Assessing the Impact of Agricultural Transport Costs on Industrial Agglomeration[*]

Pedro Herrera Catalán[‡]

London School of Economics

November 7, 2014

Abstract

Regarding equilibrium in industrial location, the home market effect (HME) foresees that larger regions within a country host a disproportionate share of the manufacturing sectors, and are net exporters of industrial goods. This prediction however has been arduously called into question in the literature when the transport costs of the agricultural goods are positive, and no consensus has been reached. This paper examines the impact of agricultural transport costs upon the agriculture-related firms' incentives to disperse or to agglomerate in Peru. Using distance-based tests of industry localisation developed by Duranton and Overman (2005, RES, 72,1077-1106), we find that agriculture-related industries do not exhibit a strong localisation patterns. Only 1 out of 13 industries in our sample is highly localised at short distances whereas the remaining industries are dispersed or are localised at long distances with a significantly weaker degree of intensity. Conditional logit models indicate that a high localisation at short (long) distances is explained by low (high) agricultural transport costs. Weak localisation at long distances and dispersion patterns are explained by the interplay of high and low transportation costs in the agricultural sector. These results shed light on the Economic Geography literature-related controversy concerning the role of agricultural transport costs in core-periphery models.

Keywords: Agricultural transport costs; home market effect, distance-based methods; micro-geographic data; Economic Geography

JEL Classification: R11, R12

To be presented at 61st NARSC Congress, Washington, D.C., EEUU, November 12-15, 2014

---

[*] I am especially grateful to Steve Gibbons, Henry Overman and Gabriel Ahlfeldt for insightful and valuable comments on earlier drafts. I am also grateful to participants at the 54th European Congress of the Regional Science Association International for their comments and suggestions.

[‡] Department of Geography & Environment, London School of Economics, Houghton Street London WC2A 2AE, United Kingdom. E-mail address: P.P.Herrera@lse.ac.uk

Contents

List of Tables

List of Figures

1. Introduction

The study of the impact of transport costs and barriers to trade on the agglomeration process of manufacturing sector have habitually been the main concern of the Economic Geography literature. Unfortunately, the agricultural sector has comparatively received little attention regarding its impact on industry localisation. According to Takatsuka and Zeng (2012) this is mainly due to the simplifying assumptions of formulation models, which usually consider the agricultural goods to be homogenous and transportable at zero cost in order to offset the trade imbalance in the manufacturing sector. However, this assumption is not justifiable in the real world for two reasons. First, world trade is not balanced in the manufacturing sector (Dekle et al., 2007). Second, agricultural transport costs play a key role when defining the spatial configuration of economic activity, especially in developing countries where an important part of their inhabitants is devoted to agricultural activities (Picard and Zeng, 2005).

A few Economic Geographers have researched the role of the transport costs of the agricultural good in core-periphery models. Nevertheless, no consensus had emerged about the impact of agricultural transport costs on agglomeration process. The debate is rooted in Helpman and Krugman's (1985) work, which coined the term 'home-market effect' (HME) to refer to the case when manufacturing firms prefer producing in larger markets to save transport costs. These authors however only included into their analysis the transport costs for the manufacturing sector, resulting in a continuing academic dispute about the real effect of agricultural transport costs in Economic Geography models. First, Fujita et al. (1999) find that an increase in agricultural transport costs causes dispersion as intensely as a rise in manufacturing transport costs. Second, Davis (1998) includes trade costs for manufactured and agricultural goods into a single Helpman– Krugman model. He finds that the assumption of free transport of the agricultural good is not objectionable since the HME vanishes if the agricultural good is transported with identical positive cost as the manufactured goods. Third, Yu (2005) extended Davis's (1998) result. He finds that the HME may either disappear, be reversed, or remain when the transport costs of the agricultural good are positive.

Fourth, Zeng and Kikuchiz (2005) reformulated the homogeneous-agricultural-good assumption in Davis' model to show that the HME does exist even if the transport costs of the agricultural goods are positive. Similar results are obtained by Crozet and Trionfetti (2008). Fifth, Picard and Zeng (2005) merge the study of agricultural transport costs and agricultural labour markets into a single model. These authors report the existence of a re-dispersion process through the agricultural sector, and that over-urbanization crucially depends on the values of agricultural transport costs, and on the firms' requirement for local-unskilled labour.

Industrial location patterns in developing countries can be illustrated by the example of Peru. During the last four decades, Peruvian manufacturing firms have mainly clustered in the capital of the country, core cities and coastal provinces (Herrera, 2009). However, currently, the concentration of economic activity is also occurring in rural areas around low-value-added manufacturing. Webb (2012) reports the increase of agglomeration in some Peruvian peripheral areas. He states that since 2000 the population growth rate of rural townships (2.4%) has been quite higher than those of cities (1.9%). Although this phenomenon can be somewhat explained by territorial order policies (Gonzales, 2010), it is also explained by the low accessibility of distant markets of foods and commodities (Escobal and Torero, 2005). At the beginning of 2000, as a result of the free trade agreements signed with EEUU, EU, China, Korea, Mexico and Japan, agricultural economic activities have become a key issue to Peruvian policy makers. After the Agrarian Reform applied in the 70's, whose objective was to allocate small land units to farmers, the 2000's trade policies have been mainly intended to promote the agglomeration of agricultural production so as to take advantage of the free trade agreements signed. Thus, significant changes related to trade liberalization, development of rural road networks and technological externalities are likely to modify the configuration of economic geography in both urban and rural areas in Peru.

The paper examines the impact of agricultural transport costs upon the agriculture-related firms' incentives to disperse or to agglomerate. In doing so, the distance-based tests of industry localisation developed by Duranton and Overman (2005) is used to estimate the geographic concentration of agriculture-related industries in Peru. Then conditional logit models are used to parametrically identify the effects of both the agricultural transport costs and a set of urban-rural variables and firm's characteristics on the firms' decisions on where to locate. To the best of our knowledge, there are no previous attempts to empirically measure the impact of agricultural transport costs on industry localisation. Thus, this investigation fills a gap in the empirical literature by examining the Economic Geography literature-related controversy concerning the role of agricultural transport costs in core-periphery models.

The results of the paper imply that agriculture-related industries in Peru do not exhibit strong localisation patterns. Using the distance based approach developed by Duranton and Overman (2005), we find that only 1 out of 13 industries in our sample is highly localised at short distances whereas the remaining industries are dispersed or are localised at long distances with a significantly weaker degree of intensity. Conditional logit models suggest that a high localisation at short (long) distances is explained by low (high) agricultural transport costs. Weak localisation at long distances and dispersion patterns are explained by the interplay of high and low transportation costs in the agricultural sector. Therefore, high or low levels of agricultural transport costs constitute a dispersion force that weakens the intensity of the HME, and cause dispersion of economic activities in agriculture-related industries. These results are consistent with the empirical literature that has identified patterns of dispersion of industry into hinterlands in developing countries such as Korea, Mexico and Brazil (Hanson, 1996; Henderson *et al*., 2002; Chun and Lee, 1985).

The remainder of this paper is organised as follows. Section 2 describes the micro-geographic datasets. Section 3 presents the Duranton and Overman (2005)'s distance-based test used to measure location patterns in agriculture-related industries in Peru, presenting the main results. Section 4 investigates the role of the agricultural transport costs in the geographical concentration of Peruvian agriculture-related industries using parametric conditional logit models. Section 5 concludes.

2. Data

The empirical analysis uses three databases. Firstly, it uses detailed firm-level data from the 2013 National Database of Manufacturing Firms (2013 NDMF), which are the data underlying the 2007 Annual Census of Peruvian Manufacturing Industries (2007 ACPMI). This database represents an exceptionally rich dataset since contains information for 89,268 manufacturing firms. The Ministry of Production of Peru was responsible for collecting the 2007 ACPMI during 2006-2007 and is updated on a yearly basis. For every firm, we count on information regarding its spatial reference, four-digit industry classification code (based on the Standard Industrial Classification - SIC- REV.3-1989) and number of employees. Since we study cause-effect linkages between agricultural transport costs and industrial location, we restrict our sample to agriculture-related firms, that is, manufacturing firms that mainly transform agricultural goods into manufactured goods. In doing so, the input-output matrix of the Peruvian economy is used to identify production linkages between agricultural and manufacturing sectors. In particular, we identify manufacturing industries that mostly employ agricultural goods in their production process[1]. As a result, we are left with a sample of 3,789 firms grouped in 13 industries, each of which has 10 or more firms.

To illustrate location patterns in Peruvian agriculture-related industries, Figures 1(a) to 1(c) show the geographical distribution of firms across the country for three illustrative industries: (a) Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711), (b) Tanning and Dressing of Leather (SIC 1911), and (c) Processing and Preserving of Fruit and Vegetables (SIC 1513), with each dot representing the location of a firm in each industry. The maps show that the (a) Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711) appears to be relatively more concentrated in the Lima Metropolitan Area (the capital of Peru), whereas the (b) Tanning and Dressing of Leather (SIC 1911) apparently is evenly distributed across the country. The (c) Processing and Preserving of Fruit and

[1] See Torres (2003) for a detailed presentation of the input-output matrix for the year 1994 used to restrict our sample.

