CESifo WORKING PAPERS

8649
2020

October 2020

The Comparative Advantage of Dutch Cities

Steven Brakman, Tijl Hendrich, Charles van Marrewijk, Jennifer Olsen

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8649*

The Comparative Advantage of Dutch Cities

Abstract

The trade literature often treats countries as dimensionless points, which is a strong assumption. Agglomeration or lumpiness of production factors within countries can affect the national pattern of trade. In this paper we analyze comparative advantage patterns for 22 cities and 4 regions for (a selection of) 83 sectors within The Netherlands. Our findings are as follows. First, analysis of the lens condition indicates that the regional concentration of production factors (lumpiness) does not affect the Dutch national trade pattern. This suggests that the mobility of firms and factors of production is consistent with the so-called welfare maximizing integrated equilibrium. Second, despite the fact that the lens condition is verified, comparative advantage patterns across locations differ significantly from each other. We show this by comparing location specific distributions of the Balassa-Index (BI). Third, the differences across locations of comparative advantage patterns is determined by the interaction of local skill-abundance and sector skill-intensity, in line with the predictions of the factor abundance model. Moreover, at the sectoral level, location-specific variables such as market access or density, have limited effects. Fourth, most locations that house sectors that have a strong comparative (dis-) advantage relative to the Netherlands also have a strong comparative (dis-) advantage relative to the world. Only a few locations house sectors that are locally strong, but globally weak, and vice versa. The results indicate that international trade policies and disputes, such as Brexit or the US-China trade war, can have strong local consequences.

JEL-Codes: F110, F150, R120.

Keywords: comparative advantage, cities, Heckscher-Ohlin, factor abundance.

Steven Brakman
Faculty of Economics and Business
University of Groningen / The Netherlands
s.brakman@rug.nl

*Tijl Hendrich**
CPB Netherlands Bureau of Economic Policy Analysis
The Hague / The Netherlands
t.j.m.hendrich@cpb.nl

Charles van Marrewijk
Faculty of Economics
Utrecht University / The Netherlands
j.g.m.vanMarrewijk@uu.nl

Jennifer Olsen
CPB Netherlands Bureau of Economic Policy Analysis
The Hague / The Netherlands
j.d.olsen@cpb.nl

*corresponding author

October 2020

We would like to express our gratitude to Rob Euwals, Harry Garretsen, Florian Mayneris, Gerdien Meijerink, Harro van Heuvelen and conference participants of the Urban Economic Association (UEA2020), for useful comments and suggestions. Furthermore, we thank Bram Hendriks for preparing the locational characteristics and sector concordance, Judith Bayer for preparing educational data, and Paul Verstraten for allowing us to use the wage analysis code. Any remaining errors are, of course, our own. All results in this paper are based on calculations by the authors using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information see microdata@cbs.nl. The code used to generate the results is available upon request.

1 Introduction

Traditional trade analysis assumes that countries are dimensionless points. The internal spatial economic structure of a country does not affect trade flows. Implicitly, it is assumed that the distribution of economic activity or production factors within countries is homogeneous over space, or that spatial unevenness is too small to be of any consequence for the national structure of trade. As shown by Courant and Deardorff (1992, 1993) in a Heckscher-Ohlin (factor abundance) setting, lumpiness (or the spatial *uneven* distribution of production factors within a country) can affect the national pattern of trade in complex ways (see Debaere and Demiroglu, 2003, Debaere, 2004, and Brakman and Van Marrewijk, 2013, for empirical evidence). Based on market access and agglomeration economies, a salient characteristic of the world economy is, indeed, that economic activity and factors of production are unevenly distributed across space at all levels of regional disaggregation. This suggests that trade flows are potentially affected by the regional spatial distribution of economic activity (see Brakman, Garretsen, and van Marrewijk, 2020, for a survey).

This paper links international trade patterns to the internal spatial economic structure of a small open economy: The Netherlands. We do this in three steps that increasingly zoom in on the characteristics of the smaller spatial units. We first determine whether the so-called lens condition, developed by Courant and Deardorff (1992), holds for The Netherlands.[6] If so, the national trade structure is not affected by the regional distribution of factors of production; the integrated equilibrium can be reproduced by the regional trade structure. Second, we explore whether regional trade patterns differ across locations. Even if the lens condition is fulfilled, regional trade patterns can still differ if the distribution of factors of production differs substantially between locations, but not to the extent that the lens condition is violated. Third, we link the differences in regional trade patterns to local circumstances. For example: do high-skilled industries and workers sort into specific locations, that subsequently export high-skill intensive products? If so, is this sorting process affected by local characteristics, such as density or market access? Using micro-firm data we study local trade patterns in 22 cities and 4 regions for 83 sectors in the period 2007-2017 and determine characteristics of local trade patterns. Our findings are the following.

First, we find that the lens condition is satified. Based on this condition the regional distribution of production factors does not affect the Dutch structure of trade as a whole. The results indicate that from the perspective of the lens-condition the spatial regional distribution of firms and factors of production is consistent with the welfare maximizing integrated equilibrium. Second, using local micro data on firms and workers, we establish city-region distributions of Revealed Comparative Advantage using the Balassa Index (henceforth: BI, if abbreviated). For each city-region we have a distribution of BIs for all sectors that are active

[6] Evidence on lumpiness is relatively scarce, and the existing evidence is mixed. Using the so-called lens condition for regional data, lumpiness is not a concern for Japan, the UK, and India (Debaere, 2004). Debaere and Demiroglu (2003) show that for the group of OECD countries the lens condition is not violated. Bernard, Robertson, and Schott (2010), however, show that for Mexico regional lumpiness of production factors might be important. Brakman and Van Marrewijk (2013) show that on the city level the lens condition is violated for most countries in the sample.

in that location, and identify sectors with a comparative (dis) advantage relative to the Netherlands and relative to the world. Using the so-called Harmonic Mass index as a test statistic, we establish that the distributions of the Balassa index differ significantly from each other, illustrating that comparative advantage differs across locations. Third, we link the city-region trade patterns to local characteristics. We find that the interaction of local skill-abundance and sector skill-intensity systematically explains the local trade patterns, in line with the work of Davis and Dingel (2020). Explanations that are related to local characteristics such as market access or density only have limited explanatory power. Fourth and finally, we identify city-sector combinations that have a comparative advantage relative to both the Netherlands and the world; these are the sectors on which Dutch exports rely intensely. We also identify city-sectors that have a comparative advantage relative to the Netherlands but not relative to the world (and vice versa). This implies that a strong regional position of a sector does not always translate to a strong position internationally and, similarly, that some national exports are so strong that even weaker regions prove to be strong international players. Of the sectors that the Netherlands as a whole is weak in, all but two are internationally strong in at least one Dutch region, implying that there are a number of regional strengths that are not visible at the national level.

The paper is organized as follows. Section 2 introduces the dataset. For the period 2007-2017 we have micro-firm export data, factor endowments and factor intensities for 22 Dutch cities, 4 rural areas, and 83 sectors. This enables us to calculate and explain local trade patterns. Our sample covers about 90 per cent of total Dutch exports. Section 3 analyzes the lens condition and finds that this condition is not violated. Sections 4 and 5 analyze the local distributions of comparative (dis) advantage. Using the Harmonic Weighted Mass Index as a test statistic, we find that the distributions of Revealed Comparative Advantage, measured by the Balassa index, differ significantly across locations. Section 6 analyzes local trade patterns and shows that the interaction between local factor abundance and sector skill-intensities is the main explanation for the local structure of trade flows. Section 7 presents the sectors in which The Netherlands is especially strong or weak; that is, those sectors that have a comparative (dis)advantage relative the rest of the Netherlands as well as relative to the world. The more marginal sectors are also identified; sectors that have a comparative (dis) advantage relative to the Netherlands but not relative to the world, and vice versa. Section 8 concludes.

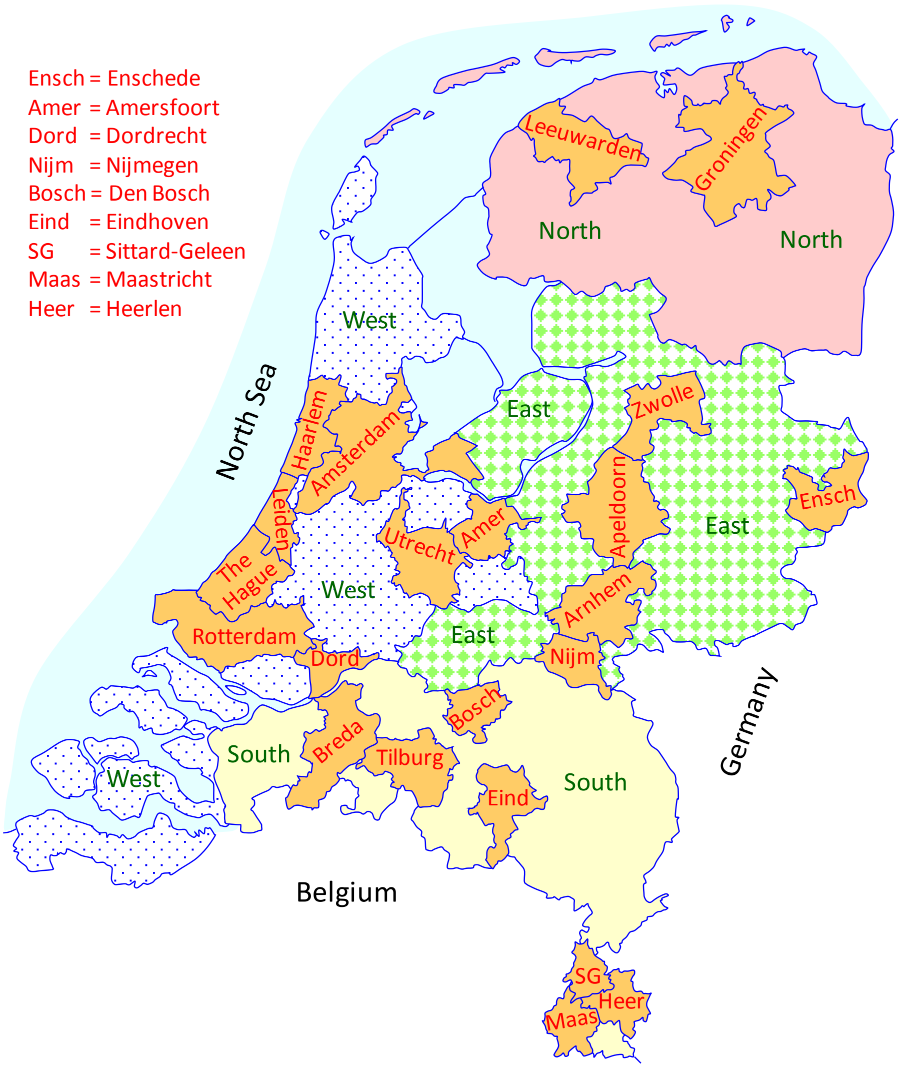
2 Data

We construct a disaggregated data set of Dutch exports at the location- and sector level using administrative data. This enables us to calculate revealed comparative advantage, factor intensities and factor endowments for Dutch locations and sectors.

2.1 Spatial Units

We consider two types of spatial units: City districts and rural regions. Statistics Netherlands defines 22 city districts, which consist of municipalities with city-status as well as surrounding municipalities that are determined to be economically dependent on the city.[7] Municipalities that do not form part of any city district are grouped by NUTS 1 region, forming four rural regions: North, South, West and East. To avoid any ambiguity, we use the term ‘cities’ when referring to the 22 city districts, the term ‘regions’ when referring to the 4 rural regions, and the term ‘cities-regions’ when referring to all 26 locations (both cities and regions).

Figure 1 Dutch Locations; 22 Cities and 4 Regions



Source: constructed by authors. Based on CBS 2005, *Grootstedelijke agglomeraties en stadsgewesten afgebakend*

Figure 1 provides an overview of the location of the cities-regions within the Netherlands, while Table 1 provides an overview of their size in terms of (working) population. The main cities are located in the western part of the country, close to the sea. This includes the three largest cities: Amsterdam, Rotterdam, and The Hague (in that order), each with a population of more than one million people. As Figure 1 shows, there are 8 cities located in the West (of

[7] For a description (in Dutch), see: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/gemeente/gemeenten-en-regionale-indelingen/niet-landelijk-dekkende-indelingen

which Amsterdam is the largest), 7 cities located in the South (of which Eindhoven is the largest), 5 cities in the East (of which Arnhem is the largest), and 2 cities in the North (of which Groningen is the largest). Taken together, the 22 cities account for about 56 per cent of the Dutch population. The 4 regions take care of the remaining 44 per cent of the population.

Table 1 Overview of population and working population; ranked by population size, 2017

Location	Code	Population		Working population	
		Size	% total	Size	% total
West	NL3	2,420	14.3	1,080	12.9
East	NL2	2,004	11.9	957	11.4
South	NL4	1,768	10.5	819	9.8
Amsterdam	SG10	1,586	9.4	863	10.3
North	NL1	1,182	7.0	548	6.5
Rotterdam	SG14	1,172	6.9	581	6.9
The Hague	SG13	1,062	6.3	542	6.5
Utrecht	SG09	661	3.9	373	4.4
Haarlem	SG11	422	2.5	207	2.5
Eindhoven	SG19	421	2.5	222	2.6
Groningen	SG01	363	2.1	209	2.5
Arnhem	SG06	362	2.1	193	2.3
Leiden	SG12	346	2.0	175	2.1
Breda	SG16	325	1.9	170	2.0
Enschede	SG04	316	1.9	169	2.0
Tilburg	SG17	302	1.8	161	1.9
Amersfoort	SG08	289	1.7	149	1.8
Dordrecht	SG15	287	1.7	135	1.6
Nijmegen	SG07	261	1.5	151	1.8
Heerlen	SG21	247	1.5	123	1.5
Apeldoorn	SG05	214	1.3	103	1.2
Den Bosch	SG18	205	1.2	109	1.3
Maastricht	SG22	183	1.1	93	1.1
Zwolle	SG03	182	1.1	96	1.1
Leeuwarden	SG02	174	1.0	94	1.1
Geleen/Sittard	SG20	148	0.9	74	0.9

Source: author calculations; size in thousands.

As indicated in Figure 1 the regions consist of the NUTS 1 areas excluding the cities located there. Table 1 shows the distribution of the whole population and the working population across cities-regions. The rural regions West, East, and South (in that order) have the largest populations, followed by Amsterdam. This reflects the fact that even rural areas in the Netherlands are densely populated. At the national level, the working population is about half of the total population. There are substantial deviations in the distribution of the population and the working population across cities-regions, both in relative and absolute terms. Not surprisingly, the share of the working population is smaller than the share of the total population for all four rural regions. Both in absolute and relative terms, the discrepancy is biggest for the West region, which contains 14.3 per cent of the Dutch population compared to 12.9 per cent of the Dutch working population. In contrast, almost all cities (with the exception of Dordrecht and Apeldoorn) have a higher share of the working population than of the total population. In absolute terms, the differences are largest in Amsterdam, Utrecht, and Groningen. In relative terms, the difference is largest in Nijmegen, Groningen, and Utrecht.

2.2 Firms

Throughout this study, export is measured in terms of annual export-revenue generated by Dutch exporting firms. Our initial data set draws on a complete registry of anonymized Ducth firm level value-added tax (VAT) statements for the period 2007-2017, and as such covers all generated export revenues. We restrict our final dataset to exporting firms.[8] Firms are excluded from the final sample if the reporting is incomplete or illogical, such as zero or negative revenue or missing location-data.

Table 2 Overview of exporting firm size, 2007-2017

Year	Exporting firms (× 1000)	# Local branches (mean)	# Firm employees (mean)	Share employees 90th percentile (per cent)	Share exports 90th percentile (per cent)
2007	33	1.5	29	67	90
2008	33	1.5	31	68	90
2009	33	1.5	31	68	90
2010	34	1.7	33	69	89
2011	36	1.7	32	68	89
2012	37	1.7	32	69	89
2013	37	1.7	31	70	89
2014	38	1.7	34	72	89
2015	40	1.7	33	72	88
2016	41	1.7	33	72	89
2017	42	1.7	34	73	89

Summary information on exporting firms is provided in Table 2. Our final data set comprises 33 to 42 thousand exporting firms per year. This is around 3.5 per cent of all Dutch firms, accounting for about 80 to 90 per cent of all Dutch exports in any given year. Exports are highly concentrated within this group of firms, with 90 per cent of the export-revenue generated by the top 10 per cent of firms (the 90th percentile in Table 2), which is in line with previous studies (see, for example, Bernard et al., 2012). As indicated in Table 2, in 2017 the

[8] In line with the reporting convention of Statistics Netherlands, we label a firm 'exporter' if its annual export revenue exceeds 5000 euro.

average firm employed 34 workers (measured in full time equivalents), while 73 per cent of the workers are employed by the top 10 per cent of firms in terms of size.[9] The average firm has about 1.7 local branches, so most firms are small and only active in one location. From 2007 to 2017, the average firm size has been increasing (from 29 to 34 workers), while the share of workers in the largest firms has increased as well (from 67 to 73 per cent).

2.3 Sectors and Exports

Exports for cities-regions and sectors are constructed using registry data from Statistics Netherlands for the period 2007-2017.[10] Because export data are reported at the firm level, we match (national) firm exports with cities-regions in four steps.

First, we use the General Business Register (GBR) and its local counterpart (LGBR) to collect information on the sector classification and location (municipality) of all local branches for all Dutch firms, including self-employed workers. Each firm and branch is assigned a sector code according to the Dutch coding system (SBI 2008) of which the first two digits correspond to the international NACE rev. 2 classification.[11] By this definition, our data set contains a total of 83 different two-digit sectors. It should be noted Statistics Netherlands may assign local branches a different sector than their parent firm if their main economic activities differ.[12]

Second, we match the branch location- and sector- data to the firm level VAT data described in section 2.2.

Third, we allocate the annual export revenues of each firm to its local branches, using wages as weights. Each branch b can be associated with a share of the total firm export revenue E_b equal to the percentage of wages earned in that specific branch. This is calculated on an annual basis as given in equation (1), where E_b and E_F are the export revenues of a branch and its firm, w_b and w_F are the branch and firm wage sums, and $fb \in F$ is the set of branches belonging to the same firm as branch b.

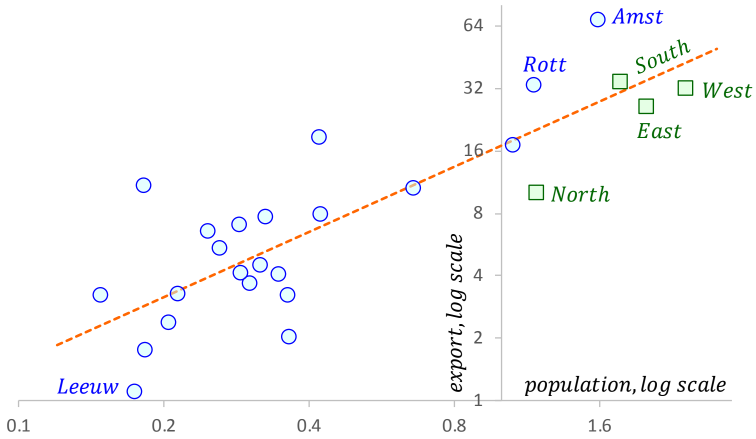
(1) $$E_b = E_F \frac{w_b}{\sum_{fb \in F}(w_{fb})} = E_F \frac{w_b}{w_F}$$

Wages are sourced from a monthly registry of all jobs performed at each Dutch firm, which also contains information about the municipalities in which jobs are performed.[13] If a firm

[9] With the very largest firms employing in the tens-of-thousands.
[10] For privacy reasons these data are not publicly available. The Netherlands Bureau of Statistics manages these micro data, which can be obtained for research purposes upon request: https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research
[11] SBI = Standaard Bedrijfs Indeling (Standard Firm Classification); see Appendix A for a list of sectors. We discard SBI sectors 97 to 99, since these are unused categories not containing any firms.
[12] Consider, for instance, a web retailer with an IT development branch and a logistics branch in a different location.
[13] Each wage record represents the monthly wage of a single job of a single worker performed at a firm. If workers are employed at multiple firms they have multiple records. Wage data are categorized by the type of work (such as: employed, civil servant, self-employed, temporary) and the type of wage (wage, unemployment benefits, old-age pension). We remove records if they contain unemployment benefits or old-age pensions, keeping records of wages for overtime, payments in kind and bonuses. Information on the location of jobs is

has multiple local branches within a municipality, we cannot allocate wages to a specific branch. We therefore aggregate firm branches at the municipal level. Aggregated branches are assigned to the same sector as the largest firm branch in the municipality.

