



LUND UNIVERSITY
School of Economics and Management

Department of Economics

Commuting patterns, infrastructure and the location of economic activity

*A study of sector concentration and
municipality specialisation in Sweden*

NEKP01: Master essay II

Spring 2012

Author: Cecilia Carlsson

Supervisor: Joakim Gullstrand

Abstract

This thesis investigates the impact of commuting and the quality of infrastructure on the degree of sector concentration and municipality specialisation in Sweden. Data on seven sectors and 286 Swedish municipalities for the years 2001 and 2003 is used in the empirical study performed using Ordinary Least Squares (OLS). The robustness of the results is tested with respect to changes in the variable-transformation method as well as the measurement of some of the included variables. Although partly affected, the obtained results in general seem to be qualitatively quite robust to the performed changes.

Both the direction and the magnitude of commuting influence the concentration of economic activity, whereas the degree and type of municipality specialisation is mainly affected by the direction of the occurring commuting flows. In general, however, the results indicate that interactions between municipality and industry characteristics have a larger impact on the type of municipality specialisation than on the sector concentration of economic activity. Partly in contrast with theoretical predictions, sectors for which links to other producers are of large importance tend to be relatively less concentrated in municipalities characterised by large market potential. Such municipalities also tend to be less specialised in sectors for which links to consumers and producers are of relatively large importance.

Key words: commuting, infrastructure, concentration, specialisation, New Economic Geography, OLS

List of abbreviations

CBD	Central Business District
CES	Constant elasticity of substitution
EG	Ellison-Glaeser index
FE	Fixed effects-method
GC	Gini coefficient
GDP	Gross Domestic Product
GRP	Gross Regional Product
NEG	New Economic Geography
NLS	Non-linear Least Squares
R^2	Coefficient of determination
SCB	Statistiska Centralbyrån (Statistics Sweden)
SKL	Sveriges Kommuner och Landsting (Swedish Association for Local Authorities and Regions)
VIF	Variance Inflation Factor

Table of contents

1. Introduction	7
2. Theoretical foundation	9
2.1. The importance of factor endowments	9
2.2. The standard New Economic Geography-model	10
2.3. Commuting in a New Economic Geography-model	11
2.4. Infrastructure in a New Economic Geography-model	13
2.5. Implications for this thesis	15
3. Earlier empirical studies	17
3.1. The importance of natural advantages and spillover effects	17
3.2. The importance of factor endowments and links between economic agents	18
3.3. The importance of infrastructure	21
4. Methodology	22
4.1. From theory to empirics	22
4.2. Incorporating commuting and infrastructure into the model	23
4.3. Estimation method	23
4.4. Econometric problems and countermeasures	25
5. Data and empirical specification	27
5.1. Variable overview	27
5.1.1. Dependent variables	27
5.1.2. Variables to capture size differences	28
5.1.3. Municipality characteristics	29
5.1.4. Industry characteristics	32
5.1.5. Interaction variables	34
5.2. Sample	36
5.3. Empirical specification	38
6. Results	39
6.1. Basic specification	39
6.1.1. Sector concentration	39
6.1.2. Municipality specialisation	42
6.2. Robustness tests	45
6.2.1. Alternative measure of commuting	45
6.2.1.1. Sector concentration	45
6.2.1.2. Municipality specialisation	47

6.2.2.	Employment as the dependent variable	49
6.2.2.1.	Sector concentration	49
6.2.2.2.	Municipality specialisation	51
6.2.3.	Standardisation as variable transformation method	53
6.2.3.1.	Sector concentration	53
6.2.3.2.	Municipality specialisation	56
7.	Discussion	59
8.	Conclusion	62
9.	References.....	63
10.	Appendices.....	66
10.1.	Appendix 1: Descriptive statistics	66
10.2.	Appendix 2: Industry intensities.....	67
10.3.	Appendix 2: Alternative estimation results.....	68

List of tables

Table 5.1. Dependent variables28

Table 5.2. Municipality characteristics.....32

Table 5.3. Industry characteristics.....34

Table 5.4. Interaction variables36

Table 6.1. Concentration using normalised variables41

Table 6.2. Specialisation using normalised variables.....44

Table 6.3. Concentration using the alternative commuting variable.....46

Table 6.4. Specialisation using the alternative commuting variable.....48

Table 6.5. Concentration using the alternative dependent variable50

Table 6.6. Specialisation using the alternative dependent variable.....52

Table 6.7. Concentration using standardised variables55

Table 6.8. Specialisation using standardised variables57

Table A1. Descriptive statistics.....66

Table A2. Industry intensities67

Table A3. Concentration using normalised variables.....68

Table A4. Specialisation using normalised variables.....70

1. Introduction

In a report written for the Swedish Ministry of Finance, Eliasson et al. (2007, p.94-98) extensively portrait and analyze the trends and development of migration and commuting within Sweden during the years 1974-2005. The authors show that the share of workers that commute to reach their working place in another region has increased from about 18 percent in 1974 to 30 percent in 2004.¹ Potential explanations to this noticeable development are the rapid technological development in the area of infrastructure and communications together with the increase in the supply of public transportation in many regions.

It seems reasonable to believe that the tendency of an increasing propensity for people to commute have consequences for a wide range of phenomena in the economy. Focusing on one of these, this thesis aims at investigating the impact of commuting patterns on the location of economic activity. Several potential aspects matter for this relationship. For example, one could argue that if people consume a share of their income in their home municipality, commuting patterns can affect location decisions for industries that want to locate close to its consumers. Furthermore, differences in commuting patterns can be seen as an indication of differences in the quality of infrastructure between municipalities, thereby affecting the location of production by inducing differences in the cost of production between different regions.

The research question at issue in this thesis can therefore be formulated as follows:

- What is the impact of differences in commuting patterns on the location of economic activity?

The reasons for this topic being of interest are manifold. First, the degree of concentration of production naturally affects the distribution of income across regions, and the aim to promote regional development and to act against regional divergence of income is on the agenda of many regional, national and international policymakers. As an example, regional spending for the years 2007-2013 amounts to one third of the total EU budget, which is equivalent to approximately 350 billion euro (EU1, 2012-03-02). Second, concentration of economic activity in turn has consequences for the economic performance of countries and regions. For example, estimations by Ciccone and Hall (1996) indicate that higher employment density is associated with higher labour productivity, and Baldwin and Forslid (2000) theoretically show that concentration of production can be favourable to economic growth, although the higher growth rate comes at the cost of some people in society being worse off.

¹ Commuting is measured as the share of employed workers commuting over an administrative border (municipality or region).

Third, the degree, or even the absence, of labour mobility in terms of migration is a discussed topic within theoretical models as well as in debates about labour market-related issues. Although commuting is not as extensive as migration, it should be of interest to assess the importance and consequences also of this type of labour mobility.

This thesis aims at assessing the impact of commuting patterns on the location of economic activity in terms of both sector concentration and municipality specialisation. To do so, the performed empirical study makes use of a sample consisting of data covering seven sectors and 286 Swedish municipalities. However, due to limitations when it comes to the amount of available data, the data set consists of two years only.

The disposition of the thesis is as follows. Section 2 presents relevant theoretical models and section 3 gives an overview of performed empirical studies. Section 4 and 5 contain a description of the methodology and the data, respectively. The results of the performed study are presented in section 6 and discussed in section 7. Section 8 concludes the thesis.

2. Theoretical foundation

The following section aims at presenting the theoretical foundation forming the basis for the empirical study. First, the main insights of the Heckscher-Ohlin theory regarding the importance of factor endowments are presented. Thereafter, a New Economic Geography-model aiming at explaining the emergence of central and peripheral regions is presented. Finally, two extensions of this model, which incorporate commuting and infrastructure, respectively, are described.

2.1. The importance of factor endowments

One of the core paradigms in traditional trade theory is the Heckscher-Ohlin model, which explains trade and production patterns by differences in the supply of production factors between countries.²

In its most basic version, the Heckscher-Ohlin model consists of two countries, two goods and two factors of production. It is assumed that the goods are produced in a perfectly competitive environment and under constant returns to scale. Furthermore, preferences are the same in both countries and factors of production are mobile within sectors in a country but not between countries.

In the Heckscher-Ohlin model, it is the relative, rather than the absolute, supply of factors of production that determine the relative price of these factors. Also, the production technology is the same in all countries but differ between sectors and the sector using much of one factor in relation to the other is intensive in that factor. By this reasoning, one country is always abundant in one of the production factors while the other country is abundant in the other factor, even if one country has a larger supply of both production factors in absolute terms. Similarly, if one sector is intensive in one particular factor of production, the other sector must be intensive in the other factor, also when one sector requires more input of both factors in absolute terms.

Since technology is assumed to be the same in both countries, it is the interaction between the countries' abundance of factors of production and the intensity of factors used in production that determines the production pattern. For each good, production will be concentrated in the country that is abundant in the production factor intensively used in the production of the good, but concentration will not be total. Under the assumption of perfect competition, both countries will still produce both goods, although in different proportions.

² This section builds on Senior Nello (2005), chapter 4, which gives a brief introduction and Krugman and Obstfeld (2009), who present a more extensive overview.

This basic set-up of the Heckscher-Ohlin model has been modified in many different ways. Extensions include versions expanding the model to include a continuum of goods as in Dornbusch et al (1980), adding features of variable returns to scale in production as in Young (1991) and incorporating the Heckscher-Ohlin theory into models of New Economic Geography³ as in Epifani (2005). For the purpose of this thesis, however, the insights of the basic Heckscher-Ohlin model regarding the importance of factor abundance and factor intensity in production are sufficient.

2.2. The standard New Economic Geography-model

In an influential article, Krugman (1991, p. 485-490) develops a theoretical two-region, two-sector model to analyze the driving forces behind the partition of the economy into a core- and a periphery region. The two sectors are agriculture and manufacturing, and it is assumed that production in the agricultural sector exhibits constant returns to scale and is intensive in the immobile factor land, whereas the manufacturing sector uses relatively little land in production and is characterised by increasing returns to scale. Also, of large importance is the assumption that transport costs of agricultural products are assumed to be equal to zero, whereas transporting the manufacturing output implies transport costs of the iceberg form, so that a share of the value of the product “disappears” through transportation. In addition, in the long run, labour can move between regions, whereas peasants cannot, and in their decision of whether to move or not, labour care about real wages. Furthermore, the existence of pecuniary externalities in terms of supply- and demand-linkages, rather than technological spillover effects are assumed to be of crucial importance for the location decision of firms, and it is the implied lower transport costs of locating close to demand and supply that make these pecuniary externalities of large importance.

By developing a theoretical model based on a Dixit and Stiglitz-monopolistic competition setting, Krugman (1991, p. 490-493) is able to identify some of the crucial features determining the development of the economy when it comes to the settlement of labour and thereby the concentration of production of the sector characterised by increasing returns to scale. In the short run, labour is immobile and cannot move between regions. In this situation, there are two opposing effects in action affecting the relative wages of labour. One is the so called home market-effect, which implies a higher demand for labour in the concentrated region. This in turn implies that wages will also be higher. On the other hand, the second effect, called the competition effect, suggests that in the periphery region, labour is scarce and therefore its price, the wage, is higher. In the long run, however, when labour is mobile, it will care about which region offers the highest real wage. Since a larger supply of the

³ The basic and more elaborate New Economic Geography-models are of large interest for this thesis, but the most relevant extensions focus on the importance of infrastructure and commuting rather than on the importance of factor endowments. See section 2.2, 2.3 and 2.4 for a discussion of the models of interest.

manufacturing good in the concentrated region implies a lower price of this good, the real wage will be higher in the concentrated region. Hence, this price effect together with the home market effect therefore promote divergence across the two regions, while the competition effect works in favour of convergence.

Moreover, Krugman (1991, p. 494-498) analyses the stability of an equilibrium characterised by all workers being concentrated to one region and is able to show that the likelihood of such an outcome is dependent of some characteristic features of the economy. First, of large importance is the size of transport costs. It is argued that transport costs of intermediate size are favourable for divergence and a relatively uneven location of economic activity, since costless transport of the produced goods makes the manufacturing firm indifferent to locating in one region or the other and high transport costs makes it costly for the firm to concentrate its production to one region only. Hence, both low and high transport costs therefore act to promote convergence and a relatively even location of economic activity, across regions. Second, also the share of income spent on the manufacturing good is of importance. The larger this share, the more likely that divergence will occur since a high demand for the manufacturing product makes it more profitable for the firm to locate close to the large market in the concentrated region. Third, the higher the elasticity of substitution between goods, the smaller is the scope for firms to make use of internal economies of scale. Therefore, the higher this substitutability, the more likely it is that the economy is characterised by convergence.

Hence, to conclude, an outcome characterised by divergence is more likely to occur when transport costs are of medium size, when the share of income spent on the manufacturing good is large and when the elasticity of substitution is low.

2.3. Commuting in a New Economic Geography-model

Gruber (2010) builds on Krugman (1991) when analysing the impact on the location of production and people when labour is not only able to move between regions but can also choose to commute between them.

The setup of the model is somewhat different to previous models that aim at incorporate commuting into a New Economic Geography-model. These include Tabuchi and Thisse (2006), whose setting implies not only that individuals can commute to work, but that they are in fact forced to. In the model, each individual “consumes” one unit of immobile land and commute to work in the central business district (CBD). In this setting, high transport costs are associated with a high degree of agglomeration of production.

Also Murata and Thisse (2005) model the importance of commuting costs and their interaction with transport costs and are able to show that, as opposed to the “standard” result in New Economic Geography-models, high transport costs, when combined with low commuting costs, can result in agglomeration.

Ottaviano et al. (2002) analyse the importance of transport and commuting costs, and are able to show that high commuting costs are associated with a lower degree of optimal agglomeration. In addition, commuting costs combined with low enough transport costs is favourable to divergence between regions (Gruber, 2010, p.110-113)

Returning to the model by Gruber (2010, p. 114-119), there are a number of important assumptions that should be emphasised. Just as in the setting of Krugman (1991) described above, it is assumed that the economy consists of two regions. To allow the incorporation of commuting patterns, there are, however, three sectors in the modified New Economic Geography-model. Just as in Krugman (1991) there is one agricultural good which is costlessly tradable and which is produced using immobile land. Furthermore, production of the agricultural good is characterised by constant returns to scale and takes place in a competitive environment. There is also one manufacturing good, which is costly to transport and therefore implies costs of the iceberg type. Production of this good is intensive in unskilled labour and land and is produced using a constant elasticity of substitution-technology (CES) in a monopolistically competitive environment. The third good is non-tradable services, which use skilled labour intensively in production and which are produced with constant returns to scale in a competitive environment. Also, workers spend a fraction of their income on services in the region where they work, and the remaining share of consumption of services is undertaken in the region of residence.

Naturally, of large importance is also the assumption of workers being able not only to move between regions, but also to commute between them. What choice individuals will make when it comes to the question of moving or commuting depends decisively on the costs of commuting and on the cost of housing, which affect the net real wage. Also, it is assumed that commuting is costly and that the costs are of iceberg form, just as the transport costs for the manufacturing good described above. This setting leaves the worker with the choice of three potential situations: she could either (i) live and work in one region, (ii) live in one region, work in the other region and commute to work or (iii) move to work in the other region.⁴ Furthermore, three different scenarios are analysed depending on the assumptions regarding the mobility of workers. These are characterised by 1) only unskilled labour

⁴ There is the fourth opportunity for the individual to move to the other region and commute to work in the region she just emigrated from, but this scenario is disregarded by the author as it is not feasible within the current setup of the model.

being mobile, 2) only skilled labour being mobile or 3) both unskilled and skilled labour being mobile. In all three settings, however, both unskilled and skilled labour can commute to work.

Empirical research has found that the propensity to move and commute tend to increase with the education level.⁵ The corresponding scenario in the model by Gruber (2010, p.120-132) would thus be the second, in which skilled labour is mobile. In this setting, the model predicts strong agglomeration of mobile skilled labour into one region being the only stable outcome of the model. However, economic activity is not fully concentrated in one region. This is due to the fact that unskilled workers, who cannot move between regions, still spends a share of their income on the non-tradable services in their home region.

The scenario in which only unskilled labour, which is intensively used in the production of the manufacturing good, can migrate results in an outcome similar to the model by Krugman (1991). Gruber shows that in this setting, the two regions tend to diverge for intermediate values of transport and commuting costs. For the third scenario, in which all workers are mobile, the outcome is characterised by strong, but not full, agglomeration of labour into one region and skilled workers both agglomerate and commute to a larger extent than unskilled labour.

In summary, the model by Gruber (2010) enables an analysis of the effects of commuting possibilities for the agglomeration and location of production and the concentration of people's place of living. There is a relatively strong tendency of agglomeration of labour into one region, but the possibility to commute and the demand for non-tradable services results in some economic activity remaining in the periphery.