Vegetables (SIC 1513) seems to be geographically concentrated in the North Coast and Central Coast of the country with apparently unevenly distributed of firms in the rest of regions. However, whether or not these 3 industries are localised in specific areas or dispersed across the country results far from evident. In Section 3, we examine detailed location patterns for all manufacturing industries in our sample by using spatial point pattern techniques.

Figure 1. Geographical distribution of firms for three illustrative industries

(a) Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711)

(b) Tanning and Dressing of Leather (SIC 1911)

(c) Processing and Preserving of Fruit and Vegetables (SIC 1513)

Secondly, we use farm-level data from the 2012 National Census of Agriculture and Livestock (2012 NCAL) to calculate agricultural production of field crops according to rural micro-geographic areas. These rural areas correspond to administrative areas that are officially defined for implementing agricultural censuses in Peru, and comprise on average 100 farming units, all of which are geo-coded. Since most farmers in Peru usually manage livestock in addition to crops, livestock farming is also considered when calculating farming production. The input-output matrix of the Peruvian economy indicates that 15 types of agricultural crops and 4 livestock species are mainly used in agriculture-related industries. These products account for 31% and 92% of the total agricultural production and livestock population in the country[2].

Table 1 shows farming production according to the 13 agriculture-related industries used in our analysis. We present data on production of field crops and livestock species, and the number of rural micro-geographic areas from which these farming outputs are gathered. Given that each firm in the 2013 NDMF and each rural micro-geographic area

[2] We focus on rural micro-geographic areas that concentrate large amounts of farming production. In this way, we discard small quantities of crops and livestock species that are primarily intended for direct human consumption. Furthermore, by focusing on large quantities of farming production, we increase the likelihood that farming production is transported from production zones to manufacturing firms or consumer markets, which does not distort our estimation of agricultural transport costs. The criteria used to determine rural micro-geographic areas with relatively large amounts of farming production is: (i) greater than or equal to 10 hectares for agricultural crops (except for barley and sugar cane), and (ii) greater than or equal to 100 head of livestock for livestock species. For barley and sugar cane, the criterion is greater than or equal to 1 hectare, because these crops mostly are grown in smaller areas.

in the 2012 NCAL is geo-coded, by merging both databases we can spatially connect every agricultural production zone with a specific firm location at a high level of precision. This allows us to connect agricultural supply with industrial demand and consumer markets, which enables to calculate the transport costs of the agricultural goods.

Table 1. Farming production by four-digit agriculture-related industries

Agriculture				
SIC	**Industry**	**Agricultural Crops**	**Area planted (hectares)**	**N° of rural micro-geographic areas**
1513	Processing and preserving of fruit and vegetables	Avocado, asparagus, mango, orange and apple	166,181	1,764
1514	Manufacture of vegetable and animal oils and fats	Oil palm	26,525	209
1531	Manufacture of grain mill products	Rice and wheat	188,915	2,745
1533	Manufacture of prepared animal feeds	Hard yellow corn	236,910	4,632
1542	Manufacture of sugar	Sugar cane	1,579	212
1543	Manufacture of cocoa, chocolate and sugar confectionery	Cocoa	138,368	2,100
1549	Manufacture of other food products n.e.c.	Coffee	420,105	3,478
1552	Manufacture of wines	Grapevine	40,681	482
1553	Manufacture of malt liquors and malt	Barley grain	43,138	5,188
1711	Preparation and spinning of textile fibres; weaving of textiles	Cotton	25,428	353
Livestock Farming				
SIC	**Industry**	**N° Livestock Species**	**Head of livestock**	**N° of rural micro-geographic areas**
1511	Production, processing and preserving of meat and meat products	Beef, ovine, llama and porcine cattle	20,016,625	25,380
1520	Manufacture of dairy products	Beef cattle	4,304,890	13,478
1911	Tanning and dressing of leather	Beef and ovine cattle	14,030,955	22,598

Source: Author’s elaboration based on the 2012 National Census of Agriculture and Livestock (2012 NCAL)

Thirdly, we generate a database containing agricultural transport costs by using the Peruvian Accessibility Model developed by IFPRI (2009). Agricultural transport costs are interpreted on the notion of accessibility and comprise two components (I and II). On the one hand, component I entails the costs related to transporting agricultural goods from agricultural production zones to regional or local consumer markets. These markets are approximated as a conglomerate of people with more than 20,000 inhabitants but less than 50,000 inhabitants[3]. On the other hand, component II involves the costs of transporting agricultural production from production zones to firms' locations. Both components ensure that the agricultural production not intended for direct human consumption at the production region is traded. Formally, these two components of the agricultural transport costs are defined as the ease with which a consumer market (component I) or firm location (component II) may be reached from a particular agricultural production zone by taking into account the distance travelled on different types of roads and an impedance factor that reflects the travel speeds on roads of different types and qualities, the slope of the terrain and the presence of natural obstacles. As a result, an accessibility indicator expressed as a weighted average of the distance travelled on each type of road is obtained, where the weights allocated are directly proportional to the impedance factor. See IFPRI (2009) for a detailed description of the Peruvian Accessibility Model.

The accessibility indicator measured in terms of time is used to approximate the agricultural transport costs in Peru. Technically, this indicator calculates the least cost path surface to access to the nearest consumer market or firm location from a specific agricultural production zone for each of the 13 industries in our sample, using GIS[4]. The estimation of the agricultural transport costs considers three variables: (i) *transportation infrastructure:* roads of different types (first order roads, second order roads, dirt road

[3] This classification corresponds to the category of small cities as defined by the National Statistics Office of Peru (INEI, 2007).

[4] The interpretation of the agricultural transport costs only in terms of time may be limited. However, their importance lies in the very fact that they capture in an accurate manner the physical transport costs in the Sierra and Jungle of the country, which are characterised by highly heterogeneous geography and accessibility problems. Since most of the agricultural production zones in Peru are located in these regions, the calculation of the agricultural transport costs based on time can be considered non-negligible.

tracks and walking trails) and navigable rivers. This information comes from the Ministry of Transport and Communications of Peru; (ii) *land slope:* used to allocate different walking travel speeds (on horseback, on footpath, and off footpath) to different options of walking travel (dirt road tracks, walking trails, and no paths). Data on elevation comes from the SRTM 90m Digital Elevation Database; and (iii) *natural barriers:* that considers innavigable rivers. Hydrology data comes from the National Authority of Water Resources of the Ministry of Agriculture of Peru.

Table 2 presents average transport costs for the agricultural goods according to their components I and II. As we can see, there is a wide variation in transportation costs in the agricultural sector, both among industries and between components. Firstly, an inter-intra industry comparison indicates that the highest transport costs correspond to transporting agricultural goods for the manufacture of oils and fats (industry SIC 1514) and meat products (industry SIC 1511). It can take until 29 hours to transport oil palm from agricultural production zones to consumer markets, whereas it may take up to 45 hours to bringing livestock species to industrial production centres. Interestingly, agricultural goods used by industry SIC 1514 also exhibit the highest average transport cost in both components. A firm within this industry can take on average 12 hours to obtain agricultural inputs from agricultural production areas. This however stands in contrast to the transports costs for a firm within industry SIC 1511, which does not incur transport costs to get agricultural products. This information on transport costs shows the marked differences in the inter- and intra-industry variability between both components of agricultural transport costs.

Secondly, an inter-component examination indicates that component II exhibits the highest transport costs and the highest average transport costs. Whereas in component I the highest agricultural transport cost amounts to 29 hours, in component II such a cost is 45 hours. In addition, in component II the highest average transportation cost is equivalent to 12 hours whereas in component I is around 2 hours. Thus, we can state that the costs related to transporting agricultural goods to industrial production centres turn out to be more important in the agricultural sector.