Figure 2 Population and Exports of Dutch Cities-Regions, 2017



Source: author construction; population in million, exports in billion euro; both scales in logs; dashed line is regression with slope 1.04, it explains 69.1 per cent of the variance; regions denoted by squares; Amst = Amsterdam; Rott = Rotterdam; Leeuw = Leeuwarden.

Fourth, and finally, we aggregate the municipal data for all sectors within the 22 cities and 4 regions as discussed in section 2.1 and illustrated in Figure 1. Our data set thus contains export information for 22 cities, 4 regions, and 83 two-digit sectors for a period of 10 years. In view of the relatively small geographical units of our study (cities and regions within a country, rather than the country as a whole), our measure of revealed comparative advantage can be volatile on an annual basis (see for a discussion Hinloopen and van Marrewijk, 2001). To avoid excess volatility, we therefore split annual data into two time periods, namely 2007-2012 and 2012-2017, and calculate the average BI of a sector for each sub-period. The first period includes the Great Recession and most of the subsequent recovery. Our discussion in the paper focuses on the most recent second period (2012-2017), using the first period as a robustness check for our main conclusions. Graphs frequently use the most recent year (2017) for illustration purposes. Figure 2, for example, shows that larger cities-regions in terms of population in 2017 also tend to have larger export flows (this explains 69 per cent of the variance in exports), starting with Amsterdam, followed by South, Rotterdam, and West.

2.3 Skills, Abundance, and Intensity

In terms of factors of production, our analysis focuses on human capital in terms of skills from schooling for cities-regions as well as sectors.[14] We use annual registry data on the highest attained level of education for Dutch citizens to identify three skill levels for general schooling, labelled high-, medium-, and low-skilled (with sub-indices *high*, *med*, and *low*,

only collected for the month of December in any given year. This implies that the wages of about 30 per cent of jobs not held in the month of December cannot be matched to a municipality.

[14] Unfortunately, we have no reliable information on the capital abundance in regions-cities, nor on the capital intensity in sectors, so, like land, this factor of production is excluded from the analysis.

respectively).[15] In addition, we differentiate between technical and non-technical types of schooling using the same classification (identified by subindices *tec-high*, *tec-med*, and *tec-low*).[16] Note, that the regular skills classification shares add to one, while this is not the case for technical skill shares; we therefore use the sub-index *tech* to refer to the sum of *low*, *medium*, and *high* technical skill.

Table 3 Skill abundance in cities-regions; ranked by high-skilled, per cent, 2017

City-region	General schooling share		Technical schooling share	
	High-skilled	Low-skilled	High-skilled	Low-skilled
Utrecht	44.5	22.4	7.6	1.8
Nijmegen	39.3	24.1	6.2	2.4
Groningen	38.4	20.1	5.2	2.0
Amsterdam	37.0	26.4	5.0	2.2
Leiden	36.9	25.3	6.3	2.3
Den Bosch	35.5	27.5	5.8	2.9
Eindhoven	34.3	27.3	9.3	3.2
Amersfoort	34.2	27.5	5.7	2.6
Zwolle	33.9	25.3	4.3	2.7
Haarlem	33.8	27.1	5.0	2.5
Maastricht	33.0	27.3	4.7	3.0
Breda	32.2	27.3	5.0	2.9
The Hague	32.1	30.4	7.1	2.5
Arnhem	30.9	29.3	5.3	2.8
Tilburg	30.4	28.9	4.2	3.3
Enschede	27.5	30.0	5.8	3.2
West	26.8	29.9	4.4	3.4
Geleen-Sittard	26.7	31.8	5.1	3.7
Apeldoorn	26.7	30.9	4.4	3.3
Leeuwarden	26.6	27.2	3.4	3.1
Rotterdam	26.1	34.0	4.0	3.0
East	24.2	31.6	4.3	3.9
South	24.1	32.3	4.5	4.5
Dordrecht	23.6	32.5	4.0	3.7
North	20.8	32.0	2.9	4.2
Heerlen	20.5	36.5	3.6	4.6

Source: author calculations; exports as per cent of Dutch total; skill distribution as per cent of working population (15 to 75 years), based on place of residence.

[15] Citizens that have no registered education are excluded from our analysis, as are citizens that are not of working age as defined by Statistics Netherlands (15-75 years old).
[16] For this we use the categorization of education programmes created by Statistic Netherlands. We consider a category ‘technical’ if it falls within “Natural Sciences, Mathematics, and Statistics”, "Information and Communication Technologies" or "Engineering, Manufacturing, and Construction".

We analyse differences in human skills from two perspectives: From a cities-regions perspective (we refer to this as the *abundance* of skills in a location), and from a sector perspective (we refer to this as the *intensity* of skills in a sector). Citizens are assigned to cities-regions using their registered home addresses, and to sectors using their work locations from firm-level job data.[17] Our analysis in section 6 focuses on the *interaction* between abundance and intensity for determining revealed comparative advantage, which is consistent with the Heckscher-Ohlin (factor abundance) trade model.

Regarding skill abundance, Table 3 provides some information for 2017. The table is ordered by the share of high-skill workers, starting with Utrecht, Nijmegen, and Groningen (which are relatively abundant in high-skill workers) and ending with Dordrecht, North, and Heerlen. The highest share for Utrecht is 44.5 per cent, the lowest share for Heerlen is 20.5 per cent. It is worth noting that relatively large shares of high-skill workers are not only reserved for the largest cities. Mid-sized cities that house a university, such as Nijmegen, Groningen, and Leiden also have a relatively large high-skill share. In general, cities are more abundant in high skilled workers than (rural) regions (our data have similar characteristics as found for other countries, see for example Glaeser and Resseger, 2010). Obviously, if the share of high-skilled workers is relatively high, the share of low-skilled workers tends to be low (correlation is -0.89). Table 3 also provides information on the abundance of technical skills. Regarding technical high-skilled workers, Eindhoven ranks highest (9.3 per cent), followed by Utrecht and The Hague, while Heerlen, Leeuwarden, and North (2.9 per cent) rank lowest. Although there is substantial variation in the ordering of high-skilled workers and technical high-skilled workers, the correlation between these two variables is strongly positive (0.68). Nonetheless, it is clear that the cities-regions are diverse in terms of the skill abundance of their inhabitants.

Table 4 Sector skill intensity; Top 5 and Bottom 5 by high-skilled workers, 2017

		Size	General schooling		Technical schooling	
SBI	Sector	# work	High-skill	Low-skill	High-skill	Low-skill
Top 5 sectors high-skilled intensity						
85	Education	515	80.8	3.3	5.7	1.0
72	R&D	31	78.2	3.1	41.8	0.5
60	Program & broadcasting	8	73.4	2.9	2.5	0.4
64	Financial institutions	86	71.8	3.5	8.6	0.4
70	Holding companies	116	71.1	4.4	16.1	0.7
Bottom 5 sectors high-skilled intensity						
49	Land transport	139	10.5	29.8	1.6	6.2
56	Food services	504	10.4	33.8	0.9	3.1
81	Facility management	172	9.5	46.8	1.7	5.0
96	Wellness; funeral activity	59	9.5	24.0	0.8	2.0
80	Security & investigation	34	8.9	19.0	1.2	2.0

[17] Note that the sectors of employees are defined on the firm (national) level and not the local branch level in this part of the analysis.

Source: author calculations; size by working population (thousands); skill distribution as percent of working population; see Appendix A for more complete sector description.

Regarding skill intensity, Table 4 provides some information for 2017 for the top and bottom five sectors in terms of high-skill worker shares, where workers are counted in full time equivalents (fte). The top-5 sectors employ around 80 per cent of high-skill workers, starting with (not surprisingly) education and R&D. The bottom-5 sectors have around 10 per cent of high-skilled workers, including wellness and security. A large sector in terms of the number of workers at the top is education (515 thousand), whereas the food sector is large at the bottom (504 thousand). There are substantial differences in terms of the technical intensities of the sectors. Education and R&D both require about 80 per cent of high-skill workers, but in terms of technical skills, education requires only 5.7 per cent of technical high-skilled workers compared to 41.8 per cent for R&D (more than seven times as much). When comparing Table 2 and Table 3, it is clear that the sectors are more diverse in terms of their skill intensities than the cities-regions are in terms of their skill abundance. We return to this issue in section 3.

2.4 International Exports

At the *national* level, our export data come from the UN Comtrade database, which classifies exports by product category using the Harmonized System (HS2017). In order to compare international product data with our regional sector data we perform a concordance analysis, matching HS2017 product classifications with Dutch sector classifications (SBI2008), as explained in section 4.3.

3 Lumpiness and trade[18]

3.1 Theory

The relationships between urbanization and the potential effects on trade flows can best be explained by an Edgeworth-box (see Figure 3). For expositional reasons we assume that the country under consideration is small, such that goods prices are given. The figure – made popular by Dixit and Norman (1980) – depicts a perfectly integrated country, in which there are no distortions, two factors of production – High skill workers H and low skill workers L – and two goods, X and Y, produced under constant returns to scale. The country consists of two areas, I and II. Moreover, consumer preferences are identical and homothetic. The (given) amount of L is depicted on the horizontal axis, and the (given) amount of H along the vertical axis, where the use of endowments in area I is measured from the O origin and the use of endowments in area II is measured (upside down) from the O^* origin. If the endowments are distributed over the two areas, given world prices, this determines the production levels of goods X and Y, the country's income level, the demand for goods X and Y, and thus its internal trade flows (all welfare maximizing under standard circumstances).

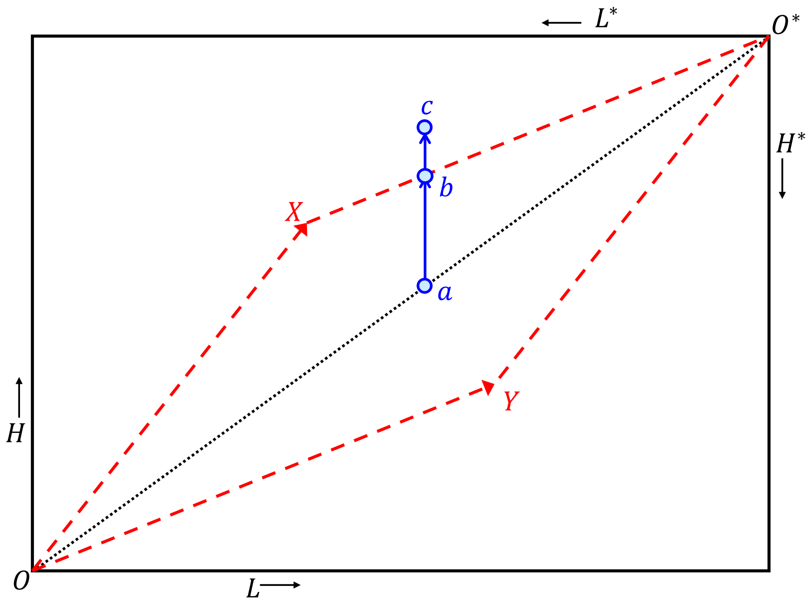
Figure 3 depicts the integrated equilibrium. Total supply in the integrated equilibrium is characterized by OX of good X and OY of good Y (with an appropriate unit of measurement).

[18] This section is partially based on Brakman and Van Marrewijk (2013); see also Courant and Deardorff (1992).

The slope of the vectors indicates that we have assumed that the production of good *X* is relatively *H* intensive. If we perform a vector summation on *OX* and *OY*, total factor use in both sectors is exactly equal to the total amount of available factors of production, *L* and *H*.

The central question that is answered in Figure 3 is: can the welfare maximizing integrated equilibrium be reproduced once the country is split into two separate areas with given factor endowments? The answer is: 'yes', as long as the distribution of production factors in a country is not too different, namely within the factor price equalization (*FPE*) set; *OXO*Y*. For spatial distributions outside the *FPE* set the answer is 'no' (see Dixit and Norman, 1980 for a detailed explanation). If we, for example, redistribute *L* and *H* such that we follow the arrow starting in point *a*, production of *X* increases and *Y* decreases in area *I*, and the production of *X* decreases and *Y* increases in area *II*. These are standard Rybczynski effects in both areas. Along the arrow *ab* the integrated (within country) equilibrium can be reproduced and the redistribution of workers has no effect on the trade flows of the country with the outside world.

Figure 3 The Integrated Equilibrium and Lumpiness



The two areas within the country do trade with each other; the high skill abundant area exporting the high skill intensive good, and the low skill abundant area exporting the low skill intensive good. This is possible until one or both areas are completely specialized. As drawn, at point *b* area *I* still produces both *X* and *Y*, but area II is completely specialized in *Y*. The total amounts of both *X* and *Y* correspond to the integrated equilibrium. If we follow the arrow from the point of complete specialization, say from *b* to *c*, the amount of *X* in *I* increases, but without the accompanying decrease of *X* in *II*. The amount of *Y* decreases in both areas. This is caused by the Rybczynski effect in *I* (given goods prices), and a further reduction of the production of *Y* in *II*, which is specialized in *Y*. This unambiguously raises the supply of *X*, and reduces that of *Y*, thus potentially influencing the country's trade patterns. As a result, outside the FPE parallelogram *OXO*Y* the country's trade pattern is affected by the lumpy distribution of factors of production.

Note, that outside the FPE set trade patterns are difficult to establish. If we, for example, move horizontally instead of vertically from *a* and apply a similar reasoning as above, we create an excess supply of good *Y*, instead of an excess supply of *X*. Introducing a second country in which lumpiness also matters makes the determination of trade patterns even more difficult. The combination of lumpiness in both countries might strengthen predictions by the HO model (if the abundant factors in both countries are the lumpy factors) or might go in the opposite direction (if the relatively scarce factors are the lumpy factors).[19]

It is relatively easy to generalize Figure 3 into a country with many areas, and many goods/sectors in a two production factor world, giving rise to the so-called *lens condition* (Deardorff, 1994, Debaere, 2004, Debaere and Demiroglu, 2003). We can rank factor intensities of all sectors according to decreasing high skill / low skill intensities above the diagonal (and vice versa below the diagonal) and concatenate the corresponding vectors of factor intensity. Following a similar procedure we can concatenate the vectors of relative factor endowments in each area. If the line of relative factor intensities in the sectors encloses the line of relative factor endowments in the areas, the integrated equilibrium can be reproduced. This is called the *lens condition* because if we introduce a large number of goods and areas the two concatenations look like lenses (see below).[20]

Figure 4 The Lens Condition

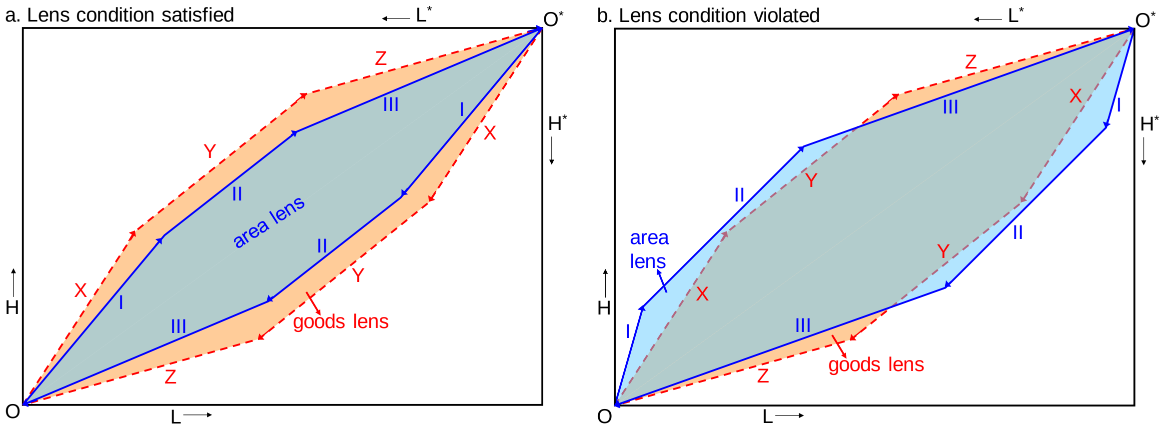


Figure 4, illustrates the condition for a three goods (*X*, *Y*, and *Z*) and three area (*I*, *II*, and *III*) example. In Figure 4a the lens condition is *satisfied*: the factor endowment lens for the areas is a subset of the (factor use) goods lens, indicating that the empirical distribution of the factors of production across the various areas within the country does not influence the country's overall trading position. In Figure 4b the lens condition is *violated*: the factor endowment lens for the areas is *not* a subset of the goods lens, indicating that the empirical distribution of the factors of production across the various areas within the country *does* influence the country's overall trading position.

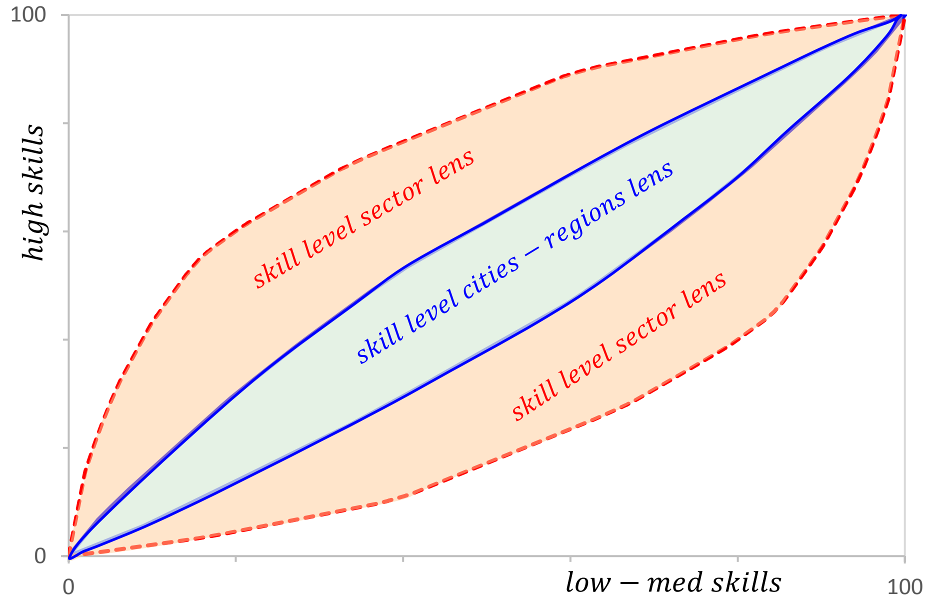
[19] See Courant and Deardorff (1992) for a discussion on the difficulties to determine trade patterns.
[20] See Debeare and Demiroglu (2003) for a more detailed discussion of the lens condition.

3.2 The Lens Condition for Dutch Cities

The empirical question we need to answer in light of the above discussion is thus whether the lens condition is satisfied, or not, for Dutch cities. We have information available on factor distributions and factor intensities for different labour skills. We identify three skill levels (low, medium, and high) and two skill types (technical and non-technical), see section 2. The modest number of $2 \times 3 = 6$ factors of production already presents us with a large number of possible lens conditions in 2-dimensional space, in particular if we also combine factors of production.[21] To streamline the analysis, we focus on the lens condition for 2017 in two steps, by first discussing the general skills and then go into more detail for the different technical skills.[22]

Nationally, in 2017, 29.5 per cent of the Dutch working population had a high skill level, 41.0 per cent had a medium skill level and 29.5 per cent had a low skill level. In the period 2007-2017, the share of the working population with a high skill level has been rising by 2.9 percentage points and with a medium skill level with 2.2 percentage points. This obviously implies that the share with a low skill level has been declining by 5.1 percentage points in this period. For our skill level lens discussion, we combine the low and medium skills and compare with the high skills. For any lens we construct, we normalize each factor of production to range from 0 to 100.

Figure 5 Dutch General Skills Lens Condition, 2017



[21] There are 15 combinations of the 6 production factors. If we look only at the levels there are 3 more combinations, while if we only look at the types there is 1 more combination. If we combine factors of production, as we do in Figure 4 and Figure 5, more combinations are possible, but some of these would make no sense. For example, it seems reasonable to compare high skill levels relative to a combination of low & medium skills or low skill levels relative to a combination of high & medium skills, but not to compare medium skill levels relative to a combination of high & low skills. Viewed this way, the combinations provide an additional 6 possibilities (4 at the production factor level and 2 at the education level) for a total of 25 possible combinations for each year.

[22] Results for other years are similar.

