detailed investigation would be required. However, the scope of this thesis does not allow for that.

Table 10: Firm level fixed effects on Value-Added growth, using z-scores

	(1) All Sectors	(2) Manufacturing	(3) High Tech	(4) Low & Medium Tech
loglvasq	0.0413^{***}	0.0369^{***}	0.0391^{***}	0.0334^{***}
	(0.000894)	(0.00377)	(0.00610)	(0.00403)
loglva	-1.255^{***}	-1.188^{***}	-1.304^{***}	-1.114^{***}
	(0.0126)	(0.0652)	(0.117)	(0.0665)
stdlunrelvar	0.0250^{**}			
	(0.00795)			
stdlrelvar	0.194^{***}			
	(0.00858)			
stdlmuv		0.00315		
		(0.0225)		
stdlmrv		0.107^{***}		
		(0.0179)		
stdlhtmuv			-0.0120	
			(0.0440)	
stdlhtmrv			0.0954^{*}	
			(0.0483)	
stdllmtmuv				0.00195
				(0.0222)
stdllmtmrv				0.0780^{***}
				(0.0157)
Constant	7.175^{***}	7.621^{***}	8.707^{***}	7.203^{***}
	(0.0467)	(0.289)	(0.569)	(0.280)
Observations	322500	19738	5386	21030
R^2	0.461	0.369	0.404	0.369

Robust standard errors in parentheses. Excluding firms with 5 or less employees
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As a next step, the manufacturing sector is divided into high-tech and medium

& low tech sectors (Model (3) and (4)). The estimators reveal a stronger effect of related variety in high-tech sectors than in non-high-tech sectors. This supports the hypothesis that related variety plays a greater role in knowledge intense industries and furthermore reflects the findings of Hartog et al. [2012], who emphasize the relevance of related variety for high-tech sectors. On the other hand, the model does not provide any evidence for a significant effect of unrelated variety in manufacturing industries.

8 Conclusion

This thesis uses random and fixed effects models to investigate the impact of related and unrelated variety on firms' value-added growth. This section summarizes the main findings by following the hypotheses proposed in the beginning. After that, limitations and problems are discussed.

Looking at all sectors jointly, I find evidence for a positive relationship between related variety and value-added growth on a firm-level. The estimations for the manufacturing sector as a whole and for the high-tech and low & medium tech sectors come to a similar result. This supports the existence of Jacobs externalities. The regional co-location of related sectors facilitates inter-industry learning, which in turn stimulates innovation activity and growth. Besides that, the estimations for unrelated variety offer mostly insignificant results. Only one specification in the fixed effects model suggests a positive effect on firm value-added growth. Hypothesis 2 can therefore not be confirmed. It can be concluded from this that there are no beneficial knowledge spillovers between firms that don't share any similarities. This goes in line with the notion of cognitive proximities as coined by Boschma [2005], after which too different firms are not able to learn from each other. It can be noted here that Castaldi et al. [2015] provide evidence that unrelated variety incites radical innovations in the US. Since radical innovations are rather rare, it is a possibility that the economy of Norway is too small to find such relationships. Furthermore, I find that population density and the degree of centrality in a region have a positive effect on the value-added growth of its firms. The results are significant for all specifications tested and provide evidence for a beneficial effect of urbanization economies on firm performance. After analysing the variety effects on all sectors jointly, I focussed on manufacturing industries. The estimations confirm the hypothesis that related variety has a greater impact on firm performance in knowledge-intense manufacturing industries. This goes in line with the empirical findings of Hartog et al. [2012] for Finland and Bishop and Gripaios [2010] for the UK.

The consistency of the results through different model specifications emphasizes the robustness of the results. Nevertheless, the relationships analyzed in this thesis are complex and difficult to capture. Consequently, my research design is connected to several limitations that have to be addressed. Firstly, the large amount of firm-level data suggested an analysis on firm level. However, the data lacks certain useful variables. First, a more suitable dependent variable would have been innovation activity (measured by number of patents, R&D expenditure or responses to surveys) as it allows to measure a more direct effect of knowledge spillovers resulting from economic variety. Amongst the independent variables it is not possible to measure R&D expenditure. This is a drawback, as Hartog et al. [2012] argue that it 'shows a regions ability to adapt to innovations produced somewhere else'. The consequence of these shortcomings is a low goodness of fit in the models.