Table 2. Agricultural transport costs by four-digit agriculture-related industries

SIC	Industry	Agricultural Transport Costs (Time in hours)							
		Component I				Component II			
		Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Agriculture									
1513	Processing and preserving of fruit and vegetables	1.02	1.42	0.03	18.45	1.17	1.22	0.00	10.18
1514	Manufacture of vegetable and animal oils and fats	0.73	0.67	0.44	5.29	11.89	9.68	0.34	45.12
1531	Manufacture of grain mill products	1.14	1.99	0.03	12.46	1.85	1.80	0.00	9.68
1533	Manufacture of prepared animal feeds	0.96	1.03	0.21	4.48	1.84	1.21	0.20	5.07
1542	Manufacture of sugar	0.80	0.33	0.30	1.83	1.71	1.50	0.30	5.62
1543	Manufacture of cocoa, chocolate and sugar confectionery	0.80	1.37	0.03	12.63	5.61	2.09	0.06	14.46
1549	Manufacture of other food products n.e.c.	1.84	2.44	0.00	12.88	4.35	7.82	0.04	41.57
1552	Manufacture of wines	1.57	2.41	0.01	11.80	1.19	1.63	0.01	10.30
1553	Manufacture of malt liquors and malt	1.34	0.79	0.02	2.76	4.47	11.38	0.05	42.09
1711	Preparation and spinning of textile fibres; weaving of textiles	0.63	0.20	0.02	3.42	2.94	2.56	0.02	24.06
Livestock Farming									
1511	Production, processing and preserving of meat and meat products	1.08	2.23	0.02	29.21	0.75	2.02	0.00	21.23
1520	Manufacture of dairy products	1.16	1.99	0.01	19.68	0.85	1.58	0.00	15.36
1911	Tanning and dressing of leather	0.62	1.09	0.03	16.83	0.42	1.06	0.00	11.84
	Mean	*1.05*	*1.38*	*0.09*	*11.67*	*3.00*	*3.50*	*0.08*	*19.74*
	Std. Dev.	*0.37*	*0.77*	*0.14*	*8.08*	*3.13*	*3.59*	*0.12*	*14.28*
	Min	*0.62*	*0.20*	*0.00*	*1.83*	*0.42*	*1.06*	*0.00*	*5.07*
	Max	*1.84*	*2.44*	*0.44*	*29.21*	*11.89*	*11.38*	*0.34*	*45.12*

Source: Author's calculations based on the estimation of the Peruvian Accessibility Model (IFPRI, 2009)

3. Location Patterns of Agriculture-related Industries in Peru

In this section, we present the methodology developed by Duranton and Overman (2005) (hereafter D-O index) to identify localised agriculture-related industries in Peru. The D-O index is based on the kernel density of the distribution of bilateral distances across all firms in an industry, and compares that distribution to a counterfactual one that is gained under the assumption of spatial randomness. An industry is defined to be significantly localised or dispersed, respectively, if its distribution of bilateral distances significantly deviates from randomness. Essentially, the D-O index consists of three steps. First, it computes the bilateral distances between all firms in an industry and then estimates a kernel density function (K-densities) of the distance distribution. Second, it simulates counterfactual location distributions by assuming that all firms in the industry are randomly allocated. Third, it constructs confidence interval bands to test whether an industry exhibits localisation or dispersion. These three steps are shortly sketched below.

3.1. Calculation of Kernel Density Function (K-densities)

This step implies to compute the density of bilateral distances between all pair of firms in an industry. For each industry A, with n firms, the Euclidean distance between every pair of firms is calculated, which generates $\frac{n(n-1)}{2}$ bilateral distances. Then, kernel-smoothed distributions (K-densities) of such bilateral distances are estimated. Denote by $d_{i,j}$ the Euclidean distance between firms i and j, and given n firms, the estimator of the density at distance d is:

$$\widehat{K}_A(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} f\left(\frac{d-d_{ij}}{h}\right) \tag{1}$$

Where h is the bandwidth and f is the (Gaussian) kernel function[5].

3.2. Constructing Counterfactuals

The appropriate counterfactuals of randomly located pseudo industries are constructed in order to compare them to the K-densities. In doing so, we assume that the set of all existing sites currently occupied by manufacturing firms in the industrial branches SIC 15, SIC 17 and SIC 19 of the 2013 NDMF database represent the set of all potential sites for any agriculture-related firm[6]. This is because; these sites correspond to locations of manufacturing firms that are entirely involved in agriculture-related activities. By defining the counterfactual sample in this way, we ensure two essentials. First, that an agriculture-related firm is allocated to sites currently occupied by their peers. Second, that a firm is not allocated to areas restricted for planning or zoning constraints.

Then, the counterfactuals are randomly constructed drawing locations from the overall potential sample of sites (16,756 observations) and then computing the set of bilateral distances. That is, in each simulation we randomly draw locations of the same number as the number of firms in the corresponding manufacturing industry, and then compute the bilateral distances of the sites and estimate the K-density. This process guarantees that the complete patterns of agglomeration in a particular manufacturing sector are controlled altogether. We run 1,000 Monte Carlo simulations for each agriculture-related industry.

[5] As defined in Duranton and Overman (2005), a Gaussian kernel with optimal bandwidth is used to calculate bilateral densities. The optimal bandwidth is $1.06sn^{-0.2}$, where n is the observed number and s is the standard deviation (Klier and McMillen, 2008). For details, see Silverman (1986).

[6] These industrial branches correspond to: SIC 15: Manufacture of food products and beverages; SIC 17 Manufacture of textiles; and SIC 19 Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear.

3.3. Identification of Location Patterns of Industries

At this step, we analyse whether a particular industry exhibits dispersion or localisation by comparing the K-densities with that of the counterfactual distribution. This procedure involves constructing global confidence bands containing 95% of the randomly drawn K-densities to test whether an industry is globally localised through assessing the statistical significance of departures from randomness[7]. For the 1,000 simulations for each agriculture-related industry, we construct two-sided confidence bands containing 95% of the randomly drawn K-densities. That is, by using the 1,000 trials, the upper global confidence interval is calculated in such a way that 95% of the 1,000 simulations lie above the lower band and another 95% of the 1,000 simulations lie below the upper band. The interpretation of the global confidence bands is the following: if a higher K-density at short distances than the density of randomly drawn distributions is observed, the industry is said to exhibit localisation. Analogously, if a lower K-density at short distances than the density of randomly drawn distributions is observed, the industry tends to exhibit dispersion. We construct global confidence bands between $[0, 180]$ km.

By means of this procedure, the upper global confidence interval $\overline{K}_A(d)$ and the lower global confidence interval $\underline{K}_A(d)$ of industry A are obtained. If $\widehat{K}_A(d) > \overline{K}_A(d)$ for at least one $d \epsilon [0, 180]$ industry A is defined as localised in global terms at the 5% confidence level. On the other hand, if $\widehat{K}_A(d) < \underline{K}_A(d)$ for at least one $d \epsilon [0, 180]$, industry A is defined as dispersed in global terms. Indexes of global localisation and global dispersion are defined as equation (2) and (3), respectively:

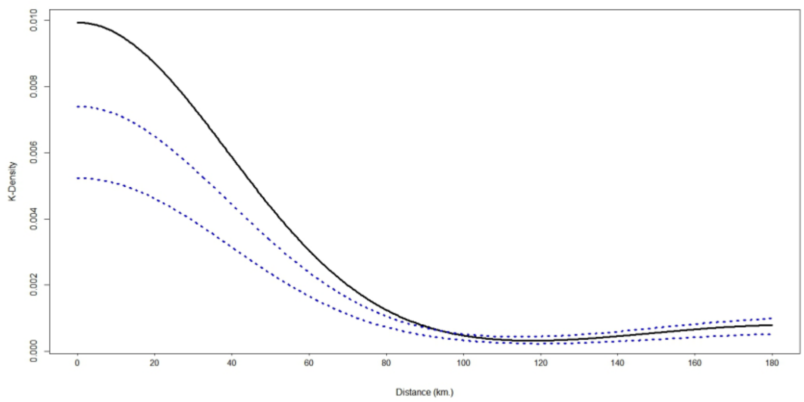
$$\Gamma(d) \equiv max\left(\widehat{K}_A(d) - \overline{K}_A(d), 0\right) \qquad (2)$$

[7] Duranton and Overman (2005) also constructed local confidence intervals for each distance. However, as stated by the authors, testing localisation based on local confidence bands only allows us to make statements at a specific distance. Therefore, we focus on global confidence bands to detect localisation or dispersion over the whole range of distances in Peruvian agriculture-related industries.

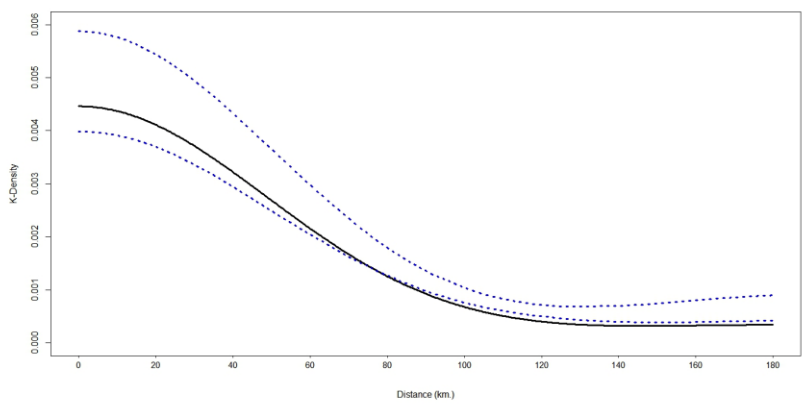
$$\Psi(d) \equiv \begin{cases} max\left(\underline{K}_A(d) - \widehat{K}_A(d), 0\right) & if \quad \sum_{d=0}^{d=180} \Gamma(d) = 0 \\ \\ 0 & otherwise \end{cases} \tag{3}$$

For illustration, we present the K-densities and the corresponding two-sided confidence intervals of the three formerly introduced industries in Figures 2(a) to 2(c). The solid lines in these figures indicate the observed distribution of distances in the industry (K-densities as estimated according to equation (1)) whereas the upper (lower) dashed line(s) plots the upper (lower) global confidence band.