Source: authors; see section 2 for data sources; low-med = combination of low and medium skill level.

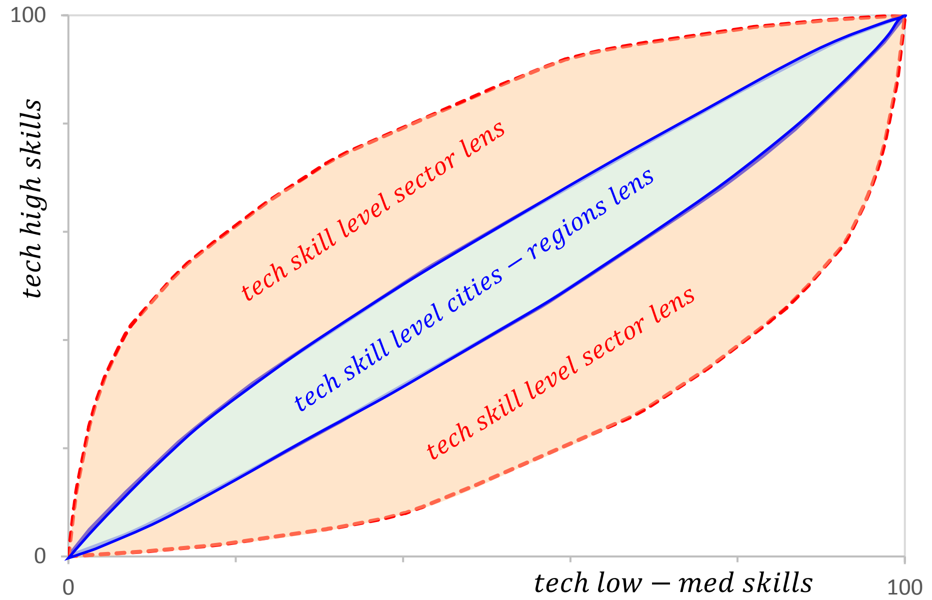
Of the 26 cities-regions, Heerlen has the lowest share of high skill workers (20.5 per cent), while Utrecht has the highest share (44.5 per cent). To create the area / cities-regions lens, we order the locations in terms of high skill relative to low skill abundance (both rising and falling) and create vectors with a length proportional to the number of workers in that location (which ranges from about 93 thousand in Maastricht to more than 1 million in West). The result is illustrated in Figure 5 under the label ‘skill level *cities-regions* lens’. With high skill on the vertical axis and low-med skill on the horizontal axis, the steepest slope of the cities-regions lens (for Utrecht) is 1.91, which is 3.1 times steeper than the flattest slope of 0.61 (for Heerlen). The difference is thus substantial, but not enormous, making the skill level cities-regions lens not too wide (see Figure 5).

Of the 83 Dutch sectors of production, sector 80 (security and detection) has the lowest share of high skill workers (8.9 per cent), while sector 85 (education) has the highest share (80.8 per cent). To create the goods / sector lens, we order the sectors in terms of high skill relative to low-med skill intensity (both rising and falling) and create vectors with a length proportional to the number of workers in that sector (which ranges from about 805 for sector 12 [tobacco] to about 1.4 million for sector 78 [temporary employment agencies]). The result is illustrated in Figure 5 under the label ‘skill level *sector* lens’. With high skill on the vertical axis and low-med skill on the horizontal axis, the steepest slope of the sector lens (for sector 85) is 7.87, which is 42.9 times steeper than the flattest slope of 0.18 (for sector 80). The difference is thus much larger than for the area lens, which in combination with all the other sectors of production creates a fairly wide sector lens.

Figure 5 depicts both the skill level sector lens and the skill level cities-regions lens. Since the sector lens is much wider than the cities-regions lens, it immediately follows that the lens condition is *satisfied*. This is in contrast to the conclusion in Brakman and van Marrewijk (2013). We return to this in the next sub-section. For now, we go one step deeper by analysing the lens condition for both the type and level of skill, where we focus on technical workers.

Nationally, 15.4 per cent of the Dutch working population had technical schooling in 2017, a decline by 0.4 percentage points relative to 2007. Of the workers with technical schooling in 2017, about 20.8 per cent had a low skill level, 46.9 per cent had a medium skill level, and 32.3 per cent had a high skill level. Relative to the total Dutch working population, this translates to 3.2 per cent with a low technical skill level, 7.2 per cent with a medium technical skill level, and 5.0 per cent with a high technical skill level. Please keep in mind, therefore, that the sum of low-, medium-, and high technical skill levels does not add up to 100 per cent (but to 15.4 per cent nationally). For our technical skill level lens discussion, we combine (as above) the low and medium technical skill levels and compare with the high technical skill level.

Figure 6 Dutch Technical Skills Lens Condition, 2017



Source: authors; see section 2 for data sources; tech = technical; low-med = combination of low and medium skills

Of the 26 cities-regions, North has the lowest share of technical high skill workers (2.9 per cent), while Eindhoven has the highest share (9.3 per cent). Ordering the locations in terms of rising and falling abundance of technical high skill workers relative to technical low-med skill workers in combination with the number of workers at each location allows us to derive the 'tech skill level *cities-regions* lens' as illustrated in Figure 6. With technical high skill on the vertical axis and technical low-med skill on the horizontal axis, the steepest slope of the cities-regions lens (for Eindhoven) is 1.96, which is 3.4 times steeper than the flattest slope of 0.58 (for North). This difference is similar to what we found for the cities-regions lens in Figure 5, although in combination with the other locations the resulting cities-regions lens is somewhat smaller (compare Figure 6 with Figure 5).

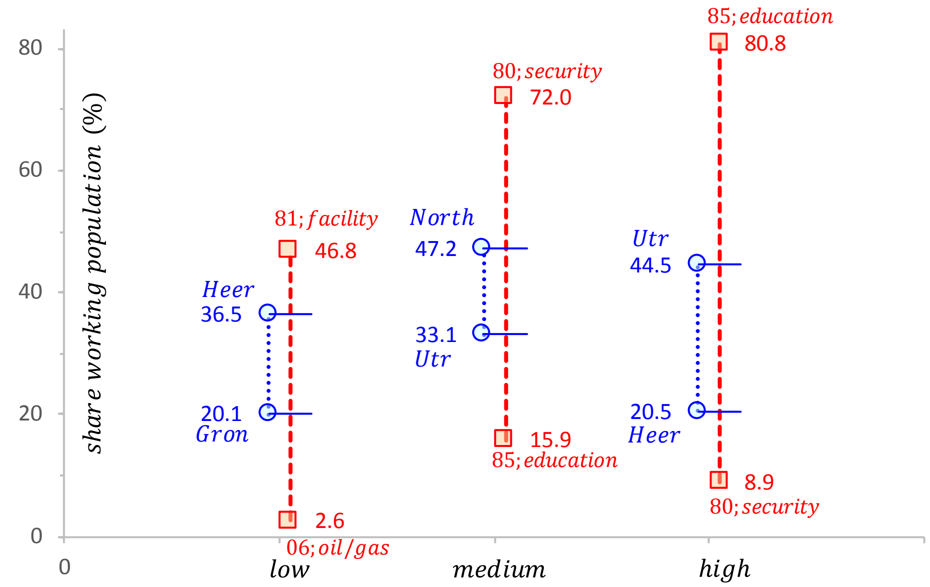
Of the 83 Dutch sectors of production, sector 87 (nursing care with guidance for overnight stay) has the lowest share of technical high skill workers (0.6 per cent) and sector 71 (architects, engineers and technical design & advice) has the highest share (44.9 per cent). Using a similar procedure as before, we create the 'tech skill level *sector* lens' in Figure 6. With technical high skills on the vertical axis and technical low-med skills on the horizontal axis, the steepest slope of the sector lens (for sector 71) is 13.53, which is 140 times steeper than the flattest slope of 0.10 (for sector 87). As shown in Figure 6, this is much wider than the cities-regions lens and the lens condition is again easily *satisfied*.

3.3 Explanation and Implications

The analysis in section 3.2 shows that the lens condition is satisfied for the general skills level and the technical skills level. A similar picture and conclusion arises for all other possible combinations.[23] This section explains from an analytical perspective why this is the case, and why it differs (regarding Figure 5) from an earlier analysis involving Dutch cities (Brakman and van Marrewijk, 2013). We conclude by pointing out what the main implications are for our analysis of the comparative advantage of Dutch cities-regions.

From an analytical perspective, the cities-regions lens can only be a subset of the sector lens if this holds close to the respective origins of the Edgeworth-boxes. This requires that the minimum slope of the sector lens is lower than the minimum slope of the cities-regions lens, while the maximum slope of the sector lens is larger than the maximum slope of the cities-regions lens. These slopes are determined by the shares of factor abundance in locations for the cities-regions lens and the shares of factor intensities in sectors for the sector lens, so the slope requirements translate to share requirements.

Figure 7 Skill Level Ranges for Dutch Cities-Regions and Sectors, 2017



Source: authors; see section 2 for data sources; see Appendix A for sector names; Heer = Heerlen; Utr = Utrecht; Gron = Groningen.

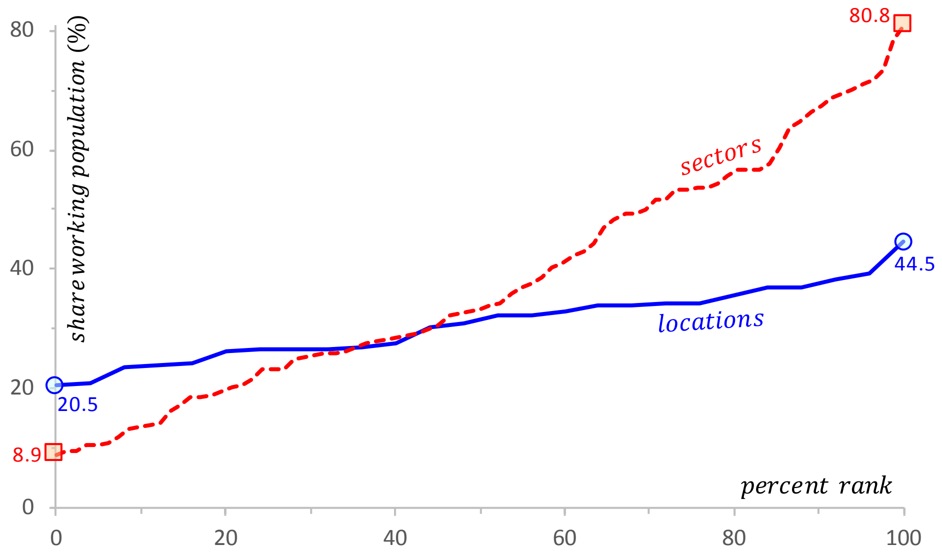
Figure 7 provides the range of skill level shares (of working population in per cent) for Dutch cities-regions and sectors in 2017 for low, medium, and high skill levels. In all cases, the cities-regions shares are strictly in between the sector shares and the cities-regions range constitutes only a modest fraction of the sector range (about 1/3rd). For the high skill levels, this translates directly to the differences in slopes and slope ratios illustrated in Figure 5 and discussed in section 3.2, where the sector lens starts off much wider than the cities-regions lens. Similar remarks hold for the other skill levels. When we go one level deeper and look at the shares for different technical or non-technical skill levels for cities-regions and sectors, the conclusion is similar, but somewhat stronger (see Appendix B). The cities-regions shares are strictly in between the sector shares and the cities-regions range constitutes only a small fraction of the sector range (about 27 per cent for non-technical workers and only 17 per cent

[23] This amounts to 6 combinations in total: 3 skill levels (high, medium, low) x 2 skill types (technical and non-technical)

for technical workers). In all cases, therefore the sector lens contains the cities-regions lens close to the origins and is much wider, as illustrated in Figure 5 and Figure 6.

Note that the technical analysis close to the origins is a necessary but not sufficient condition for the lens condition to be satisfied because the violation could, in principle, also occur more towards the center of the (Edgeworth) box in Figure 5 or Figure 6 (see, for example, Figure 4b). This situation does not arise in our data set because we have detailed factor intensity information available for 83 different sectors wich differ substantially in their factor shares. This is illustrated in Figure 8 for the high skill share rank distribution for sectors and locations in 2017. There are 26 locations ordered from lowest per cent rank (0) to highest (100) with high skill shares from 20.5 to 44.5 per cent (see also Figure 7). Similarly, there are 83 sectors ordered from lowest to highest with high skill shares from 8.9 to 80.8 per cent (see again also Figure 7). The point is that there are many sectors with different sector shares over a wide range. As a consequence, the sector lens gradually moves from high to low slopes (or vice versa), which ensures that the sector lens is fairly wide (as in Figure 5 and Figure 6) and strictly contains the cities-regions lens over the entire domain.

Figure 8 High Skill Share Rank Distribution, 2017



Source: authors; see section 2 for data sources.

In contrast to our results in Figure 5, Brakman and van Marrewijk (2013) show that the lens condition is violated for high skill versus low-med skill workers (under the labels 'high-skill' and 'labour') for the Netherlands. Although a different time period and a different number of locations (16 rather than 26) may play a role in the different findings, the most important reason is without doubt the use of better and more detailed sector data in the current study.[24]

[24] The lack of more detailed data forces Brakman and van Marrewijk (2013) to group all rural areas together in one artificial region, whereas in this paper we distinguish four regions based on detailed micro level data. Using more aggregate input-output data has the advantage that it is available for several countries, but the disadvantage that the number of sectors is limited and the factor intensities are not known for all sectors. As a result, many sectors are *assumed* to have the same factor intensities, such that only 18 different factor intensities remain (see Table A1 in Brakman and Van Marrewijk, 2013 for a complete overview). In the current data set, the share of

An economic explanation for the fulfillment of the lens condition in the Netherlands is *labour mobility*. Violation of the lens condition implies factor price *in*equality. If factors of production (in this case different types of labour) respond to these factor price differences, the labour distribution and composition adjusts. In other words, the cities-regions lens is *endogenous* as a result of mobility of factors of production.[25] There are many potential obstacles to mobility within countries, based on distance, cultural-, religious-, language differences, or amenities (such as climate), and legal restrictions. In a small country like the Netherlands there are no legal restrictions to factor mobility, distance plays a minor role, the climate is similar throughout the country, everyone speaks the same language and has a similar culture, while religious obstacles for migration seem to be minor. We should therefore not be surprised if the cities-regions lens adjusts through migration flows to become a subset of the sector lens.

There are three main (policy) implications for our finding that the lens condition is satisfied.

First, it suggests that the distribution of factors of production across Dutch cities does not to affect aggregate trade flows for the Netherlands as a whole: the so-called integrated equilibrium can be reproduced by the current spatial distribution of factors of production. In *this* sense the spatial distribution is optimal. Mobility of production factors in The Netherlands, notably labour, is sufficient to ensure that the cities-regions lens is within the sector lens. By implication, from the perspective of the integrated equilibrium, there is no need for the Dutch government to stimulate certain types of labour to move from one location to another or to try to alleviate other types of barriers, such as a low accessibility of regions or a low market access.

Second, it does *not* imply that the distribution of factors of production across Dutch cities is not important for trade. In terms of Figure 3, along the arrow *ab*, the redistribution of high skills and low skills over regions has no effect on the national structure of trade, but the city-region contribution changes; some cities-regions contribute more to the exports or imports of a commodity than others. The spatial concentration of production factors *does* affect the local contribution to the overall trade pattern. This raises the question whether specific spatial charateristics contribute to a locations' export structure. More specifically, in the next sections we will show that spatial differences in the structure of trade are significant; each city has its own trade structure. This arises from sorting of sectors and (skilled) workers in certain locations on the basis of factor abundance and sorting of labour-types into regions on the basis of factor intensity. A direct implication is that – at least on the short term-exogenous foreign shocks to trade, such as Brexit or the China/EU – USA trade war, impact Dutch cities-regions in different ways and thus provide relevant information for location-specific policies. An import tariff on paper products for example would hit Maastricht particularly hard (see Table 7). The current US administration has since 2017 introduced

---

high skill workers, for example, for these sectors ranges from 13.6 per cent for sector 16 (wood) to 40.2 per cent for sector 20 (chemicals), see Appendix C. Because of the current more detailed information, the sector lenses are much wider and the lens condition is always satisfied.

[25] In a long-run perspective, the sector lens is also endogenous as it changes in response to R&D efforts, but these changes are likely to require more time than adjustments of the area lens because of migration.

many new tariffs these will affect especially those locations that have a comparative advantage in those products (see for an up to date list of US tariffs: https://hts.usitc.gov/).

Third, and most important for our analysis below: the fact that the lens condition is satisfied implies that the Dutch production structure is inside the factor price equalization (FPE) area of the integrated equilibrium (see Figure 3). This, in turn, motivates why we can assume that the production technology used in different sectors is the same across different cities-regions within the Netherlands. This is important in the analysis below when we use national sector intensity information in different cities-regions when explaining local comparative advantages.

4 Revealed Comparative Advantage

The discussion below proceeds in three steps. First, we determine the comparative advantage of Dutch cities-regions relative to the Netherlands based on the Balassa index. Second, we determine the comparative advantage of The Netherlands relative to the World. Third, we determine the comparative advantage of Dutch cities-regions relative to the World.

4.1 Step One: Comparative Advantage of Dutch Cities-Regions Relative to The Netherlands

Our measure of revealed comparative advantage is the Balassa index, denoted by BI. As noted above, for city c and sector i in period t this is a normalized export share relative to an appropriate reference group of countries (ref), see the first equality sign in equation 2.

$$(2) \qquad BI_{i,t}^{c,NL} = \frac{export\ share_{i,t}^{c}}{export\ share_{i,t}^{ref}} = \frac{E_{i,t}^{c}/E_{t}^{c}}{E_{i,t}^{NL}/E_{t}^{NL}}$$

Exports E; for city-region c and sector i in period t is denoted by $E_{i,t}^{c}$. Total exports for city-region c in period t is denoted by E_{t}^{c} and is simply the sum over all sectors: $E_{t}^{c} \equiv \sum_i E_{i,t}^{c}$. This implies that the export *share* for sector i in city-region c at time t is equal to: $export\ share_{i,t}^{c} = E_{i,t}^{c}/E_{t}^{c}$. As explained in section 2, we calculate the average of yearly BIs for the two subperiods, 2007-12 and 2012-17, for 83 different sectors at the city-region level, in order to avoid too much volatility.

If the Netherlands is the reference group in equation 2 $ref = NL$ in the sup-index of the equation. It implies that if $BI_{i,t}^{c,NL} > 1$, for example for the export of tobacco from Groningen, then Groningen has a revealed comparative advantage *within* the Netherlands (which is the case).[26] It is the most direct way to determine a location's relatively strong and weak sectors within a country. Obviously, Dutch exports for sector i in period t, denoted by $E_{i,t}^{NL}$, is simply

---

[26] Before proceeding with our empirical analysis of the Balassa Indices, we exclude city-sectors that do not satisfy the Hillman-condition (Hillman, 1980). This condition evaluates the correspondence between Revealed Comparative Advantage and comparative advantage based on pre-trade relative prices (which are not observed). In general, the Hillman condition is violated if a country has an extreme market share in the supply of a particular commodity in combination with a 'high enough' degree of export specialization. The condition is satisfied for virtually all our city-sectors, 'Extraction of crude petroleum and natural gas' in Groningen being the only notable exception. This finding is in accordance with Hinloopen and van Marrewijk (2008), who find that Hillman violations are small in number and occur mostly in natural-resource intensive sectors.

the sum over all cities-regions: $E_{i,t}^{NL} \equiv \sum_c E_{i,t}^{c}$. Similarly, total Dutch exports in period t, denoted by E_t^{NL}, is then the sum over all sectors: $E_t^{NL} \equiv \sum_i E_{i,t}^{NL}$. The refence export *share* for sector i at time t is thus: $export\ share_{i,t}^{ref} = E_{i,t}^{NL}/E_t^{NL}$, hence the second equality sign in equation 2.

Note that the Balassa index for cities defined in equation 2 differs slightly from the regular definition at the country level as it focuses on exports outside of the Netherlands as a whole, instead of exports outside the city-region only. In this respect, the country thus serves as a double benchmark, which has advantages when we go to step 3 in section 4.3. In addition, note that the range of $B_{i,t}^{c,NL}$ (which starts at zero) is limited from above by the inverse of a city-region's trade share. The maximum for a good i is reached if city-region c is the only city-region that exports good i in period t, in which case: $E_{i,t}^{c} = E_{i,t}^{NL}$ and $B_{i,t}^{c,NL} = \frac{E_{i,t}^{c}/E_t^{c}}{E_{i,t}^{NL}/E_t^{NL}} = \frac{E_{i,t}^{NL}/E_t^{c}}{E_{i,t}^{NL}/E_t^{NL}} = \frac{E_t^{NL}}{E_t^{c}} = \frac{1}{s_{c,t}} > 1$, where $s_{c,t}$ is the share of city c in Dutch exports in period t.