2.4. Infrastructure in a New Economic Geography-model

Martin and Rogers (1995, p. 336-339) model the impact of infrastructure in a New Economic Geography-model and explicitly discuss the relation and interaction between domestic and international infrastructure. The term infrastructure is here used in a wide sense, comprising not only transport and telecommunications but also, for example, institutions such as the judicial system and law and order. The authors stress that in contrast to the model by Krugman (1991), in which transport costs are associated with international trade, the quality of infrastructure in the model by Martin and Rogers affects trade both between and within countries. Still, however, costs related to infrastructure take the iceberg form described above.

⁵ See for example Eliasson et al. (2007), p. 82-103.

Previous models have incorporated infrastructure by including it as a factor of production in the production function, but modelling the importance of infrastructure in this way does not allow the assessment of its importance for regional integration. In the model by Martin and Rogers (1995) described below, the quality of infrastructure is therefore assumed to affect the cost of a good, and is of the, by now known, iceberg form. By acting through this cost channel, infrastructure affects economic activity mainly through its impact on the price of the good and thereby on the demand.

In the two-country-, two-good-model developed by Martin and Rogers (1995, p. 339-344), one good is produced with constant returns to scale and acts as a numeraire. Furthermore, this good does not exhibit any infrastructure costs.⁶ The other good is a composite good consisting of a number of differentiated products, all of which are produced with identical technologies using capital and labour, and where production requires a fixed amount of capital for each good. The location of production affects the infrastructure costs associated with the good, since production abroad implies that both foreign and domestic infrastructure costs matter. Also, it is assumed that labour can move between sectors but not between countries.

Of large importance for the analysis of the model is the distinction and interaction between domestic and international infrastructure. Examples of the former include the effectiveness of public administration and transport infrastructure that ease trade within a country, whereas examples of the latter are the effectiveness of trade administration between countries and the building and/or existence of a harbour or international airport.

The authors show that in autarky when no trade takes place between countries, initial capital endowments determine the location of production. However, following an improvement in international infrastructure, for example in terms of trade integration, firms will tend to locate in countries where domestic infrastructure is of good quality. The reason is that goods produced in the country with domestic infrastructure of relatively less good quality will be charged a higher price. Therefore, demand for this domestic product will be lower and demand for the foreign product produced under better infrastructure conditions and with lower costs, will be higher. To take advantage of the high demand and to be able to benefit from the returns-to-scale technology characterising production, firms will move to the country that provides good-quality infrastructure.

This relocation is facilitated by international infrastructure being of good quality. Hence, the propensity to relocate, and thereby the likelihood of production being concentrated, is larger when international infrastructure is of good quality. In contrast, in the case of relatively bad-quality

⁶ This set-up of the model is similar both to the model by Krugman (1991) and Gruber (2010). See sections 2.2 and 2.3 above.

international infrastructure, firms will choose to locate near the demand for their products and concentration will not be as extensive as in the case of good-quality international infrastructure. Furthermore, concentration is more likely to occur in the country with the largest initial stock of capital. In this country, competition is fiercer, but the negative impact of the existence of competing firms is compensated by the higher demand for the produced resulting from a higher income level.

In summary, therefore, economic activity tends to concentrate in countries where domestic infrastructure is of good quality, in which income is high and in which markets are large, and the propensity to agglomerate increases with the quality of international infrastructure.⁷

2.5. Implications for this thesis

The reviewed theoretical models above form the basis for the empirical study performed in section 6. Building on the theory, the study will incorporate the interaction between production technology and the supply of production factors as predicted by the Heckscher-Ohlin theory and the importance of closeness to consumers and other producers for sectors characterised by strong forward- and backward linkages as emphasised by Krugman (1991).

Including the consumer link together with a variable measuring the degree of commuting will also enable the study to investigate the prediction of the model by Gruber (2010) of commuting together with consumers' consumption patterns resulting in some sectors being less concentrated than others.

Furthermore, it seems reasonable to believe that the degree of commuting is related to the quality of the available infrastructure, and commuting can therefore also be seen as a proxy for infrastructure in this setting. Acknowledging that infrastructure is a broad concept covering not only aspects related to the transportation of people but also for example telecommunications, electricity and the legal system, it is evident that treating commuting as a proxy for infrastructure cannot capture all relevant aspects. However, due to it being such a broad concept, it is not an easy task to appropriately capture the characteristics of infrastructure in usable data. Commuting can therefore act as an indicator of the overall level of infrastructure in a certain geographical region. In addition, since the aim of this study is to assess the concentration of economic activities and not to in detail investigate the properties of different kinds of infrastructure, the choice of commuting as a proxy for infrastructure seems acceptable for the purpose of the study within this theses. When given this interpretation, including a variable measuring the degree of commuting on the municipality level can thereby also be used to test

⁷ For a discussion about the impact of cross-border infrastructure on regional inequalities, see for example Puga (1998).

the prediction of the model by Martin and Rogers (1995) of economic activity tending to concentrate in municipalities characterised by good-quality infrastructure.^{8 9}

⁸ In the current setting, municipality-level infrastructure and national infrastructure would correspond to the distinction between international and domestic infrastructure as in Martin and Rogers (1995).

⁹ Accounting for the importance of factor endowments, links between economic agents and the importance of infrastructure broadly corresponds to the classification of the driving forces behind the location of economic into first-, second- and potentially also third nature geography. See for example Overman et al (2001) or Naudé (2009) for a description of this classification.

3. Earlier empirical studies

The following section aims at presenting previous empirical research of relevance for this thesis.

During the research process, no studies directly assessing the impact of commuting patterns on the location of economic activity have been found. The following section therefore presents empirical studies performed within related research fields. Several of these studies also use a similar methodology and empirical specification to the one that will be employed in the study within this thesis.

First, studies assessing the importance of natural advantages and knowledge spillovers for the concentration of economic activity are presented. Second, several studies investigating the importance of factor endowments and links between economic agents are reviewed and finally, a study examining the importance of infrastructure for economic activity is briefly presented. Naturally, due to space and time limits, the following section is not a comprehensive description but rather an overview of some empirical findings related to the empirical study to be performed within this thesis.

3.1. The importance of natural advantages and spillover effects

Ellison & Glaeser (1999, p. 313-316) aim at investigating if and to what extent agglomeration patterns and the concentration of economic activity can be attributed to differences in natural advantages, that is to the endowments of for example farmland and timberland and to variables such as unionisation and prices of electricity and natural gas. The impact of such natural advantages on the state share of industry employment is estimated using Non-linear Least Squares (NLS) using data on state employment shares in manufacturing industries on the four-digit level. The model specification includes 16 interactions of variables representing the endowment of natural advantage of the state and the intensity of the industry in that factor. The results vary somewhat depending on the exact specification of the model and the number of included dummy variables for two-digit- and three-digit industries, but the authors conclude that about 20 percent of the observed and measured concentration of production can be ascribed to only a small number of explanatory variables capturing natural advantages. This set of explanatory variables is not comprehensive, however, and the authors argue for the case of at least 50 percent of the geographical concentration of industries being due to natural advantages.

Braunerhjelm and Johansson (2003, p. 58-61) investigate the development over time and the determinants of the spatial concentration of manufacturing and services in Sweden using data on all establishments in 1975 and 1993. Using the so called Ellison-Glaeser (EG) index and the Gini-coefficient (GC), respectively, as the dependent variable, the obtained estimation results show that

manufacturing has become more concentrated over time, whereas the service sector has become more widely spread. Furthermore, the results indicate that spatial concentration is positively affected by plant size, but, as opposed to what one might expect, the results reveal no indication of concentration being more pronounced in knowledge-intensive industries for which spillover effects are believed to be of relatively large importance.

3.2. The importance of factor endowments and links between economic agents

Midelfart-Knarvik et al. (2000, p. 31-32) aim to investigate the development of the location of production over time and to assess what forces in the economy are the main determinants of the location of production. The authors stress the importance of the interactions of these country- and industry characteristics as the determinants of the location of production. Although country- and industry characteristics are included also as separate variables, the interactions, which capture variations in the two dimensions represented by country- and industry characteristics, are the variables of main interest in the performed empirical study. Guided by theory and the statistical significance of their estimates, they argue that four country characteristics and six industry characteristics are of largest interest. The former are market potential, agricultural production, education and researchers and scientists, whereas the latter consist of input and output linkages, economies of scale, agricultural input, non-manual workers in production and the R&D intensity.

The data covering 36 industries in the manufacturing sector in 14 EU countries between 1970 and 1997 is divided into time periods consisting of four years. For each time period, the impact of the interaction variables is estimated using OLS¹⁰ with the logarithm of the country share of total activity of the specific industry as the dependent variable.

Depending on the time period in question, the included variables can explain up to 83 percent of the variation in the dependent variable. The conditioning of the variables on their respective standard deviations makes the interpretation of the obtained coefficients somewhat difficult. It implies that each coefficient is an estimation of the elasticity of the production share to the specific variable one standard deviation above the cut-off point for that variable.¹¹ Nevertheless, it is noticeable that forward and backward linkages and the supply of educated labour, researchers and scientists seem to have increased in importance over time, with the estimated coefficients being significantly positive at the ten percent-level. Furthermore, the importance of increasing returns to scale in production in

¹⁰ For a somewhat more detailed description of the methodology and the empirical specification, see section 4.1. For a comprehensive discussion, the interested reader is referred to Midelfart-Knarvik et al. (2000) and Midelfart-Knarvik et al. (2002).

¹¹ For more details on this matter, see Midelfart-Knarvik et al. (2002), p. 17-19.

inducing firms to locate in central regions has decreased over time, and the coefficient on the interaction between agricultural production and agricultural input is insignificant in all but the latest time period, in which it is significant at the ten percent significance level. (Midelfart-Knarvik et al., 2000, p.34-38).

A similar study performed by Midelfart-Knarvik et al. (2002, p. 16-22) also focuses on the importance of interaction variables combining country- and industry characteristics. The data used is similar to the study by Midelfart-Knarvik et al. (2000), covering 36 manufacturing industries in 14 EU countries but limited to the period 1980 to 1997. However, rather than using four-year averages, the study is performed separately for time periods of four years, and the dependent variable is a “double relative” measure of output, which controls for differences in the size of countries and industries.¹² Interactions include the country characteristics supply of agricultural endowment, skilled labour and researchers, the elasticity of market potential, relative market potential and access to suppliers and the industry characteristics consist of intensity in agricultural goods, R&D, skill and intermediate goods and transport costs and the share of output going to industry.

In general, the findings reported by Midelfart-Knarvik et al. (2000) are quite well confirmed by the empirical estimations by Midelfart-Knarvik et al. (2002). As in the former, the estimates are conditioned on the standard deviation of the corresponding variable and depending on the specification, 14-17 percent of the variation in the dependent variable can be explained by the included explanatory variables.

The results point to some interesting tendencies. For example, the importance of countries’ supply of researchers and scientists for industries intensively using this type of labour in production has increased over time, whereas the importance of access to skilled labour remains constant. The interaction representing agriculture and supplier access respectively, are significant only in the latest time period, whereas the interaction of relative market potential and the share of output going to industry is significant in all time periods, although it is steadily decreasing over time. Also, the results suggest that the elasticity of the production share is largest for the interaction between skill endowment and skill intensity with a reported coefficient of around 1.5-1.6 depending on the time period, followed by the interaction between supply of researchers and research intensity and the interaction between supplier access and intensity in intermediate goods. In addition, it is noticeable that the coefficient of the importance of market potential combined with transport costs is only significant in the pooled sample, and in that case, it is negative. Although counterintuitive, this would imply that industries which are intensive in transport locate in countries with low market potential.

¹² For more details on such measures, see for example Overman et al. (2001)

According to the authors, lack of appropriate data and the possibility that transport costs are more important on the local than on the country level might be two potential explanations. Finally, the sensitivity of the results is controlled in several ways, and the results seem to be robust to various changes in the specification as well as the estimation method (Midelfart-Knarvik et al., 2002, p.17-24).

13

Using Swedish data, a similar study using data on 10 different sectors¹⁴ in 284 municipalities has been undertaken by Gullstrand and Hammarlund (2007, p. 43-44 and p. 163-165). The development over time of the specialisation of Swedish municipalities is described, and an empirical study aiming at investigating the determinants of the location of production is undertaken. In this study, the dependent variable is the sector share of employment of total employment in each municipality and all variables are normalised to facilitate comparison of estimated coefficients for variables which are originally measured in different units. The empirical specification of the estimated model is based on Midelfart-Knarvik et al. (2000) and Midelfart-Knarvik et al. (2002) described above and the main focus is on the size and significance of the interaction variables. The results of the estimations are reported for a large number of such interactions, which are similar in type to the ones included by Midelfart-Knarvik et al. (2000) and Midelfart-Knarvik et al. (2002).

The estimations are performed separately for the years 1993, 1995, 2001 and 2003 and depending on the specification, the explanatory variables can explain between 53 and 59 percent of the variation in the dependent variable. The estimated coefficients for the interaction terms indicate that the interaction between market potential and the share of inputs in production together with the interaction between market potential and product differentiation have the largest impact on the dependent variable, with reported coefficients in the range of 0.3. For the former, the estimated coefficient also displays an increase in size over time. Noticeable is also that several of the interactions involving the supply of and industry intensity in primary production are insignificant and those that are statistical significant are of relatively modest size compared to the coefficients of other interaction terms. Furthermore, with one exception, all interaction terms involving market potential are significant, although of varying size.

Noticeable is that the interaction of market potential and economies of scale in production is significantly negative, which, as opposed to what one would expect, would imply that industries characterised by a high degree of economies of scale in production tend to locate not in but rather outside municipalities characterised by a large market. Furthermore, the interactions between market

¹³ For more details on the performed robustness tests, see Midelfart-Knarvik, 2002, p. 22-23.

¹⁴ As opposed to Midelfart-Knarvik et al. (2000) and Midelfart-Knarvik (2002), the data does not only include manufacturing sectors but also for example services and agriculture, forestry, fishing and extraction of minerals.

potential and consumer link and economies of scale, respectively, become insignificant when a variable measuring the degree of product differentiation is included. These results would point to the consumer link- and the economies of scale-variables mainly capturing the importance of the degree of product differentiation in the different industries. A further interpretation of the results imply that industries characterised by a strong link to domestic consumers produce goods of relatively high degree of product differentiation and are located in areas with relatively much economic activity. Industries characterised by economies of scale in production, on the other hand, produce goods characterised by a low degree of differentiation and locate outside of the central area (Gullstrand and Hammarlund, 2007, p. 62-71 and 163-165).

3.3. The importance of infrastructure

The empirical research assessing the impact of infrastructure on the location decision on firms seems to be rather limited. However, Fox and Murray (1990, p. 413 and p. 418-426) investigate the impact on interregional business location and start-up decisions of a large number of factors related to public infrastructure and public policies in the short- and long run. The used data set covers all sectors in the economy and consists of pooled time-series, cross-section data for counties in the U.S state Tennessee during the years 1980 to 1986 and the estimations are performed using firm entry as the dependent variable. Explanatory variables include, among others, short-term policies such as different kinds of taxes and expenditures on education, streets and highways and long-term variables such as the existence of a highway, a major in-county rail line and the distance to the nearest airport.

The results of the estimations performed using both OLS and Tobit estimation indicate that small firms are affected by a larger number of factors than large firms. Also, long-term policies, such as variables capturing the effects of interstate highways, railway infrastructure and educational policy, are the most robust throughout the study. However, even for these variables, the results are mixed both when it comes to the statistical significance of the estimated coefficients and when it comes to their estimated signs. In addition, the authors stress that the estimated coefficients in general are small, and that large-scale policies are required for such policies to be able to have an impact the location decisions of firms.

4. Methodology

This section presents and discusses methodological aspects of relevance for the empirical study to be performed within this thesis. First, the theoretical model underlying the empirical specification is described, and the inclusion of commuting and infrastructure into this model is discussed. The following section discusses the estimation method. Finally, the data sample as well as the empirical specification used in the econometric analysis are presented.

4.1. From theory to empirics

The study to be performed within this thesis is based on the theoretical model explicitly developed in Midelfart-Knarvik et al. (2002) and implemented in the empirical studies by Midelfart-Knarvik et al. (2000), Midelfart-Knarvik et al. (2002) and Gullstrand and Hammarlund (2007).

The model developed in Midelfart-Knarvik et al. (2002) depicts an economy consisting of a number of countries, industrial sectors and primary factors. Constant returns to scale and perfect competition characterise all sectors and each sector produces a number of varieties of differentiated goods. Trade of goods is associated with transport costs of the iceberg form. Setting up the model in this way implies that the value of production in each country is determined by the geographical location of demand, by prices of input goods and by supply of factors of production.