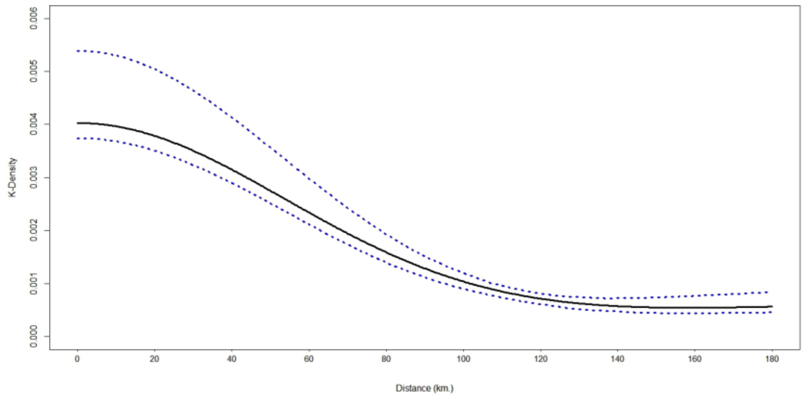
Figure 2. Firm K-densities and global confidence bands for three illustrative industries



(a) Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711)



(b) Tanning and Dressing of Leather (SIC 1911)



(c) Processing and Preserving of Fruit and Vegetables (SIC 1513)

The three agriculture-related industries showed in Figure 2 display three different geographical patterns. Figure 2(a) for the Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711) depicts an industry that is localised at short distances because its K-density is above the upper global confidence band for distances within the 0-95 km range. However, it is dispersed at longer distances. For every distance within the range of 0-95 km, the K-density is above the upper global confidence band, which provides us with evidence that this industry is localised. In particular, this industry has a large cluster in the capital of the country (the Lima Metropolitan Area) where around 66% of firms are located as can be seen in Figure 1(a). Firms within this agglomeration are located close to each other, which justifies the high density at short distances. In fact, this industry together with other related textile industries form one of the largest clusters in Peru.

On the other hand, Figure 2(b) for the Tanning and Dressing of Leather (SIC 1911) depicts an industry that is significantly dispersed at long distances. For every distance within the range of 80-180 km, the K-density lies below the lower global confidence band. As Figure 1(b) illustrates, the location pattern of this industry is much more evenly distributed. Last, Figure 2(c) for the Processing and Preserving of Fruit and Vegetables (SIC 1513) illustrates an industry that is neither significantly localised nor significantly dispersed. For every distance within the range of 0-180 km, the K-density falls between the confidence bands and never above the upper global or lower confidence band. So, this industry does not exhibit any location pattern within 180 km. Indeed, the location pattern of this industry is not significantly distinct from one that would be obtained by randomness.

3.4. Results

We find that 7 out of 13 (54%) agriculture-related industries deviate from randomness. Thus, these industries are globally localised at a 5% confidence level. On the other hand, 2 out of 13 (31%) industries are identified to be globally dispersed[8]. Table 3 presents the indices of global localisation and global dispersion corresponding to these 9 industries[9]. Interestingly enough, we find that two of the most traditional manufacturing industries in Peru exhibit the strongest spatial localisation patterns: (i) Preparation and spinning of textile fibres; weaving of textiles (SIC 1711), and (ii) Manufacture of wines (1552). Both industries share the common feature of being two of the oldest industries in the country, and heavily dependent on local workforce. It is worth mentioning the case of the industry SIC 1711, which is highly concentrated at a regional scale as shown in Figure 1(a). Moreover, it is noteworthy that the localisation pattern of this industry seems to be not only representative of Peru, since the industry SIC 1711 was also identified among the ten most localised industries in the UK in Duranton and Overman (2005). This last finding would suggest similar agglomeration patterns in this industry in developed and less developed countries.

[8] The remaining 4 agriculture-related industries are neither significantly localised nor significantly dispersed. Thus, we do not present their results. These industries are: (i) Processing and preserving of fruit and vegetables (SIC 1513); (ii) Manufacture of prepared animal feeds (SIC 1533); (iii) Manufacture of sugar (SIC 1542), and (iv) Manufacture of malt liquors and malt (SIC 1553).

[9] Similar to Duranton and Overman (2005), we construct cross-industry indices to calculate the indices of global localisation and global dispersion for each industry A at any given distance, respectively: $\Gamma(d) \equiv \sum_A \Gamma_A(d)$ and $\Psi(d) \equiv \sum_A \Psi_A(d)$. These indices allow us to know the extent of localisation and dispersion across all industries for each distance.

On the other hand, as reported in Table 3, industries related to production of tanned leather (SIC 1911), and manufacture of oils and fats (SIC 1514) exhibit a dispersed location pattern. This suggests that these industries incur higher transport costs and are relatively more dependent on natural resources than other agriculture-related industries in our sample. Note moreover that Duranton and Overman (2005) find that four industries belonging to the food-related branch SIC 15[10] are among the ten most dispersed industries in the UK. Since the Peruvian industry SIC 1514 belongs to the same industrial branch, our findings on dispersed industries are consistent with Duranton and Overman's results.

Table 3. Localised and dispersed four-digit agriculture-related industries

SIC	Industry	N° of firms	Γ or ψ
Localised			
1711	Preparation and spinning of textile fibres; weaving of textiles	487	0.329
1552	Manufacture of wines	221	0.316
1549	Manufacture of other food products n.e.c.	467	0.048
1520	Manufacture of dairy products	538	0.036
1543	Manufacture of cocoa, chocolate and sugar confectionery	252	0.029
1511	Production, processing and preserving of meat and meat products	404	0.007
1531	Manufacture of grain mill products	470	0.007
Dispersed			
1911	Tanning and dressing of leather	363	0.021
1514	Manufacture of vegetable and animal oils and fats	99	0.010

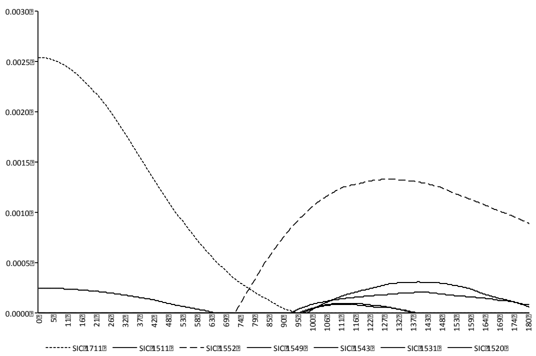
Source: Author's calculations based on the estimation of the distance-based tests of industry localisation developed by Duranton and Overman (2005)

However, it should be noted that although a larger share of agriculture-related industries is found to be globally localised, in strict, only the industry SIC 1711 is highly localised at short distances. The industry SIC 1552 is highly localised but at long distances whereas the remaining five localised industries in Table 3 are weakly localised and mostly at long distances. These irregular patterns of industrial activity are

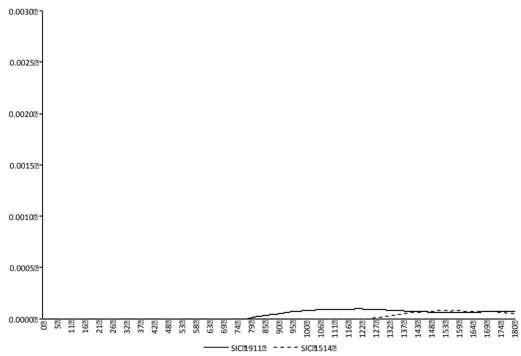
[10] SIC 15: Manufacture of food products and beverages.

shown in Figure 3(a) and Figure 3(b), which depict the indices of global localisation and global dispersion by agriculture-related industry. Figure 3(a) shows that 5 out of 7 localised industries falls far below from the indices calculated for the two most localised industries (SIC 1711 and SIC 1552, represented by the dotted and dashed lines, respectively). These marked differences in the intensity of localisation patterns represent one of the most striking feature of industrial localisation in agriculture-related industries in Peru. On the other hand, Figure 3(b) shows the indices of global dispersion for the two dispersed industries. As we see, both industries follow a similar dispersion pattern but at different distances. Both industries first begin to disperse at long distances and then stabilise, without displaying any particular pattern from there onwards.