4.2 Step Two: Comparative Advantage of The Netherlands Relative to the World

To determine the relatively strong and weak sectors for the Netherlands as a whole, we apply the regular Balassa index using UN Comtrade data with the *world* as reference group. At the Dutch national level, the export share of sector i in period t is equal to $E_{i,t}^{NL}/E_t^{NL}$. We denote world variables with a sup-script w, hence the world export share of sector i in period t is equal to $E_{i,t}^{w}/E_t^{w}$ and the Balassa index at the national level is provided in equation 3.

$$(3) \qquad BI_{i,t}^{NL,w} = \frac{E_{i,t}^{NL}/E_t^{NL}}{E_{i,t}^{w}/E_t^{w}}$$

To determine the comparative advantage for Dutch cities-regions in step one, we use 83 SBI sectors. To determine the comparative advantage for the Netherlands as a whole in step 2, we use international trade data based on the Harmonized System (HS, 2017 classification). For goods sectors we can make an adequate concordance between these two data sets using the NACE and ISIC classifications as intermediate steps. As summarized in 5, this means that we can classify 5,304 HS2017 products at the 5-digit level to correspond to 42 SBI goods sectors, which is then used to determine comparative advantage for these 42 sectors for the Netherlands as a whole.[27] Unfortunately, we cannot perform a sufficiently reliable similar correspondence exercise for services sectors, which means that step 2 in this sub-section and step 3 in the next sub-section can only be performed for the 42 goods sectors listed in Table 5.

Table 5 Overview of included SBI sectors and number of HS2017 concordance subsectors

SBI code	# HS products	SBI code	# HS products	SBI code	# HS products
01	288	18	27	32	382
02	39	19	21	33	518
03	71	20	862	35	2

[27] A summary and complete overview of the concordance is available upon request.

06	4	21	4	38	4
08	37	22	50	41	4
09	53	23	148	43	8
10	537	24	364	58	12
11	18	25	124	59	3
12	7	26	211	71	1
13	470	27	75	74	2
14	276	28	295	79	1
15	48	29	47	90	6
16	90	30	74	95	4
17	95	31	21	96	1

Source: author calculations; concordance includes 42 SBI sectors with 5304 corresponding HS2017 products; see Appendix A for SBI sector description; the table lists the number of HS2017 products corresponding to a given SBI sector; for example: SBI sector 08 consists of 37 HS2017 products.

4.3 Step Three: Comparative Advantage of Dutch Cities Relative to The World

Once we have determined the comparative advantage of Dutch cities relative to the Netherlands $BI_{i,t}^{c,NL}$ in step 1 and the comparative advantage of the Netherlands as a whole relative to the world $BI_{i,t}^{NL,w}$ in step 2, simple multiplication suffices to determine the comparative advantage of Dutch cities relative to the world $BI_{i,t}^{c,w}$ for the sectors listed in Table 4:

$$(4) \qquad BI_{i,t}^{c,w} = BI_{i,t}^{c,NL} \cdot BI_{i,t}^{NL,w} = \left(\frac{E_{i,t}^{c}/E_{t}^{c}}{E_{i,t}^{NL}/E_{t}^{NL}}\right)\left(\frac{E_{i,t}^{NL}/E_{t}^{NL}}{E_{i,t}^{w}/E_{t}^{w}}\right) = \frac{E_{i,t}^{c}/E_{t}^{c}}{E_{i,t}^{w}/E_{t}^{w}}$$

City-region c thus has a revealed comparative advantage relative to the world in sector i in period t if $BI_{i,t}^{c,w} = BI_{i,t}^{c,NL} \cdot BI_{i,t}^{NL,w} > 1$. Note that it is possible for city-region c to be relatively strong in sector i within the Netherlands, but not relative to the world, namely if $BI_{i,t}^{c,NL} > 1$ and $BI_{i,t}^{c,w} < 1$ (note: this requires $BI_{i,t}^{NL,w} < 1$). Similarly, it is possible for city c to be relatively weak in sector i within the Netherlands, but strong compared to the world, namely if $BI_{i,t}^{c,NL} < 1$ and $BI_{i,t}^{c,w} > 1$ (note: this requires $BI_{i,t}^{NL,w} > 1$). We return to these issues in section 7.

5 Comparative Advantage of Dutch Cities-Regions relative to the Netherlands

The information in this section is based on average Balassa indices per sector and location over the years 2012-2017 relative to The Netherlands.[28] We use the period 2007-2012 for comparison throughout our discussion. We start with an overview of the main characteristics of the Balassa index in section 5.1, followed by a discussion on the comparability of the distribution for different locations in section 5.2 and an introductory analysis on the link between factor abundance and comparative advantage in section 5.3. A more detailed

[28] If a sector does not export from a location in all years it is excluded from the analysis; we thus only have non-zero Balassa indices, which allows us to use the log of the Balassa index for some of our analysis.

analysis of the connection between abundance, intensity, and comparative advantage is provided in section 6.

5.1 Characteristics of the Balassa Index

Table 6 provides summary statistics for the Balassa index of Cities-Regions relative to The Netherlands. In the period 2012-2017, the average Balassa index is 1.17 and the median is 0.53, which indicates the overall distribution is skewed to the right, with skewness of 8.47 and a maximum of 44.2. The share of sectors with a Balassa index higher than one (and thus a revealed comparative advantage) is 32.2 per cent. All these observations are similar to the findings in Hinloopen and van Marrewijk (2001) and to the 2007-2012 period, see Table 6.

Table 5 also indicates that there are differences for cities and regions. Since there are only 4 regions and 22 cities, the city distribution is close to the overall distribution just described. In contrast, for the regional distribution the mean and median are higher than for cities (1.32 and 0.95 compared to 1.13 and 0.46, respectively) and the skewness is much lower (4.48 compared to 8.44). As a consequence, almost half of the sectors in regions has a revealed comparative advantage ($BI > 1$) compared to less than 30 per cent at the city level. As we show in Figure 9, this is related to the number of sectors included in the analysis, which is higher for regions than for cities. The distribution in regions is thus less skewed and sector in rural regions are more likely to have a comparative advantage compared to the Netherlands as a whole. Similar observations hold for the 2007-2012 period, see Table 6.

Figure 9 illustrates the differences in distribution characteristics for the individual cities-regions in four panels. The horizontal axis always displays the share of sectors with a revealed comparative advantage ($BI > 1$) in percentage points. The vertical axis in panels *a*-*d* depict the relationship with the mean, median, maximum, and the number of included sectors, respectively. Cities are identified by balls and regions by squares. Panels *a* and *b* also display the mean and median for the cities and regions separately.

There is no relationship between the share of the Balassa index above one and the maximum observation for a city-region, see panel *c* (share of variance explained is 0.0 per cent). The maximum is particularly high for Groningen (sector 06; oil and gas extraction, and results in a violation of the Hillman condition) and Den Bosch (sector 65; insurance and pensions). By far the highest share of explained variance (82 per cent) is for the median Balassa index, see panel *b*. The relationship with the mean is less tight (31.4 per cent, see panel *a*) because it is influenced by the maximum observations. Finally, if the number of included sectors is high, the share of sectors with revealed comparative advantage ($BI > 1$) is also high (see panel *d*), which explains why this is particularly high for regions. The next section analyses the extent to which the distribution differs between locations.

Table 6 Summary Statistics of the Balassa Index; Cities-Regions relative to NL

a. Period 2012-2017

Variable	Regions only	Cities only	Cities-Regions
Number of observations	326	1607	1933
Average Balassa index	1.32	1.13	1.17
Median Balassa index	0.95	0.46	0.53

Skewness	4.48	8.44	8.47
Maximum Balassa index	15.6	44.2	44.2
Number $BI > 1$	158	464	622
Per cent $BI > 1$	48.5	28.9	32.2

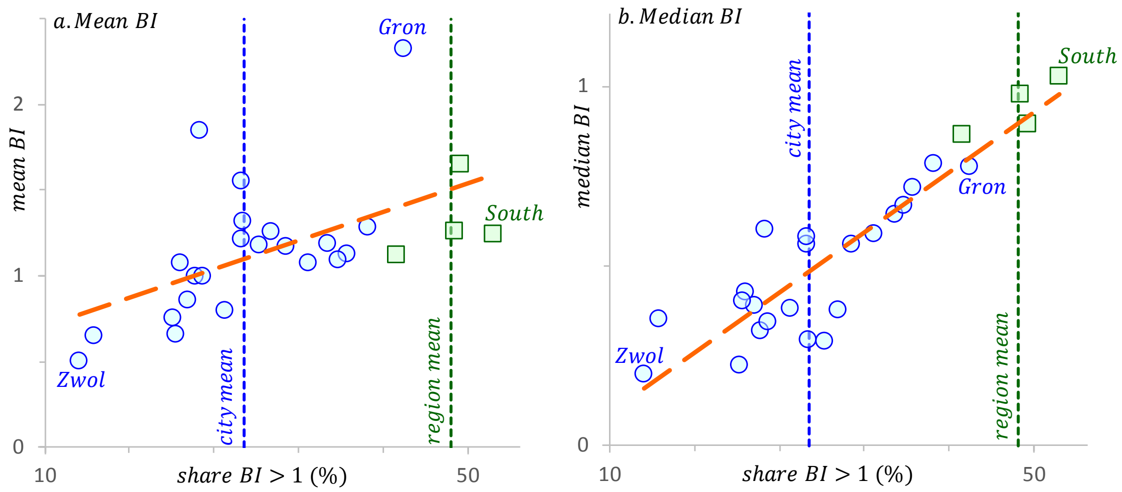
b. Period 2007-2012

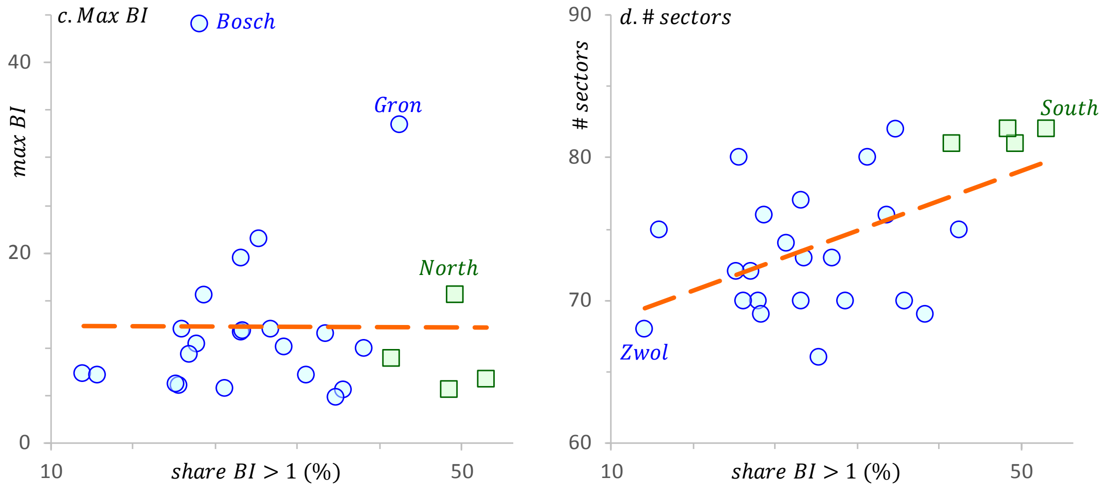
Variable	regions	cities	all
Number of observations	323	1563	1886
Average Balassa index	1.42	1.25	1.28
Median Balassa index	0.99	0.46	0.54
Skewness	3.20	10.67	10.95
Maximum Balassa index	14.2	69.9	69.9
Number $BI > 1$	160	459	619
Per cent $BI > 1$	49.5	29.4	32.8

Source: author calculations; see section 2 for data info; see section 4 for calculation details.

Table 7 provides an overview of the two strongest export sectors as identified by the Balassa index for each city-region. For readers familiar with the economic structure of the Netherlands, some of these are in accordance with expectations. Groningen and North, for example, have sector 06 (extraction of oil and gas) as their strongest sector. Similarly, Amsterdam (where the national airport is located) is strong in sector 79 (travel) and Rotterdam (where the largest port is located) is strong in sector 50 (water transport). Other locations are strong in sectors that upon reflection are understandable. The West region (along the coast) is strong in sector 03 (fishing) and the South region (with lots of trees) is strong in sector 02 (foresty).

Figure 9 Balassa Index Cities-Regions Characteristics, 2012-2017





Source: authors; see section 2 for data info; BI = Balassa Index; Zwol = Zwolle; Gron = Groningen; Bosch = Den Bosch; share of variance explained is 31.4; 82.0; 0.0; and 29.6 per cent in panels *a-d*, respectively.

Other observations are perhaps more remarkable. Take the textile sector, for instance. Despite a decades long decline of the number of workers in the textile sector, the East region is still strong in textiles (sector 13). Eindhoven is strong in sectors 28 (machinery not elsewhere classified) and 26 (computers and electronics), while Utrecht is strong in sector 85 (education) and Dordrecht is strong in sector 81 (facility management). We will argue below that this is related to the skill requirements (intensity) of sectors in combination with the skill abundance in locations. Sectors 26 and 28 are relatively intensive in the use of technical high skill workers and Eindhoven is relatively abundant in technical high skill workers. Similarly, sector 85 is most intensive in the use of high skill workers and Utrecht is most abundant in high skill workers, while sector 81 is most intensive in low skill workers and Dordrecht is relatively abundant in low skill workers.

Table 7 Cities-Regions Strongest Export Sectors; relative to NL, 2012-2017

City-Region	Strongest SBI sector	Second strongest SBI sector
North	06; Extract gas	08; Mining
East	75; Veterinary act	13; Textiles
West	03; Fishing	36; Collect water
South	12; Tobacco prod	02; Forestry
Groningen	06; Extract gas	86; Human health
Leeuwarden	17; Paper prod	25; Fabricated metal
Zwolle	70; Holding comp	74; Industrial design
Enschede	31; Furniture	30; Other transport eq
Apeldoorn	24; Basic metals	17; Paper products
Arnhem	16; Wood prod	84; Public services
Nijmegen	27; Electrical eq	86; Human health
Amersfoort	73; Advertising	65; Insurance & pension
Utrecht	88; Social work	85; Education
Amsterdam	79; Travel agencies	63; Information services
Haarlem	09; Mining support	18; Printing

Leiden	21; Pharmaceutical prod	72; R&D
The Hague	41; Construction	93; Sports and recreation
Rotterdam	50; Water transport	52; Warehousing
Dordrecht	19; Man oil prod	81; Facility man
Breda	42; Civil engineering	22; Rubber & plastic
Tilburg	15; Leather	29; Motor vehicles
Den Bosch	65; Insurance & pension	37; Sewerage
Eindhoven	28; Machinery n.e.c.	26; Man computers
Geleen-Sittard	96; Wellness & funeral	29; Motor vehicles
Heerlen	11; Beverages	96; Wellness & funeral
Maastricht	17; Paper prod	82; Other business serv

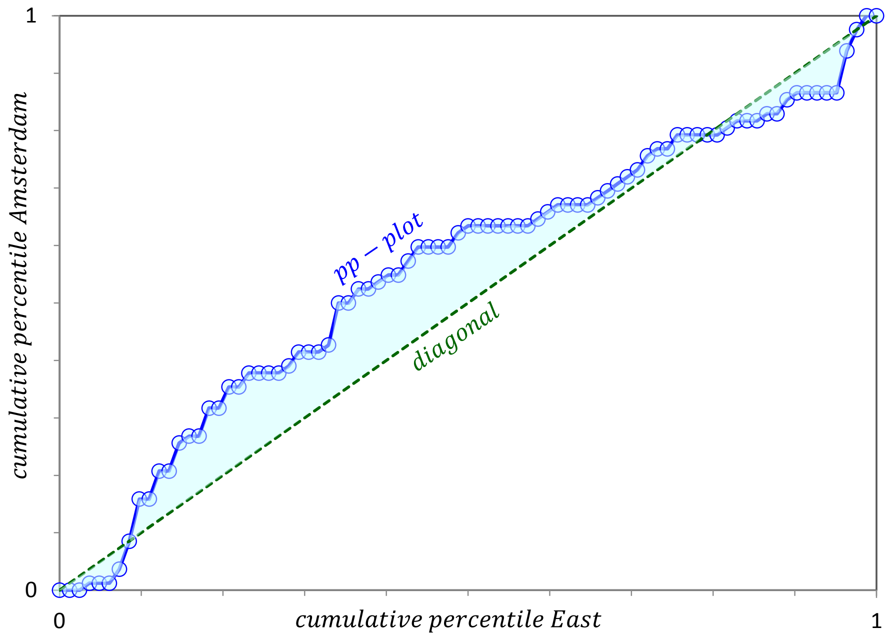
Source: author calculations; see section 2 for data details; ranking based on Balassa index; see Appendix A for more detailed sector description.

5.2 Comparing Distributions

The firm-sector export data allows us to calculate the Balassa index for each sector in each city-region to identify strong export sectors, as discussed in section 5.1. For a given city-region we have a sector-based distribution of Balassa indices, which allows us to order or rank the sectors in terms of revealed comparative advantage within the city-region. The question arises to what extent we can compare values of the Balassa index in different locations. Can we conclude, for example, that a Balassa index of 4 for a sector in Amsterdam is stronger than a Balassa index of 2 for another sector in East? To be able to do so, the observations should be based on a similar underlying distribution. We thus require a test to conclude whether the observations from two different locations are drawn from the same distribution, or not.[29]

[29] Differences can subsequently be related to explanatory variables to explain Balassa index patterns (see also Deardorff, 2011 and Kowalski and Bottini, 2011).

Figure 10 PP-plot and HWM index; East and Amsterdam, 2012-2017



Source: authors; see section 2 for data details; HWM index is based on shaded area between diagonal and pp-plot

The method we use is the 2-sample Harmonic Weighted Mass (HWM) index developed by Hinloopen, Wagenvoort, and van Marrewijk (2012). The essence of this method is the comparison of two entire distributions based on so-called probability–probability plots for the (empirical) cumulative distributions.[30] This is illustrated for East and Amsterdam in the period 2012-2017 in Figure 10. If the draws are from the same underlying distribution, the expected value of the pp-plot coincides with the diagonal. The HWM index takes the deviation between the actual pp-plot and the diagonal (the shaded area in Figure 10) corrected for the number of observations as a measure to determine if the underlying distributions are the same (if the corrected area is sufficiently small) or not (if the corrected area exceeds a critical value). In the case of East and Amsterdam, the underlying distributions are *not* the same (beyond the one per cent significance level).

The table in Appendix D provides a summary of all possible bilateral comparisons (at the 10 per cent significance level). The conclusion is that, in general, the Balassa index distributions differ from each other. For example, in the period 2012-2017 the Balassa index distribution of Groningen differs from the Balassa index distribution in all other locations (100 per cent of the cases) while The Hague (which is among the locations with the highest number of similarities with other distributions) still differs in 76 per cent of all comparisons. On average, the distributions are significantly different for 89 per cent of all cases in the 2012-

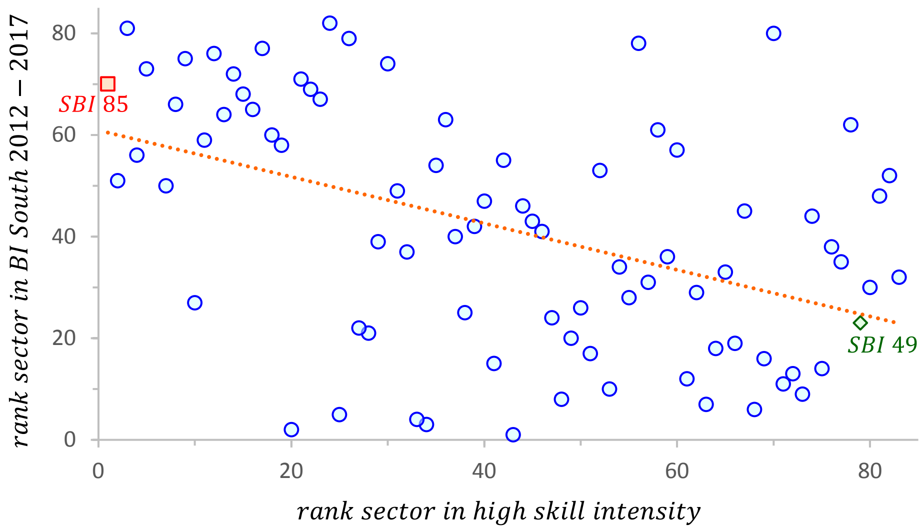
[30] The HWM-index has many attractive properties for applied research: it is not susceptible to outliers in the data, is scale-invariant and there is no need for discrete approximations, such as in applications using Markov transition matrices. Hinloopen, Wagenvoort, and van Marrewijk (2012) also analytically derive exact, finite-sample critical values for the HWM-index, which makes it more attractive than (variants) of kernel estimates.