Another issue is that the assignment of firms to regions leads to bias. Firms might be located on the border of an economic region, while the research design assumes that they only interact with firms in the region they are assigned to. In other words, any firm interactions that reach across borders are not captured. Furthermore, the data set reveals only one location for each company, even if it operates various plants and departments at different places. Consequently, employees are associated to the location of their firm's headquarter, regardless of their actual location. This leads to a certain degree of bias in the variety measures. Moreover, firms operating in different fields only have one NACE code. Multidivisionality is therefore not captured and might lead to further bias. However, the majority of analysed firms are small in size and therefore I only expect very few firms to operate in different regions and NACE sectors.

Further possible research could address these issues and shortcomings by using different data, variables and estimation methods. Besides that, it could be argued that even though I discussed spillover effects as a channel to stimulate growth, the actual mechanics of 'knowledge spillovers' have not been analysed. Even though Hartog et al. [2012] expect labour mobility as a main source, it has not been addressed in literature regarding related and unrelated variety. This could possibly be done by looking into different industrial clusters and by using survey based data. Additionally, the large data set used in this thesis, covering 280.000 firms over a 10-year time range, offers many opportunities for further research. In particular, it allows to analyse the effects of related and unrelated variety for different sectors and regions. On this basis, one could for example analyse whether the co-location in clusters has a beneficial effect on the growth of the regions.

Understanding these relationships can be important for Norwegian policy makers. The high price level and high wage costs [Aarstad et al., 2016] make it difficult to export industrial goods to international markets. Therefore, it is a challenge for regions to come up with strategies to ensure sustainable economic growth. The findings of this thesis and many research papers suggest that such a long-term strategy would be to encourage firms from related industries to co-locate in a region. This strategy is already pursued in many regions, such as in the petroleum cluster in Rogaland, the mechanical engineering cluster around Kongsberg or in the co-location of furniture and fish processing industries in the north of Norway.

References

Jarle Aarstad, Olav A Kvitastein, and Stig-Erik Jakobsen. Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. *Research Policy*, 45(4):844–856, 2016.

Monica Andini, Guido De Blasio, Gilles Duranton, and William C Strange. Marshallian labour market pooling: Evidence from italy. *Regional Science and Urban Economics*, 43(6):1008–1022, 2013.

Roberto Antonietti and Giulio Cainelli. The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis. *The Annals of Regional Science*, 46(3):577–600, 2011.

Bjørn Asheim, Arne Isaksen, Jerker Moodysson, and Markku Sotarauta. The changing and diverse roles of ris in the globalizing knowledge economy: a theoretical re-examination with illustrations from the nordic countries. *Dynamic Geographies of Knowledge Creation and Innovation*, pages 227–249, 2011.

Mohsen Attaran. Industrial diversity and economic performance in us areas. *The Annals of Regional Science*, 20(2):44–54, 1986.

Mohsen Attaran and Martin Zwick. Entropy and other measures of industrial diversification. *Quarterly Journal of Business and Economics*, pages 17–34, 1987.

Paul Bishop and Peter Gripaios. Spatial externalities, relatedness and sector employment growth in great britain. *Regional Studies*, 44(4):443–454, 2010.

Ron Boschma. Proximity and innovation: a critical assessment. *Regional studies*, 39(1):61–74, 2005.

Ron Boschma and Simona Iammarino. Related variety, trade linkages, and regional growth in italy. *Economic geography*, 85(3):289–311, 2009.

Ron Boschma, Asier Minondo, and Mikel Navarro. Related variety and regional growth in spain. *Papers in Regional Science*, 91(2):241–256, 2012.

Niels Bosma, Erik Stam, and Veronique Schutjens. Creative destruction and regional productivity growth: evidence from the dutch manufacturing and services industries. *Small Business Economics*, 36(4):401–418, 2011.

Carolina Castaldi, Koen Frenken, and Bart Los. Related variety, unrelated variety and technological breakthroughs: an analysis of us state-level patenting. *Regional studies*, 49(5):767–781, 2015.