Figure 3. Indices of global localisation and dispersion by agriculture-related industry and distance



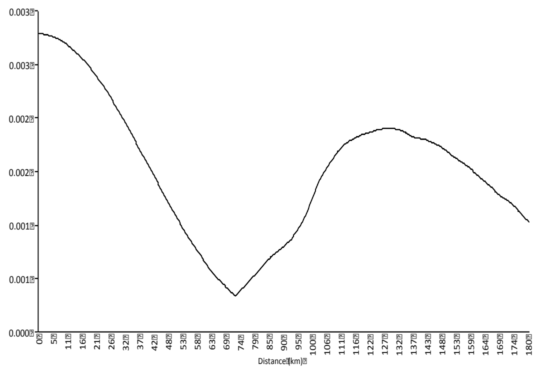
(a) Global localisation



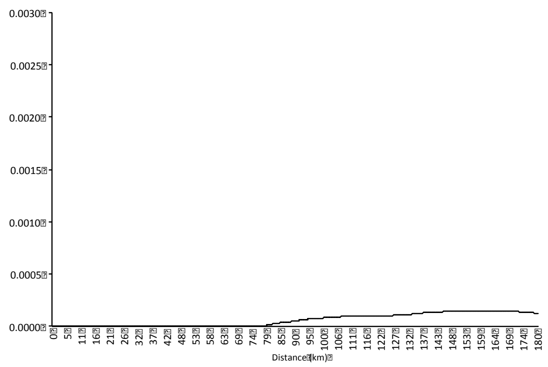
(b) Global dispersion

Figures 4(a) and 4(b) depict the same indices of global localisation and global dispersion of Figure 3 but aggregated for the whole localised and dispersed industries. From these graphs we can draw two main conclusions. First, localisation of industries does not only occur at shorter distances, as it has been found in most studies of industrial location. As Figure 4(a) illustrates, global localisation index exhibits high values for distances below 70 km and also at intermediate distances of 100-140 km. That is, the extent of localisation is great at small and long distances. In particular, we can identify one conventional (Type I) and two non-conventional localisation patterns (Type I and II): (i) Type I: highly localised at short distances (SIC 1711); and (ii) Type II: highly localised at long distances (SIC 1552) and (iii) Type III: weakly localised mostly at long distances (SIC 1549, SIC 1520, SIC 1543, SIC 1511, and SIC 1531). Therefore, we can say that localisation of agriculture-related industries takes place within small and large areas in Peru. Moreover, dispersion shows no clear pattern, which is in line with Duranton and Overman's results. The index of global dispersion is stable over the range of distances from 80-180 km, as can be seen in Figure 4(b).

Figure 4. Indices of global localisation and dispersion by distance



(a) Global localisation



(b) Global dispersion

These findings allow us to reach our main conclusion in this section: agriculture-related industries do not exhibit a strong localisation patterns in Peru. Only one industry is highly located at short distances whereas the remaining industries are localised at long distances with a significantly weaker degree of intensity. This locational phenomenon can be explained by the high level of concentration of the Peruvian economy in a small number of regions and core cities. So, for example, Lima, the capital of the country, concentrates around the 30% of the country's population and accounts for more than two-thirds of the nation's gross domestic product. Most of the Peru's imports and exports pass through the port of Callao (located near the Lima

Metropolitan Area), which constitutes the nation's most important port. These factors have contributed to the strong localisation pattern in and around Lima, which may have weakened the rise of localisation patterns in other regions. Nevertheless, this strong urbanization pattern cannot be the only justification for the unusual localisation patterns in agriculture-related industries in Peru. In next section, we explore the roots of this uncommon geographic concentration using parametric conditional logit models.

4. Conditional Logit Models: The Role of Agricultural Transport Costs

In this section, we present conditional logit models so as to investigate the probability that a particular agriculture-related firm chooses to locate in a specific area. These potential site choices correspond to the rural micro-geographic areas defined in Section 2. Our main interest is to explain the three types of localisation patterns and the dispersion patterns identified in Section 3 with positive transport costs for the agricultural goods[11]. That is to say, we try to capture the importance agricultural transport costs as a determinant of agriculture-related firm location in Peru. In addition, we include other explanatory variables such as the attributes of the choice alternatives (for example, average altitude of the micro-geographic area) as well as characteristics of the firms making the choices (such as number of employees). One advantage of conditional logit models is that they allow us to exploit micro-geographic data. With micro data we can explore in detail the factors driving location choices of agriculture-related firms in Peru.

[11] Since the following 4 industries (i) Processing and preserving of fruit and vegetables (SIC 1513), (ii) Manufacture of prepared animal feeds (SIC 1533), (iii) Manufacture of sugar (SIC 1542), and (iv) Manufacture of malt liquors and malt (SIC 1553) showed neither localisation nor dispersion, these were not considered in this explanatory analysis.

4.1. Location Choice Model

We model the location decision of agriculture-related firms as a conditional logit problem where the dependent variable is the rural micro-geographic area chosen by each manufacturing firm. Following McFadden's (1974) conditional logit model, the probability that firm i selects locational choice j (rural micro-geographic area) among a set of J_i locational alternatives is given by:

$$P_{ij} = \frac{\exp(\beta' X_{ij})}{\sum_{j=1}^{J_i} \exp(\beta' X_{ij})} \quad (4)$$

where β is a vector of unknown parameters indicating the effect of independent variables on the probability of choosing one particular location over another. X_{ij} are independent variables (covariates) that may change with firms, locational choice, or both. That is, the expected utility of a locational choice may depend on characteristics of the site alternatives, characteristics of the manufacturing firms making the choices, and variables that are particular to a combination of firm and locational choice. It should be noted that, one of our main predictors (the amount of time it would take firm i located at j to obtain agricultural goods from a particular agricultural production zone[12]) belongs to the latter category[13]. Now, if we let d_{ij} be the dependent variable that takes the value 1 (one) if manufacturing firm i selects locational choice j and the value 0 (zero) otherwise, the log-likelihood function for the conditional logit model to be estimated can be written as:

$$LL_{clm} = \sum_{i=1}^{N} \sum_{j=1}^{J_i} d_{ij} ln P_{ij} \quad (5)$$

where N is the total number of agriculture-related firms.

[12] This variable corresponds to the component II of the agricultural transport costs.

[13] In general, for each variable X_k, there are J values of the variable for each manufacturing firm, but only the single parameter β_k.

The specification of the location choice model involves defining the set of rejected locational alternatives. That is, we identify the locational choices that have not been taken by manufacturing firms (which have zero value in the independent variable ($d_{ij} = 0$)). Given that our firm-level data contains 3,789 agriculture-related firms, in particular each firm faces 3,789 potential locational choices (rural micro-geographic areas)[14] [15]. However, similar to Klier and McMillen (2008), we follow Ben-Akiva and Lerman (1985) and randomly choose five rejected choices when estimating the conditional logit models. That is, each firm is matched with five randomly chosen rural micro-geographic areas, which are both different from the firm's current micro-geographic area and different from each other. As a result, our dataset contains $6n$ observations for the 9 agriculture-related industries under study (19,806 observations), where the dependent variable equals one (1) for the first observation for each manufacturing firm, and zero (0) for the next five observations for each firm. In both cases, the explanatory variables comprise the characteristics for the chosen firm site and the randomly chosen rejected locations, respectively[16].

Table 4 shows descriptive statistics for independent variables used in the conditional logit models. Variables include agricultural transport costs (components I and II), population density, rural micro-geographic areas' average altitude, and the proportion of workers in manufacturing jobs. We present sets of statistics for the 9 agriculture-related industries in Table 3, and statistics for samples of randomly chosen alternative locations for the same industries. Table 4 indicates that rural micro-geographic areas with firms in industry SIC 1711 (localisation pattern Type I) are less likely than randomly chosen alternatives to have higher agricultural transport costs (component I and II) whereas the

[14] The option for one firm may involve a micro-geographic area that already has another firm.
[15] Since rural micro-geographic areas have been defined for gathering agricultural data, some agricultural-related firms do not necessarily locate in sites that are circumscribed to these micro-geographic areas. Hence, in these cases we use district boundaries as reference areas to implement the conditional logit models. This case corresponds to 21.8% of our firm sample, that is, 827 out of 3,789 firms.
[16] The samples of randomly chosen alternative locations were obtained by means of a random sampling without replacement from the set of all existing sites currently occupied by agriculture-related firms in Peru (3,789 observations).

opposite is the case for firms in industry SIC 1552 (localisation pattern Type II). In addition, the probability of having rural micro-geographic areas with higher agricultural transportation costs depends on contradictory effects in firms belonging to industries of the localisation pattern Type III. As for dispersion patterns, we can observe that micro-geographic areas are less likely than randomly chosen alternatives to have higher agricultural transport costs in firms within industry SIC 1911, and this likelihood is dependent on opposite effects in firms within industry SIC 1514. The three remaining variables concerning population density, areas' average altitude, and the proportion of workers in manufacturing jobs show a much greater statistical variability.

4.2. Location Choice Results

Table 5 shows the conditional logit model results. For firms belonging to the localisation pattern Type I (industry SIC 1711) the results imply that a rural micro-geographic area is more likely to be chosen as a firm location if it presents low agricultural transport costs of the component II. That is to say, the probability of an existing agriculture-related firm in this industry is higher when the micro-geographic area enables firms to gain access to agricultural goods (cotton in this case) in less time. In addition, the likelihood of an existing firm is higher when the micro-geographic area is located at a high altitude, has a high population density and a high proportion of workers in manufacturing jobs. Thus, we can say that an agriculture-related firm appears to be highly localised at short distances whether it incurs low transport costs to obtain agricultural products. The component II of the agricultural transport costs is not statistically significant.