Midelfart-Knarvik et al. (2002) proceed by theoretically deriving expressions for the demand of goods, input prices and a number of other characteristics of the model. Also, a numerical simulation of the model reveals its most important properties. These include the result that countries abundant in a specific factor will host production in sectors which use this factor intensively, and countries with a relatively low price of intermediate goods will host production in sectors which use this input good intensively in production. Furthermore, the numerical simulation indicates that industries that are intensive in transport will locate in countries with large market potential, and industries which output to a large extent is used as intermediate goods will locate where demand for the good is high. Noticeable is also that the result of the numerical simulation reveals the property of the existence of a cut-off point when it comes to the above described abundance of factors and the importance of prices of intermediate goods and market potential. This implies that there exists an industry with a factor intensity equal to this cut-off value, and which location does not depend on the factor abundance/intermediate goods prices.¹⁵

¹⁵ See by Midelfart-Knarvik et al. (2000), p.32-33 and/or Midelfart-Knarvik et al. (2002), p. 6-7 for more details on this property.

Hence, the model is able to capture both the predictions by the Heckscher-Ohlin theory of production locating in the country being abundant in the factor of production intensively used in production and the predictions by New Economic Geography-model of the importance of transport costs and the size of markets as additional important determinants of production.¹⁶

To enable an empirical estimation of the model, it is log-linearised around a reference point. This implies that the concentration measure can be expressed as a function of the supply side, that is interactions between input shares and input values, and the demand side, that is interactions between industry characteristics and the elasticity of market potential with respect to these characteristics. In the econometric implementation and empirical estimation based on the model, the dependent variable is the double relative measure of output taking differences in the size of both sectors and countries into account. The set of explanatory variables consists of the single country- and industry characteristics and the interactions between these.^{17 18}

4.2. Incorporating commuting and infrastructure into the model

The aim of this thesis is to assess the importance of commuting for the location of economic activity. As mentioned in the introductory section, one can imagine commuting patterns to affect the location of production through several different channels. For the purpose of this thesis, focus is mainly on commuting as a proxy for the quality of infrastructure and on the effects of commuting on the location of demand.

In terms of the theoretical model described above, the degree of commuting will be treated as an additional municipality characteristic, and as such, it will be interacted with certain industry characteristics to investigate what impact commuting patterns and the quality of infrastructure have on the degree of sector concentration and municipality specialisation.¹⁹

4.3. Estimation method

Data on the Gross Regional Product (GRP) on the sector- and municipality level is not officially published by Statistics Sweden, and is for the empirical study in this thesis only available for the years 2001 and 2003.

¹⁶ See section 2 and for example Krugman (1991).

¹⁷ In Midelfart-Knarvik et al. (2000) and Gullstrand and Hammarlund (2007), additional variables accounting for the differences in the size of countries/municipalities and in the size of the different sectors are included. This will also be the case in the study in this thesis. See section 5.3 for the empirical specification of the model to be estimated.

¹⁸ For a description of the results of this and similar empirical studies, see section 3.2.

¹⁹ See section 5 for more information on the variables and the data.

The availability of data from more than one year for the same cross-sectional units would potentially enable the study to make use of panel data estimation methods, and initially, the intention was to perform the empirical study using the fixed-effects within estimator as well as the fixed-effects first-differences estimator²⁰ and to perform estimations using the Ordinary Least Squares (OLS) estimator to control the sensitivity of the results. However, the short time period covered in the sample and the relatively modest variation in the included variables makes the appropriateness of the use of the fixed effects-method on the available data sample questionable.

For this reason, the estimations will be performed using OLS on the pooled data sample, which implies that repeated observations for the same cross-sectional unit are treated as separate observations, and the additional information from recognising that they belong to the same unit is not explicitly accounted for in the estimation (Cameron, A. C. and Pravin, K. T., 2005, p.702-703). In addition, OLS estimations will be performed on two subsamples comprising the years 2001 and 2003, respectively, to enable an analysis of any tendencies and trends occurring over time.²¹

Compared to using the fixed effects-method, the use of OLS on the pooled data and the separate subsamples have the disadvantage of not being able to account for (un)observed heterogeneity not captured by the included explanatory variables. Recognising that this may pose a problem for the econometric study, it can be argued to be of minor importance in the current setting. Rather than developing a new model that aims at including all potential determinants, the aim of this thesis is to expand an existing theoretical and empirical model to take account of the additional impact of commuting. Furthermore, the use of OLS rather than the fixed effects-method has some advantages. Time-invariant variables, which cannot be estimated using the fixed effects-method (Verbeek, 2008, p.357-360), can be estimated by OLS, and, as mentioned above, estimations can be performed for separate years in order to detect trends and tendencies taking place over time. In addition, OLS can also be favoured for comparability reasons, since the previous studies presented in section 3.2 all make use of OLS as the main estimator.

The variables used in the empirical study will be transformed via normalisation and standardisation, respectively, before they are used in the estimations. The use of such variable-transformation methods enables comparison of the size of coefficients also for variables which are initially measured in different units.^{22 23} Moreover, by using both normalisation and standardisation, the sensitivity of the results with respect to the chosen transformation method is tested.

²⁰ For more information on the fixed effects-method, see for example Verbeek, 2008, p.355-375.

²¹ The limited availability of data and the fact that the available data is from two years relatively close in time, however, makes an extensive analysis of any such tendencies or trends being beyond the scope of this thesis.

²² The terms normalisation and standardisation are sometimes used interchangeably in the literature. In this thesis, however, the terms are used as representing separate transformation methods.

Transforming the variables according to the procedures described above, however, also affects the interpretation of the coefficients.²⁴ For an interpretation of the direct relationship between the original variables, one would need to make a retransform in order for the estimated coefficients to illustrate the impact of the variables measured on their original scales, which is cumbersome, especially in the presence of interaction variables. In the presentation and discussion of the results in sections 6 and 7, focus is therefore on the comparison of the size of the obtained coefficients both across time and between the included variables.

4.4. Econometric problems and countermeasures

Several countermeasures are undertaken to mitigate and remedy potential problems arising in the estimation procedure due to the characteristics of the data sample.

Considering the short time period for which data is available, non-stationarity should not be a reason for concern in the current setting, and heteroscedasticity and autocorrelation are taken account of by the use of robust standard errors. Reverse causality, which implies that the explanatory variable is endogenous and correlated with the error term, and which results in OLS being biased and inconsistent (Verbeek, 2008, p. 138-140), is mitigated by the use of lagged explanatory variables.²⁵

Since the purpose of this thesis is to expand an existing model to account for the impact of commuting and infrastructure on the location of economic activity rather than building an entirely new model, potential problems of omitted variables should not be as worrying in this context as they could be in other settings. Since potential extensions of the empirical model and the general way of modelling are not within the scope of this thesis, potential omitted variables will not be extensively discussed in relation to the empirical study.

Multicollinearity, which implies that two or more of the included explanatory variables can be approximated by a linear relationship, is, however, a reason for concern in the current setting. The correlation between the explanatory variables results in large variance of the estimators, which, in turn, often results in statistically insignificant coefficients as the estimation procedure cannot distinguish the impact of the separate correlated explanatory variables on the dependent variable. In addition, multicollinearity can result in coefficient estimates of the opposite sign compare to what economic

²³The normalisation method implies that each observation is given a value between zero and one, while standardisation results in transformed variables with the property of having an expected value of zero and a standard deviation of one. For more information, see for example Han et al, 2011, p.114 and Washington et al. 2010, p.79-80.

²⁴ For standardised variables, the coefficients illustrate the estimated impact on the dependent variable per unit change of one standard deviation in the explanatory variable (Washington et al. (2010, p.79-80).

²⁵ See section 5 for a more detailed description on the variables and the data.

theory would predict, and estimates are often sensitive to respecification of the model so that only a small change in the empirical specification can result in substantial changes of the estimation results (Gujarati, 2006, p.363-377).

Estimating the model as initially intended results in values of the Variance Inflation Factor (VIF) of approximately 2000, which implies that the variance is 2000 times larger than it would have been in the absence of multicollinearity (Verbeek, 2008, p.43-45). There exist different rules of thumb when it comes to what constitutes threshold values above which multicollinearity becomes a serious problem. Values of 10 up to 40 have been suggested, but O'Brien (2007) argues that a high value of the variance inflation factor might not be a problem in itself but that also other aspects simultaneously affecting the variance of the estimates need to be analysed. Still, however, the obtained value of 2000 far exceeds the suggested threshold values and several measures are therefore undertaken to remedy the problem of multicollinearity. These include the exclusion of some aspects intended to be included in the analysis as well as the merging of some of the explanatory variables.²⁶ By undertaking these measures, the obtained values of the variance inflation factor ranges between 18.16 and 21.19 in the different specifications. Recognising that this value is still relatively high compared to some of the suggested threshold values, it can be argued acceptable considering the large number of interaction variables included in the empirical model.²⁷ In addition, further changing or excluding variables from the model implies the risk of estimating a model not relying on theoretical foundations.

²⁶ See sections 5.1.3 and 5.1.4 for more information on these issues.

²⁷ See section 5 for more details on the included variables and the model.

5. Data and empirical specification

The choice of which variables to include in the empirical study is based both on the theory described in section 2 and on the previously performed empirical studies summarised in section 3. This section contains a description of the included variables and the sample as well as an empirical specification of the model underlying the econometric analysis.

5.1. Variable overview²⁸

5.1.1. Dependent variables

As briefly mentioned in section 4.3, data on Gross Regional Product (GRP) on the sector- and municipality level is available only for the years 2001 and 2003. For those two years, however, two different measures of GRP are available, both of which are calculated by allocating the value added of the reporting firms to the different work stations of the firms.²⁹ The first, given the suffix 1 in the presentation of the results in section 6 below, allocates the value added according to the number of employed at each work station, whereas the second, named by the suffix 2, uses information on the average value added per employed and sector on the national level to adjust the average valued added per employed in each municipality (Gullstrand and Hammarlund, 2007, p. 146).³⁰ The dependent variable in the empirical study performed within this thesis will be calculated using both these measures of GRP, and in addition, the sensitivity of the results will be controlled for by using the number of employed rather than GRP as a measure of economic activity.

When investigating the location of economic activity, one can apply two related, but different, measures. The first focuses on the degree of specialisation of a geographical unit, for example a country, region or municipality, and is measured as the amount of a certain kind of activity related to the total activity within that same geographical unit. The second measure focuses on the degree of concentration of economic activity and is measured as the share of total economic activity, potentially for a specific sector, that is located within a geographical unit.

The difference thus lies in the choice of reference value to which economic activity is related. When specialisation is the topic of interest, economic activity is related to the total activity of the geographical unit, whereas, when focus is on concentration, it is related to the total economic activity

²⁸ For descriptive statistics of the variables, see Appendix 1.

²⁹ This is necessary since one firm can have work stations in more than one municipality.

³⁰ The obtained values are similar in size, but both measures are included as a robustness check in the empirical study.

of the particular sector (Gullstrand and Hammarlund, 2007, p.40-41 and Overman et al., 2001, p.20-21).

The reviewed theoretical models within the New Economic Geography-framework³¹ analyse the location decision of firms, which is related to the concept of concentration described above. For this reason, the dependent variable of main interest in the study is the sector activity in a specific municipality expressed as a share of total sector activity in Sweden, henceforth called **concentration**.

However, if a specific type of industry tends to be geographically concentrated to a specific municipality, this would also affect the degree and type of specialisation of that municipality. Therefore, the impacts of commuting patterns and the quality of infrastructure will also be assessed with respect to the degree of municipality specialisation. In that case, the dependent variable is the sector activity in a specific municipality expressed as the share of total economic activity in that municipality, henceforth called **specialisation**.

Table 5.1. Dependent variables

Dependent variables	
Variable	Measure(s)
Concentration	Sector activity in a specific municipality as a share of total sector activity in Sweden.
Specialisation	Sector activity in a specific municipality as a share of total economic activity in that municipality.

5.1.2. Variables to capture size differences

In accordance with the empirical model estimated in Midelfart-Knarvik et al. (2000) and Gullstrand and Hammarlund (2007), variables to control for the size of the municipality and the size of the sector, respectively, will be included in the regression model. **Population share** is measured as the municipality population in relation to the total population in Sweden and **sector share** is measured as the share of the total value added of each sector in the value added in the total economy consisting of all sectors. These are intended to capture the presumption that larger municipalities have a larger share of total economic activity and that a sector that constitutes a relatively large share of the total economy should also be represented by a relatively large share in each municipality. To be consistent with other

³¹ See section 2.

explanatory variables for which data can only be used for certain years³², these two variables are measured as the average during the years 1994-1996 and 1996-1998 for 2001 and 2003, respectively.

5.1.3. Municipality characteristics

The included variables aiming at capturing municipality characteristics are summarised in table 5.2 below. These include factor supply- and market size aspects as well as the commuting variable.

Supply of agricultural and forest land is measured as the number of hectares in each municipality that is used for production within the agricultural and forestry sector in relation to the total area of the municipality.³³ Due to data limitations, this variable is measured as the average supply between 1994-1996 and 1996-1998 for 2001 and 2003, respectively. Lack of data is also the reason to the variable supply of minerals being measured as a dummy variable for North of Sweden, where North of Sweden consists of the regions Dalarna, Gävleborg, Västernorrland, Jämtland, Västerbotten and Norrbotten.³⁴ To avoid issues of multicollinearity arising in the estimation process, one single variable, called **agrimineralsupply**, is used to measure the supply of agricultural land, forest land and minerals. Since the supply of agricultural and forest land is measured as a share whereas mineralsupply is captured by a dummy variable taking the value one or zero, the variables are transformed before they are merged.

Supply of labour and supply of highly educated labour are measured as the size of the population in relation to the area of the municipality and the share of the population with at least three years of post-secondary education, respectively. The reason for measuring these variables with respect to the population, that is the number of people with the municipality as the reported place of residence, rather than with respect to the number of daytime workers within the municipality, is twofold. First, since commuting patterns are explicitly accounted for by another included explanatory variable, measuring the supply of labour as the number of daytime workers would result in two variables partly capturing the same aspect, and would thereby potentially lead to multicollinearity among these variables. Second, data limitations impose problems when it comes to finding information on the education level of the daytime population. Such data is, however, available for the residence population and for consistency, it seems appropriate to measure the supply of labour and the supply of highly educated labour with respect to the same reference group. Again, for time consistency with other included explanatory variables, the data is retrieved for the years 1994-1996 and 1996-1998 for 2001 and 2003, respectively. Initially, the intention was to include labour supply and supply of highly educated labour

³² This especially holds for the supply of agricultural and forest land and the supply of high-skilled labour. For more information, see section 5.1.3.

³³ Agricultural land is the sum of arable land and pasture land. For a definition overview, see for example the project Markinfo (2007-02-10) by the Swedish University of Agricultural Sciences.

³⁴ This is also the approach used in Gullstrand and Hammarlund (2007, p. 163).

as two separate explanatory variables. As in the case with supply of agricultural land, forest land and minerals, severe multicollinearity arises when both of them are included. To solve this problem, the variable measuring the supply of highly educated labour is merged into the variable labour supply, so that **labour supply** is a weighted measure of population density constructed so that municipalities having a relatively large share of highly educated labour obtain higher reported values. Recognising that one would ideally want to incorporate labour supply and the supply of highly educated labour as two separate variables, this approach is necessary for reliable estimations to be implementable. The intention was also to include dummy variables representing municipalities with different characteristics. However, since the two classifications considered³⁵ both are based on population density as the main characteristic, including such dummy variables would result in multicollinearity between the dummy variables and the labour-supply variable. Nevertheless, by inclusion of the labour-supply variable, an interpretation of the results for municipalities with differences in population density will still be feasible.

Market potential is an indicator of the potential demand for goods and services. As described in Overman et al. (2001, p.11-13) and Combes and Overman (2003, p. 10-12), market potential for a region, in this study for each municipality, is often measured as the sum of economic activity in all other regions deflated by the distance to these regions, and economic activity can be measured for example in terms of GDP. In the study in this thesis, total income of the municipality is used as a proxy for economic activity when calculating the market potential for each municipality, as in Gullstrand and Hammarlund (2007). The reason for this choice is twofold. First, what matters for firms for which it is important to locate close to the demand for their product is presumably the income of those demanding the product rather than the economic activity in terms of total production, or value added, of the municipality. Second, since total production and sector production already enter the study through other variables, it might be more appropriate to measure market potential using total income rather than value added as a measure of economic activity. Potential multicollinearity and/or reverse causality are thereby, at least to some extent, mitigated. The used data on total income is calculated as the average for the years 1994-1996 and 1996-1998 for the years 2001 and 2003, respectively. As already repeatedly noted, the choice of years to include in the calculation of the average is made for consistency of the choice of time period for other variables, which are limited by data availability. The choice of basing the variables on lagged values is also made to avoid reverse causality and endogeneity.

³⁵ The two considered classifications originate from the Swedish Natural Rural Development Agency (Glesbygdsverket, 2002) and the Swedish Association for Local Authorities and Regions (SKL, Sveriges Kommuner och Landsting, 2009-10-14).