Jeroen Content and Koen Frenken. Related variety and economic development: a literature review. *European Planning Studies*, 24(12):2097–2112, 2016.

Nicola Cortinovis and Frank Van Oort. Variety, economic growth and knowledge intensity of european regions: a spatial panel analysis. *The Annals of Regional Science*, 55(1):7–32, 2015.

Michiel De Bok and Frank Van Oort. Agglomeration economies, accessibility, and the spatial choice behavior of relocating firms. *Journal of Transport and Land Use*, 4(1):5–24, 2011.

Pınar Falcıoğlu. Location and determinants of productivity: The case of the manufacturing industry in turkey. *Emerging Markets Finance and Trade*, 47(sup5): 86–96, 2011.

Koen Frenken, Frank G van Oort, Thijs Verburg, and Ron A Boschma. Variety and regional economic growth in the netherlands. 2005.

Koen Frenken, Frank Van Oort, and Thijs Verburg. Related variety, unrelated variety and regional economic growth. *Regional studies*, 41(5):685–697, 2007.

Edward L Glaeser, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer. Growth in cities. *Journal of political economy*, 100(6):1126–1152, 1992.

Matte Hartog, Ron Boschma, and Markku Sotarauta. The impact of related variety on regional employment growth in finland 1993–2006: high-tech versus medium/low-tech. *Industry and Innovation*, 19(6):459–476, 2012.

Jane Jacobs. *The economy of cities*. Vintage Books, 1969.

Alexis P Jacquemin and Charles H Berry. Entropy measure of diversification and corporate growth. *The journal of industrial economics*, pages 359–369, 1979.

Stig-Erik Jakobsen, Grete Rusten, and Arnt Fløysand. How green is the valley? foreign direct investment in two norwegian industrial towns. *Canadian Geographer/Le Géographe canadien*, 49(3):244–259, 2005.

Sandra Kublina. Related variety, unrelated variety and regional growth: Development patterns and role of absorptive capacity. 2015.

Francesca Mameli, Simona Iammarino, and Ron Boschma. Regional variety and employment growth in italian labour market areas: services versus manufacturing industries. 2012.

Alfred Marshall. Principles of economics (london, 1920). *Book VI*, pages 618–19, 1920.

Gunnar Nerheim. *Norsk oljehistorie: En gassnasjon blir til. B. 2*. Leseselskapet, 1996.

Luigi Pasinetti. *Structural economic dynamics*. Cambridge University Press, 2006.

Francesco Quatraro. Knowledge coherence, variety and economic growth: Manufacturing evidence from italian regions. *Research Policy*, 39(10):1289–1302, 2010.

Bjørnar Sæther, Arne Isaksen, and Asbjørn Karlsen. Innovation by co-evolution in natural resource industries: The norwegian experience. *Geoforum*, 42(3): 373–381, 2011.

Amir Sasson and Atle Blomgren. Knowledge based oil and gas industry. 2011.

Pier Paolo Saviotti and Koen Frenken. Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, 18(2):201–218, 2008.

Claude Elwood Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

Robert M Solow. A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94, 1956.

Sam Tavassoli and Nunzia Carbonara. The role of knowledge variety and intensity for regional innovation. *Small Business Economics*, 43(2):493–509, 2014.

Gerben Van der Panne. Agglomeration externalities: Marshall versus jacobs. *Journal of Evolutionary Economics*, 14(5):593–604, 2004.

Sofia Wixe. The impact of spatial externalities: Skills, education and plant productivity. *Regional studies*, 49(12):2053–2069, 2015.