Table 4. Descriptive statistics

Variables	Localisation patterns														Dispersion patterns				Units
	Type I: Highly localised at short distances		Type II: Highly localised at long distances		Type III: Weakly localised mostly at long distances														
	SIC 1711		SIC 1552		SIC 1549		SIC 1520		SIC 1543		SIC 1511		SIC 1531		SIC 1911		SIC 1514		
	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	Firm location	Random Site	
Agricultural transport costs: Component I	0.627	0.802	1.574	0.320	1.845	0.519	1.156	0.659	0.805	0.811	1.080	2.144	1.138	0.903	0.618	2.888	0.729	1.343	Time in hours
	(12.145)	(66.623)	(144.461)	(11.624)	(146.579)	(19.538)	(119.194)	(23.758)	(82.237)	(23.274)	133.922	(197.431)	(119.176)	(40.568)	(65.269)	(183.308)	(40.181)	(78.713)	
Agricultural transport costs: Component II	2.939	9.612	1.187	0.550	4.353	1.134	0.847	0.774	5.612	1.222	0.752	2.321	1.851	2.143	0.419	2.533	11.893	2.747	Time in hours
	(153.425)	(967.0958)	(98.051)	(42.514)	(469.146)	(60.284)	(94.871)	(28.135)	(125.448)	(57.834)	(121.232)	(85.677)	(108.245)	(103.158)	(63.368)	(115.921)	(580.917)	(209.146)	
Population density	7.404	0.034	2.135	4.299	1.933	10.078	1.918	1.164	6.067	11.959	2.790	4.112	1.687	3.877	3.399	0.016	3.935	0.018	1000 Inhabitants per sq.km.
	(8.421)	(0.057)	(5.052)	(5.447)	(0.000)	(9.757)	(5.239)	(2.294)	(7.474)	(10.454)	(6.441)	(8.176)	(4.944)	(4.637)	(6.058)	(0.006)	(7.079)	(0.009)	
Average altitude	985.1	1150.9	636.3	614.2	1616.1	247.4	1879.9	978.1	898.4	1626.8	1654.2	2450.7	1334.1	1615.8	1370.2	1675.5	514.5	1422.6	Metres above mean sea level
	(1381.199)	(1534.818)	(750.113)	(955.829)	(7.602)	(209.622)	(1571.000)	(1508.689)	(1226.286)	(1882.550)	(1679.554)	(1218.468)	(1495.097)	(1281.514)	(1416.151)	(1839.437)	(889.547)	(833.325)	
Proportion manufacturing	5.38	0.30	2.26	6.32	2.29	7.66	0.84	6.91	1.91	5.76	1.14	0.68	0.82	2.33	4.12	0.67	2.11	5.38	Proportion of workers in manufacturing jobs
	(8.162)	(0.228)	(6.834)	(8.420)	(0.168)	(5.844)	(1.513)	(7.842)	(4.444)	(8.432)	(2.999)	(0.811)	(1.756)	(3.305)	(0.177)	(0.671)	(5.139)	(9.513)	
Number of observations	**487**	**2435**	**221**	**1105**	**467**	**2335**	**538**	**2690**	**252**	**1260**	**404**	**2020**	**470**	**2350**	**363**	**1815**	**99**	**495**	

*/ Standard deviations are in parentheses for the continuous variables.
Source: Author's elaboration based on the 2013 National Database of Manufacturing Firms (2013 NDMF); National Statistics Office of Peru (INEI, 2007); and IFPRI (2009).

The model specification for location decisions of firms belonging to the localisation pattern Type II (industry SIC 1552) is contrary to the localisation model Type I. These findings indicate that a rural micro-geographic area is more likely to be chosen as a firm location if it exhibits high agricultural transport costs of the component I and II. In other words, a firm is likely to be highly agglomerated at long distances if the micro-geographic area where it is located present high agricultural transport costs. These results suggest that high transport costs in the agricultural sector configure an unusual localisation pattern in agriculture-related industries characterised by geographic concentration of firms at long distances. Exogenous variables regarding population density, areas' average altitude, and the proportion of workers in manufacturing jobs exert a dominant negative effect on industrial localisation.

As for the localisation model Type III, the 5 industries that comprise this group show dissimilar localisation patterns. However, we can identify two industrial location behaviours. On the one hand, for firms within industries SIC 1520, SIC 1511 and SIC 1531, a rural micro-geographic area is more likely to have a firm weakly localised at long distances if it has high agricultural transportation costs of the component I and low agricultural transport costs of the component II. On the other hand, for firms within industries SIC 1549 and SIC 1543 we find that these firms are more likely to be located in micro-geographic areas with high agricultural transport costs of the component II and low transportation costs of the component I[17]. These ambivalent industrial location patterns suggest that both high and low transport costs in the agricultural sector operate as a dispersion force, which weaken the intensity of the localisation pattern. Variables concerning population density, areas' average altitude, and the proportion of employment in manufacturing exert different effects on industrial localisation. Finally, the specification for dispersed location patterns (industries SIC 1911 and SIC 1514) is similar to the localisation model Type III. For firms within these industries, a rural micro-geographic area is more likely to have a firm that follow a dispersed pattern if it

[17] Although the negative coefficient of the component I of the agricultural transport costs in industry SIC 1549 is not statistically significant.

has high agricultural transportation costs of the component I and low agricultural transport costs of the component II. Therefore, we can say that the locational equilibrium of dispersed industries is defined by the interplay of high and low agricultural transport costs. That is to say, high or low levels of transportation costs in the agricultural sector cause dispersion of economic activities in agriculture-related industries.

4.3. Industry Localisation in Developing Countries

Most Economic Geography models assume that manufacturing goods are produced under an increasing-returns-to-scale (IRS) technology and monopolistic competition and are traded with positive transport costs, whereas the agricultural good is produced under a technology of constant returns to scale (CRS) and is costlessly transported across countries or regions. Under these assumptions, the usual outcome in these models is the dispersion of economic activities in the CRS sector and the agglomeration of firms in the IRS sector. In other words, the HME introduced by Helpman and Krugman (1985) always appears as a direct result of transportation cost savings in the IRS sector. However, once we consider that both the IRS and CRS sectors are subject to positive transport costs, such theoretical prediction is empirically unsatisfactory. This is because; the existence of manufacturing and agricultural transport costs cause wage disparities among regions and countries, which in turn may lead to dispersion of economic activities (Davis, 1998).

The logical sequence of the HME in a context of two regions within a country is as follows. Without the free-trade assumption in the agricultural sector, wages are not equalised in large and small markets, which enables the appearance of spatial wage disparities. This wage gap activates the interplay of two opposite forces: agglomeration and dispersion forces that define the existence of different spatial equilibrium of industrial location (Picard and Zeng, 2005).

Table 5. Conditional logit for the location of agriculture-related industries

Variables	Localisation patterns							Dispersion patterns		Units
	Type I: Highly localised at short distances	Type II: Highly localised at long distances	Type III: Weakly localised mostly at long distances							
	SIC 1711	SIC 1552	SIC 1549	SIC 1520	SIC 1543	SIC 1511	SIC 1531	SIC 1911	SIC 1514	
Agricultural transport costs: Component I	-0.0018	0.0532*	-0.0006	0.0171*	-0.0206*	0.0146*	0.0031*	0.0135*	0.0253**	Time in hours
	(0.0014)	(0.0083)	(0.0099)	(0.0020)	(0.0068)	(0.0012)	(0.0008)	(0.0040)	(0.0013)	
Agricultural transport costs: Component II	-0.0004*	0.0243*	0.1280*	-0.0216*	0.0399*	-0.0328*	-0.0018*	-0.0350*	-0.0372*	Time in hours
	(0.0001)	(0.0050)	(0.0245)	(0.0024)	(0.0119)	(0.0019)	(0.0006)	(0.0045)	(0.0011)	
Population density	0.0020*	-0.0001**	-0.0010*	0.0003*	-0.0004*	-0.0000**	0.0000	0.0200*	-0.0102	Inhabitants per sq.km.
	(0.0007)	(0.0000)	(0.0002)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0064)	(0.0050)	
Average altitude	0.0002*	-0.0004*	0.0044*	0.0001*	-0.0033*	-0.0006*	-0.0003*	0.0001	-0.1608	Metres above mean sea level
	(0.0000)	(0.0001)	(0.0015)	(0.0000)	(0.0009)	(0.0001)	(0.0001)	(0.0001)	(0.0049)	
Proportion manufacturing	0.2628*	-0.1703*	0.1905	-0.3801*	-0.2874*	0.3215**	-0.3124*	0.5124*	-0.3095**	Proportion of workers in manufacturing jobs
	(0.0645)	(0.0262)	(0.1599)	(0.0363)	(0.0828)	(0.1644)	(0.0458)	(0.0876)	(0.0080)	
Pseudo-R^2	0.6082	0.5106	0.9610	0.3818	0.9348	0.5097	0.1230	0.7818	0.6015	
Number of observations	2922	1326	2802	3228	1512	2424	2820	2178	594	

*/ Standard errors are given in parentheses: * $p < 0.01$; ** $p < 0.05$
Source: Author's calculations based on the estimation of the conditional logit models

First, a higher agricultural transport cost act as a dispersion force through the labour market. This larger transport costs increase the price of agricultural products in regions with large manufacturing sectors (large markets) and reduce them at the production regions (small markets). Because individuals in large markets demand less of the relatively more expensive agricultural good, farmers experience an increment in the supply of agricultural goods, which reduces their prices even more. Consequently, wage income of farmers declines more in agricultural regions, which attracts manufacturing firms and encourages dispersion of industrial activities (Picard and Zeng, 2005). Second, the agglomeration force operates on the manufacturing sector. When manufacturing transport costs are small and skilled workers locate in large markets, they earn higher wages and increase their consumption of manufactured goods. This attracts additional skilled workers and creates an incentive for more firms to locate in the same region due to increasing demand. Then, competition increases, the price of manufactured goods reduces and this expands even more the consumption of manufactured goods. As a result, economic activity tends to agglomerate in the region with high manufacturing share. Ultimately, however, the HME is determined by the tension between these two opposite forces (Davis, 1998).