Finally, the **degree of commuting** of each municipality is a composite variable based on the number of cross-border in-commuters and the number of cross-border out-commuters. As argued in section 2.5, the degree of commuting is included not only as a measure of commuting, but can also be given the interpretation of measuring the quality of infrastructure. As such, it is intended to capture the theoretical predictions of municipalities with better infrastructure hosting relatively more economic activity. The degree of commuting is measured as the average of the share of in-commuters to total municipality population and the share of out-commuters to total municipality population during 1994-1996 and 1996-1998 for the years 2001 and 2003, respectively. As for several of the variables described above, the choice of time period is made for consistency with other explanatory variables and to avoid endogeneity. At first, the intention was to include several separate variables in order to capture a richer picture of the commuting patterns. In that case, in-commuting and out-commuting measured as the number of people commuting into and out of the municipality in relation to the total municipality population, respectively, and the interaction between these two variables would have been included. However, for similar reasons as in the case of not including dummies representing different types of municipalities, the larger set of commuting variables is excluded to avoid multicollinearity among the explanatory variables. Instead, as described above, commuting is in the basic specification included as a single variable based on the number of in-commuters and out-commuters in relation to the population size.

The sensitivity of the results with respect to this commuting variable is, however, tested by the use of an alternative variable capturing commuting patterns. While the degree of commuting, as described above, mainly captures the magnitude of commuting in relation to the population, one might also want to include a measure taking into account that different effects might arise for municipalities characterised by a large degree of in-commuting and a large degree of out-commuting, respectively. To take account of such differences in commuting flows, the robustness of the results is tested by measuring commuting using the variable **commutingratio**, which is defined as the logarithm of the ratio of in-commuters to out-commuters.

As mentioned in the description of the variables above, some of them are measured using data from previous years. In the table 5.1 below, variables that are lagged in this way are marked with an asterisk. The choice of including lagged rather than current values of some of the explanatory variables is made in order to avoid potential problems of reverse causality in terms of the location of production affecting explanatory variables such as commuting patterns, the supply of labour, the supply of highly educated labour, and not (only) vice versa. Also, it seems reasonable to believe that location decisions of firms are a time-consuming process and that aspects that influence the decision do so with a time lag. Including past rather than current values of the explanatory variables can therefore be motivated on several grounds.

Table 5.2. Municipality characteristics

Municipality characteristics	
Variable	Measure(s)
Supply of agricultural land, forest land and minerals	Composite variable consisting of a dummy variable to capture mineral supply and the number of hectares of agricultural and forest land in relation to the total municipality area *
Supply of labour	Population in relation to the area of the municipality, weighted to account for differences in education level *
Market potential	The sum of the ratio of economic activity over distance for all other municipalities in the sample, where economic activity is measured as total income. *
Degree of commuting	Average of the share of in-commuters to total municipality population and the share of out-commuters to total municipality population. *
Commutingratio	The logarithm of the ratio of in-commuters to out-commuters *

*using data from previous years

5.1.4. Industry characteristics^{36 37}

Table 5.3 below provides a summary of the variables representing industry characteristics.

Agrimineralintensity is measured as the value of inputs coming from agriculture, forestry and mineral extraction in relation to the total value of all inputs used in production. Labour intensity in production is measured as the share of total value added in production paid as wages and intensity in highly educated in employment is calculated as the share of employed with at least three years of post-secondary education. Note that the latter is based on data from previous years and is calculated as the average of the share in 1994-1996 and 1996-1998 for 2001 and 2003, respectively. Again, the choice of years is made in order to avoid reverse causality and endogeneity and for time consistency with other explanatory variables. In addition, the existence of a break in the way of measuring and valuing education in the year 2000 results in weak comparability of data before and after this year. For consistency with the measurement of municipality characteristics described in section 5.1.3, **labour**

³⁶ With the exception of intensity in highly educated in employment and the degree of economies of scale in production, all industry characteristics are calculated from input-output tables for the year 2000 provided by Statistics Sweden (SCB) and are thus constant over time.

³⁷ The calculated values of the industry intensities can be found in Appendix 2.

intensity in production is a weighted measure of labour intensity in production with higher reported values for sectors intensively using highly educated labour in production.

Economies of scale in production is proxied by the number of work stations per total sector output in all municipalities. Just as for the value added in each sector, data on the average number of work stations in each sector and in each municipality is available only for the years 2001 and 2003. This is somewhat unfortunate since it might be the case that the location of economic activity affects the number of work stations, thereby resulting in reverse causality. However, since it is the number of work stations per total output among all municipalities rather than the number of work stations for each municipality that is included, the risk of reverse causality and endogeneity should not be too large.

Consumer link is measured as the share of output sold to domestic consumers and is intended to capture differences when it comes to the importance for firms in different sectors to locate close to the final demand for their products. Since demand comes not only from consumers demanding the good for final use but also from producers using the good as an input in production, also the forward producer link, measured as the share of output used as input goods, should be taken into account. To avoid multicollinearity, the forward producer link is, however, not included as a separate explanatory variable. Instead, together with the backward producer link, which is measured as the total value of inputs as a share of the total value of production, it constitutes the variable **producer link**, which is measured as the average of the forward- and backward producer link.

Initially, the intention was also to include a variable measuring the transport intensity in production as the share of total inputs coming from the transport sector. Since the intensity in transport is highly correlated with the producer link, however, this variable is excluded from the model. Although unfortunate, the measure is necessary in order for the model to be able to be estimated and an analysis of the importance of transports can still be made with the help of the other included explanatory variables.

Table 5.3. Industry characteristics

Industry characteristics	
Variable	Measure(s)
Intensity in agricultural-, forestry- and mineral inputs	Share of inputs from agriculture, forestry and minerals of total inputs *
Labour intensity in production	Composite variable comprising the share of total value added going to wages and the share of highly educated labour in total employment *
Economies of scale in production	Number of work stations per total output.
Consumer link	Share of output going to domestic consumers *
Producer link	Average of backward and forward producer links, measured as the share of output used as input goods and as the total value of inputs as a share of the total value of production, respectively *

*using data from previous years

5.1.5. Interaction variables

As described in section 3 presenting previous research, much of the research undertaken assesses the importance of factor endowments, production technology and links between industries and consumers by interacting country or municipality characteristics with industry characteristics, thereby accounting for location patterns in two dimensions.³⁸ This will be the approach used also in the empirical study in this thesis. Table 5.4 below provides a summary of the included interactions between municipality- and industry characteristics.

The interactions **agrimineral supply* agrimineral intensity and supply of labour* labour intensity in production** are included to capture the implications of the Heckscher-Ohlin theory of sectors intensively using a specific factor in production tending to locate in countries or municipalities abundant in this factor.

The following three interactions all include market potential as the municipality characteristic and are similar to interactions included in Gullstrand and Hammarlund (2007) and Midelfart-Knarvik et al.

³⁸ See for example Midelfart-Knarvik et al. (2002), p.1-3.

(2000). The inclusion of **market potential*producer link** is intended to capture the importance for firms intensively using input goods in production and firms selling a large share of their output as input goods to locate close to areas with much economic activity.³⁹ The interaction **market potential*consumer link** captures the importance for firms facing a large demand from domestic consumers to locate close to that demand, thereby avoiding transport costs. **Market potential*economies of scale** illustrates the presumption that industries characterised by increasing returns to scale in production can benefit from concentrating production to a limited number of work stations, and that such concentration is likely to occur in areas with large market potential.

At most three interactions are included to capture the importance of commuting patterns for the location of production of firms. The interaction **degree of commuting*labour intensity** is intended to capture the importance of commuting and good-quality infrastructure resulting in a larger and/or more diversified labour force for firms intensively using labour in production. Thereby, it serves as a complement to the included variable labour supply, which is measured with respect to the population. To investigate the prediction by Gruber (2010) of some sectors being less concentrated in the presence of commuting, the interaction **degree of commuting*consumer link** is included. This interaction can also be interpreted as investigating whether sectors characterised by strong consumer links tend to be more or less concentrated due to good-quality infrastructure resulting in relatively low transport costs. For sectors characterised by strong producer links, this aspect is captured by the interaction **degree of commuting*producer link**.

³⁹ It would be perhaps even more appropriate to include a separate variable measuring the proximity to other suppliers. As argued above, however, it might be inappropriate to base the calculation of market potential/supplier proximity on the value added of each municipality, since value added already enters the model through other variables. In addition, including two variables measuring economic activity but with respect to income and production, respectively, is likely to result in highly correlated variables and thereby multicollinearity among the included explanatory variables.

Table 5.4. Interaction variables

Interactions between municipality- and industry characteristics	
Municipality characteristic	Industry characteristic
Heckscher-Ohlin theory	
Supply of agricultural land, forest land and minerals *	Intensity in agricultural-, forestry- and mineral inputs *
Labour supply *	Labour intensity in production *
Consumer and producer links	
Market potential *	Producer link *
Market potential *	Consumer link *
Market potential *	Economies of scale in production
Commuting and infrastructure	
Commuting *	Labour intensity *
Commuting *	Consumer link *
Commuting *	Producer link *

*using data from previous years

5.2. Sample

The data used in the empirical study is collected from the official data provided by Statistics Sweden (SCB). The sample used in the econometric analysis consists of 286 municipalities (kommuner). During the 1990s, four new municipalities that previously belonged to a neighbouring municipality were created, so that the total number of municipalities in 2003 amounted to 290. To be able to obtain a data sample consisting of an equal number of observations for all variables, the four new municipalities are in this study treated as if they had not been created. The corresponding data is therefore added to the original municipalities to which they belonged in the beginning of the 1990s.⁴⁰

Several arguments can be made in favour of the choice of municipalities as the cross-sectional unit of analysis in the study. First, basing the study on data on the municipality level rather than on larger regions implies a large number of observations, which should improve the quality of the econometric analysis. Second, basing the study on the smallest cross-sectional unit possible to some extent mitigates the potential problem of underestimating the concentration of production within that unit, since the larger the geographical unit, the larger is the scope for concentration to take place within that

⁴⁰ The four concerned municipalities are Nykvarn (code 140), Knivsta (code 330), Bollebygd (code 1443) and Lekeberg (code 1814) which are merged with Södertälje (code 181), Uppsala (code 380), Borås (code 1490) and Örebro (code 1880), respectively.

unit.⁴¹ Third, changes in commuting patterns are likely to be more pronounced if the unit of analysis is municipalities rather than larger areas. For example, the division of Sweden into a number of “functional labour market regions” (LA-regioner) is partly based on commuting patterns. Using such LA-areas as the unit of analysis would therefore miss out on existing patterns and changes. Fourth, from a local tax-base point of view, changes and tendencies in the location of production and the place of residence of people are of interest primarily on the municipality level since income taxes in Sweden are mainly collected at the local rather than at the regional or national level.⁴² Although taxes are not taken into account in the analysis in this thesis, choosing municipalities as the unit of analysis implies that the results of the study could potentially also be of practical interest for policy makers.

The data used in the study covers the seven sectors agriculture, forestry, fishing, mineral extraction, low-technology manufacturing, high-technology manufacturing and services. The choice of included sectors might seem arbitrary but can be motivated by the presumed differences between the sectors when it comes to for example factor intensities in production, the importance of consumer and producer links and the effect of commuting patterns. Furthermore, it should be noted that the included sectors do not add up to the total value-added of all production in each municipality. The analysis is therefore not extensively covering the whole economy, but the choice of limiting the number of included sectors is due to the aim of this study to investigate the importance of commuting patterns and infrastructure for sectors with different characteristics. Naturally, when value added of a specific sector is expressed as a share, however, total value added including all sectors is used as the reference value.

As mentioned above, data limitations on the sector-specific value added on the municipality level is the main reason for the sample containing data for two years only. The missing values for some sectors and municipalities are treated as having a reported value equal to zero. In general, treating a missing value as a reported value equal to zero should be made with caution, but recognising the comprehensiveness of the available dataset, it seems appropriate to treat a missing observation on value added as implying that value added is equal to zero. This also makes the data set strongly balanced.

Including 286 municipalities, two years and seven sectors results in a total number of observations equal to 4004. This is the number of observations included in the estimations using OLS on the whole sample. For natural reasons, the number of observations is restricted to being equal to 2002 in the estimations performed separately for the years 2001 and 2003.

⁴¹ See for example Overman et al. (2001) and Gullstrand and Hammarlund (2007) for a discussion about the measurement of concentration and specialisation.

⁴² For a description and overview of the tax system, see for example Skatteverket (2010).

5.3. Empirical specification

The basic empirical specification of the model forming the basis for the econometric analysis takes the following form:

$$\begin{aligned}
 \text{Share of economic activity}_{ij} = & \beta_0 + \beta_1 \ln \text{Sectorshare}_i + \beta_2 \ln \text{Populationshare}_i + \\
 & \beta_3 \text{Agrimineralsupply}_i + \beta_4 \text{Laboursupply}_i + \beta_5 \text{Marketpotential}_i + \\
 & \beta_6 \text{Commuting}_i + \beta_7 \text{Agriminerallintensity}_j + \beta_8 \text{Labourintensity}_j + \\
 & \beta_9 \text{Economies of scale}_j + \beta_{10} \text{Consumerlink}_j + \beta_{11} \text{Producerlink}_j + \\
 & \beta_{12} \text{Agriminerallinteraction}_{ij} + \beta_{13} \text{Consumerlinkinteraction}_{ij} + \\
 & \beta_{14} \text{Ecoscaleinteraction}_{ij} + \beta_{15} \text{Producerlinkinteraction}_{ij} + \\
 & \beta_{16} \text{Labourinteraction}_{ij} + \beta_{17} \text{Producerlinkcominteraction}_{ij} + \\
 & \beta_{18} \text{Consumerlinkcominteraction}_{ij} + \beta_{19} \text{Labourcominteraction}_{ij} + \varepsilon_{ij}
 \end{aligned}
 \tag{5.1}$$

The empirical specification of the model as stated above underlies the estimation using both normalised and standardised variables. Municipality characteristics are denoted with the subscript i and industry characteristics with the subscript j . The double subscript ij indicates that the variable varies both across industries and municipalities.

The municipality- and industry characteristics are denoted with the same name as in the previous section. Furthermore, *Agriminerallinteraction* represents *agrimineralsupply*agriminerallintensity* and *Consumerlinkinteraction* is the abbreviation for *market potential*consumer link*. *Market potential*economies of scale* is represented by *Ecoscaleinteraction*, whereas *market potential*producer link* is abbreviated with *Producerlinkinteraction* and *supply of labour* labour intensity in production* with *Labourinteraction*. Furthermore, *Producerlinkcominteraction* represents *degree of commuting*producer link* and *Consumerlinkcominteraction* is the abbreviation for *degree of commuting*consumer link*. Finally, *Labourcominteraction* represents the *interaction degree of commuting*labour intensity*. In addition, the specification includes the constant β_0 and the error term ε_{ij} .

The dependent variable, *Share of economic activity_{ij}*, represents both concentration, which implies that sector activity in a specific municipality is expressed as a share of total sector activity in Sweden, and specialisation, which means that sector activity in a specific municipality is expressed as the share of total economic activity in that municipality. Similarly, the term commuting is referring to both measures of commuting described in section 5.1.3.

6. Results

The results of the empirical model as specified above are presented below. First, the results for the estimations of the basic specification using normalised variables with sector concentration and municipality specialisation, respectively, as the dependent variable are presented. The following section is devoted to robustness tests of the results with respect to the commuting variable, the measurement of the dependent variables and the variable-transformation method, respectively. Throughout the results section, the main focus is on the results regarding the interaction variables combining municipality- and industry characteristics, which thereby capture differences in the location of production in two dimensions.

6.1. Basic specification

For both concentration and specialisation as the dependent variable, the results of estimations using the two different ways of measuring GRP described in section 5.1.1 are very similar. For this reason, only the results based on the first measure are presented in this section. The results from the estimations of the basic specification using the alternative second measure can be found in appendix 3.

6.1.1. Sector concentration

The results for the estimations assessing the determinants of concentration using normalised variables according to the basic empirical specification are presented in columns 1, 2 and 3 in table 6.1 below.

Of the variables representing the size of the sector and the municipality, respectively, *populationshare* is positive and highly significant, whereas *sectorshare* is insignificant in all specifications. Noticeable is that one of few significant municipality characteristics is the variable *commuting*, which in the current estimation is significant at the five-percent level in the estimations for 2003 and the whole sample and on the ten-percent level for 2001. The estimated coefficient is positive in all three specifications, and would thus point to a high degree of commuting and/or good-quality infrastructure resulting in a higher degree of geographical concentration of production. However, since the coefficient for 2003 is smaller than for 2001, the results indicate that the importance of infrastructure for the sectoral concentration of production has decreased over time. Intuitively reasonable is also that the estimated coefficient for the whole sample is an approximate average of the estimated coefficients for the years 2001 and 2003 separately.