Appendices

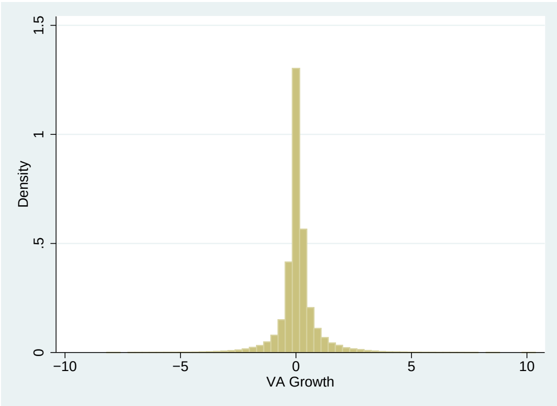
```
. xtdescribe, patterns(20)

  orgnr:  8.100e+08, 8.100e+08, ..., 9.997e+08              n =     280374
   year:  2008, 2009, ..., 2017                              T =         10
          Delta(year) = 1 year
          Span(year)  = 10 periods
          (orgnr*year uniquely identifies each observation)

Distribution of T_i:   min      5%      25%       50%       75%     95%     max
                         1       1        3         8        10      10      10

     Freq.  Percent    Cum. |  Pattern
 ---------------------------+------------
   124814     44.52   44.52 |  1111111111
    24428      8.71   53.23 |  .........1
    21995      7.84   61.07 |  ........11
    19251      6.87   67.94 |  .......111
    17585      6.27   74.21 |  ......1111
    16557      5.91   80.12 |  .....11111
    15392      5.49   85.61 |  ....111111
     9582      3.42   89.03 |  ...1111111
     7504      2.68   91.70 |  ..11111111
     7219      2.57   94.28 |  .111111111
     1286      0.46   94.74 |  111111111.
      791      0.28   95.02 |  1.........
      648      0.23   95.25 |  11.1111111
      630      0.22   95.47 |  1.11111111
      578      0.21   95.68 |  11111111.1
      575      0.21   95.88 |  1111.11111
      575      0.21   96.09 |  111111.111
      565      0.20   96.29 |  11........
      564      0.20   96.49 |  1111111.11
      537      0.19   96.68 |  111.111111
     9298      3.32  100.00 | (other patterns)
 ---------------------------+------------
   280374    100.00         |  XXXXXXXXXX
```

Figure 11: Panel data set description taken from Stata

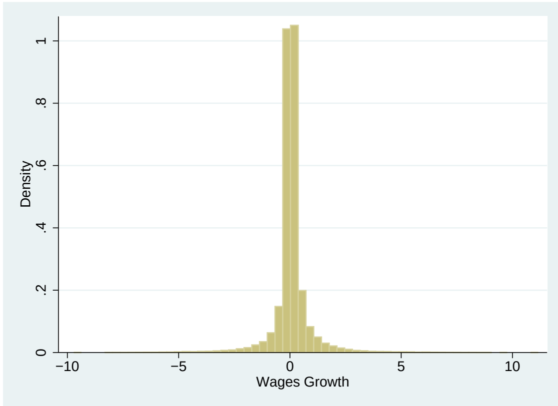
Table 11: Random Effects, excl. 4 largest agglomerations and firms with 0 employees

	(1) lndva	(2) lndva	(3) lndva	(4) lndva
loglvasq	0.0539***	0.0538***	0.0540***	0.0540***
	(0.000575)	(0.000576)	(0.000575)	(0.000574)
loglva	−1.058***	−1.057***	−1.054***	−1.054***
	(0.00820)	(0.00820)	(0.00819)	(0.00819)
lunrelvar	0.00441	−0.00326	0.00273	0.00287
	(0.00735)	(0.00738)	(0.00736)	(0.00733)
lrelvar	0.0368***	0.00830	0.0218*	0.00139
	(0.00705)	(0.00841)	(0.00853)	(0.00875)
density		0.0154***	0.0150***	
		(0.00149)	(0.00189)	
north		0.0532***	0.0544***	0.0780***
		(0.00686)	(0.00691)	(0.00740)
west		0.00653	0.00822	0.0315***
		(0.00563)	(0.00560)	(0.00599)
educ			−0.000346	−0.00150***
			(0.000350)	(0.000366)
oilshock			−0.0194***	−0.0186***
			(0.00162)	(0.00162)
centrality				0.000372***
				(0.0000315)
Constant	4.640***	4.623***	4.585***	4.400***
	(0.0367)	(0.0370)	(0.0379)	(0.0397)
R^2 within	0.425	0.425	0.423	0.423
R^2 between	0.0924	0.0924	0.0921	0.0922
R^2 overall	0.124	0.124	0.125	0.125