Our results in Section 4.2 support the existence of the HME in industries belonging to the localisation pattern Type I, II and III, and the absence of the HME in dispersed industries. In other words, our findings suggest that the agglomeration force outweigh the dispersion force in industries highly and weakly localised whereas in dispersed industries the opposite is the case. Indeed, we find that positive transport costs for the agricultural goods affect the degree of intensity of the HME at varying distances. Low agricultural transport costs cause a strong HME at short distances whereas high agricultural transport costs lead to the rise of a strong HME but at long distances. Moreover, a weak HME at long distances and the absence of the HME are caused by the interplay of high and low agricultural transportation costs. These last results confirm Davis (1998)'s findings, where a relatively higher wage in the larger market caused by positive agricultural transport costs weakens the HME and ultimately may nullify it.

These varying locational equilibria in agriculture-related industries are obtained because of the removal of two theoretical assumptions from our empirical approach. Firstly, we eliminated the assumption of homogeneous agriculture. We include 15 varieties of agricultural goods and 4 types of livestock into our analysis in order to capture the fact that different regions produce diverse crops according to their local particularities of land and natural endowments. For instance, while the north coast of the country has specialised in the production of rice and cotton, the jungle region has concentrated on the cultivation of coffee and cocoa. These geographical peculiarities also are representative of manufacturing firms in these regions, which define to some extent the exogenous degree of specialisation patterns in agriculture-related industries. As Fujita et. al. (1999) observe, the assumption of homogeneity of agricultural production is the most natural and simplest assumption to impose but is empirically indefensible. We give heed to this statement and include different varieties of farming products into our empirical strategy.

Second, we removed the assumption of costless agricultural transportation. In the real world, agricultural goods are costly to transport, as the same as manufactured goods. As showed in Table 2 agriculture-related industries incur significant positive transport costs. The presence of positive transport costs for the agricultural goods allows us to capture two real-world effects. First, lower prices of agricultural goods at the production region and higher prices of them at the consumer markets. Second, higher firm's production costs for agriculture-related products, which in turn increase the price of manufactured goods and their export prices in international markets. Both effects enable obtaining wage differentials among regions, which determine different kinds of locational equilibria. Undoubtedly, these transmission channels are particularly relevant in developing countries with a high proportion of their population involved in agricultural activities. As Fujita et al., (1999) note, agricultural transport costs operate as a brake on urbanisation process. Our results support this hypothesis and also defend the idea that agricultural transport costs are crucial in determining the spatial configuration of economic activity in developing countries.

5. Conclusions

In the Economic Geography literature, most theoretical studies have approached agglomeration of the industrial sector by assuming free transport of the agricultural good. This assumption however involves the theoretical shortcoming of equal wages in large and small markets within a country, which makes it not possible the examination of the HME. Consequently, the study of locational equilibrium when the transport costs of the agricultural good are positive has been overshadowed in significance by Economic Geography models. In fact, this theoretical flaw has profound implications when analysing the economic geography of developing economies with a high proportion of their population involved in agricultural activities. This theoretical aspect of the HME has not been empirically addressed until now.

In this paper, we examine the impact of agricultural transport costs upon the firms' incentives to disperse or to agglomerate in agriculture-related industries in Peru. By doing so, we employ a combination of nonparametric and parametric methods. First, using a nonparametric procedure developed by Duranton and Overman (2005), we find that agriculture-related industries do not exhibit strong localisation patterns. Only 1 out of 13 industries in our sample is highly localised at short distances whereas the remaining industries are dispersed or are localised at long distances with a significantly weaker degree of intensity.

We then explore cause-effect linkages between agricultural transport costs and industrial location using parametric conditional logit models. We find that the probability that a rural micro-geographic area has a firm highly localisation at short (long) distances is higher if it has low (high) agricultural transport costs. Likewise, weak localisation at long distances and dispersion patterns are explained by the interplay of high and low transportation costs in the agricultural sector. Therefore, both high and low levels of agricultural transport costs operate as a dispersion forces that weakens the intensity of the HME, and ultimately cause dispersion of economic activities in agriculture-related industries. These results argue in favour of the Economic Geography literature that supports the positive relationship between agricultural transport costs and dispersion of economic activities.

References

Ben-Akiva, M. and S. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA: MIT Press.

Chun, D.H. and Lee, K.S. (1985). “Changing Location Patterns in Population and Employment in the Seoul Region”. World Bank Report No. UDD-65.

Crozet, M., and F. Trionfetti (2008) “Trade costs and the home market effect”. Journal of International Economics 76, 309–21

Davis, D.R. (1998). “The home market, trade, and industrial structure”. American Economic Review 88, 1264–1276.

Dekle, R., J. Eaton and S. Kortum (2007) “Unbalanced Trade”. American Economic Review, Vol. 97, No. 2, pp. 351–355.

Duranton, G. and Overman, H. G. (2005). “Testing for Localization Using Microgeographic Data”. Review of Economic Studies, 72, 1077–1106.

Escobal, J. and Torero, M. (2005). “Measuring the impact of asset complementarities: the case of rural Peru”. Cuadernos de Economia 42. Pp. 1-26.

Fujita, M., Krugman, P.R. and Venables, A.J. (1999). The Spatial Economy: Cities, Regions and International Trade. MIT Press, Cambridge (Mass.).

Gonzales, E.V. (2010). “Decentralization, divergence and regional development in Peru”. In Tello, M.D. and Rodriguez, J. (Eds.) Options of Economic Policy in Peru 2011-2015 (pp.175-204). Lima: Publishing Fund of the Pontifical Catholic University of Peru.

Hanson, G.H. (1996). “Agglomeration, dispersion, and the pioneer firm”. Journal of Urban Economics 39, 255– 281.

Helpman, E., and P. Krugman (1985) Market Structure and Foreign Trade (Cambridge, MA: MIT Press)

Henderson, J.V.; Shalizi, Z. and Venables, A.J. (2002). “Geography and Development”. World Bank, Working Papers on Infrastructure, Telecoms, Power, Water, No. 2456.

Herrera, P. (2009). “Analyzing the business climate at the municipal level”. Journal of the Competence and the Intellectual Property, INDECOPI - Year 4, Number 8, spring: 115-141

IFPRI (2009). “Constructing a typology of rural communities of the Peruvian highlands using stochastic profit frontier estimation”. International Food Policy Research Institute, Washington.

Klier, T. and McMillen, D. P. (2008). “Evolving Agglomeration in the U.S. Auto Supplier Industry”. Journal of Regional Science 48 (1): 245–267.

McFadden, D. (1974). “Conditional logit analysis of qualitative choice behaviour”. In Frontiers in Econometrics , ed. P. Zerembka, 105–142. New York: Academic Press.

Picard, P. M. and Zeng, D.-Z. (2005) “Agricultural sector and industrial agglomeration”. Journal of Development Economics, 77, 75_106.

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall.

Takatsuka, T. and Zeng, D.-Z. (2012). “Trade Liberalization and Welfare: Differentiated-good vs. Homogeneous-good Markets”. Journal of the Japanese and International Economics, Vol. 26, No. 3, 308-325.

Torres, J. (2003). “Industrial clusters in Peru”. Department of Economics, Pontifical Catholic University of Peru. Working Paper N° 228. Lima.

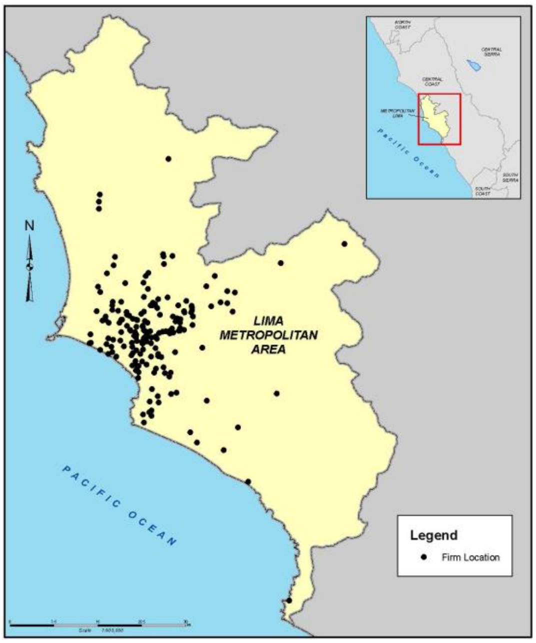
Webb, R. (2012) “Poverty and population dispersion”. Final Report PM01-2008. Economic and Social Research Consortium. Lima.