Furthermore, most of the included interaction variables are insignificant. The exceptions are *producerlinkinteraction* and *labourinteraction*, which are negative and positive, respectively, in all

specifications, and producerlinkcominteraction, which is negative throughout the estimations. Hence, sectors that are characterised by strong links to other producers do not tend to concentrate in municipalities characterised by large market potential, which must be seen as a somewhat surprising and counterintuitive result. This tendency also seems to hold over time as the estimated coefficients are of similar size in the different specifications. Neither do sectors for which links to other producers are important tend to concentrate in municipalities characterised by a high degree of commuting and good-quality infrastructure. In this case, the coefficient is slightly decreasing from 2001 to 2003. Also for labourinteraction, the estimated coefficient decreases over time, and sectors using labour intensively in production therefore tends to be less overrepresented in municipalities with high population density in 2003 than in 2001.

To test the sensitivity of the results with respect to the number and choice of included interaction variables, columns 4, 5 and 6 in table 6.1 present the results for estimations in which the insignificant variable consumerlinkcominteraction has been excluded.⁴³ As can be seen in the table, there are some small changes in the size of some of the estimated coefficients, but in general, no substantial changes occur. Due to space limitations, an extensive table including the results of the estimations excluding the insignificant variables labourcominteraction and consumerlinkcominteraction together with labourcominteraction, respectively, are not presented here. It can be noted, however, that the exclusion of these variables result in changes similar to the ones described for the estimations excluding consumerlinkcominteraction.

⁴³ In the econometric study performed, estimations were also performed using a larger number of interactions involving the degree of commuting, but the resulting high degree of multicollinearity made exclusion of some interactions necessary. The choice of which variables to exclude is based on theoretical considerations and the statistical significance of the interactions. The procedure to find the most suitable empirical model is thereby influenced by the general-to-specific modelling approach. See Campos et al. (2005) and Verbeek (2008, p.59-60) for more information on this topic.

Table 6.1. Concentration using normalised variables

Normalised variables	Dependent variable: Concentration1					
	(1)	(2)	(3)	(4)	(5)	(6)
	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)	0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)
Ln_Populationshare	0.018*** (0.006)	0.019*** (0.003)	0.019*** (0.003)	0.018*** (0.006)	0.019*** (0.003)	0.019*** (0.003)
Agrimineralsupply	0.001 (0.002)	-0.000 (0.001)	0.001 (0.001)	0.001 (0.002)	-0.000 (0.001)	0.000 (0.001)
Laboursupply	-0.053*** (0.020)	-0.049** (0.020)	-0.051*** (0.014)	-0.055*** (0.020)	-0.050** (0.020)	-0.052*** (0.014)
Marketpotential	0.052 (0.038)	0.061 (0.037)	0.057** (0.027)	0.055 (0.036)	0.063* (0.036)	0.060** (0.026)
Commuting	0.371* (0.215)	0.263** (0.123)	0.313*** (0.121)	0.353* (0.204)	0.252** (0.116)	0.299*** (0.115)
Agrimineralintensity	-0.000 (0.007)	-0.001 (0.006)	-0.000 (0.004)	-0.000 (0.007)	-0.001 (0.006)	-0.000 (0.004)
Labourintensity	-0.005 (0.010)	-0.005 (0.007)	-0.005 (0.006)	-0.004 (0.010)	-0.005 (0.007)	-0.004 (0.006)
Economies of scale	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.003)	-0.005 (0.004)	-0.004 (0.004)	-0.005 (0.003)
Consumerlink	0.004 (0.005)	0.004 (0.003)	0.004 (0.003)	0.004 (0.005)	0.003 (0.003)	0.003 (0.003)
Producerlink	0.016 (0.013)	0.016 (0.012)	0.016* (0.009)	0.016 (0.013)	0.016 (0.012)	0.015* (0.009)
Agrimineralinteraction	0.001 (0.003)	0.004 (0.003)	0.002 (0.002)	0.001 (0.003)	0.004 (0.003)	0.003 (0.002)
Consumerlinkinteraction	-0.023 (0.023)	-0.019 (0.022)	-0.021 (0.016)	-0.029 (0.022)	-0.023 (0.022)	-0.025 (0.016)
Ecoscaleinteraction	0.037 (0.024)	0.031 (0.022)	0.033** (0.016)	0.037 (0.024)	0.031 (0.021)	0.034** (0.016)
Producerlinkinteraction	-0.153** (0.065)	-0.155** (0.065)	-0.154*** (0.046)	-0.156** (0.064)	-0.156** (0.064)	-0.155*** (0.045)

Labourinteraction	0.134*** (0.047)	0.121*** (0.047)	0.127*** (0.033)	0.136*** (0.046)	0.123*** (0.046)	0.129*** (0.033)
Producerlinkcominteraction	-0.484** (0.209)	-0.386*** (0.126)	-0.432*** (0.119)	-0.444*** (0.165)	-0.360*** (0.097)	-0.399*** (0.093)
Consumerlinkcominteraction	-0.042 (0.068)	-0.027 (0.046)	-0.034 (0.040)			
Labourcominteraction	0.108 (0.155)	0.123 (0.094)	0.116 (0.087)	0.073 (0.139)	0.100 (0.082)	0.087 (0.077)
Constant	0.509*** (0.004)	0.510*** (0.003)	0.509*** (0.002)	0.509*** (0.004)	0.509*** (0.003)	0.509*** (0.002)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	21.07	21.07	21.07	21.19	21.19	21.19
R-squared	0.177	0.195	0.183	0.176	0.195	0.182
Adjusted R-squared	0.169	0.188	0.179	0.168	0.187	0.178

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.1.2. Municipality specialisation

Columns 1-3 in table 6.2 below contain the results from the basic estimation of the determinants of municipality specialisation using normalised variables.

Compared to the estimations assessing the determinants of sector concentration, a considerably larger number of explanatory variables are significant. Commuting is, however, insignificant in all specifications and there is thereby no indication of differences in the quality of infrastructure and the magnitude of commuting itself to affect the specialisation of municipalities.

The extensive significance of the included interaction variables is noticeable. Positive coefficients are obtained for ecoscaleinteraction and labourinteraction whereas the estimated coefficients for producerlinkinteraction, consumerlinkinteraction and agrimineralinteraction are negative. The results thereby indicate that municipalities characterised by large market potential tend to be relatively less specialised in sectors for which producer- and consumer links are of large importance. By similar reasoning, municipalities characterised by large market potential are relatively more specialised in sectors characterised by a low degree of economies of scale in production,^{44 45} and sectors

⁴⁴ Since the degree of economies of scale in production is measured as the average number of workstations per total output, a large value implies a large number of workstations and hence a lower degree of economies of

characterised by using labour intensively in production contribute to the specialisation in labour-intensive sectors of labour-abundant municipalities.

When it comes to the interaction variables including commuting as the municipality characteristic, positive coefficients are obtained for labourcominteraction, whereas the estimated coefficients for producerlinkcominteraction are negative. Thereby, municipalities characterised by a high degree of commuting/good-quality infrastructure tend to be relatively less specialised in sectors characterised by strong producer links and relatively more specialised in sectors intensively using labour in production.

Again, columns 4-6 contain the results from the estimations performed after excluding the insignificant variable consumerlinkcominteraction. Noticeable is that the coefficients for producerlinkcominteraction and labourcominteraction tend to decrease in absolute value, and the estimated impact of interactions between municipality- and industry characteristics is therefore to some extent dependent on the model specification. The statistical significance of the variables is, however, in general preserved and the exclusion of consumerlinkcominteraction does not result in any large qualitative changes of the results.

scale in production. A negative coefficient thus implies that a larger degree of economies of scale in production results in a higher degree of sector concentration/municipality specialisation, whereas a positive coefficient implies that sector concentration/municipality specialisation is enhanced by a low degree of economies of scale.

⁴⁵ This is a somewhat counterintuitive result. It is, however, partly in line with previous empirical research (see for example Gullstrand and Hammarlund, 2007, p. 42). Comparisons should, however, be made with some caution since the number and choice of included sectors might vary between empirical studies.

Table 6.2. Specialisation using normalised variables

Normalised variables	Dependent variable: Specialisation1					
	(1)	(2)	(3)	(4)	(5)	(6)
	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.512*** (0.020)	0.524*** (0.019)	0.508*** (0.014)	0.512*** (0.020)	0.524*** (0.019)	0.508*** (0.014)
Ln_Populationshare	0.008 (0.012)	0.008 (0.009)	0.008 (0.007)	0.008 (0.012)	0.008 (0.009)	0.008 (0.007)
Agrimineralsupply	0.026*** (0.007)	0.021*** (0.007)	0.024*** (0.005)	0.026*** (0.007)	0.021*** (0.007)	0.024*** (0.005)
Laboursupply	-0.103*** (0.023)	-0.092*** (0.021)	-0.097*** (0.016)	-0.104*** (0.023)	-0.094*** (0.022)	-0.099*** (0.016)
Marketpotential	0.260*** (0.050)	0.191*** (0.044)	0.222*** (0.033)	0.263*** (0.050)	0.195*** (0.043)	0.226*** (0.033)
Commuting	0.053 (0.076)	0.056 (0.065)	0.055 (0.050)	0.037 (0.050)	0.033 (0.041)	0.035 (0.033)
Agrimineralintensity	-0.408*** (0.017)	-0.438*** (0.016)	-0.416*** (0.012)	-0.408*** (0.017)	-0.438*** (0.016)	-0.416*** (0.012)
Labourintensity	-0.513*** (0.024)	-0.563*** (0.021)	-0.528*** (0.016)	-0.512*** (0.023)	-0.562*** (0.021)	-0.527*** (0.016)
Economies of scale	-0.090*** (0.020)	-0.081*** (0.019)	-0.080*** (0.014)	-0.090*** (0.020)	-0.081*** (0.019)	-0.080*** (0.014)
Consumerlink	0.203*** (0.013)	0.229*** (0.012)	0.212*** (0.009)	0.203*** (0.013)	0.228*** (0.012)	0.212*** (0.009)
Producerlink	0.589*** (0.027)	0.633*** (0.024)	0.599*** (0.018)	0.588*** (0.027)	0.632*** (0.024)	0.599*** (0.018)
Agrimineralinteraction	-0.025*** (0.008)	-0.015* (0.008)	-0.020*** (0.006)	-0.025*** (0.008)	-0.015* (0.008)	-0.020*** (0.006)
Consumerlinkinteraction	-0.311*** (0.042)	-0.252*** (0.038)	-0.274*** (0.028)	-0.316*** (0.042)	-0.259*** (0.037)	-0.280*** (0.028)
Ecoscaleinteraction	0.699*** (0.093)	0.602*** (0.086)	0.645*** (0.065)	0.700*** (0.092)	0.602*** (0.086)	0.646*** (0.065)
Producerlinkinteraction	-0.354*** (0.067)	-0.269*** (0.060)	-0.306*** (0.045)	-0.356*** (0.067)	-0.272*** (0.060)	-0.309*** (0.045)

Labourinteraction	0.155*** (0.038)	0.138*** (0.035)	0.144*** (0.026)	0.157*** (0.038)	0.140*** (0.035)	0.146*** (0.026)
Producerlinkcominteraction	-0.326* (0.178)	-0.294* (0.153)	-0.308*** (0.117)	-0.290** (0.130)	-0.242** (0.108)	-0.263*** (0.084)
Consumerlinkcominteraction	-0.039 (0.093)	-0.055 (0.078)	-0.047 (0.060)			
Labourcominteraction	0.273** (0.126)	0.241** (0.107)	0.254*** (0.082)	0.241** (0.097)	0.196** (0.080)	0.215*** (0.062)
Constant	0.293*** (0.008)	0.315*** (0.007)	0.302*** (0.005)	0.293*** (0.008)	0.315*** (0.007)	0.302*** (0.005)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	21.07	21.07	21.07	21.19	21.19	21.19
R-squared	0.783	0.824	0.803	0.783	0.824	0.803
Adjusted R-squared	0.781	0.822	0.802	0.781	0.822	0.802

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2. Robustness tests

6.2.1. Alternative measure of commuting

6.2.1.1. Sector concentration

Columns 1, 2 and 3 in table 6.3 below contain the results of the estimations assessing the determinants of concentration using the alternative measure of commuting. As described in section 5.1.3, this variable, named *commutingratio*, focuses on the direction of the occurring commuting flows rather than on their magnitude.

Compared to the estimation results using the degree of commuting in assessing the determinants of sector concentration, some changes are noticeable. Although the estimated coefficient for commuting as a single variable is substantially smaller than before, it is significantly positive at higher significance levels compared to the basic commuting variable. Also, it should be noted that the significance of *producerlinkcominteraction* has been “replaced” by the significance of *labourcominteraction*, which displays a positive coefficient. Hence, sectors using labour intensively in production tend to concentrate in municipalities for which the inflow of commuters is relatively large compared to the outflow.

This tendency is even more pronounced in the estimations excluding the insignificant variable consumercominteraction, presented in columns 4-6. For the rest of the estimated coefficients, however, the results are very similar to the results of the estimations in columns 1-3.

Again, due to space limitations, estimations excluding further interactions are not extensively reported in table format. In general, excluding the insignificant variable producerlinkcominteraction reveal very similar results to the ones reported in table 6.3, but, just as described in section 6.1.2 for the estimations assessing municipality specialisation using the basic commuting variable, the magnitude of the remaining interaction variables changes somewhat due to such exclusion. In the current setting, this is true for the variable labourcominteraction, which estimated coefficient displays a decrease in absolute value.

Table 6.3. Concentration using the alternative commuting variable

	Dependent variable: Concentration1					
	(1)	(2)	(3)	(4)	(5)	(6)
Normalised variables	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)	0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)
Ln_Populationshare	0.015*** (0.006)	0.017*** (0.003)	0.016*** (0.003)	0.015*** (0.006)	0.016*** (0.003)	0.016*** (0.003)
Agrimineralsupply	0.001 (0.002)	-0.001 (0.001)	0.000 (0.001)	0.001 (0.002)	-0.000 (0.001)	0.000 (0.001)
Laboursupply	-0.066*** (0.022)	-0.059*** (0.020)	-0.062*** (0.015)	-0.065*** (0.022)	-0.058*** (0.020)	-0.061*** (0.015)
Marketpotential	0.121*** (0.039)	0.108*** (0.037)	0.114*** (0.027)	0.120*** (0.039)	0.108*** (0.037)	0.113*** (0.027)
Commutingratio	0.007*** (0.002)	0.006** (0.002)	0.007*** (0.002)	0.007*** (0.002)	0.006*** (0.002)	0.007*** (0.002)
Agrimineralintensity	-0.000 (0.007)	-0.001 (0.006)	-0.000 (0.004)	-0.000 (0.007)	-0.001 (0.006)	-0.000 (0.004)
Labourintensity	-0.006 (0.010)	-0.006 (0.007)	-0.005 (0.006)	-0.006 (0.010)	-0.006 (0.007)	-0.006 (0.006)
Economies of scale	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.003)	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.003)
Consumerlink	0.004	0.004	0.004	0.005	0.004	0.004

	(0.005)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)
Producerlink	0.016	0.015	0.015*	0.016	0.016	0.015*
	(0.014)	(0.012)	(0.009)	(0.014)	(0.012)	(0.009)
Agrimineralinteraction	0.001	0.004	0.003	0.001	0.004	0.002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)
Consumerlinkinteraction	-0.039*	-0.030	-0.034**	-0.035	-0.027	-0.031**
	(0.023)	(0.022)	(0.016)	(0.022)	(0.022)	(0.016)
Ecoscaleinteraction	0.035	0.030	0.032*	0.034	0.029	0.031*
	(0.025)	(0.022)	(0.016)	(0.025)	(0.022)	(0.017)
Producerlinkinteraction	-0.234***	-0.209***	-0.219***	-0.235***	-0.210***	-0.220***
	(0.068)	(0.065)	(0.047)	(0.069)	(0.066)	(0.047)
Labourinteraction	0.145***	0.130***	0.137***	0.143***	0.129***	0.135***
	(0.047)	(0.047)	(0.033)	(0.047)	(0.046)	(0.033)
Producerlinkcominteraction	-0.046	-0.073	-0.060	-0.063	-0.086	-0.075
	(0.128)	(0.087)	(0.075)	(0.117)	(0.082)	(0.070)
Consumerlinkcominteraction	0.034	0.028	0.031			
	(0.066)	(0.051)	(0.041)			
Labourcominteraction	0.132	0.137**	0.134**	0.164**	0.163***	0.163***
	(0.094)	(0.068)	(0.056)	(0.064)	(0.049)	(0.039)
Constant	0.507***	0.508***	0.507***	0.507***	0.508***	0.507***
	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	18.73	18.73	18.73	19.06	19.06	19.06
R-squared	0.139	0.166	0.149	0.138	0.165	0.149
Adjusted R-squared	0.130	0.158	0.145	0.130	0.158	0.145

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2.1.2. Municipality specialisation

The results of the estimations assessing the determinants of specialisation using the alternative measure of commuting are presented in table 6.4 below.