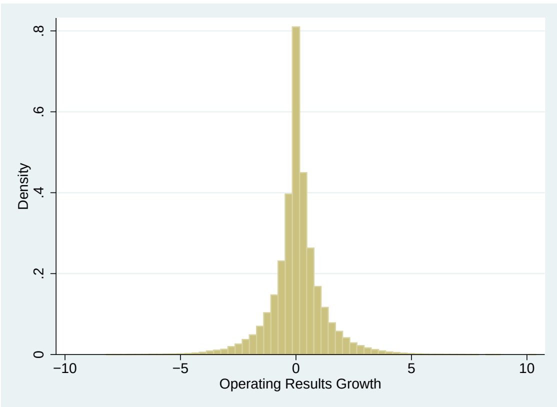
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$



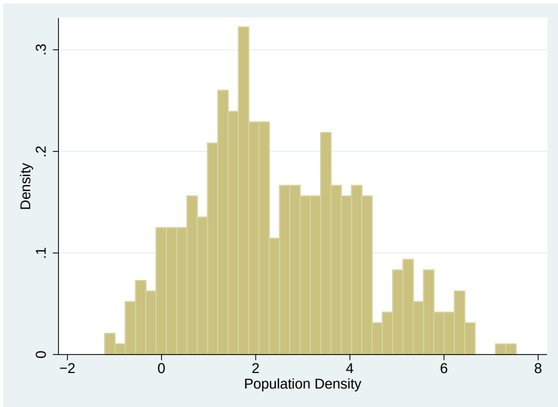
(a) Value-Added



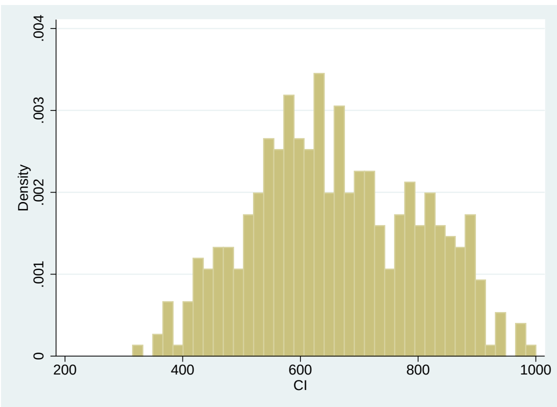
(b) Wage Costs



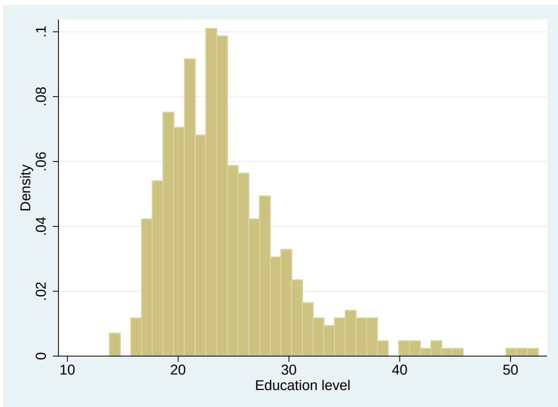
(c) Operation Profits



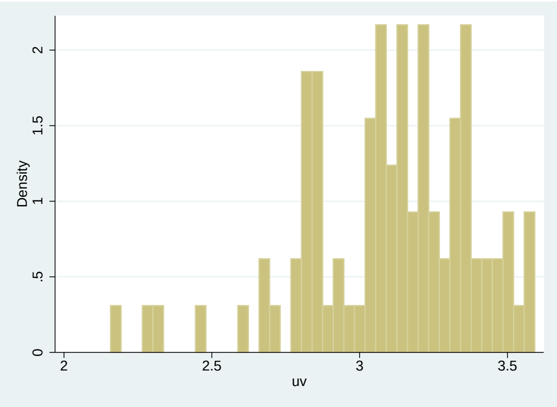
(d) Population Density

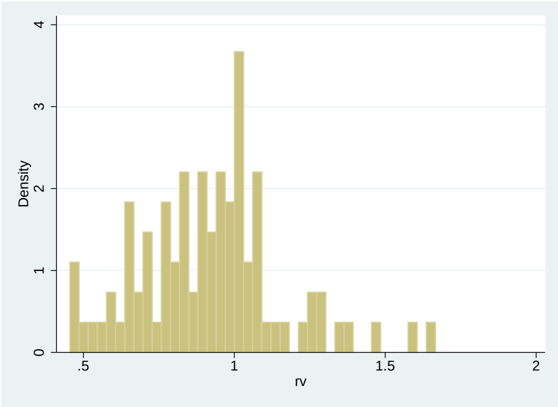


(e) Centrality Index



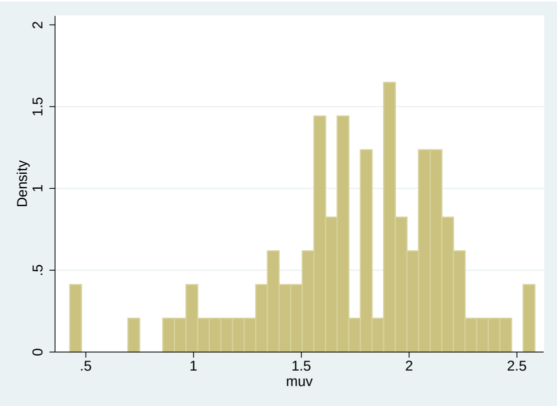