Yu, Z. (2005) “Trade, Market Size, and Industrial Structure: Revisiting the Home Market Effect”. Canadian Journal of Economics, February 38(1), 255-72.

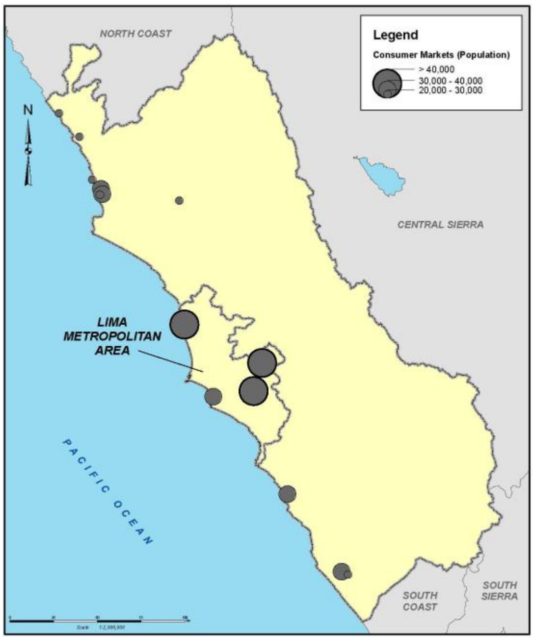
Zeng, D.-Z. and Kikuchi, T. (2005). “The Home Market Effect and the Agricultural Sector”. ERSA conference papers ersa05p135, European Regional Science Association.

Appendix A. Localised industry: Preparation and Spinning of Textile Fibres; Weaving of Textiles (SIC 1711)

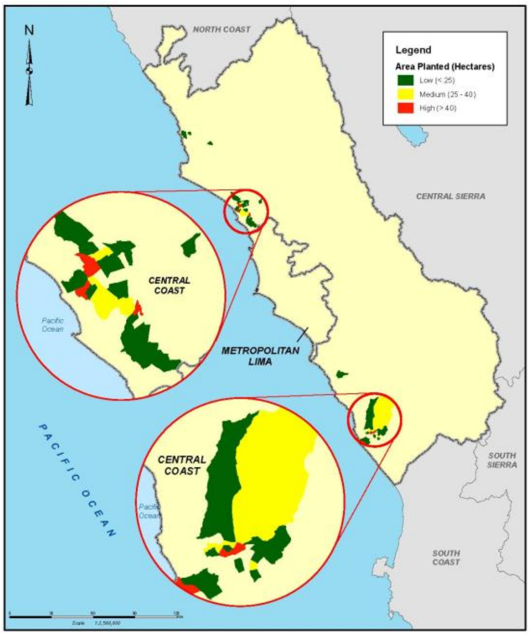
Zoomed area in Figure 1(a): Lima-Region



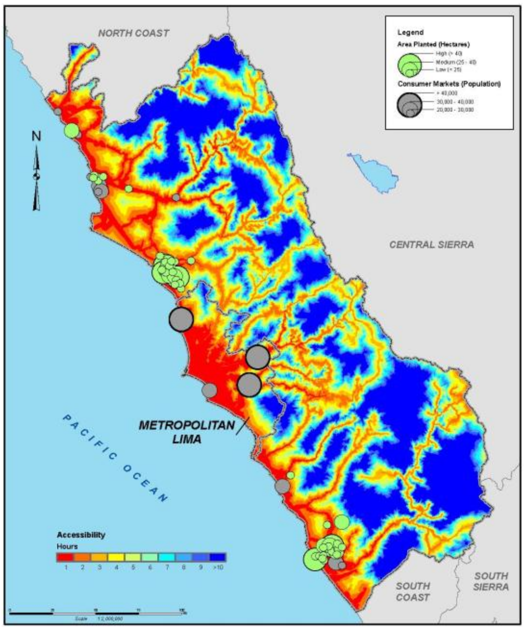
(a) Geographical distribution of firms



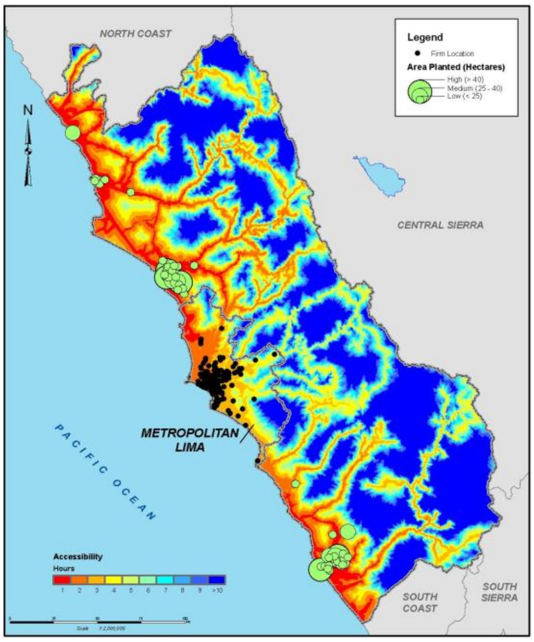
(b) Regional and local consumer markets



(c) Agricultural production of cotton



(d) Agricultural transport costs: Component I

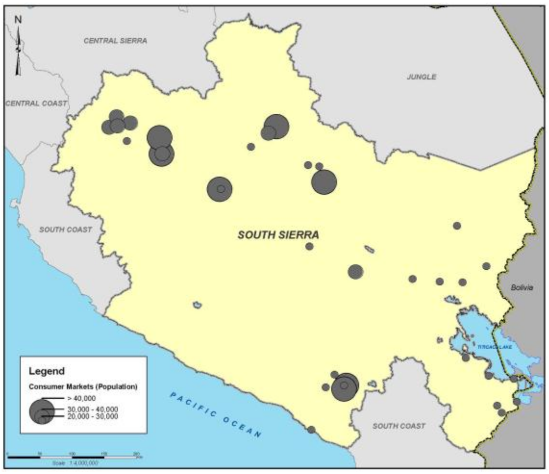


(e) Agricultural transport costs: Component II

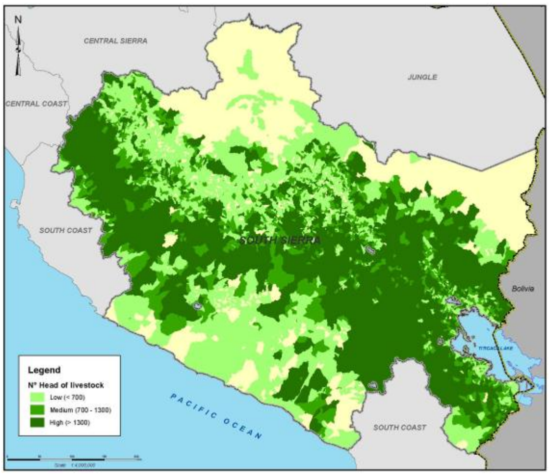
Appendix B. Dispersed industry: Tanning and Dressing of Leather (SIC 1911)

Zoomed area in Figure 1(b): South Sierra

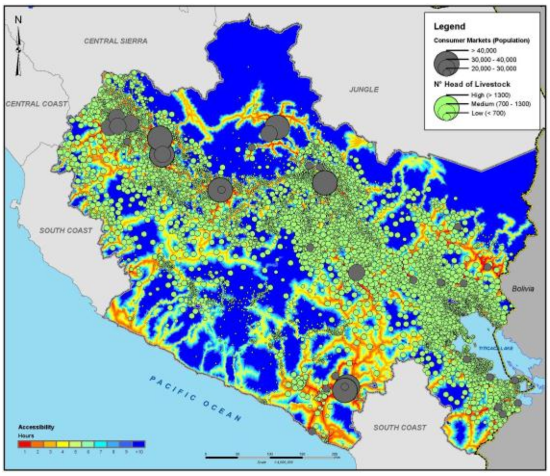
(a) Geographical distribution of firms



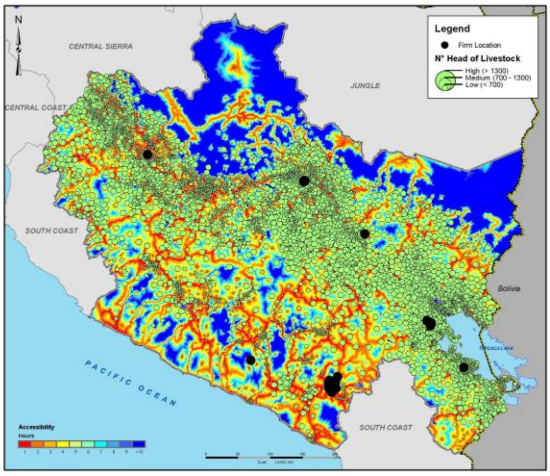
(b) Regional and local consumer markets



(c) Head of livestock: Beef and ovine cattle



(d) Agricultural transport costs: Component I



(e) Agricultural transport costs: Component II