2017 period and for 87 per cent of the cases in the 2007-2012 period. We therefore conclude that it is not appropriate to compare the value of the Balassa index for corresponding sectors in different cities-regions, because the same value could represent a very different ranking and a different value a similar ranking. It is only appropriate to compare the order within the same city-region, or the ranking of sectors in different cities-regions (see also sections 5.3 and 6).[31]

5.3 The Sorting of Sectors and workers across Cities-Regions

How can we explain the strengths and weaknesses of a sector across city-regions? In section 5.1 we already suggested that this is related to the skill intensity of sectors and the skill abundance in cities-regions. An explanation closely related to the Heckscher-Ohlin trade model. We develop this line of reasoning in this section in two steps. The first step is to show the link between sector skill intensity and a city-region's strong export sectors in terms of a correlation for that city-region. The second step is to show the link between this correlation and a city-region's relative factor abundance. Section 6 provides a more detailed statistical analysis of these interaction effects.

Figure 11 Rank Correlation Balassa Index and High Skill Share; South, 2012-17



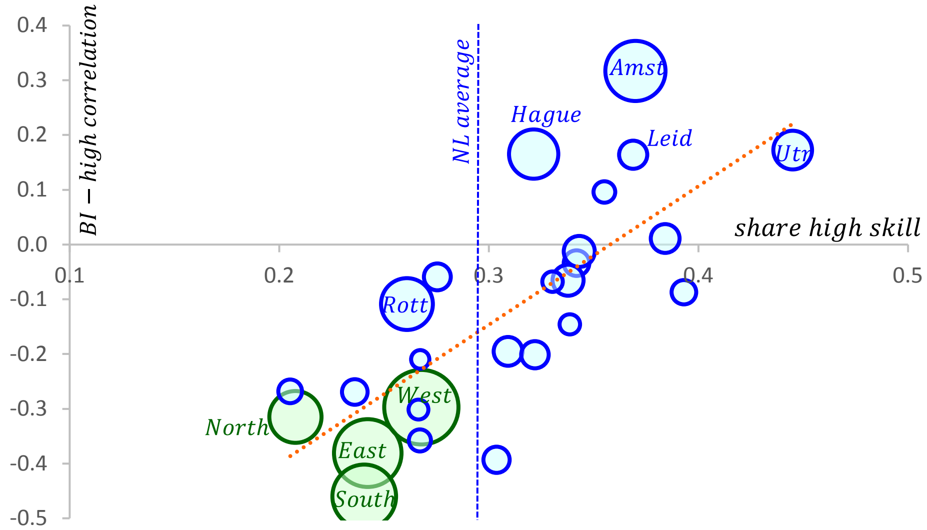
Source: authors; see section 2 for data info; rank correlation is -0.4596; dotted line is a trendline.

The first step (relating strong sectors to skill intensity) is illustrated for South in Figure 11. We motivated in section 5.2 why we prefer to compare the ranking of sectors in different locations over comparing the value of the Balassa index in different locations. The vertical axis in Figure 11 therefore shows the rank of a sector in South's Balassa index (where 1 is highest, the strongest export sector), while the horizontal axis shows the rank of that sector in terms of high skill intensity. Sector 85 (education), for example, has rank 1 in terms of high skill intensity and rank 70 in terms of South's Balassa index, while sector 49 (land transport) has rank 79 in terms of high skill intensity and rank 23 in terms of South's Balassa index. The figure illustrates that, in general there is a negative association for South between its strong

[31] Furthermore, the BI is a ratio and values are susceptible to nominator and denominator effects; a sector with the same rank in two cities might have very different values.

export sectors and high skill intensity (rank correlation is -0.4596). In South, therefore, relatively strong export sectors tend to have relatively low high skill intensity (the $BI - high$ correlation is negative). Please note that this outcome is not robust, because if we perform a similar calculation for other cities-regions in some cases the $BI - high$ correlation is positive and in other cases this correlation is negative. This brings us to the next step.

Figure 12 Relationship between Share High Skill and BI-High Correlation, 2012-17



Source: authors; see section 2 for data info; Amst = Amsterdam; Utr = Utrecht; Rott = Rotterdam; Leid = Leiden; Hague = The Hague; bubbles proportional to population size; dotted line is a trendline; share of (unweighted) variance explained is 57.7 per cent.

Step 2, relating correlations between strong sectors and skill intensity to factor abundance, is illustrated in Figure 12. On the vertical axis we depict the correlation between strong sectors (the Balassa index) and high skill intensity in a city-region (the $BI - high$ correlation). This is related on the horizontal axis to the share of high skill workers in a city region. Moreover, the vertical dashed line indicates the average share of high skill workers for The Netherlands as a whole for reference, the sloping dotted line is a trendline, and the size of the bubbles is proportional to population. Figure 12 illustrates that the cities-regions with a relatively *low* share of high skill workers, like South, the other regions and Rotterdam, have a *negative* $BI - high$ correlation, while cities-regions with a relatively *high* share of high skill workers, like Utrecht, Amsterdam, Leiden, and The Hague, have a *positive* $BI - high$ correlation.[32] The share of high skills available in cities-regions explains 57.7 per cent of the variance in $BI - high$ correlation in Figure 12. In other words, this correlation is related to the factor abundance in a city-region (in this case of high skill workers), which indicates that strong export sectors are related to the interaction of sector skill intensity with a city-region's factor abundance.

This section has illustrated, using high skill workers as an example, that strong sectors in a city-region are determined by the interaction between sector skill intensity and city-region

[32] Note that Rotterdam stands out in the 'Randstad' area (which also includes Utrecht, Amsterdam, Leiden, and The Hague), as it is relatively low skill abundant instead of high skill abundant.

skill abundance. The next section will analyse this in more detail, extending the analysis to include other skills and technical skills, to show that low skill intensive sectors thrive in low skill abundant cities-regions and technical skill intensive sectors thrive in technical skill abundant cities-regions. As such the analyses links factor abundance, skill intensity, and revealed comparative advantage at the city-region level within The Netherlands.

6 Abundance, Intensity, and Comparative Advantage

As explained in section 2.3, we identify two types of human skills (general and technical) at three levels (high, medium, and low). We have information available regarding the availability of skills in cities-regions and regarding the use of skills in sectors. In section 5.3 we argued that the distribution of strong sectors across cities-regions is related to the interaction between the abundance of skills in cities-regions and the intensity of the use of skills in sectors. Before we can analyse this interaction in more detail in section 6.2, we have to specify how we define abundance and intensity in section 6.1.

6.1 Abundance and Intensity

A city-region is *abundant* in a certain skill if this skill is *relatively* widely available in that city-region. Similarly, a sector is *intensive* in a certain skill if this skill is intensively used in *relative* terms. To determine the abundance of skills in cities-regions or the intensity of skills in sectors we therefore need an appropriate measure to compare to, for which we take the average of this skill for the Netherlands as a whole. As indicated in section 2.3, we use simple mnemonics for our variables. We use $abun$ for abundance, int for intensity, and $share$ for shares. In addition, we use sub-indices $high$, med, and low for high-, medium-, and low general skill levels, sub-indices $tec-high$, $tec-med$, and $tec-low$ for high-, medium, and low technical skill levels, and a sub-index t for time. Finally, we use a sup-index c for cities-regions, a sup-index i for sectors, and a sup-index NL for The Netherlands as a whole.

(5) $$Abun_{high}^{c} = \frac{1}{6}\left\{\left(\sum_{t=2012}^{2017} share_{high,t}^{c}\right) - \left(\sum_{t=2012}^{2017} share_{high,t}^{NL}\right)\right\}$$

(6) $$Int_{high}^{i} = \frac{1}{6}\left\{\left(\sum_{t=2012}^{2017} share_{high,t}^{i}\right) - \left(\sum_{t=2012}^{2017} share_{high,t}^{NL}\right)\right\}$$

Equation *(*5) provides the definition of high skill abundance for cities-regions in 2012-2017 and equation *(*6) provides a similar definition of high skill intensity for sectors. Similar definitions apply to all other types and levels of skills. The definition implies that abundance is *positive* in a city-region if the period-average share in the city-region is *higher* than the period-average share for the Netherlands as a whole, and negative otherwise. Similarly, intensity is *positive* for a sector if the period-average share in the sector is *higher* than the period-average share for the Netherlands as a whole, and negative otherwise.

Table 8 Abundance of Cities-Regions and Intensity of Sectors; Correlation, 2012-2017

a. Correlation Abundance in Cities-Regions; 26 observations, 2012-2017

	$Abun_{low}$	$Abun_{med}$	$Abun_{high}$	$Abun_{tec-low}$	$Abun_{tec-med}$
$Abun_{low}$	1.000				
$Abun_{med}$	0.466	1.000			
$Abun_{high}$	-0.895	-0.811	1.000		

$Abun_{tec-low}$	0.849	0.673	-0.899	1.000	
$Abun_{tec-med}$	0.722	0.802	-0.881	0.919	1.000
$Abun_{tec-high}$	-0.453	-0.718	0.661	-0.500	-0.550

b. Correlation Intensity for Sectors; 83 observations, 2012-2017

	Int_{low}	Int_{med}	Int_{high}	$Int_{tec-low}$	$Int_{tec-med}$
Int_{low}	1.000				
Int_{med}	0.603	1.000			
Int_{high}	-0.887	-0.903	1.000		
$Int_{tec-low}$	0.669	0.461	-0.627	1.000	
$Int_{tec-med}$	0.267	0.387	-0.367	0.794	1.000
$Int_{tec-high}$	-0.429	-0.468	0.502	-0.050	0.284

Source: author calculations; see section 2 for data details; shaded cells have negative correlation.

Table 8 provides the correlation matrices for cities-regions abundance (panel *a*) and sector intensity (panel *b*). In absolute value the correlation tends to be higher for abundance than for intensity.[33] This indicates that there is more variation within sectors than for cities-regions, which is in line with our lens discussion in section 3. We also observe that with one exception the correlation sign is the same for abundance and intensity. The exception is the correlation between technical medium and technical high skills, which is negative for cities-regions (abundance) and positive for sectors (intensity). All correlations between low and medium skills for both general and technical types, as well as general-technical cross-correlations, are positive. Except for the one exception mentioned above, all correlations between high skills and either medium or low skills for both general and technical types, as well as general-technical cross correlation, are negative. Finally, the correlation between high skills and technical high skills is positive.

6.2 Interaction and Comparative Advantage

Our OLS regression analysis focuses on the interaction between factor abundance for cities-regions and factor intensity for sectors to determine a city-region's strong sectors. We do not analyse the causality of the sorting problem if sectors decide to produce in cities-regions abundant in the skills that the sector uses intensively, or if workers decide to locate in cities-regions with sectors that intensively use their skills.[34]

$$(7) \qquad rank_i^c\left(BI_{i,t}^{c,NL}\right) = \beta_0 + \beta_1 Abun_{high}^c \times Int_{high}^i + .. + controls + \varepsilon_{ict}$$

Table 9 Abundance, Intensity, and Comparative Advantage, 2012-2017

a. OLS regression on rank of Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun_{high}^c \times Int_{high}^i$	-258.0***					
	(4.77e-10)					
$Abun_{med}^c \times Int_{med}^i$		-776.6***				
		(2.64e-08)				

[33] The off-diagonal average absolute value is 0.720 for abundance and 0.513 for intensity.
[34] Note, the sector with the highest BI has rank=1, the sector with the lowest BI, the highest rank in a location (the number of exporting sectors differs per location).

$Abun^c_{low} \times Int^i_{low}$			-438.1*** (0.000205)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				-772.2** (0.0175)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					-1,767*** (0)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						-13,312*** (0)
R-squared	0.175	0.171	0.163	0.159	0.181	0.191

b. Probit regression on Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun^c_{high} \times Int^i_{high}$	15.89*** (2.42e-08)					
$Abun^c_{med} \times Int^i_{med}$		42.37*** (1.32e-05)				
$Abun^c_{low} \times Int^i_{low}$			31.22*** (0.000276)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				52.64** (0.0211)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					121.5*** (0)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						890.9*** (0)

c. OLS regression on natural logarithm Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun^c_{high} \times Int^i_{high}$	20.72*** (3.88e-09)					
$Abun^c_{med} \times Int^i_{med}$		60.05*** (8.06e-08)				
$Abun^c_{low} \times Int^i_{low}$			33.84*** (0.00247)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				41.29 (0.305)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					104.9*** (2.93e-08)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						845.2*** (1.60e-10)
R-squared	0.241	0.236	0.231	0.227	0.237	0.243

Source: author calculations; see section 2 for data details; robust pvalues in parentheses; *** p<0.01; ** p<0.05; * p<0.1; all regressions have 1886 observations and include sector fixed effects and city-region fixed effects.

Our regression estimates aim to explain strong export sectors as identified by the Balassa index in three ways. First, we use the *rank* of sector i within city-region c, $rank_i^c\left(\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,NL}}\right)$, to determine the location-specific order of strong export sectors. This eliminates any location-specific distribution issues related to the *value* of the BI as discussed in section 5.2. Our focus is on determining the interaction effects for city-region abundance and sector intensity, such as parameter β_1 for high skills in equation (7), taking into consideration spatial specific control variables as discussed below. Second, we analyze strong export sectors that have a revealed comparative advantage, $\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,NL}} > 1$. A probit analysis determines the probability that $\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,NL}} > 1$ in a structure similar to equation (7). Third, we use $ln\left(\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,NL}}\right)$ as a measure of export strength for sector i in city-region c in period t. Note that this variable suffers from the limitations discussed in section 5.2, and is provided as a robustness check only.

Table 9 provides an overview of the individual interaction effects for all three types of analyses. To focus on the abundance-intensity interaction effects, we correct for sector fixed effects and cities-regions fixed effects. In general, the abundance-intensity interaction effects are strong and highly significant: all estimated coefficients are significant at the 1 per cent level, except for high technical skills, which is significant at the 5 per cent level for the rank and probit analysis (panels *a* and *b*) and not significant for the log analysis (panel *c*). For general schooling, the high skill variable has most explanatory power. In contrast, for technical schooling, the low skill variable has most explanatory power. The conclusion based on the Table 9 is that location specific comparative advantage is to a large extent explained by the combination of location specific factor abundance and skill intensity. This interaction also determines the probability of export success as is illustrated by the Probit results. The results are in accordance with the Heckscher-Olin theory.

6.3 Locational characteristics and Comparative Advantage

Next we look at how locational characteristics affect the BI rank of sectors across locations, in an OLS regression of the rank of a specific sector in a city-region and location specific characteristics:

$$(8) \qquad rank_i^c(BI_i^c) = \beta_0 + \beta_1 LocationCharacteristics^c + .. + controls + \varepsilon_{ic}$$

Table 10 Descriptive statistics, locational chaacteristics and controls, 2012-2017

Variable	mean	sd	min	max	p5	p95
cityhigh	0.297	0.0585	0.197	0.434	0.200	0.381
citytechigh	0.0500	0.0135	0.0287	0.0911	0.0332	0.0740
markaccexp	7.672	1.687	3.395	9.832	3.982	9.497
density	938.3	535.6	120.4	2,359	235.4	1,596
schiphroad	104.3	53.33	15.02	197.9	28.42	188.7
portroad	111.2	64.49	9.750	252.2	26.77	237.5

Number of observations: 1886

We aggregate the 83 sectors of the analysis into larger categories in order to have enough observations per category. First we aggragate sectors into two broad categories: Goods and Services.[35] Next, we aggregate sectors into five broader sector categories as defined by Statistics Netherlands.[36] In general, locational characteristics only marginally contribute to local comparative advantage.

The location specific explanatory variables are: (i) city-region population density, (ii) market access, (iii) the road distance to the nearest main international airport (Schiphol), and (iv) the road distance to the nearest main international Port (Rotterdam). We include the proportion of

[35] The ‘Goods’ category consists of sector codes 01 (Agriculture and related service activities) up to and including 32 (Manufacture of other products n.e.c.). All other sectors belong to the ‘Services’ category in the analysis, see also Appendix A.
[36] Note that in order to analyze the effect of locational qualities on a sector in a city-region the locational characteristics also have to vary across city-regions.

high-skilled (*cityhigh*) and technical high-skilled workers (*citytechigh*) in a city-region as controls. We also include the average BI rank (*avgBIrank*) of a city-region, to control for the fact that the number of exporting sectors differs across locations, implying that the maximum value of the rank number per location differs. For measuring distances, we use the economic centroid of each city-region as a reference point.[37]

Table 11 Goods, Services and regional Comparative Advantage, 2012 - 2017

	Goods			Services		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
density	0.010***	0.006**	0.006*	-0.005***	-0.003	-0.003
	(7.23e-10)	(0.02)	(0.05)	(0.00002)	(0.1)	(0.2)
cityhigh			62***			-31**
			(0.006)			(0.04)
citytechigh			-146			73
			(0.2)			(0.3)
markaccexp		-3.7***	-3.4**		1.9**	1.7*
		(0.005)	(0.01)		(0.03)	(0.06)
schiphroad		-0.1***	-0.06		0.05**	0.03
		(0.001)	(0.1)		(0.02)	(0.3)
portroad		-0.04	-0.06		0.02	0.03
		(0.3)	(0.1)		(0.4)	(0.3)
avgBIrank	1.1***	1.3***	1.5***	0.9***	0.9***	0.7***
	(0.00003)	(0.00001)	(7.62e-07)	(2.88e-07)	(0.00002)	(0.0003)
Observations	623	623	623	1,263	1,263	1,263
R^2	0.081	0.100	0.111	0.034	0.039	0.042

Source: author calculations; BI rank is the dependent variable; constant included, not reported; robust p values in parentheses; *** p<0.01; ** p<0.05; * p<0.1

The market access variable *markaccexp* is defined as a population-weighted exponential distance decay function, and is a proxy for the size of the local- and nearby markets.[38] The *density* measure is the average population per square km in a city-region, and captures local costs-of-living aspects as well as possible (knowledge) spill-overs. The distance to Schiphol (*schiphroad*) and the port of Rotterdam (*portroad*) captures the distance of a city-region to the main international airport and international port in the Netherlands by road in km. Both variables are positively correlated with market access (see Appendix F). Table 10 presents some descriptive statistics for the locational characteristics and controls.

The results are presented in Table 11. In general, the explanatory power of the relationship between export performance and locational characteristics is lower than the results for equation (7) presented in Table 9. Services tend to have a better BI rank in high-skilled

[37] The economic centroid is the average of the spatial centroids of the municipalities within a city-region weighted by the number of jobs in each municipality.

[38] Distance decay is $exp(-d_{ij})$, where d_{ij} is the distance between the economic centres of cities-regions i and j in 100 km.

regions (see footnote 34).[39] The opposite is true for Goods. This is not surprising, as the service-industry more intensively uses high-skill workers and benefits from knowledge spill-overs. Density is *negatively* correlated to the BI ranking of Goods sectors: these sectors perform better in more spaceous/less populated areas. Market access benefits Goods more than Services. This could be related to higher transportation costs for goods than for services. This also could explain the positive effect of the proximity to the international airport, Schiphol, compared to services. sector. However, this does not apply for proximity to the international port of Rotterdam, which does not benefit export performance of goods or services.

Table 12 presents the results for a further categorisation of sectors into five broad categories. These categories are the five largest administrative sector-categories (categories C, J, M and N) as maintained by Statistics Netherlands. Appendix A indicates how the 83 sectors in our sample are allocated to the five broad categories. The results are displayed in Table 12.[40]

The results for *Density* in Table 12 are less pronounced than for the results presented in Table 11. High Density is not beneficial for any of the sector categories. Manufacturing (category C) tends to receive better BI ranks in *less* dense areas, which again could reflect a need for space in manufacturing. Consistent with the results in Table 11, *Market access* benefits manufacturing, which could be related to transportation costs in this sector. *Market access* is correlated with *lower* ranks for consultancy, research and other specialised business services (sector M) and renting and leasing of tangible goods and other business support services (sector N).