In general, the differences between the estimations employing the two different commuting variables are smaller for municipality specialisation compared to the results for sector concentration. Two

changes are, however, worth noting. The first relates to the estimated coefficient of producerlinkcominteraction, which is smaller in absolute value in the estimations using the alternative variable commutingratio. The second regards commutingratio itself, which, when included as a single variable, is significantly negative in all specifications. Hence, municipalities characterised by a large inflow of commuters in relation to the corresponding outflow tend to be characterised by a more diversified production structure rather than being specialised in a specific type of production.

Table 6.4. Specialisation using the alternative commuting variable

Normalised variables	Dependent variable: Specialisation1					
	(1) 2001	(2) 2003	(3) Whole sample	(4) 2001	(5) 2003	(6) Whole sample
Ln_Sectorshare1	0.512*** (0.020)	0.524*** (0.019)	0.508*** (0.014)	0.512*** (0.020)	0.524*** (0.019)	0.508*** (0.014)
Ln_Populationshare	0.014 (0.012)	0.013 (0.009)	0.013* (0.008)	0.014 (0.012)	0.013 (0.009)	0.014* (0.008)
Agrimineralsupply	0.027*** (0.007)	0.021*** (0.007)	0.024*** (0.005)	0.027*** (0.007)	0.021*** (0.007)	0.024*** (0.005)
Laboursupply	-0.095*** (0.023)	-0.086*** (0.021)	-0.091*** (0.016)	-0.096*** (0.023)	-0.087*** (0.022)	-0.092*** (0.016)
Marketpotential	0.258*** (0.050)	0.191*** (0.045)	0.221*** (0.034)	0.258*** (0.050)	0.191*** (0.045)	0.222*** (0.034)
Commutingratio	-0.021*** (0.008)	-0.019** (0.008)	-0.020*** (0.005)	-0.022*** (0.008)	-0.019** (0.008)	-0.020*** (0.005)
Agrimineralintensity	-0.408*** (0.017)	-0.438*** (0.016)	-0.415*** (0.012)	-0.408*** (0.017)	-0.438*** (0.016)	-0.416*** (0.012)
Labourintensity	-0.513*** (0.024)	-0.563*** (0.021)	-0.528*** (0.016)	-0.513*** (0.023)	-0.563*** (0.021)	-0.527*** (0.016)
Economies of scale	-0.090*** (0.019)	-0.081*** (0.019)	-0.080*** (0.014)	-0.090*** (0.019)	-0.081*** (0.019)	-0.080*** (0.014)
Consumerlink	0.203*** (0.013)	0.229*** (0.012)	0.212*** (0.009)	0.203*** (0.013)	0.228*** (0.012)	0.212*** (0.009)
Producerlink	0.589*** (0.027)	0.633*** (0.024)	0.599*** (0.018)	0.588*** (0.027)	0.632*** (0.024)	0.599*** (0.018)
Agrimineralinteraction	-0.025*** (0.008)	-0.015* (0.008)	-0.020*** (0.006)	-0.025*** (0.008)	-0.015* (0.008)	-0.020*** (0.006)

Consumerlinkinteraction	-0.314*** (0.044)	-0.255*** (0.039)	-0.277*** (0.029)	-0.317*** (0.044)	-0.260*** (0.039)	-0.281*** (0.029)
Ecoscaleinteraction	0.699*** (0.090)	0.601*** (0.084)	0.645*** (0.063)	0.699*** (0.090)	0.602*** (0.084)	0.645*** (0.063)
Producerlinkinteraction	-0.369*** (0.070)	-0.283*** (0.063)	-0.321*** (0.047)	-0.368*** (0.070)	-0.281*** (0.063)	-0.320*** (0.047)
Labourinteraction	0.157*** (0.037)	0.140*** (0.035)	0.146*** (0.026)	0.158*** (0.037)	0.141*** (0.035)	0.148*** (0.026)
Producerlinkcominteraction	-0.242** (0.109)	-0.211** (0.091)	-0.224*** (0.071)	-0.230** (0.101)	-0.191** (0.083)	-0.208*** (0.065)
Consumerlinkcominteraction	-0.024 (0.081)	-0.040 (0.066)	-0.032 (0.051)			
Labourcominteraction	0.278** (0.124)	0.245** (0.103)	0.258*** (0.080)	0.255** (0.106)	0.208** (0.087)	0.228*** (0.068)
Constant	0.304*** (0.008)	0.325*** (0.008)	0.312*** (0.006)	0.304*** (0.008)	0.325*** (0.008)	0.312*** (0.006)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	18.73	18.73	18.73	19.06	19.06	19.06
R-squared	0.784	0.824	0.803	0.784	0.824	0.803
Adjusted R-squared	0.782	0.822	0.802	0.782	0.822	0.802

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2.2. Employment as the dependent variable

6.2.2.1. Sector concentration

The results from the estimations investigating the determinants of sector concentration with the dependent variable measured using the number of employed are displayed in table 6.5 below. The signs and significance of the obtained coefficients are very similar to the results in table 6.1 presenting the results based on GRP as the dependent variable. The change of calculation method for the dependent variable does therefore not imply any large qualitative changes of the results. However, there is a general tendency of the coefficients from the estimation using employment being larger in absolute value, and the explanatory power (R^2) of the model is also notably larger. Furthermore, commuting is significant only in some specifications and the exclusion of the insignificant variable

consumerlinkcominteraction results in a comparably somewhat smaller estimated coefficient of producerlinkcominteraction.

To not overfill this section with regression tables, the estimation results using employment as the dependent variable and the alternative variable commutingratio to capture commuting flows are not presented here. However, with the exception of producerlinkcominteraction, which is insignificant, the tendencies regarding coefficient significance are the same as in table 6.5 but with somewhat larger significant estimates in absolute value. Noticeable is also that commutingratio is positive and significant in all specifications, and the value of the coefficient is somewhat larger than the corresponding value for the estimations using GRP.

Table 6.5. Concentration using the alternative dependent variable

	Dependent variable: Concentration (employment)					
	(1)	(2)	(3)	(4)	(5)	(6)
Normalised variables	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.000 (0.009)	0.000 (0.009)	-0.001 (0.006)	0.000 (0.009)	0.000 (0.009)	-0.001 (0.006)
Ln_Populationshare	0.040*** (0.007)	0.039*** (0.006)	0.039*** (0.005)	0.040*** (0.007)	0.039*** (0.006)	0.039*** (0.005)
Agrimineralsupply	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Laboursupply	-0.108** (0.042)	-0.097** (0.039)	-0.103*** (0.028)	-0.108*** (0.042)	-0.099*** (0.038)	-0.104*** (0.028)
Marketpotential	0.074 (0.086)	0.083 (0.071)	0.078 (0.055)	0.075 (0.081)	0.087 (0.069)	0.081 (0.053)
Commuting	0.801 (0.563)	0.688* (0.406)	0.741** (0.340)	0.796 (0.535)	0.669* (0.389)	0.728** (0.324)
Agrimineralintensity	-0.001 (0.010)	-0.001 (0.010)	-0.001 (0.007)	-0.001 (0.010)	-0.001 (0.010)	-0.001 (0.007)
Labourintensity	-0.007 (0.012)	-0.008 (0.012)	-0.007 (0.008)	-0.006 (0.012)	-0.008 (0.012)	-0.006 (0.009)
Economies of scale	-0.008 (0.008)	-0.007 (0.007)	-0.007 (0.005)	-0.008 (0.008)	-0.007 (0.007)	-0.007 (0.005)
Consumerlink	0.005 (0.006)	0.005 (0.005)	0.005 (0.004)	0.005 (0.006)	0.005 (0.005)	0.005 (0.004)

Producerlink	0.029 (0.019)	0.030 (0.019)	0.029** (0.013)	0.029 (0.019)	0.030 (0.019)	0.028** (0.013)
Agrimineralinteraction	0.005 (0.005)	0.006 (0.005)	0.005* (0.003)	0.005 (0.005)	0.006 (0.004)	0.006* (0.003)
Consumerlinkinteraction	-0.034 (0.051)	-0.029 (0.044)	-0.031 (0.033)	-0.035 (0.049)	-0.035 (0.042)	-0.035 (0.032)
Ecoscaleinteraction	0.059 (0.048)	0.047 (0.040)	0.052* (0.031)	0.059 (0.048)	0.047 (0.040)	0.052* (0.031)
Producerlinkinteraction	-0.279** (0.142)	-0.268** (0.122)	-0.272*** (0.092)	-0.280** (0.140)	-0.270** (0.121)	-0.273*** (0.091)
Labourinteraction	0.274*** (0.103)	0.252*** (0.095)	0.262*** (0.070)	0.274*** (0.103)	0.255*** (0.095)	0.264*** (0.070)
Producerlinkcominteraction	-0.871* (0.524)	-0.847** (0.356)	-0.858*** (0.309)	-0.860** (0.424)	-0.804*** (0.294)	-0.830*** (0.252)
Consumerlinkcominteraction	-0.012 (0.161)	-0.046 (0.110)	-0.030 (0.095)			
Labourcominteraction	0.036 (0.390)	0.142 (0.280)	0.093 (0.234)	0.026 (0.364)	0.104 (0.268)	0.068 (0.221)
Constant	-0.003 (0.006)	-0.002 (0.005)	-0.003 (0.004)	-0.003 (0.006)	-0.002 (0.006)	-0.003 (0.004)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.89	20.89	20.89	21.01	21.01	21.01
R-squared	0.259	0.262	0.259	0.259	0.261	0.259
Adjusted R-squared	0.252	0.255	0.255	0.252	0.254	0.255

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2.2.2. Municipality specialisation

As for concentration, the results for municipality specialisation, which are presented in table 6.6 below, are very similar to the estimation results obtained when GRP underlies the calculation of the dependent variable. The obtained coefficients are of the same sign, but, just as for concentration, their estimated impact is somewhat larger than their corresponding impact on GRP. Worth emphasising is also that the exclusion of the insignificant variable consumerlinkcominteraction results in smaller estimated coefficients for producerlinkcominteraction and labourcominteraction.

Again, the results of the estimations employing the alternative commuting variable are not extensively presented, but as in the estimations using GRP as the dependent variable, the estimated coefficient of commutingratio is significantly negative in all specifications, and compared to that estimation, it is even larger in absolute value. Similarly, the obtained coefficients for producerlinkcominteraction and labourcominteraction are larger in absolute value compared to the corresponding results for GRP as the dependent variable.

Table 6.6. Specialisation using the alternative dependent variable

	Dependent variable: Specialisation (employment)					
	(1)	(2)	(3)	(4)	(5)	(6)
Normalised variables	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.892*** (0.025)	0.901*** (0.026)	0.881*** (0.018)	0.892*** (0.025)	0.901*** (0.026)	0.881*** (0.018)
Ln_Populationshare	0.017 (0.013)	0.019 (0.013)	0.018* (0.009)	0.017 (0.013)	0.019 (0.013)	0.018* (0.009)
Agrimineralsupply	0.037*** (0.010)	0.035*** (0.010)	0.036*** (0.007)	0.037*** (0.010)	0.035*** (0.010)	0.036*** (0.007)
Laboursupply	-0.136*** (0.032)	-0.134*** (0.031)	-0.135*** (0.022)	-0.139*** (0.032)	-0.137*** (0.031)	-0.138*** (0.023)
Marketpotential	0.361*** (0.068)	0.287*** (0.060)	0.319*** (0.046)	0.366*** (0.067)	0.293*** (0.060)	0.324*** (0.045)
Commuting	0.104 (0.115)	0.093 (0.098)	0.098 (0.075)	0.077 (0.070)	0.063 (0.059)	0.071 (0.046)
Agrimineralintensity	-0.728*** (0.021)	-0.761*** (0.021)	-0.733*** (0.015)	-0.728*** (0.021)	-0.761*** (0.021)	-0.733*** (0.015)
Labourintensity	-0.894*** (0.028)	-0.965*** (0.029)	-0.913*** (0.020)	-0.893*** (0.028)	-0.964*** (0.029)	-0.912*** (0.020)
Economies of scale	-0.163*** (0.027)	-0.143*** (0.027)	-0.143*** (0.019)	-0.163*** (0.027)	-0.143*** (0.027)	-0.143*** (0.019)
Consumerlink	0.337*** (0.016)	0.376*** (0.017)	0.350*** (0.012)	0.336*** (0.016)	0.375*** (0.017)	0.350*** (0.012)
Producerlink	1.021*** (0.032)	1.082*** (0.033)	1.033*** (0.023)	1.020*** (0.032)	1.081*** (0.033)	1.032*** (0.023)
Agrimineralinteraction	-0.036*** (0.011)	-0.031*** (0.012)	-0.033*** (0.008)	-0.035*** (0.011)	-0.031*** (0.011)	-0.033*** (0.008)

Consumerlinkinteraction	-0.476*** (0.061)	-0.394*** (0.052)	-0.423*** (0.040)	-0.484*** (0.060)	-0.403*** (0.052)	-0.432*** (0.039)
Ecoscaleinteraction	1.086*** (0.143)	0.932*** (0.125)	1.000*** (0.096)	1.087*** (0.143)	0.933*** (0.125)	1.001*** (0.096)
Producerlinkinteraction	-0.482*** (0.089)	-0.390*** (0.078)	-0.427*** (0.059)	-0.486*** (0.089)	-0.394*** (0.078)	-0.431*** (0.059)
Labourinteraction	0.210*** (0.054)	0.199*** (0.051)	0.201*** (0.037)	0.214*** (0.054)	0.203*** (0.051)	0.205*** (0.037)
Producerlinkcominteraction	-0.485* (0.275)	-0.434* (0.233)	-0.456** (0.180)	-0.425** (0.190)	-0.367** (0.162)	-0.393*** (0.125)
Consumerlinkcominteraction	-0.063 (0.142)	-0.070 (0.121)	-0.066 (0.093)			
Labourcominteraction	0.393** (0.191)	0.352** (0.161)	0.367*** (0.125)	0.340** (0.139)	0.294** (0.121)	0.312*** (0.091)
Constant	0.197*** (0.009)	0.225*** (0.010)	0.207*** (0.007)	0.196*** (0.009)	0.225*** (0.010)	0.206*** (0.007)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.89	20.89	20.89	21.01	21.01	21.01
R-squared	0.854	0.864	0.858	0.854	0.864	0.858
Adjusted R-squared	0.853	0.863	0.858	0.853	0.863	0.858

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2.3. Standardisation as variable transformation method

Estimations using standardised variables are performed for the basic specification as well as for the alternative specifications using commutingratio and employment, respectively and combined. Due to space limitations, however, only the results from the basic specification are presented in table format. Nevertheless, any observable discrepancies or tendencies obtained in the alternative specifications are briefly discussed in words below.

6.2.3.1. Sector concentration

Table 6.7 below presents the results from the estimations using standardised variables in the estimation of the determinants of concentration. Although the coefficients of standardised and normalised variables are not directly comparable, the tendencies regarding the interaction variables

reported for the normalised variables seem to hold also in the current specification. As in the previous results for normalised variables, labourcominteraction is insignificant throughout all specifications and the negative coefficient of producerlinkcominteraction is larger than the estimated coefficient for producerlinkinteraction. Noticeable is also that, in contrast to previous results, agrimineralinteraction is now partly significant.

Also the results for the specification using the alternative variable commutingratio are similar to the obtained results for normalised variables. Commutingratio is significantly positive in all specifications, but the estimated coefficient is smaller than the corresponding estimated impact of the basic commuting variable. Moreover, labourcominteraction is now significantly positive in all but one specification, and both its size and significance seem to increase when the insignificant variable consumerlinkcominteraction is excluded. Similar to the results for normalised variables, producerlinkcominteraction is insignificant in all specifications and producerlinkinteraction displays a smaller coefficient in absolute value compared to the estimations using the basic commuting variable.

For employment as the dependent variable, the obtained results are very similar to the results of the estimations using GRP. Noticeable is, however, that just as when using normalised variables, the impact of commuting is weaker when employment forms the basis for the calculation of the dependent variable, with significant estimates obtained in four of six specifications.

Finally, the results of the estimations employing both alternative variables commutingratio and employment reveal a significantly positive estimate of commutingratio in all specifications. On the contrary, the coefficient for producerlinkcominteraction is insignificant throughout the estimations. Both these results are in line with the reported tendencies for the estimations using GRP as the dependent variable. The estimated impact of the significant variables is also in general slightly larger for employment compared to GRP. However, in contrast to those results, labourcominteraction is significant in only one specification.