(f) Education Level

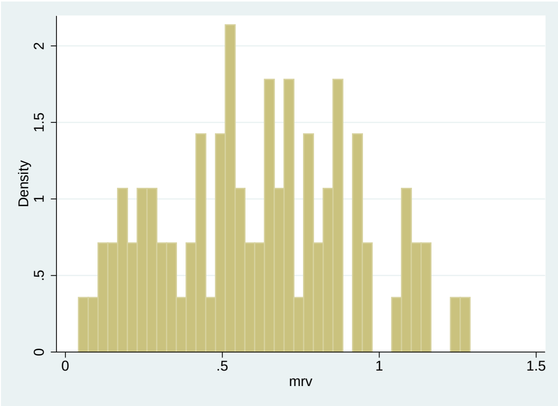




(a) Unrelated Variety 2017

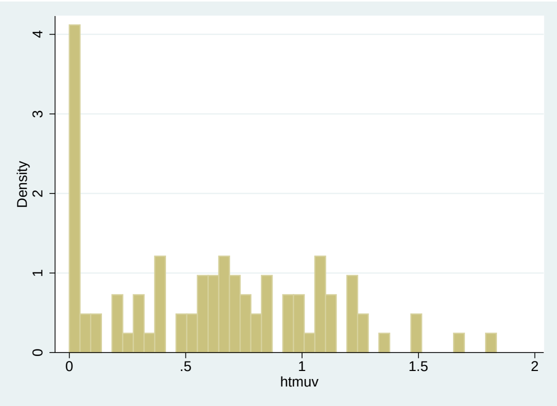
(b) Related Variety 2017

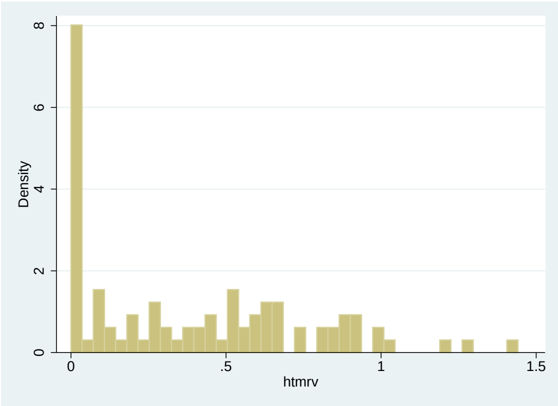




(c) Manufacturing Unrelated Variety 2017

(d) Manufacturing Related Variety 2017





(e) High-Tech Manufacturing Unrelated Variety 2017

(f) High-Tech Manufacturing Related Variety 2017

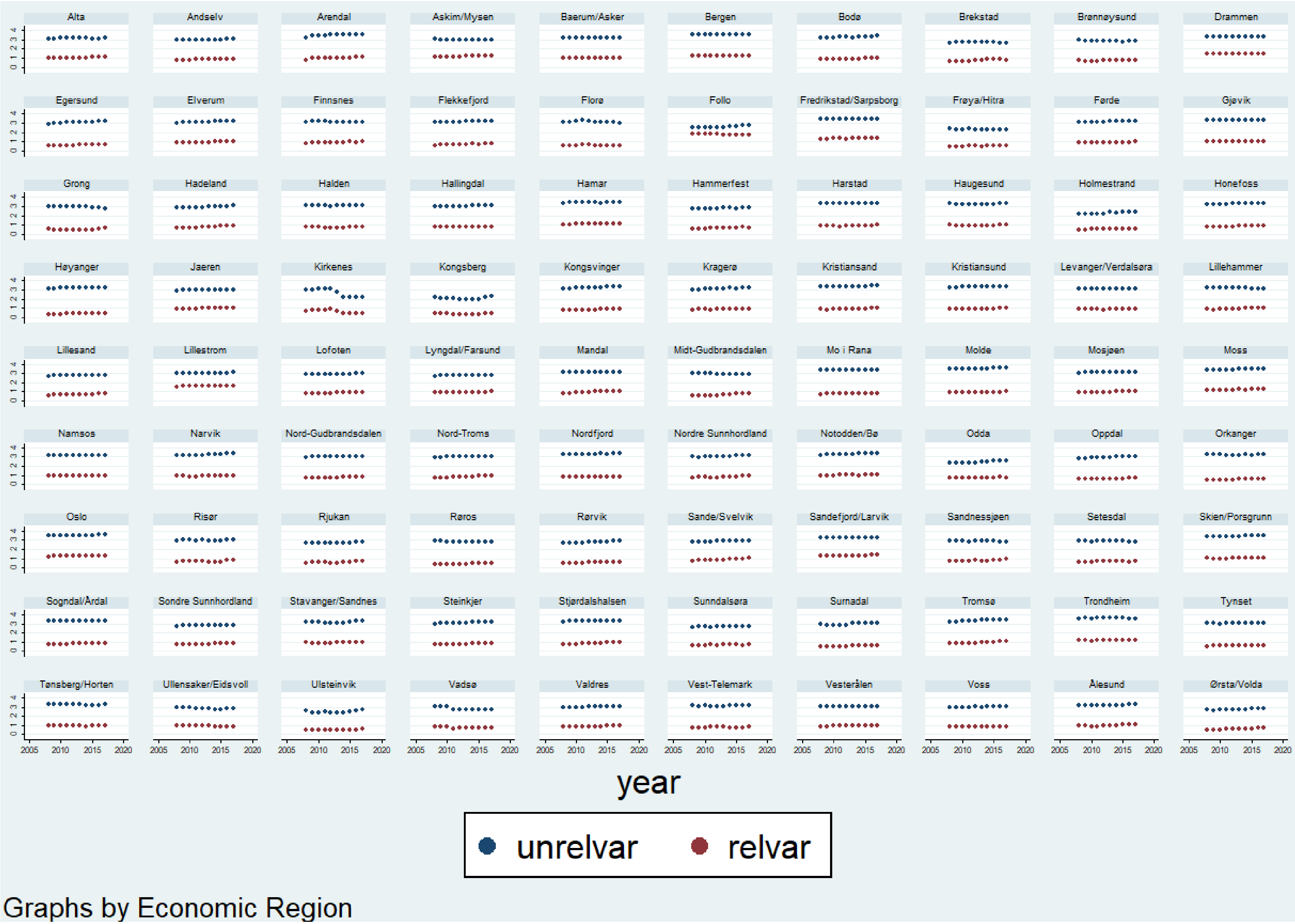


Figure 14: Related and Unrelated Variety by Economic Region

Figure 15: Descriptive Statistics Standardized Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
stdlunrelvar	1607856	0	1	-4.649358	1.197072
stdlrelvar	1607856	0	1	-2.843335	2.731589
stdlmuv	137850	0	1	-3.681685	1.517091
stdlmrv	137850	0	1	-2.869806	2.10038
stdlhtmuv	1603953	0	1	-2.221457	2.08066
stdlhtmrv	1603953	0	1	-1.763839	2.112441
stdllmtmuv	1603953	0	1	-3.517278	1.518494
stdllmtmrv	1603953	0	1	-2.621433	2.475732