Proximity to the international airport Schiphol is positively correlated with higher ranks for manufacturing and for transportation and storage (sector H). It should not be surprising that the transport industry thrives around Schiphol. Access to the international port of Rotterdam appears particularly beneficial for manufacturing, whereas is it associated with lower ranks for transportation and storage. This may result from the location of this sector's offices (where revenues accrue to), and the location of foreign customers (who generate export revenues), as opposed to the location of this sector's daily activities.

Finally, the results for the controls, high intensity and high skills, are not robust.

We conclude from the analyses in this section that the results for the relationship between local export performance and locational characteristics (such as market access and density) are not very robust.

Table 12 Sector categories and regional Comparative Advantage, 2012 - 2017

	Sector C Manufacturing	Sector H Transport	Sector J Info-com	Sector M Consultancy	Sector N Rent-lease

[39] Note that care should be taken in comparing the magnitude of the effects in our analyses, as variables are measured at different scales.

[40] Some categories contain more sectors than others. Thus, the results for the 'smaller' categories have a low statistical power compared to categories with more observations, increasing the probability of type II errors. This is why we have opted for the five largest categories.

Variables	(1)	(2)	(3)	(4)	(5)
cityhigh	62^{***}	-89	-92^{**}	0.03	2.9
	(0.006)	(0.1)	(0.04)	(1.0)	(0.9)
citytechigh	-138	420*	261	-188	-123
	(0.2)	(0.06)	(0.3)	(0.3)	(0.5)
density	0.005^{*}	-0.008	-0.006	-0.001	0.002
	(0.07)	(0.2)	(0.4)	(0.9)	(0.7)
markaccexp	-4.0^{***}	0.9	4.2	4.1^{*}	5.7^{**}
	(0.005)	(0.8)	(0.1)	(0.09)	(0.03)
schiphroad	-0.06^{*}	-0.2^{***}	0.2^{**}	0.1^{**}	0.09
	(0.07)	(0.002)	(0.02)	(0.03)	(0.1)
portroad	-0.07^{*}	0.2^{**}	0.06	0.02	0.09
	(0.08)	(0.02)	(0.5)	(0.8)	(0.2)
avgBIrank	1.5^{***}	0.7	1.3^{**}	0.7	-0.3
	(0.000001)	(0.4)	(0.03)	(0.2)	(0.6)
Observations	556	109	143	172	154
R^2	0.126	0.180	0.248	0.091	0.051

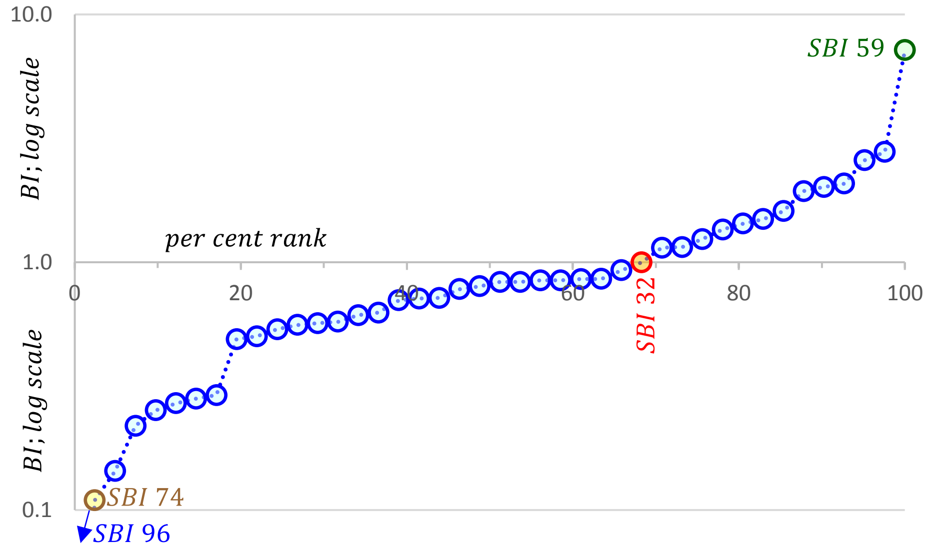
Source: author calculations; BI rank is the dependent variable; constant included, not reported; robust p values in parentheses; *** p<0.01; ** p<0.05; * p<0.1; for brief reference names: see Appendix A.

7 Comparative Advantage of Dutch Cities relative to the World

The main objective of calculating the Balassa Index is to determine which sectors are relatively strong or weak at a certain location. In this section we classify as 'weak' all sectors with a Balassa index below one and as 'strong' all sectors with a Balassa index above one.

A crucial aspect of calculating the Balassa index is choosing the appropriate group of reference countries for determining strong and weak sectors. As discussed in section 4.3, equation *(*4): $BI_{i,t}^{c,w} = BI_{i,t}^{c,NL} \cdot BI_{i,t}^{NL,w}$ we can determine strong and weak sectors for Dutch cities relative to the *world*, rather than for The Netherlands only. To do so, we limit attention to 42 sectors (see Table 4 and the discussion in section 4.2). We first determine strong and weak sectors for the Netherlands as a whole relative to the world in section 7.1, that is $\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{NL,w}}$. In section 7.2 we determine strong and weak sectors for Dutch cities relative to the world, that is $\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,w}}$. The results for $\mathrm{BI}_{\mathrm{i,t}}^{\mathrm{c,NL}}$, are already presented in section 5. This enables us to identify sectors in City-regions that are strong to the world and to The Netherlands. We can also identify sectors that are weak relative to one benchmark, but strong relative to the other; the switching sectors.

Figure 13 Balassa Index; per cent rank, Netherlands relative to World, 2012-2017



Source: authors; see section 2 for data info; see Appendix A for sector description; 42 SBI sectors.

7.1 The Netherlands relative to the World

Using the methodology outlined above, we determine strong and weak sectors for the Netherlands as a whole relative to the World for 42 SBI sectors. We focus the discussion on the period 2012-2017, but results for 2007-2012 are similar.[41] The Balassa index ranges from a miminum of 0.004 to a maximum of 7.18. This is illustrated in Figure 13 using a log scale for the ordered sectors by per cent rank. There are 14 strong sectors (1/3rd of the total) with a Balassa index above one, starting from SBI 32 (other manufactures) up to SBI 59 (movies and television), see Table 13 for all strong sectors. Equivalently, there are 28 weak sectors (two-thirds of the total), with SBI 96 (wellness) and SBI 74 (industrial design) as weakest sectors.

Table 13 Strong SBI sectors; BI Netherlands relative to World, 2012-2017

SBI	BI	# HS	Short name	SBI	BI	# HS	Short name
59	7.18	3	Movies & television	95	1.49	5	Repair computer
12	2.78	7	Tobacco	11	1.43	19	Beverages
62	2.58	10	ICT support	41	1.36	3	Construction
19	2.08	22	Refined oil	21	1.24	122	Pharmaceutical
10	2.01	537	Food products	58	1.15	15	Publishing
01	1.94	289	Agriculture	26	1.14	257	Computers
20	1.61	717	Chemicals	32	1.00	297	Oth manufactures

Source: authors; see section 2 for data info; SBI refers to sector number; BI is the Balassa Index; # HS refers to the number of 5-digit HS sub-sectors; see Appendix A for sector description.

People familiar with the Dutch economy will not be surprised to see that refined oil, food products, agriculture, and chemicals are among the strong export sectors. The relatively high score for movies & television (mainly because of television), tobacco, ICT support, computer

[41] All SBI sectors with $BI > 1$ in 2007-2012 also had $BI > 1$ in 2012-2017, while two marginally weak sectors in 2007-2012 (SBI 26 and 58, with $BI = 0.98$ and $BI = 0.97$, respectively) switched to strong in 2012-2017.

repair, and construction may be more surprising. People looking for the export of flowers in the list of strong sectors will be disappointed to see that it is not there. It is, actually, included in the *agriculture* sector, which combines no less than 289 HS products. Food products and Chemicals include even more HS sub-sectors (537 and 717, respectively). It is thus important to keep in mind that the concordance table is lop-sided in terms of the number of HS sub-sectors an SBI sector represents. It is for that reason that movies & television, tobacco, and ICT support can more easily score high Balassa indices as they represent a concentration of HS sub-sectors only (namely 3, 7, and 10, respectively). Similarly for the low-scoring SBI sectors in Figure 13 at the other extreme, where wellness consists of only 1 HS sub-sector and industrial design of only 3 HS sub-sectors. With these caveats in mind, it is now time to discuss the strong and weak sectors for Dutch cities relative to the world.

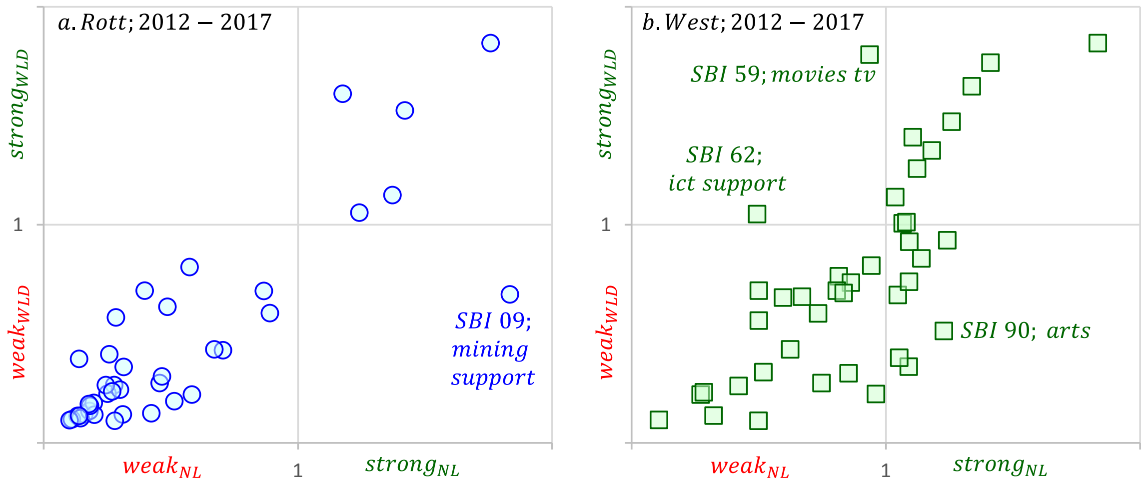
7.2 Dutch Cities relative to the World

As explained in section 4.3, we can calculate the comparative advantage of a Dutch city c relative to the world w for sector i in period t by multiplication: $BI_{i,t}^{c,w} = BI_{i,t}^{c,NL} \cdot BI_{i,t}^{NL,w}$. As a city for a sector can be either strong or weak relative to either the Netherlands and the World, we arrive at four logical possibilities identified by weak ($BI < 1$) and strong ($BI > 1$) with a sub-index NL relative to the Netherlands and WLD relative to the world. Table E1 in Appendix E provides a complete overview of all strong sectors for all locations. We follow Brakman and van Marrewijk (2017) by presenting the information in a table-like graph using a monotone transformation (to ensure all variables range from 0 to 2) that does not affect the weak or strong classification.[42] We say a 'reference switch' occurs if a sector either goes from weak to strong or from strong to weak. Table E2 in Appendix E provides a complete overview of all sectors that reference switch from weak to strong or from strong to weak for all locations. Figure 14 shows the classification for the locations with the lowest and highest number of reference switches.

Four cities, namely Eindhoven, Geleen-Sittard, Groningen, and Rotterdam have only one reference switch. In all cases it involves a reference switch from $strong_{NL}$ to $weak_{WLD}$. It is a different sector for each city. This is illustrated in panel *a* of Figure 14 for Rotterdam, which shows one observation (SBI 09, mining support) in the off-diagonal parts and all other observations in the diagonal parts. There are 34 sectors both $weak_{NL}$ and $weak_{WLD}$ and 5 sectors both $strong_{NL}$ and $strong_{WLD}$. For Rotterdam (and the other cities above) it thus does not really matter if the sector classification is relative to the world or relative to the Netherlands.

[42] More specifically, if $BI_i \leq 1$ the transformed variable is $0.1 + 0.9BI_i$, where the 0.1 avoids cluttering the diagram at the lower-left corner). If $BI_i > 1$ the transformed variable is $1 + ln(BI_i)/\left(1.2max_j\left(ln(BI_j)\right)\right)$, where the 1.2 avoids cluttering the diagram at the upper-right corner.

Figure 14 Strong and weak sectors, relative to NL and WLD; Rotterdam and West, 2012-17



Source: authors; see section 2 for data info; Rott = Rotterdam; see Appendix A for sector description; 40 sectors for Rotterdam and 41 sectors for West.

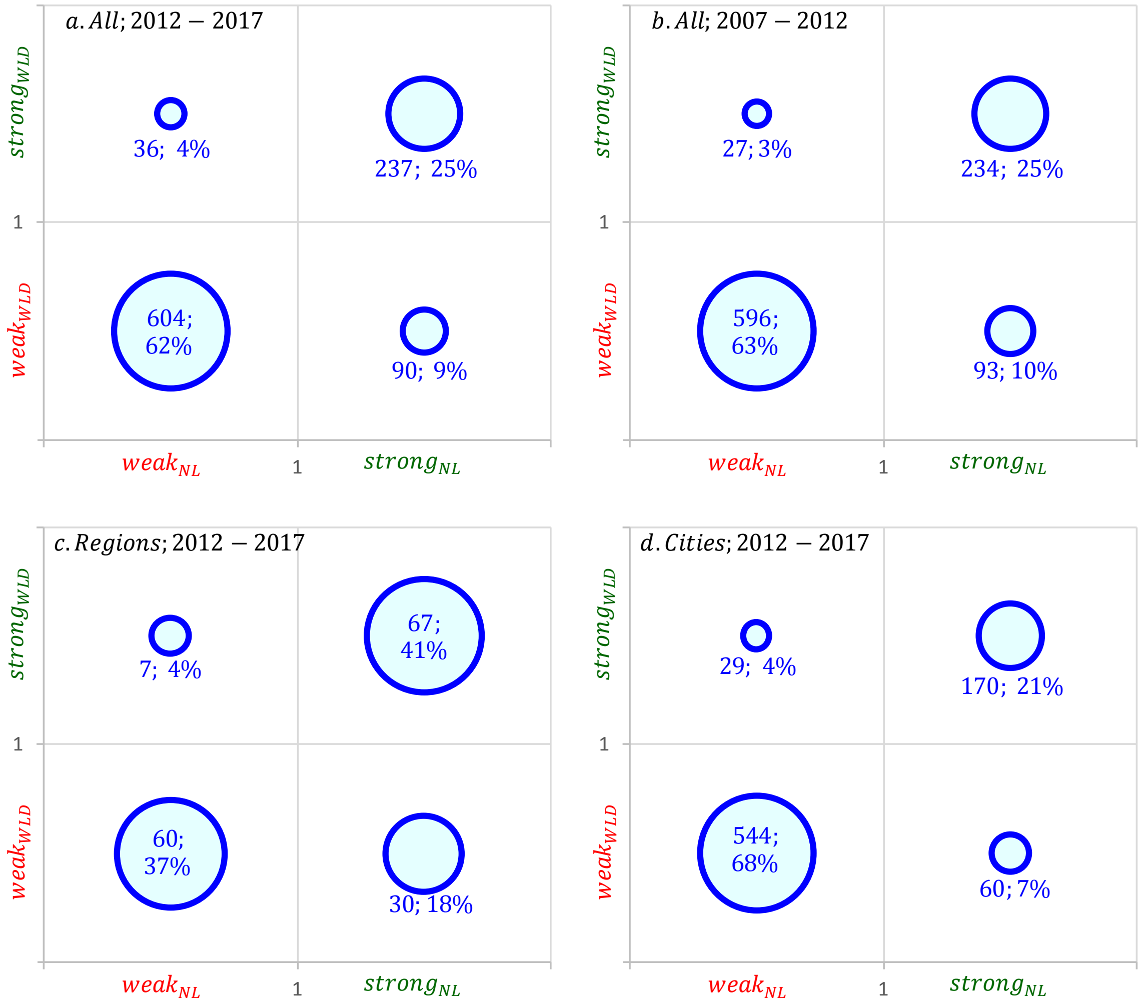
Two regions, namely West and South, have no less than 10 reference switches, most of which (namely 8 and 9, respectively) are from $strong_{NL}$ to $weak_{WLD}$. This is illustrated in panel *b* of Figure 14 for West, which shows two sectors (SBI 59 and 62) in the upper-left corner ($weak_{NL}$ and $strong_{WLD}$) and 8 sectors in the lower-right corner ($strong_{NL}$ and $weak_{WLD}$). There are 21 sectors both $weak_{NL}$ and $weak_{WLD}$ and 10 sectors both from $strong_{NL}$ and $strong_{WLD}$. Since a reference switch occurs for about one-quarter of all sectors for these two regions it *does* matter if the classification is relative to the world or relative to the Netherlands.

The two examples in Figure 14 are special because they show the lowest and highest number of reference switches. They are, however, good illustrations of two general principles.

- First, both panels suggest that the number of reference switches is higher from $strong_{NL}$ to $weak_{WLD}$ than the other way around.
- Second, the panels suggest that reference switches occur more frequently for regions than for cities.

Both suggestions are true, and they are also connected (as discussed in section 7.3). It implies that the comparison benchmark is more important for regions than for cities. We illustrate this in Figure 15 by showing the number of sectors (and per cent of the total) in each part of a panel with centered bubbles proportional to the number of sectors. Panel *a* shows this for all locations in 2012-2017, panel *b* for all locations in 2007-2012, panel *c* for all regions in 2012-2017, and panel *d* for all cities in 2012-2017.

Figure 15 Strong and weak sectors, relative to NL and WLD



Source: authors; see section 2 for data info; see Appendix A for sector description; 42 sectors; 4 locations for regions; 22 locations for cities; 26 locations for all; bubbles proportional to number of sectors in that panel; number of sectors and per cent of total listed in each panel.

As is clear from panel *a* in Figure 15, the majority of the observations (62 per cent) are in the $weak_{NL} - weak_{WLD}$ part of the diagram, followed by 25 per cent of the observations in the $strong_{NL} - strong_{WLD}$ part of the diagram. Most sectors (87 per cent of the total) therefore have the same classification relative to both the Netherlands and the world. Only 13 per cent of the sectors reference switches in classification, most of them from $strong_{NL}$ to $weak_{WLD}$ (9 per cent of the total), rather than the other way around (4 per cent of the total). This confirms the first principle above. Panel *b* shows that this finding is stable over time: there is at most one percentage point deviation in the 2007-2012 period compared to the 2012-2017 period.

Panels *c* and *d* of Figure 15 shows that there are differences between regional locations and city locations. Since there are 22 cities compared to 4 regions, the distribution for cities is more similar to the overall distribution, with deviations in the opposite direction for regions. We note that for regions there are more observations in the upper-right (strong-strong) corner and fewer observations in the lower-left (weak-weak) corner. Most importantly, there are more observations in relative terms (twice as many, namely 18 versus 9 per cent) in the lower-right ($strong_{NL}$ and $weak_{WLD}$) corner. The next section analyses this finding.

7.3 Explaining the reference switching suggestions

Our discussion of the reference switching suggestions first focuses on the relative abundance of reference switches from $strong_{NL}$ to $weak_{WLD}$ and then on the difference between regions and cities. Basic information regarding the reference switching sectors for Dutch cities is provided in Table 13, where the left part of the table focuses on reference switches from $strong_{NL}$ to $weak_{WLD}$ and the right part focuses on reference switches from $weak_{NL}$ to $strong_{WLD}$. In both cases, the top part provides summary information and the bottom part information by sector.