Table 6.7. Concentration using standardised variables

Standardised variables	Dependent variable: Concentration1					
	(1)	(2)	(3)	(4)	(5)	(6)
	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.000 (0.080)	0.000 (0.068)	-0.004 (0.054)	0.000 (0.080)	0.000 (0.068)	-0.004 (0.054)
Ln_Populationshare	0.135*** (0.041)	0.136*** (0.019)	0.135*** (0.022)	0.135*** (0.041)	0.136*** (0.019)	0.135*** (0.022)
Agrimineralsupply	0.012 (0.031)	-0.042* (0.025)	-0.016 (0.020)	0.011 (0.031)	-0.043* (0.025)	-0.017 (0.020)
Laboursupply	-0.245*** (0.094)	-0.220** (0.093)	-0.233*** (0.066)	-0.254*** (0.094)	-0.226** (0.093)	-0.240*** (0.066)
Marketpotential	0.209 (0.165)	0.244 (0.160)	0.228** (0.116)	0.227 (0.158)	0.256 (0.157)	0.242** (0.113)
Commuting	1.123* (0.651)	0.800** (0.373)	0.950*** (0.367)	1.070* (0.620)	0.765** (0.353)	0.906*** (0.348)
Agrimineralintensity	0.000 (0.129)	-0.001 (0.101)	0.005 (0.084)	0.000 (0.129)	-0.001 (0.101)	0.005 (0.084)
Labourintensity	-0.082 (0.165)	-0.085 (0.109)	-0.078 (0.102)	-0.071 (0.165)	-0.078 (0.111)	-0.069 (0.102)
Economies of scale	-0.076 (0.067)	-0.069 (0.062)	-0.069 (0.046)	-0.077 (0.067)	-0.069 (0.062)	-0.070 (0.046)
Consumerlink	0.069 (0.076)	0.059 (0.049)	0.061 (0.046)	0.062 (0.075)	0.054 (0.049)	0.056 (0.046)
Producerlink	0.244 (0.208)	0.244 (0.183)	0.237* (0.140)	0.239 (0.208)	0.241 (0.184)	0.233* (0.141)
Agrimineralinteraction	0.018 (0.028)	0.058*** (0.020)	0.039** (0.017)	0.020 (0.028)	0.058*** (0.020)	0.040** (0.018)
Consumerlinkinteraction	-0.075 (0.074)	-0.063 (0.071)	-0.068 (0.051)	-0.093 (0.071)	-0.074 (0.070)	-0.082* (0.050)
Ecoscaleinteraction	0.098 (0.063)	0.083 (0.057)	0.089** (0.042)	0.100 (0.063)	0.083 (0.057)	0.090** (0.042)
Producerlinkinteraction	-0.473** (0.206)	-0.480** (0.208)	-0.475*** (0.146)	-0.482** (0.204)	-0.486** (0.206)	-0.482*** (0.145)

Labourinteraction	0.524*** (0.182)	0.477*** (0.181)	0.498*** (0.128)	0.532*** (0.181)	0.482*** (0.180)	0.505*** (0.127)
Producerlinkcominteraction	-1.025** (0.441)	-0.824*** (0.269)	-0.916*** (0.251)	-0.941*** (0.350)	-0.768*** (0.208)	-0.848*** (0.198)
Consumerlinkcominteraction	-0.074 (0.118)	-0.048 (0.080)	-0.060 (0.069)			
Labourcominteraction	0.274 (0.392)	0.316 (0.238)	0.296 (0.219)	0.185 (0.353)	0.258 (0.206)	0.224 (0.195)
Constant	-0.011 (0.020)	0.010 (0.022)	0.000 (0.014)	-0.011 (0.020)	0.010 (0.022)	0.000 (0.014)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.86	20.86	20.86	20.98	20.98	20.98
R-squared	0.177	0.196	0.182	0.176	0.195	0.182
Adjusted R-squared	0.169	0.188	0.178	0.168	0.188	0.178

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.2.3.2. Municipality specialisation

Table 6.8 below reports the results of the estimations assessing the determinants of municipality specialisation using standardised variables. As for concentration, the estimation results seem to be robust to the change of variable-transformation method. Furthermore, the estimated coefficients for the interaction variables are of the same signs as in the setting with normalised variables and as in that case, commuting is insignificant in all specifications.

The regressions including the alternative variable commutingratio reveal coefficient estimates very similar to the results obtained when using the basic commuting variable and just as for normalised variables, the estimated coefficient of commutingratio is significantly negative in all specifications.

Testing the sensitivity of the results by using employment in the estimations with standardised variables yield estimated coefficients of the interaction variables that in general are somewhat smaller compared to the setting with GRP as the dependent variable. Just as in that specification, the estimated coefficients of the basic commuting variable are insignificant.

Combing the two specification changes by using both commutingratio and employment results in significantly negative estimates of commutingratio in all specifications, and the size of the coefficient

is similar to the obtained estimates for GRP as the dependent variable. The estimated coefficients for producerlinkcominteraction and labourcominteraction are however somewhat smaller than in that setting.

Table 6.8. Specialisation using standardised variables

Standardised variables	Dependent variable: Specialisation1					
	(1) 2001	(2) 2003	(3) Whole sample	(4) 2001	(5) 2003	(6) Whole sample
Ln_Sectorshare1	1.094*** (0.043)	1.121*** (0.041)	1.087*** (0.030)	1.094*** (0.043)	1.121*** (0.041)	1.087*** (0.030)
Ln_Populationshare	0.009 (0.014)	0.008 (0.011)	0.008 (0.009)	0.009 (0.014)	0.008 (0.011)	0.008 (0.009)
Agrimineralsupply	0.043** (0.022)	0.037* (0.020)	0.041*** (0.015)	0.043** (0.022)	0.037* (0.020)	0.041*** (0.015)
Laboursupply	-0.064*** (0.015)	-0.060*** (0.015)	-0.062*** (0.011)	-0.065*** (0.015)	-0.062*** (0.015)	-0.063*** (0.011)
Marketpotential	0.142*** (0.032)	0.109*** (0.028)	0.124*** (0.021)	0.144*** (0.032)	0.113*** (0.028)	0.127*** (0.021)
Commuting	0.024 (0.035)	0.024 (0.030)	0.024 (0.023)	0.017 (0.024)	0.014 (0.020)	0.015 (0.016)
Agrimineralintensity	-1.294*** (0.053)	-1.380*** (0.049)	-1.313*** (0.036)	-1.294*** (0.053)	-1.380*** (0.049)	-1.313*** (0.036)
Labourintensity	-1.353*** (0.062)	-1.487*** (0.056)	-1.392*** (0.042)	-1.352*** (0.062)	-1.485*** (0.056)	-1.391*** (0.042)
Economies of scale	-0.221*** (0.048)	-0.200*** (0.047)	-0.195*** (0.034)	-0.221*** (0.048)	-0.200*** (0.047)	-0.196*** (0.034)
Consumerlink	0.531*** (0.034)	0.599*** (0.032)	0.555*** (0.023)	0.530*** (0.033)	0.597*** (0.032)	0.554*** (0.023)
Producerlink	1.431*** (0.066)	1.543*** (0.058)	1.458*** (0.045)	1.430*** (0.066)	1.542*** (0.059)	1.458*** (0.045)
Agrimineralinteraction	-0.027* (0.016)	-0.020 (0.014)	-0.023** (0.011)	-0.027* (0.016)	-0.019 (0.014)	-0.023** (0.011)
Consumerlinkinteraction	-0.158*** (0.022)	-0.130*** (0.020)	-0.140*** (0.015)	-0.161*** (0.022)	-0.133*** (0.020)	-0.143*** (0.015)

Ecoscaleinteraction	0.298*** (0.039)	0.258*** (0.036)	0.276*** (0.027)	0.298*** (0.039)	0.258*** (0.036)	0.276*** (0.027)
Producerlinkinteraction	-0.153*** (0.033)	-0.123*** (0.030)	-0.136*** (0.022)	-0.155*** (0.032)	-0.125*** (0.030)	-0.138*** (0.022)
Labourinteraction	0.091*** (0.022)	0.083*** (0.021)	0.086*** (0.015)	0.092*** (0.022)	0.085*** (0.021)	0.087*** (0.015)
Producerlinkcominteraction	-0.104* (0.056)	-0.095* (0.048)	-0.098*** (0.037)	-0.093** (0.043)	-0.078** (0.035)	-0.085*** (0.028)
Consumerlinkcominteraction	-0.009 (0.025)	-0.014 (0.020)	-0.012 (0.016)			
Labourcominteraction	0.105** (0.048)	0.094** (0.041)	0.098*** (0.031)	0.093** (0.039)	0.077** (0.032)	0.084*** (0.025)
Constant	-0.014 (0.010)	0.015 (0.010)	0.000 (0.007)	-0.014 (0.010)	0.015 (0.010)	0.000 (0.007)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.86	20.86	20.86	20.98	20.98	20.98
R-squared	0.782	0.823	0.802	0.782	0.823	0.802
Adjusted R-squared	0.780	0.821	0.801	0.780	0.821	0.801

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

7. Discussion

The results reported above partly support the discussion in section 2.5 about the theoretical implications for the empirical study in this thesis. When included as a single variable, commuting as a proxy for the quality of infrastructure in general contribute to the concentration of economic activity as predicted by the model by Martin and Rogers (1995). This result holds for both included measures of commuting, but is somewhat less stable for employment as the dependent variable.

However, the presumption of sectors characterised by strong consumer links concentrating in areas characterised by large market potential is in general not supported by the empirical study. Moreover, the results regarding the impact of the importance of producer links even indicate that sectors characterised by strong links to other producers locate outside areas with large market potential. Thereby, the implications of the standard New Economic Geography-model by Krugman (1991) of the importance of producer and consumer links resulting in firms concentrating in core areas, do not find any support in the empirical study.

Neither the suggestion by Gruber (2010) of commuting inducing sectors characterised by strong consumer links to be less concentrated is supported by the performed econometric analysis. The interaction variable combining commuting and the consumer link is even insignificant in all specifications. However, sectors experiencing strong links to other producers tend to be less concentrated in municipalities characterised by a large degree of commuting. A possible interpretation would be that good-quality infrastructure reduces the costs associated with locating far away from other producers, thereby resulting in less concentrated industries. By this reasoning, the importance of transport costs in the model by Krugman (1991) is to some extent supported by the empirical study.

The estimations assessing the determinants of municipality specialisation indicate that municipalities characterised by much commuting tend to be less specialised in sectors for which links to producers are of large importance, whereas no significant tendency is found for sectors characterised by strong consumer links. Furthermore, municipalities characterised by large market potential tend to be more specialised in sectors characterised by a relatively low degree of economies of scale in production and less specialised in sectors for which consumer and producer links are of relatively large importance. In this case, the results regarding producer and consumer links seem more reasonable. Municipalities characterised by a large market potential are mostly located in conurbation areas, which are also more densely populated. The associated larger demand likely also enables a more diversified production structure and, hence, less specialised municipalities.

However, in contrast to the results on concentration, there is no indication of the degree of commuting and the quality of infrastructure in itself having a statistically significant impact on the specialisation patterns of municipalities, and the predictions of the model by Martin and Rogers (1995) do therefore not apply in this case. Nevertheless, when measured in order to capture commuting flows, the commuting variable is significantly negative in all specifications. Municipalities in which the inflow of commuters dominates the outflow in general thus tend to be less specialised and thereby characterised by a relatively diversified production structure.

Noticeable is also that labour-intensive sectors tend to be overrepresented in municipalities characterised by abundance of labour and that these municipalities are also relatively specialised in labour-intensive sectors. For these sectors, the predictions by the Heckscher-Ohlin theory are thereby confirmed for both sector concentration and municipality specialisation.

Furthermore, whereas municipalities characterised by much commuting seem to be relatively specialised in labour-intensive production in both specifications of the commuting variable, a similar tendency appears for concentration only when commuting is measured by the ratio of in- to out-commuters. Hence, for labour-intensive sectors, the direction of commuting flows rather than the mere magnitude of these flows are of importance in determining the concentration of economic activity.

The performed robustness tests with respect to the variable commuting, the measurement of the dependent variable and the variable-transformation method, respectively, partly affect the obtained results in the empirical study. In general, using employment as the basis for the calculation of the dependent variables does not qualitatively change the obtained results, but the estimated coefficients are larger in absolute value and the empirical model can explain the variation in the dependent variable to a larger extent than in the corresponding setting with GRP. Also the number and choice of included interaction variables affect the size of the coefficients to some extent, but the sign and significance of the obtained estimates in general seem to be robust also to such changes. The specification change affecting the results to the largest extent is the use of two different commuting variables, focusing on the quantity and the direction of the commuting flows, respectively. When changing from the former to the latter, the importance of producer links tend to decrease and the importance of labour intensity in production tend to increase when it comes to both sector concentration and municipality specialisation.

In light of the fact that most interaction variables tend to decrease in importance over time, it is interesting to note that the explanatory power (R^2) in general is higher in the estimations for 2003 compared to 2001. Although it should be kept in mind that the limited availability of data does not enable a more extensive analysis of any trends and tendencies occurring over time, a cautious

interpretation would be that some of the variation in concentration and specialisation, respectively, explained by the interaction between municipality supplies and industry intensities in 2001 can be explained by the single- and/or the size variables in 2003. Worth emphasising is also the considerable difference in the ability of the model to explain the determinants of sector concentration and municipality specialisation. Depending on the exact specification, the empirical model is able to explain 14-25 and 78-87 percent, respectively, of the variation in concentration and specialisation. The difference might seem large, but is in line with the reported R^2 -value in similar empirical studies.⁴⁶

The with regard to theoretical predictions partly contradictory results obtained by the performed empirical study might also to some extent be due to the presence of multicollinearity among the explanatory variables. Although the undertaken measures to reduce the degree of multicollinearity among the explanatory variables succeeded in substantially reducing the degree of this correlation, multicollinearity is still present in the performed empirical study. As described in section 4.4, this can result not only in statistical insignificance of obtained estimates but also in reversed signs of estimated coefficients and a model sensitive to small data changes.

For this reason, and due to the restrictions imposed by the limited data availability, further research is needed before for any general conclusions can be drawn. Analysing the impact of commuting patterns on sector concentration and municipality specialisation using panel data methods would require data for a larger number of years, and desirably also for a larger number of sectors. In addition, such data would enable an empirical study to more thoroughly investigate tendencies and trends occurring over time. The assessment of the determinants of concentration of economic activity and the specialisation of geographical units would likely also benefit from a further development of the theoretical model underlying empirical studies similar to the one performed within this thesis. Interesting extensions could for example cover the inclusion of potential non-linear relationships as predicted for transport costs in the model by Krugman (1991). Using spatial econometric methods to take potential spatial dependence among the geographical units into account might also be a fruitful approach to be employed by future research.

⁴⁶ See section 3 for a review of previous empirical research within the field. Note, however, that comparisons should be made with some caution since the sample and the empirical specification of the estimated models might vary between the studies.

8. Conclusion

The aim of this thesis is to assess the impact of commuting patterns and the quality of infrastructure on the degree of sector concentration and municipality specialisation.

The results of the estimations using Ordinary Least Squares (OLS) on seven different sectors in 286 Swedish municipalities for the years 2001 and 2003 indicate that the inflow of commuters and the overall magnitude of commuting, which can act as a proxy for the quality of infrastructure, in itself contributes to the concentration of economic activity. Moreover, municipalities dominated by an inflow of commuters are characterised by a relatively more diversified production structure.

Overall, the empirical model is able to explain differences in municipality specialisation to a larger extent than it can explain variations in sector concentration. Somewhat surprisingly, sectors characterised by strong links to other producers do not tend to concentrate in municipalities characterised by large market potential. Such municipalities also tend to be relatively less specialised in sectors characterised by strong producer and consumer links and relatively more specialised in sectors characterised by a low degree of economies of scale in production.

Sectors characterised by strong producer links also tend to be less concentrated in municipalities characterised by a large overall degree of commuting, and sectors using labour intensively in production tend to concentrate in municipalities in which the inflow of commuters is relatively large compared to the outflow. Similarly, municipalities characterised by much commuting in terms of both magnitude and direction are more specialised in labour-intensive sectors and less specialised in sectors for which links to other producers are important. The empirical study also reveals that sectors using labour intensively in production are overrepresented in labour-abundant municipalities, and labour-abundant municipalities are also relatively more specialised in such sectors.

The obtained results are qualitatively robust to the use of two different variables transformation methods and two alternative measurements of the dependent variables, but data covering a larger number of years is needed to enable a more extensive empirical analysis. In addition, although to a large extent remedied, multicollinearity among the explanatory variables is present in the performed empirical study. Although the in some specifications highly significant estimated impacts of the variables representing commuting aspects should not be neglected, it therefore remains a task for future research to more extensively analyse the impact of commuting-related aspects on different phenomena in the economy.