Table 14 Overview of strong-weak & weak-strong reference switches; by sector, 2012-2017

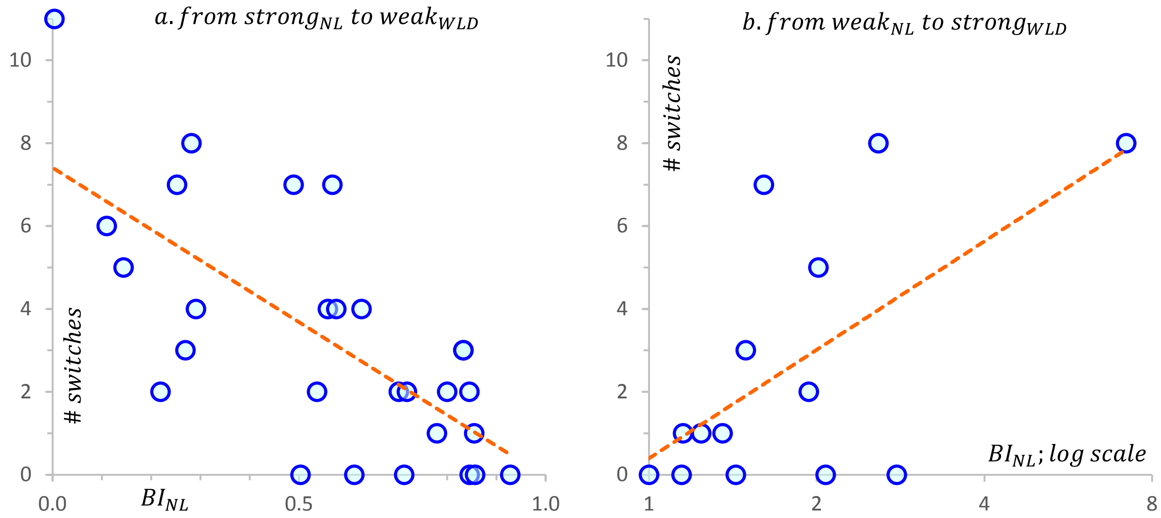
From $strong_{NL}$ to $weak_{WLD}$				From $weak_{NL}$ to $strong_{WLD}$			
# observations ($BI_i^{c,NL} > 1$)			327	# observations ($BI_i^{c,NL} < 1$)			640
# reference switches ($BI_i^{c,w} < 1$)			90	# reference switches ($BI_i^{c,w} > 1$)			36
Switches per cent of total			27.5	Switches per cent of total			5.6
SBI	$BI_i^{NL,w}$	Sector	# sw	SBI	$BI_i^{NL,w}$	Sector	# sw
96	0.004	Wellness	11	59	7.18	Movies & television	8
16	0.28	Wood	8	62	2.58	ICT support	8
14	0.57	Apparel	7	20	1.61	Chemicals	7
23	0.49	Non-metal mineral	7	10	2.01	Food products	5
90	0.25	Arts	7	95	1.49	Repair computer	3
74	0.11	Industrial design	6	01	1.94	Agriculture	2
08	0.14	Mining	5	21	1.24	Pharmaceutical	1
29	0.29	Cars	4	41	1.36	Construction	1
30	0.56	Oth transport	4	58	1.15	Publishing	1
38	0.63	Waste	4				
43	0.58	Spec construct	4				
09	0.27	Mine sup	3				
18	0.83	Printing	3				
35	0.83	Air cond	3				
02	0.70	Forestry	2				
06	0.22	Oil gas	2				
24	0.54	Basic metal	2				
25	0.72	Fabr metal	2				
27	0.85	Electrical eq	2				
28	0.80	Machinery	2				
22	0.85	Rubber plastic	1				
33	0.78	Repair machine	1				
Total number of reference switches			90	Total number of reference switches			36

Source: author calculations; see section 2 for data details; see Appendix A for sector description; # sw = number of reference switches

We already noted that there are more reference switches from $strong_{NL}$ to $weak_{WLD}$ than from $weak_{NL}$ to $strong_{WLD}$, namely 90 versus 36 reference switches. The top part of Table 11 indicates that this is actually remarkable, since there are many more observations with a Balassa index below one than above one (640 versus 327 observations). From this perspective, there is more potential for reference switches from weak to strong than the other way around, since there are many more possible reference switches. In contrast, the absolute number of reference switches is larger from strong to weak than the other way around (90 versus 36 reference switches), which makes the percent of reference switches actually occurring much higher from strong to weak than from weak to strong, namely 27.5 versus 5.6 per cent (see Table 11). One might argue that this makes the puzzle even more puzzling, but the bottom part of the table on sectors provides more insight.

SBI sector 96 (wellness) has the highest number of reference switches, namely at 11 locations from $strong_{NL}$ to $weak_{WLD}$. In fact, there are no locations anywhere in the Netherlands where wellness is strong relative to the world: the number of strong sectors in wellness thus reduces from 11 to 0! This is perfectly understandable since wellness is actually a very weak sector in the Netherlands as a whole, the Dutch Balassa index for wellnes relative to the world is only 0.004 (see Table 11) and as a consequence there is no location anywhere in the Netherlands where this is a strong sector. A similar observation holds for sector 74 (industrial design), where the number of strong locations drops from 6 to 0. The main point we should realize in this perspective is that if locations in the Netherlands are compared with the Netherlands as a whole, there must be *some* locations (at least one) that are relatively strong. From a world perspective this is no longer the case, so for some sectors the number of strong locations can become zero (or very small; please note that this logic does not work in the opposite direction). Also note that the majority of sectors at the Dutch level are weak rather than strong (28 out of 42 sectors), hence the relatively large number of strong to weak reference switches.

Figure 16 Dutch Balassa Index and number of reference switching sectors, 2012-2017



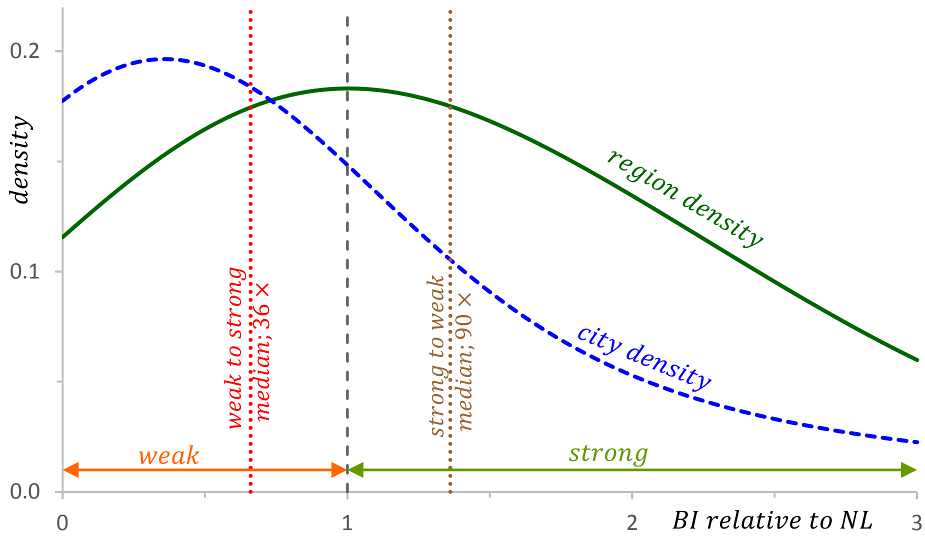
Source: authors; see section 2 for data info; horizontal axis depicts Balassa index for a sector for the Netherlands relative to the world; vertical axis depicts number of reference switching sectors at the city level, from strong to weak in panel a and from weak to strong in panel b; dashed lines are regression lines, slope is -7.46 in panel a and 3.80 in panel b, with 49 per cent of variance explained in panel a and 40 per cent in panel b.

It is worth pointing out an important observation related to the above discussion. At the national level, the Netherlands is weak relative to the world in 28 sectors and strong in 14 sectors. As we just noted, for 2 of the 28 weak sectors there is no location anywhere in the

Netherlands that is strong compared to the world. This also implies that for the other 26 weak sectors at the national level *there is at least one location* that is still strong compared to the world, even if the Netherlands as a whole is weak. There is thus enormous diversity within the Netherlands within the weak sectors even though the lens conditions is fulfilled (see section 3). Similarly, for *all* 14 sectors for which the Netherlands as a whole is strong relative to the world, there are *some* (usually many) locations within the Netherlands which are weak (already excluding locations without any activity in this sector). There is thus also enormous diversity within the Netherlands for strong sectors.

As we can see from Table 13, there are many reference switches from strong to weak in sectors with a very low Balassa index at the Dutch level (like 96, 16, 19, 74, and 08), while there are many reference switches from weak to strong in sectors where the Balassa index is high at the Dutch level (like sectors 59, 62, 20, and 10). This is, of course, as we would expect since a Dutch Balassa index below one is necessary to make a strong to weak reference switch possible and a Dutch Balassa index above one is necessary to make a weak to strong reference switch possible. This is illustrated in Figure 16, which shows the number of reference switches within a sector from strong to weak relative to the Balassa index below one in panel *a* and from weak to strong relative to the Balassa index above one in panel *b* (using a log scale). The (statistically significant) slope of the regression line is -7.46 in panel *a* and 3.80 in panel *b*, where panel *a* explains 49 per cent of the variance and panel *b* 40 per cent. Evidently, the weaker a weak sector is or the stronger a strong sector is within the Netherlands relative to the world, the more likely a reference switch may occur at the city level.

Figure 17 Balassa index kernel density relative to the Netherlands, 2012-2017



Source: authors; see section 2 for data info; median $weak_{NL}$ to $strong_{WLD}$ is 0.66; median $strong_{NL}$ to $weak_{WLD}$ is 1.36; 164 observations for regions and 803 for cities; kernel density based on normal distribution with bandwidth $(4/3)^{1/5}\hat{\sigma}n^{-1/5}$, where $\hat{\sigma}$ is estimated standard deviation and n is number of observations.

We now focus on explaining the differences in reference switches for regions and cities as illustrated in panels *c* and *d* of Figure 15, as well as the differences in weak-weak and strong-strong between regions and cities. To do so, Figure 17 provides a kernel density distribution of the Balassa index at the location level relative to the Netherlands for both regions and cities.[43] It clearly shows that the majority of the observations at the city level have a Balassa index *below* one, namely 71 per cent of the observations. In contrast, a large fraction of the observations at the regional level have a Balassa index *above* one, namely 59 per cent of the observations. This explains why the share of observations in the strong-strong quadrant is much higher for regions than for cities, and vice versa for the share of observations in the weak-weak quadrant.

Figure 17 also helps us to understand why the frequency of reference switching from weak to strong is similar for cities and regions (about 4 per cent, see Figure 15) and why reference switching from strong to weak is much more frequent for regions than for cities (18 versus 7 per cent, see Figure 15). To do so, Figure 17 also provides the median values for reference switchin from $weak_{NL}$ to $strong_{WLD}$ (about 0.66) and for reference switching from $strong_{NL}$ to $weak_{WLD}$ (about 1.36). It is clear from Figure 17 that the density around the median weak-to-strong reference switch value of 0.66 is very similar for regions and cities (hence reference switching occurs for about 4 per cent of all observations for both cities and regions). In contrast, the density is much higher for regions than it is for cities around the median strong-to-weak reference switch value of 1.36 (hence reference switching occurs for about 18 per cent of all observations for regions compared to 7 per cent for cities).

8 Conclusions

In this paper we drop the standard assumption in the international trade literature that assumes that countries are dimensionless points. The starting position is that the spatial distribution of firms and fsctors of production affect the structure of trade. As shown by Courant and Deardorff (1992, 1993) in a Heckscher-Ohlin setting, lumpiness, or spatially *uneven* distribution of production factors within a country, can affect the national pattern of trade. A striking characteristic of the world economy is, indeed, that economic activity and factors of production are unevenly distributed over space, not only between countries but also within countries. This potentially affects trade.

Using micro-firm data we study local trade patterns in 22 cities and 4 regions for 83 sectors for the period 2007-2017 and determine characteristics of local trade patterns. We start by focusing on the relationship between the national structure of trade and the regional structure of trade and subsequently zoom in on smaller spatial units to discuss city-region trade patterns and determine what local characteristics explains these patterns.

Out findings are the following. First, applying the lens condition as developed by Courant and Deardorff (1992, 1993) shows that the regional distribution of production factors is unlikely to affect the Dutch structure of trade as a whole. This is an important conclusion; from the perspective of the lens-condition the spatial regional distribution of firms and factors of production is consistent with the welfare maximizing integrated equilibrium. Second, using local micro firm data we establish city-region specific distributions of Revealed

[43] For this exercise, therefore, all observations at the regional level are combined and all observations at the city level are combined to generate the kernel density plots.

Comparative Advantage. The measure we use is the Balassa Index (BI). For each city- region we calculate the distribution of BIs for all sectors that are active in that location, and identify sectors with a comparative (dis) advantage relative to the Netherlands and relative to the world. These distributions differ significantly from each other, illustrating that comparative advantage has a local origin. Third, we find that the interaction of local factor abundance and sector skill intensity systematically explains the local trade patterns. Control variables such as local market access, density, or the distance to international airports or ports have limited explanatory power. Finally, we identify sectors that have a comparative advantage relative to the Netherlands and the world, these are the sectors on which Dutch exports rely. We also identify sectors that have a comparative advantage relative to the Netherlands but not relative to the world (and vice versa).

From a policy point of view our results indicate that: the mobility of production factors is such that it is consistent with the welfare maximizing integrated equilibrium, and that international trade shocks, such as the Brexit or the US-China trade disputes can have strong local consequences.

References

Bernard, A.B., R. Robertson, and P.K. Schott, 2010. Is Mexico a Lumpy Country? *Review of International Economics 18,* 937-950.

Bernard, A.B., J. B. Jensen, S. J. Redding, and P. K. Schott, 2012, The Empirics of Firm Heterogeneity and International Trade. *Annual Review of Economics, Vol. 4,* 283-313.

Brakman, S., H. Garretsen, and C. van Marrewijk, 2020. *An Introduction to Geographical and Urban Economics: A Spiky World*. Cambridge University Press, Cambridge, U.K.

Brakman, S., and C. van Marrewijk, 2017. A closer look at revealed comparative advantage: gross- versus value-added trade flows. *Papers in Regional Science 96(1),:* 61-92.

Brakman, S., and C. Van Marrewijk, 2013. Lumpy countries, urbanization, and trade. *Journal of International Economics*, *89*(1), 252-261.

Courant, P.N., and A. Deardorff, 1992. International Trade with Lumpy Countries. *Journal of Political Economy 100*, 198-210.

Courant, P.N., and A. Deardorff, 1993. Amenities, Nontraded goods, and the Trade of Lumpy Countries. *Journal of Urban Economics* 34, 299-317.

Davis, D. R. and Dingel, J. I. (2020). The Comparative Advantage of Cities, *Journal of International Economics* 123: 1-27.

Deardorff, A.V., 2011. Comparative advantage: the theory behind measurement. In: Kowalski, P., Stone, S.F. (Eds.), *Globalisation, Comparative Advantage and the Changing Dynamics of Trade*, OECD Publishing.

Deardorff, A.V., 1994. The Possibility of Factor Price Equalization, Revisited. *Journal of International Economics* 36, 167-175.

Debaere, P., 2004. Does Lumpiness Matter in an Open Economy? Studying International Economics with Regional Data. *Journal of International Economics* 64, 485-501.

Debaere, P., and U. Demiroglu, 2003. On the Similarity of Country Endowments. *Journal of International Economics 59*, 101-136.

Dixit, A., and V. Norman, 1980. *The theory of International Trade*. Cambridge University Press, Cambridge.

Glaeser, E.L., and M.G. Resseger, 2010. The Complementarity Between Cities and Skills. *Journal of Regional Science 50*, 221-244.

Hillman, A.L. (1980). Observations on the relation between 'revealed comparative advantage'
and comparative advantage as indicated by pre-trade relative prices. *Review of World Economics* 116, 315 – 21.

Hinloopen, J., and C. van Marrewijk, 2001. On the empirical distribution of the Balassa Index. *Review of World Economics 137*, 1–35.

Hinloopen, J., and C. van Marrewijk, 2005. Comparing distributions: the Harmonic Mass index. *Tinbergen Institute, Discussion Paper 05–122/1*.

Hinloopen, J., and C. van Marrewijk, 2008. Empirical relevance of the Hillman condition for revealed comparative advantage: 10 stylized facts. *Applied Economics* 40, 2313–2328.

Hinloopen, J., R. Wagenvoort, and C. van Marrewijk, 2012. A k-sample homogeneity test: the Harmonic Weighted Mass index. *International Econometric Review* 4(1), 17–39.

Kowalski, P., and N. Bottini, 2011. Comparative advantage and export specialisation mobility. in: Kowalski, P., Stone, S.F. (Eds.), G*lobalisation, Comparative Advantage and the Changing Dynamics of Trade*, OECD Publishing.

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

Leamer, E.E, 2012. *The Craft of Economics: Lessons from the Heckscher-Ohlin Framework*. The MIT Press, Cambridge, Mass.

Rosenthal, S.S., and W.C. Strange, 2004. Evidence on the Nature and Sources of Agglomeration Economies. In: J.V.Henderson and J-F. Thisse, *Handbook of Regional and Urban Economics*, Elsevier, North-Holland, Amsterdam.

Appendix A. Overview of Dutch Sectors

Code	Description *(brief reference name)*
A	*Agriculture, forestry and fishing (agriculture)*
01	Agriculture and related service activities
02	Forestry and logging
03	Fishing and aquaculture
B	*Mining and quarrying (mining)*
06	Extraction of crude petroleum and natural gas
08	Mining and quarrying (no oil and gas)
09	Mining support activities
C	*Manufacturing (manufacturing)*
10	Manufacture of food products
11	Manufacture of beverages
12	Manufacture of tobacco products
13	Manufacture of textiles
14	Manufacture of wearing apparel
15	Manufacture of leather, products of leather and footwear
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
17	Manufacture of paper and paper products
18	Printing and reproduction of recorded media
19	Manufacture of coke and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22	Manufacture of rubber and plastic products
23	Manufacture of other non-metallic mineral products
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computers, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29	Manufacture of motor vehicles, trailers and semi-trailers
30	Manufacture of other transport equipment
31	Manufacture of furniture
32	Manufacture of other products n.e.c.
33	Repair and installation of machinery and equipment
D	*Electricity, gas, steam and air conditioning supply (electricity)*
35	Electricity, gas, steam and air conditioning supply
E	*Water supply (water)*
36	Collection, purification and distribution of water
37	Sewerage
38	Waste collection, treatment and disposal activities; materials recovery
39	Remediation activities and other waste management

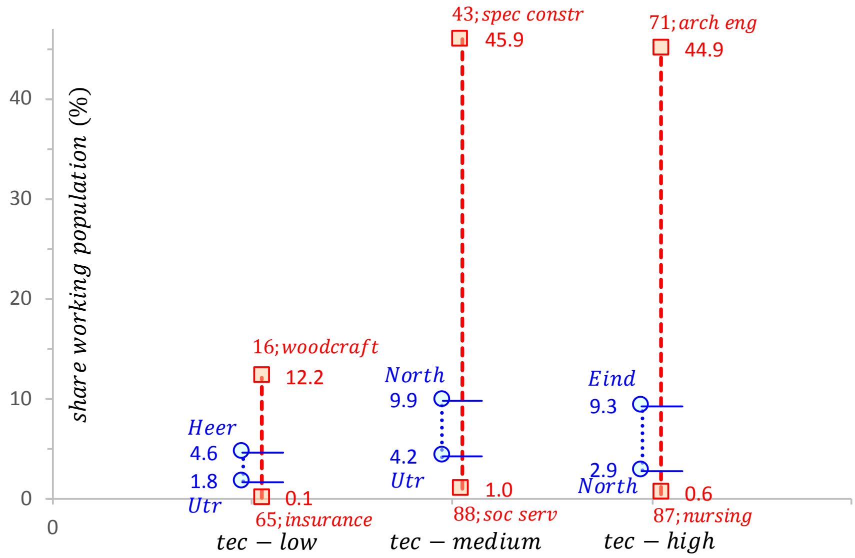
Code	Description *(brief reference name)*
F	*Construction (construction)*
41	Construction of buildings and development of building projects
42	Civil engineering
43	Specialised construction activities
G	*Wholesale and retail trade (wholesale)*
45	Sale and repair of motor vehicles, motorcycles and trailers
46	Wholesale trade (no motor vehicles and motorcycles)
47	Retail trade (not in motor vehicles)
H	*Transportation and storage (transportation)*
49	Land transport
50	Water transport
51	Air transport
52	Warehousing and support activities for transportation
53	Postal and courier activities
I	*Accommodation and food service activities (accommodation)*
55	Accommodation
56	Food and beverage service activities
J	*Information and communication (info-com)*
58	Publishing
59	Motion picture and television programme production and distribution; sound recording and music publishing
60	Programming and broadcasting
61	Telecommunications
62	Support activities in the field of information technology
63	Information service activities
K	*Financial institutions (fin-inst)*
64	Financial institutions, except insurance and pension funding
65	Insurance and pension funding (no compulsory social security)
66	Other financial services
L	*Renting, buying and selling of real estate (real estate)*
68	Renting and buying and selling of real estate
M	*Consultancy, research and other specialised business services (consultancy)*
69	Legal services, accounting, tax consultancy, administration
70	Holding companies (not financial)
71	Architects, engineers and technical design and consultancy; testing and analysis
72	Research and development
73	Advertising and market research
74	Industrial design, photography, translation and other consultancy
75	Veterinary activities
N	*Renting and leasing of tangible goods and other business support services (rent-lease)*
77	Renting and leasing of motor vehicles, consumer goods, machines, and other tangible goods
78	Employment placement, provision of temporary employment and payrolling
79	Travel agencies, tour operators, tourist information and reservation services

Code	Description *(brief reference name)*
80	Security and investigation
81	Facility management
82	Other business services
O	*Public administration, public services and compulsory social security (public admin)*
84	Public administration, public services and compulsory social security
P	*Education (education)*
85	Education
Q	*Human health and social work activities (health)*
86	Human health activities
87	Residential care and guidance
88	Social work activities without accommodation
R	*Culture, sports and recreation (culture)*
90	Arts
91	Lending of cultural goods, public archives, museums, botanical and zoological gardens and nature reserves activities
92	Lotteries and betting
93	Sports and recreation
S	*Other service activities (other services)*
94	World view and political organizations, interest and ideological organizations, hobby clubs
95	Repair of computers and consumer goods
96	Wellness and other services; funeral activities
T	*Household activities (household)*
97	Activities of households as employers of domestic personnel
98	Undifferentiated goods- and services-producing activities of private households for own use
U	*Extraterritorial organisations and bodies (extraterritorial)*
99	Extraterritorial organisations and bodies

Appendix B. Technical and Non-Technical Skill Ranges

Figure B1 provides the range of shares of working population in per cent for low, medium, and high technical skill levels for locations and sectors. For low technical skills, the cities-regions range is from 1.8 per cent in Utrecht to 4.6 per cent in Heerlen, which is strictly within the sector range from 0.1 per cent for insurance to 12.2 per cent for woodcraft. Similarly for medium and high levels of technical skill. On average, the cities-regions range is 17 per cent of the sector range, which is thus much wider.

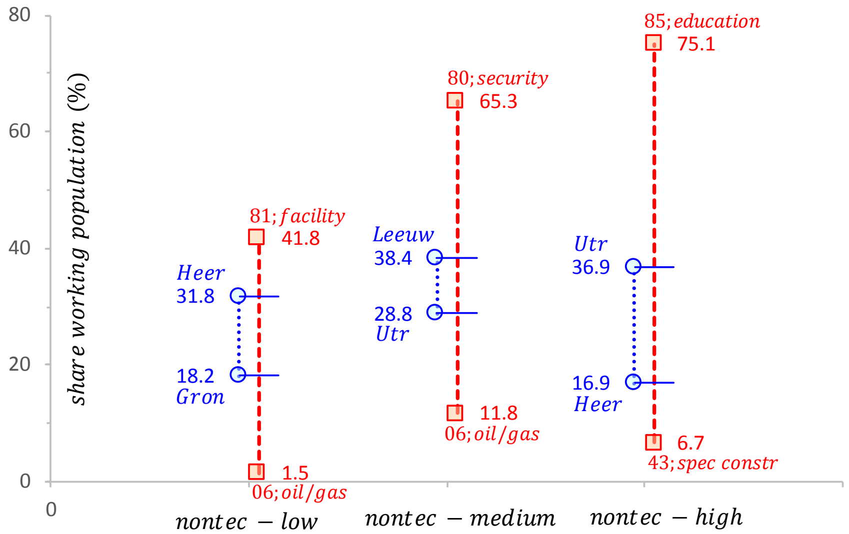
Figure B1 Technical Skill Ranges for Dutch Cities-Regions and Sectors, 2017



Source: authors; see section 2 for data sources; see Appendix A for sector names; Heer = Heerlen; Utr = Utrecht; Eind = Eindhoven.

Figure B2 provides the range of shares of working population for low, medium, and high non-technical skill levels for locations and sectors. The conclusion is the same: the cities-regions range is a strict subset of the sector range (covering 27 per cent on average in this case).

Figure B2 Non-Technical Skill Ranges for Dutch Cities and Sectors, 2017

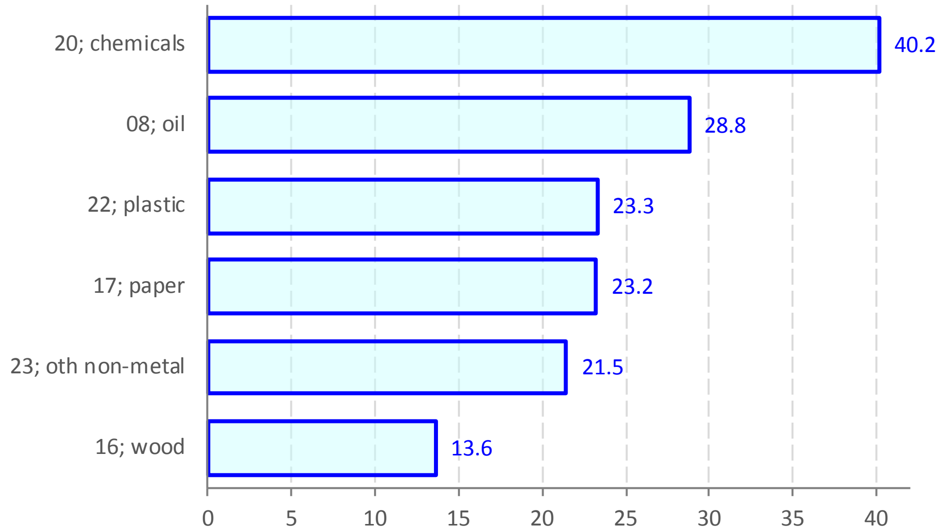


Source: authors; see section 2 for data sources; see Appendix A for sector names; Heer = Heerlen; Utr = Utrecht; Gron = Groningen; Leeuw = Leeuwarden.

Appendix C. Imposed Factor Intensity Example

Using Groningen Growth and Development Center (GGDC) data, Brakman and van Marrewijk (2013) are forced to impose the same factor intensities for many sectors. An important example is for Intermediate Manufacturing, where six sectors are imposed to have the same factor intensity of high versus low-med skill. Translated to our data set, these are sectors 08, 16, 17, 20, 22, and 23. According to our much more detailed information, these sectors have drastically different factor intensities. Figure A3 illustrates this for the share of high skill workers, which ranges from 13.6 per cent for sector 16 (wood) to 40.2 per cent for sector 20 (chemicals), which is about three times higher!

Figure A3 High Skill shares (% of working population) for six sectors



Source: authors; see section 2 for data sources; see Appendix A for sector names;

Appendix D. Comparison of Dutch BI Distributions between Locations

Balassa index 2007-12 distribution					Balassa index 2012-17 distribution				
Area	# similar	%	# different	%	Area	# similar	%	# different	%
North	0	0	25	100	North	0	0	25	100
East	1	4	24	96	East	1	4	24	96
West	0	0	25	100	West	0	0	25	100
South	1	4	24	96	South	1	4	24	96
Gron	2	8	23	92	Gron	0	0	25	100
Leeuw	3	12	22	88	Leeuw	3	12	22	88
Zwol	1	4	24	96	Zwol	1	4	24	96
Ensch	6	24	19	76	Ensch	3	12	22	88
Apel	2	8	23	92	Apel	3	12	22	88
Arnh	4	16	21	84	Arnh	5	20	20	80
Nijm	8	32	17	68	Nijm	4	16	21	84
Amer	7	28	18	72	Amer	1	4	24	96
Utr	3	12	22	88	Utr	4	16	21	84
Amst	3	12	22	88	Amst	3	12	22	88
Haar	6	24	19	76	Haar	5	20	20	80
Leid	7	28	18	72	Leid	5	20	20	80
Haag	5	20	20	80	Haag	6	24	19	76
Rott	0	0	25	100	Rott	4	16	21	84
Dord	2	8	23	92	Dord	2	8	23	92
Breda	4	16	21	84	Breda	1	4	24	96
Tilb	4	16	21	84	Tilb	6	24	19	76
Bosch	7	28	18	72	Bosch	2	8	23	92
Eind	0	0	25	100	Eind	2	8	23	92
GelSit	2	8	23	92	GelSit	6	24	19	76
Heer	2	8	23	92	Heer	1	4	24	96
Maas	4	16	21	84	Maas	5	20	20	80
average	3.2	13	21.8	87	average	2.8	11	22.2	89

Source: author calculations, based on critical values of the 2-sample HWM, 10 per cent significance.

Appendix E. Overview of Strong and Reference switching Sectors

Table E1 Overview of strong SBI sectors, 2012-2017

Location	SBI sector $strong_{NL}$	SBI sector $strong_{WLD}$
North	06; 08; 09; 10; 16; 17; 18; 20; 21; 22; 23; 24; 25; 28; 30; 31; 32; 33; 35; 38; 41; 43; 90; 96	01; 06; 08; 10; 16; 17; 20; 21; 22; 23; 24; 28; 30; 31; 32; 33; 35; 41; 43; 59; 62
East	01; 02; 10; 12; 13; 14; 15; 16; 17; 18; 19; 20; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 38; 43; 95; 96	01; 02; 10; 12; 13; 15; 17; 19; 20; 22; 24; 25; 26; 27; 28; 30; 31; 32; 38; 59; 95
West	01; 03; 06; 10; 11; 14; 16; 19; 20; 21; 22; 23; 25; 30; 35; 41; 43; 90	01; 03; 10; 11; 19; 20; 21; 22; 35; 41; 59; 62
South	01; 02; 08; 10; 11; 12; 13; 14; 15; 16; 17; 18; 19; 21; 22; 23; 24; 25; 28; 29; 30; 31; 32; 33; 38; 43; 74; 95	01; 02; 10; 11; 12; 13; 14; 15; 17; 18; 19; 20; 21; 22; 24; 25; 31; 32; 33; 95
Groningen	06; 10; 11; 12; 17; 18; 22; 23; 25; 26; 31; 32; 33; 35; 38; 43; 58; 62; 90	06; 10; 11; 12; 17; 18; 22; 23; 25; 26; 31; 32; 33; 35; 38; 43; 58; 62
Leeuwarden	01; 02; 10; 17; 22; 23; 25; 30; 32; 38; 58; 90; 95; 96	01; 10; 17; 20; 22; 25; 32; 38; 58; 62; 95
Zwolle	10; 14; 31; 33; 74	10; 31; 33
Enschede	09; 13; 14; 16; 18; 22; 26; 27; 28; 30; 31; 32; 35; 38; 43; 74; 95	10; 13; 14; 22; 26; 27; 28; 30; 31; 32; 38; 95
Apeldoorn	01; 10; 16; 17; 21; 22; 23; 24; 25; 28; 96	01; 10; 17; 21; 22; 23; 24; 25; 28; 62
Arnhem	01; 13; 14; 16; 20; 23; 26; 27; 28; 31; 33; 35; 41; 58; 96	01; 10; 13; 16; 20; 26; 31; 33; 35; 41; 58; 59; 62
Nijmegen	08; 10; 16; 23; 26; 27; 28; 29; 38; 95	10; 26; 27; 28; 95
Amersfoort	10; 18; 20; 22; 24; 25; 26; 27; 30; 31; 32; 33; 58; 90; 96	10; 18; 20; 25; 26; 27; 31; 32; 33; 58; 59; 62
Utrecht	21; 27; 32; 43; 58; 62	10; 20; 21; 27; 32; 43; 58; 59; 62
Amsterdam	14; 29; 35; 58; 59; 62; 74; 79; 90	58; 59; 62; 79; 95
Haarlem	03; 09; 14; 17; 18; 35	03; 09; 10; 17; 18; 20; 41; 59
Leiden	01; 21; 30; 32; 35; 58; 74; 95	01; 21; 30; 32; 35; 58; 62; 95
Den Haag	01; 06; 09; 18; 26; 32; 41; 43; 62; 90; 95	01; 09; 18; 21; 26; 32; 41; 43; 62; 95
Rotterdam	09; 11; 19; 20; 27; 35	11; 19; 20; 27; 35
Dordrecht	08; 10; 19; 24; 26; 30; 33; 38; 41; 43	10; 19; 20; 24; 26; 30; 33; 41; 43; 95
Breda	01; 10; 14; 22; 23; 26; 33; 59; 62; 96	01; 10; 14; 20; 22; 26; 33; 59; 62
Tilburg	02; 10; 11; 13; 14; 15; 16; 22; 23; 25; 29; 31; 32; 90; 95; 96	10; 11; 13; 14; 15; 22; 23; 25; 29; 31; 32; 58; 90; 95
Den Bosch	14; 26; 33; 41; 62; 74; 96	01; 10; 20; 26; 41; 62
Eindhoven	15; 26; 27; 28; 62; 95	15; 26; 28; 62; 95
Geleen-Sittard	10; 11; 20; 23; 29; 32; 33; 35; 38; 96	10; 11; 20; 23; 29; 32; 33; 35; 38
Heerlen	02; 08; 11; 16; 17; 20; 22; 24; 25; 32; 95; 96	02; 11; 17; 20; 22; 24; 25; 32; 62; 95
Maastricht	08; 17; 20; 23; 24; 38; 41	17; 20; 23; 38; 41; 59; 95

Source: author calculations, see section 2 for data details; see Appendix A for sector description.

Table E2 Overview of reference switching sectors, 2012-2017

Location	SBI $weak_{NL}$ & $strong_{WLD}$	SBI $strong_{NL}$ & $weak_{WLD}$
North	01; 59; 62	09; 18; 25; 38; 90; 96
East	59	14; 16; 18; 23; 29; 43; 96
West	59; 62	06; 14; 16; 23; 25; 30; 43; 90
South	20	08; 16; 23; 28; 29; 30; 38; 43; 74
Groningen		90
Leeuwarden	20; 62	02; 23; 30; 90; 96
Zwolle		14; 74
Enschede	10	09; 16; 18; 35; 43; 74
Apeldoorn	62	16; 96
Arnhem	10; 59; 62	14; 23; 27; 28; 96
Nijmegen		08; 16; 23; 29; 38
Amersfoort	59; 62	22; 24; 30; 90; 96
Utrecht	10; 20; 59	
Amsterdam	95	14; 29; 35; 74; 90
Haarlem	10; 20; 41; 59	14; 35
Leiden	62	74
Den Haag	21	06; 90
Rotterdam		09
Dordrecht	20; 95	08; 38
Breda	20	23; 96
Tilburg	58	02; 16; 96
Den Bosch	01; 10; 20	14; 33; 74; 96
Eindhoven		27
Geleen-Sittard		96
Heerlen	62	08; 16; 96
Maastricht	59; 95	08; 24

Source: author calculations, see section 2 for data details; see Appendix A for sector description.

Appendix F. Correlation and Additional Controls

Correlation matrix, 2012-2017

	1	2	3	4	5	6	7	8	9
1 Market access	1.000								
2 Density	0.382	1.000							
3 Schiphol road	-0.817	-0.548	1.000						
4 Port road	-0.863	-0.654	0.818	1.000					
5 $Abun^c_{low} \times Int^i_{low}$	-0.076	0.027	-0.120	0.090	1.000				
6 $Abun^c_{med} \times Int^i_{med}$	-0.204	-0.280	0.257	0.226	0.014	1.000			
7 $Abun^c_{high} \times Int^i_{high}$	0.077	0.141	-0.161	-0.083	0.734	0.498	1.000		
8 $Abun^c_{tec-low} \times Int^i_{tec-low}$	-0.047	-0.121	0.155	0.068	0.124	0.375	0.390	1.000	
9 $Abun^c_{tec-med} \times Int^i_{tec-med}$	-0.120	-0.311	0.275	0.196	-0.173	0.427	0.099	0.763	1.000
10 $Abun^c_{tec-high} \times Int^i_{tec-high}$	0.130	0.231	-0.156	-0.140	0.277	0.165	0.399	-0.048	-0.227

Source: author calculations; see section 2 for data details; shaded cells above 0.7 in absolute value; 1886 obs.

Additional controls

Fa. Rank analysis of Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun^c_{high} \times Int^i_{high}$	-258.0***					
	(4.77e-10)					
$Abun^c_{med} \times Int^i_{med}$		-776.6***				
		(2.64e-08)				
$Abun^c_{low} \times Int^i_{low}$			-438.1***			
			(0.000205)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				-772.2**		
				(0.0175)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					-1,767***	
					(0)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						-13,312***
						(0)
Density	0.00196	8.32e-05	0.00108	0.00183	-0.000460	0.000764
	(0.318)	(0.966)	(0.585)	(0.357)	(0.815)	(0.693)
Market access	3.346***	3.145***	2.725**	3.362***	3.684***	3.712***
	(0.00225)	(0.00529)	(0.0137)	(0.00253)	(0.000586)	(0.000466)
Schiphol road	-0.0220	0.0271	-0.0374	0.00869	0.0478	0.0306
	(0.636)	(0.566)	(0.437)	(0.856)	(0.292)	(0.489)
Port road	0.0620	0.0320	0.0612	0.0424	0.0301	0.0415
	(0.215)	(0.525)	(0.228)	(0.408)	(0.533)	(0.384)
R-squared	0.175	0.171	0.163	0.159	0.181	0.191

Source: author calculations; see section 2 for data details; robust pvalues in parentheses; *** p<0.01; ** p<0.05; * p<0.1; all regressions have 1886 observations and include sector fixed effects and city-region fixed effects.

Additional controls

Fb. Probit analysis of Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun^c_{high} \times Int^i_{high}$	15.89***					
	(2.42e-08)					
$Abun^c_{med} \times Int^i_{med}$		42.37***				
		(1.32e-05)				
$Abun^c_{low} \times Int^i_{low}$			31.22***			
			(0.000276)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				52.64**		
				(0.0211)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					121.5***	
					(0)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						890.9***
						(0)
Density	0.000470	8.91e-05	0.00111	0.000214	-0.000366	-0.000448
	(0.462)	(0.886)	(0.106)	(0.728)	(0.570)	(0.520)
Market access	2.167	1.653	3.036*	1.836	0.0650	1.245
	(0.219)	(0.347)	(0.0925)	(0.296)	(0.972)	(0.492)
Schiphol road	0.130	0.102	0.180*	0.114	0.0106	0.0792
	(0.219)	(0.334)	(0.0962)	(0.280)	(0.923)	(0.467)
Port road	-0.0359	-0.0313	-0.0484	-0.0328	-0.00289	-0.0272
	(0.370)	(0.432)	(0.230)	(0.411)	(0.944)	(0.513)

Fc. Analysis of natural logarithm Balassa Index, 2012-2017

Variable	(1)	(2)	(3)	(4)	(5)	(6)
$Abun^c_{high} \times Int^i_{high}$	20.72***					
	(3.88e-09)					
$Abun^c_{med} \times Int^i_{med}$		60.05***				
		(8.06e-08)				
$Abun^c_{low} \times Int^i_{low}$			33.84***			
			(0.00247)			
$Abun^c_{tec-high} \times Int^i_{tec-high}$				41.29		
				(0.305)		
$Abun^c_{tec-med} \times Int^i_{tec-med}$					104.9***	
					(2.93e-08)	
$Abun^c_{tec-low} \times Int^i_{tec-low}$						845.2***
						(1.60e-10)
Density	2.92e-05	0.000176	9.97e-05	5.26e-05	0.000186	0.000116
	(0.840)	(0.224)	(0.491)	(0.722)	(0.202)	(0.420)
Market access	0.406***	0.421***	0.454***	0.404***	0.385***	0.382***
	(0.000525)	(0.000388)	(0.000126)	(0.000631)	(0.000911)	(0.00101)
Schiphol road	0.00863**	0.00476	0.00974***	0.00626*	0.00392	0.00484
	(0.0112)	(0.168)	(0.00578)	(0.0734)	(0.254)	(0.151)
Port road	0.00221	0.00458	0.00232	0.00364	0.00440	0.00376
	(0.550)	(0.217)	(0.535)	(0.333)	(0.228)	(0.298)
R-squared	0.241	0.236	0.231	0.227	0.237	0.243

Source: author calculations; see section 2 for data details; robust pvalues in parentheses; *** p<0.01; ** p<0.05; * p<0.1; all regressions have 1886 observations and include sector fixed effects and city-region fixed effects.