9. References

- Baldwin, R. and Forslid, R. (2000). "The Core-Periphery Model and Endogenous Growth: Stabilizing and Destabilizing Integration", *Economica*, 67:307-324.
- Braunerhjelm, P. and Johansson, D (2003). "The Determinants of Spatial Concentration: the Manufacturing and Service Sectors in an International Perspective", *Industry and Innovation*, 10:41-63.
- Cameron, A. C. and Pravin, K. T. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Campos, J., Ericsson, N. R., and Hendry, D. F. (2005). "General-to-specific Modeling: An Overview and Selected Bibliography". FRB International Finance Discussion Paper No. 838
- Ciccone, A. and Hall, R. E. (1996), "Productivity and the Density of Economic Activity", *The American Economic Review*, 86:54-71.
- Combes, P-P. and Overman, H. (2003). "The Spatial Distribution of Economic Activities in the EU". CEPR Discussion Paper No. 3999.
- Eliasson, K., Westerlund, O., and Åström, J. (2007). *Flyttning och pendling i Sverige: bilaga 3 till Långtidsutredningen 2008*. SOU 2007:35.
- Dornbusch, R., Fischer, S. and Samuelson, P. A. (1980). "Heckscher-Ohlin Trade Theory with a Continuum of Goods". *The Quarterly Journal of Economics*, 95:203-224.
- Epifani, P. (2005). "Heckscher-Ohlin and agglomeration". *Regional Science and Urban Economics*, 35:645-657.
- Ellison, G. and Glaeser, E. (1999). "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?", *American Economic Review*, 89:311-316.
- EU1(2012-03-02)=*Regional policy: Bridging the prosperity gap*. European Union. Retrieved 2012-03-07 from http://europa.eu/pol/reg/index_en.htm.

Fox, W. F. and Murray, M. N. (1990). "Local Public Policies and Interregional Business Development", *Southern Economic Journal*, 57:413-427.

Glesbygdsverket (2002). *Glesbygdsverkets Årsbok 2002*. Stockholm: Edita Norstedts Tryckeri AB.

Gruber, S. (2010). "To Migrate or to Commute?". *Review of Economic Analysis*, 2:110-134.

Gujarati, D. N. (2006). *Essentials of Econometrics*. New York ; Singapore : McGraw-Hill/Irwin, 2006

Gullstrand, J. and Hammarlund, C. (2007). *Plats för tillväxt?: Bilaga 2 till Långtidsutredningen 2008*. SOU 2007:25.

Han, J., Kamber, M and Pei, J (2011). *Data Mining: Concepts and techniques*. San Francisco: Morgan Kaufmann.

Krugman, P. (1991). "Increasing Returns and Economic Geography". *Journal of Political Economy*, 99:483-499.

Krugman, P. and Obstfeld, M. (2009). *International Trade; Theory and Policy*, eighth edition, Boston Massachusetts: Addition-Wesley.

Markinfo (2007-02-10). *Definition: ägoslag*. Swedish University of Agricultural Sciences (SLU). Retrieved 2012-04-25 from <http://www-markinfo.slu.se/sve/mark/ago/agodef.html>.

Martin, P. and Rogers, C. A. (1995). "Industrial Location and Public Infrastructure". *Journal of International Economics*, 39:335-351.

Midelfart-Knarvik, K. H., Overman, H. G., Redding, S. and Venables, A. J. (2000). "The Location of European Industry", European Commission Economic Papers No.142.

Midelfart-Knarvik, K. H., Overman, H. G. and Venables, A. J. (2002). "Comparative Advantage and Economic Geography: Estimating the Determinants of Industrial Location in the EU". CEPR Discussion Paper No. 2618.

Murata, Y. and Thisse, J-F. (2005). "A Simple Model of Economic Geography a' la Helpman-Tabuchi". *Journal of Urban Economics*, 58:137-155.

Naudé, W. (2009). “Geography, transport and Africa’s proximity gap”, *Journal of Transport Geography*, 17:1-9.

Ottaviano, Gianmarco I.P., Takatoshi Tabuchi and Jacques-Francois Thisse (2002), “Agglomeration and Trade Revisited”, *International Economic Review* 43, 409-435.

Overman, H. G., Redding, S. and Venables, A. J. (2001). “The Economic Geography of Trade, Production and Income: A Survey of Empirics”. CEPR Discussion Papers No. 2978.

Puga, D. (2008). “Agglomeration and Cross-border Infrastructure”. IMDEA Working Papers 2008-06.

Skatteverket (2010). *Taxes in Sweden 2010*. Retrieved 2012-04-30 from <http://www.skatteverket.se/download/18.616b78ca12d1247a4b2800025728/10411.pdf>

Verbeek, M. (2008): *A Guide to Modern Econometrics*, third edition. West Sussex: John Wiley & Sons Ltd.

Washington, S. P., Karlaftis, M. G. And Mannering, F. L. (2010). *Statistical and Econometric Methods for Transportation Data Analysis*. Second edition. Boca Raton. Chapman and Hall/CRC

SKL (2009-10-14). *Äldre indelning 1999-2004*. Retrieved 2012-03-03 from http://www.skl.se/kommuner_och_landsting/om_kommuner/kommungruppsindelning/aldre_indelning

SCB: *Regionala tabeller i Statistikdatabasen* Retrieved 2012-04-08 from http://www.ssd.scb.se/bjssd/sok_link.asp?sokord1=efter+kommun&xu=c5587001&yp=duwird&lang=1

SCB: *Symmetric input-output-tables*. Retrieved 2012-03-26 from http://www.scb.se/Pages/GsaSearch____287280.aspx?QueryTerm=input%20output%202000&PageIndex=1&hl=sv

Senior Nello, S. (2005). *The European Union; Economics, Policies and History*. London: McGraw-Hill

Tabuchi, T. and Thisse, J-F. (2006). “Regional Specialization, Urban Hierarchy, and Commuting Costs”, *International Economic Review*, 47:1295–1317.

Young, L. (1991). “Heckscher-Ohlin trade theory with variable returns to scale”. *Journal of International Economics*, 31:183-190.

10. Appendices

10.1. Appendix 1: Descriptive statistics

Table A1. Descriptive statistics

Variable	Mean	Std dev	Min	Max
Concentration1	0.003	0.014	-0.331	0.318
Concentration2	0.003	0.014	-0.349	0.314
Concentration (employment)	0.003	0.013	0.000	0.327
Specialisation1	0.110	0.206	-0.282	0.946
Specialisation2	0.110	0.206	-0.282	0.946
Specialisation (employment)	0.109	0.207	0.000	0.946
Sectorshare1	0.119	0.236	0.0003	0.697
Sectorshare2	0.118	0.234	0.0003	0.691
Sectorshare_average ⁴⁷	0.118	0.235	0.0003	0.694
Populationshare	0.003	0.006	0.0003	0.082
Agrimineralsupply	22.741	39.870	0.000	560.603
Laboursupply	133.626	445.181	0.285	410.399
Marketpotential	4906.799	3165.497	858.538	35179.172
Commutingaverage	0.315	0.739	0.020	11.346
Commutingratio	0.803	0.565	0.090	3.188
Agriminerallintensity	0.102	0.084	0.006	0.205
Economies of scale1	0.089	0.153	0.001	0.458
Economies of scale2	0.089	0.153	0.001	0.460
Economies of scale_average ⁴⁷	0.089	0.153	0.001	0.459
Consumerlink	0.136	0.110	0.003	0.313
Producerlink	0.540	0.124	0.421	0.797
Agriminerallinteraction	2.310	5.593	0.000	114.924
Labourinteraction	59.476	208.892	0.046	2472.452
Consumerlinkinteraction	665.222	771.346	2.576	11011.081
Producerlinkinteraction	2652.125	1858.257	361.015	28037.803
Ecoscaleinteraction1	436.701	937.212	0.595	16129.649
Ecoscaleinteraction2	437.563	939.092	0.596	16169.209
Ecoscaleinteraction_average ⁴⁷	437.132	938.152	0.595	16149.429
Labourcominteraction	0.140	0.348	0.003	6.361
Producerlinkcominteraction	0.170	0.411	0.008	9.043
Consumerlinkcominteraction	0.043	0.133	0.000	3.551

⁴⁷ Used when the dependent variable is based on the number of employed rather than one of the two different measures of GRP.

10.2. Appendix 2: Industry intensities

Table A2. Industry intensities

Intensity	Sector						
	Agriculture	Forest	Fishing	Minerals	Low-tech	High-tech	Services
Agrimineralintensity 2001	0.205	0.194	0.009	0.170	0.113	0.006	0.014
Agrimineralintensity 2003	0.205	0.194	0.009	0.170	0.113	0.006	0.014
Labourintensity 2001	0.471	0.162	0.305	0.558	0.549	0.517	0.549
Labourintensity 2003	0.473	0.163	0.301	0.561	0.551	0.521	0.551
Consumerlink 2001	0.221	0.024	0.100	0.003	0.313	0.060	0.228
Consumerlink 2003	0.221	0.024	0.100	0.003	0.313	0.060	0.228
Producerlink 2001	0.644	0.467	0.421	0.797	0.519	0.448	0.490
Producerlink 2003	0.644	0.467	0.421	0.797	0.519	0.448	0.490
Economies of scale 2001	458.668	90.262	4.766	2.790	100.444	12.185	1797.448
Economies of scale 2003	473.563	129.255	5.161	2.787	103.388	12.325	1910.304

10.3. Appendix 2: Alternative estimation results

All estimation results using the first measure of GRP presented in section 6 have been estimated also using the second available measure. Due to space limitations, only the regression tables for the basic specification are presented in this appendix, but also for the other specifications, the results are very similar to the obtained estimates using the first measure of GRP.

Table A3. Concentration using normalised variables

Normalised variables	Dependent variable: Concentration2					
	(1) 2001	(2) 2003	(3) Whole sample	(4) 2001	(5) 2003	(6) Whole sample
Ln_Sectorshare1	-0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)	-0.000 (0.006)	0.000 (0.005)	-0.000 (0.004)
Ln_Populationshare	0.018*** (0.006)	0.019*** (0.003)	0.018*** (0.003)	0.018*** (0.006)	0.019*** (0.003)	0.018*** (0.003)
Agrimineralsupply	0.001 (0.002)	-0.000 (0.001)	0.000 (0.001)	0.001 (0.002)	-0.000 (0.001)	0.000 (0.001)
Laboursupply	-0.052*** (0.019)	-0.048** (0.019)	-0.050*** (0.014)	-0.054*** (0.019)	-0.049** (0.019)	-0.052*** (0.014)
Marketpotential	0.049 (0.037)	0.060* (0.036)	0.055** (0.026)	0.054 (0.035)	0.062* (0.036)	0.059** (0.025)
Commuting	0.379* (0.220)	0.259** (0.121)	0.315** (0.123)	0.359* (0.210)	0.247** (0.115)	0.299** (0.117)
Agrimineralintensity	-0.000 (0.006)	-0.001 (0.005)	-0.000 (0.004)	-0.000 (0.006)	-0.001 (0.005)	-0.000 (0.004)
Labourintensity	-0.005 (0.010)	-0.005 (0.006)	-0.005 (0.006)	-0.004 (0.010)	-0.005 (0.006)	-0.004 (0.006)
Economies of scale	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.003)	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.003)
Consumerlink	0.004 (0.005)	0.004 (0.003)	0.004 (0.003)	0.004 (0.005)	0.003 (0.003)	0.003 (0.003)
Producerlink	0.016 (0.013)	0.016 (0.011)	0.015* (0.009)	0.015 (0.013)	0.015 (0.011)	0.015* (0.009)
Agrimineralinteraction	0.001 (0.003)	0.004 (0.003)	0.002 (0.002)	0.002 (0.003)	0.004 (0.003)	0.003 (0.002)

Consumerlinkinteraction	-0.022 (0.022)	-0.019 (0.022)	-0.020 (0.015)	-0.029 (0.021)	-0.023 (0.021)	-0.025* (0.015)
Ecoscaleinteraction	0.035 (0.024)	0.031 (0.021)	0.033** (0.016)	0.036 (0.023)	0.031 (0.021)	0.033** (0.016)
Producerlinkinteraction	-0.146** (0.062)	-0.151** (0.063)	-0.148*** (0.044)	-0.149** (0.061)	-0.152** (0.062)	-0.150*** (0.044)
Labourinteraction	0.129*** (0.045)	0.119*** (0.046)	0.123*** (0.032)	0.131*** (0.045)	0.120*** (0.046)	0.125*** (0.032)
Producerlinkcominteraction	-0.496** (0.214)	-0.381*** (0.125)	-0.434*** (0.121)	-0.449*** (0.169)	-0.354*** (0.096)	-0.398*** (0.095)
Consumerlinkcominteraction	-0.050 (0.070)	-0.028 (0.045)	-0.038 (0.040)			
Labourcominteraction	0.112 (0.159)	0.123 (0.093)	0.118 (0.088)	0.070 (0.142)	0.100 (0.080)	0.086 (0.078)
Constant	0.524*** (0.004)	0.525*** (0.003)	0.525*** (0.002)	0.524*** (0.004)	0.525*** (0.003)	0.525*** (0.002)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.72	20.72	20.72	20.82	20.82	20.82
R-squared	0.176	0.200	0.183	0.175	0.199	0.182
Adjusted R-squared	0.168	0.192	0.180	0.167	0.192	0.179

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A4. Specialisation using normalised variables

Normalised variables	Dependent variable: Specialisation2					
	(1)	(2)	(3)	(4)	(5)	(6)
	2001	2003	Whole sample	2001	2003	Whole sample
Ln_Sectorshare1	0.500*** (0.020)	0.520*** (0.019)	0.501*** (0.014)	0.500*** (0.020)	0.520*** (0.019)	0.501*** (0.014)
Ln_Populationshare	0.008 (0.012)	0.008 (0.009)	0.008 (0.008)	0.008 (0.012)	0.008 (0.009)	0.008 (0.008)
Agrimineralsupply	0.026*** (0.007)	0.021*** (0.007)	0.023*** (0.005)	0.026*** (0.007)	0.020*** (0.007)	0.023*** (0.005)
Laboursupply	-0.104*** (0.023)	-0.092*** (0.021)	-0.097*** (0.016)	-0.105*** (0.024)	-0.094*** (0.022)	-0.099*** (0.016)
Marketpotential	0.264*** (0.051)	0.192*** (0.044)	0.222*** (0.033)	0.268*** (0.050)	0.196*** (0.043)	0.227*** (0.033)
Commuting	0.052 (0.075)	0.055 (0.064)	0.054 (0.049)	0.035 (0.049)	0.031 (0.041)	0.034 (0.032)
Agrimineralintensity	-0.401*** (0.017)	-0.434*** (0.016)	-0.411*** (0.012)	-0.401*** (0.017)	-0.434*** (0.016)	-0.411*** (0.012)
Labourintensity	-0.501*** (0.023)	-0.557*** (0.021)	-0.519*** (0.016)	-0.501*** (0.023)	-0.556*** (0.021)	-0.519*** (0.016)
Economies of scale	-0.085*** (0.020)	-0.078*** (0.019)	-0.076*** (0.014)	-0.085*** (0.020)	-0.078*** (0.019)	-0.076*** (0.014)
Consumerlink	0.202*** (0.013)	0.228*** (0.012)	0.211*** (0.009)	0.202*** (0.013)	0.227*** (0.012)	0.211*** (0.009)
Producerlink	0.573*** (0.027)	0.625*** (0.024)	0.588*** (0.018)	0.573*** (0.027)	0.625*** (0.024)	0.588*** (0.018)
Agrimineralinteraction	-0.024*** (0.009)	-0.015* (0.008)	-0.019*** (0.006)	-0.024*** (0.008)	-0.015* (0.008)	-0.019*** (0.006)
Consumerlinkinteraction	-0.313*** (0.043)	-0.254*** (0.038)	-0.276*** (0.028)	-0.318*** (0.042)	-0.261*** (0.037)	-0.282*** (0.028)
Ecoscaleinteraction	0.696*** (0.092)	0.599*** (0.086)	0.643*** (0.064)	0.696*** (0.092)	0.600*** (0.086)	0.643*** (0.064)
Producerlinkinteraction	-0.358*** (0.068)	-0.269*** (0.060)	-0.306*** (0.045)	-0.361*** (0.067)	-0.272*** (0.060)	-0.309*** (0.045)

Labourinteraction	0.156*** (0.038)	0.138*** (0.035)	0.144*** (0.026)	0.158*** (0.038)	0.141*** (0.035)	0.147*** (0.026)
Producerlinkcominteraction	-0.322* (0.175)	-0.290* (0.151)	-0.304*** (0.115)	-0.284** (0.127)	-0.237** (0.105)	-0.258*** (0.082)
Consumerlinkcominteraction	-0.041 (0.091)	-0.056 (0.077)	-0.048 (0.059)			
Labourcominteraction	0.270** (0.124)	0.238** (0.105)	0.251*** (0.081)	0.236** (0.095)	0.192** (0.079)	0.211*** (0.061)
Constant	0.293*** (0.008)	0.314*** (0.007)	0.301*** (0.005)	0.292*** (0.008)	0.314*** (0.007)	0.301*** (0.005)
Observations	2002	2002	4004	2002	2002	4004
Mean VIF	20.72	20.72	20.72	20.82	20.82	20.82
R-squared	0.779	0.822	0.800	0.779	0.822	0.799
Adjusted R-squared	0.777	0.820	0.799	0.777	0.820	0.799

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1